Journal of Memory and Language 60 (2009) 487–501

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

The effect of experience on the perception and representation of dialect variants

Meghan Sumner [a,*], Arthur G. Samuel [b]

[a] *Department of Linguistics, Margaret Jacks Hall, Bldg. 460, Stanford, CA 94305-2150, United States*
[b] *Department of Psychology, Stony Brook University, Stony Brook, NY, United States*

ARTICLE INFO

Article history:
Received 1 February 2008
Revision received 15 December 2008
Available online 23 February 2009

Keywords:
Dialect
Speech perception
Spoken word recognition
Variation
Experience
r-Dropping

ABSTRACT

The task of recognizing spoken words is notoriously difficult. Once dialectal variation is considered, the difficulty of this task increases. When living in a new dialect region, however, processing difficulties associated with dialectal variation dissipate over time. Through a series of primed lexical decision tasks (form priming, semantic priming, and long-term repetition priming), we examine the general issue of dialectal variation in spoken word recognition, while investigating the role of experience in perception and representation. The main questions we address are: (1) how are cross-dialect variants recognized and stored, and (2) how are these variants accommodated by listeners with different levels of exposure to the dialect? Three claims are made based on the results: (1) dialect production is not always representative of dialect perception and representation, (2) experience strongly affects a listener's ability to recognize and represent spoken words, and (3) there is a general benefit for variants that are not regionally-marked.

A dialect is defined by the Oxford English Dictionary as "One of the varieties of a language arising from local peculiarities of vocabulary, pronunciation, and idiom". As this definition suggests, a dialect is generally defined in terms of production. For a speaker to *have* a dialect, then, typically means that a speaker speaks a certain way (syntactically, lexically, phonologically, etc.). While such a definition of a dialect does well to describe the output of a speaker's dialect, we suggest that it does not fully specify what it means to *have* a dialect. In fact, many important questions remain. For example, how do speakers of a given dialect perceive standard and nonstandard dialectal variants? How do they store this information? Are dialect-based phonological variants treated as variants of a single lexical item, or are they stored separately, perhaps as cognates of two languages might be stored? And finally, what is the role of experience in the development of dialect perception and representation?

The examination of dialectal variation from a spoken word recognition standpoint has occurred relatively recently. The large majority of research on dialect variation has instead focused on the description of dialects, attitudes towards dialects, and the perception of vowel mergers across dialects. Research on the description of dialects and the mapping of regional dialects and their characteristics across the United States has been conducted for a number of years (Kurath, 1939), most recently culminating with the remarkable *Atlas of North American English* (Labov, Ash, & Boberg, 2006). Labov and colleagues analyzed vowel formants from over 600 talkers across the United States and Canada and used these measurements to map gross dialect boundaries, as well as sub-dialects within a dialect region. Research on the social status of dialects and attitudes towards dialects is also well-established. For example, Labov (1972) examined the use of [ɚ] versus [ə] (e.g., at the end of the word "baker") in New York City department stores. The [ə] form is traditionally analyzed as resulting from r-dropping, a variation found in a number of dialects, including the New York City dialect. Labov found a correlation between r-dropping in employees and department

* Corresponding author. Fax: +1 650 723 5666.
E-mail addresses: sumner@stanford.edu (M. Sumner), asamuel@ms.cc.sunysb.edu (A.G. Samuel).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2009.01.001

store type, with fewer employees exhibiting r-dropping in more upscale stores. There have also been a number of studies in which participants were asked to listen to recordings of speech and rate the talker on social characteristics such as status (Giles, 1970), or to make attitude judgments such as pleasantness (Preston, 1989).

In addition to these studies, a number of recent papers have examined dialect identification and categorization. For example, Clopper and Pisoni (2004) played listeners sentences from talkers from six different regions of the United States and had listeners categorize talkers by region. They found that listeners formed three broad dialect categories, and that the perceptual similarity between talker and listener dialect played a role in this categorization. Clopper and Pisoni (2007) conducted a follow-up study in which dialect regions were not pre-labeled for listeners, so listeners were free to categorize the talkers into as many unlabeled groups as they wished. Overall, listeners made more groups of talkers than predicted, supporting the notion that listeners make fine-grained distinctions between different dialects of American English. In addition, the results suggest that listeners build perceptual categories for regional dialects using social and phonological information.

The collection of spoken dialect data, along with the current interest in vowel mergers (e.g., the *pin* – *pen* merger in the US) has also led to a number of studies on the perception of merged vowels by listeners of merged and unmerged dialects. Generally, research in this area has shown that speakers of merged dialects (e.g., where *pin* and *pen* have the same pronunciation) are less able to make perceptual discriminations between the merged vowels than speakers of unmerged dialects (Bowie, 2000; Evans & Iverson, 2004; Janson & Schulman, 1983; Labov, Karan, & Miller, 1991). Some of these studies have examined the effect of dialect contact on perception. Janson and Schulman, for example, examined the perception of merged vowels in Swedish by listeners exposed to two different Swedish dialects. One dialect had four vowels, and one dialect had three vowels (as the result of a merger). Their results were consistent with the literature in that listeners from the merged dialect were unable to discriminate between merged vowels. They also found that while most of the four-vowel listeners could make a four-way discrimination, others could not. They attributed this result to exposure to merged dialects.

Bowie (2000) examined the result of dialect exposure in more detail. Bowie examined the perception of the vowels /u/ and /ʊ/, which are merged before [l] in a dialect found in Maryland. He examined two groups of listeners, natives who were lifelong residents of the town, and *exiles* who were born and raised in the town, but who lived in other regions for a part of adulthood before returning to the town. In production, all natives and exiles (but one) maintained a merged vowel in this context, but perceptually, exiles were better at discriminating the merged vowels than natives. Bowie argued that exposure to non-merged dialects resulted in the development of the discrimination ability.

While these studies have provided great detail about the discrimination ability of listeners and perceptual biases related to experience, we know little about the effect of dialectal variation on spoken language processing. For example, within-dialect variation has garnered a reasonable amount of attention lately. Researchers have examined the effects of within-dialect variation such as tapping (e.g., McLennan, Luce, & Charles-Luce, 2003; Luce, McLennan, & Charles-Luce, 2003), stop release (Deelman & Connine, 2001), schwa deletion (LoCasto & Connine, 2002), and assimilation (Gaskell & Marslen-Wilson, 1996; Gow 2001, 2003; Mitterer & Blomert, 2003). Much of this research has been aimed at understanding how listeners treat words with multiple phonetic variants in spoken word recognition and comprehension. How listeners ultimately adapt to these variants has been a difficult question to answer, though, because experience with variants is difficult to control.

There is some evidence that comprehension difficulties decrease as familiarity with a speaker of a different dialect or native language increases. In fact, Scott and Cutler (1984) have shown this to be the case for British English native speakers processing American English medial-/t/ in forms like *total*, which is produced as a tap. Scott and Cutler tested two groups of British English listeners: those who had lived in England throughout their entire lives, and those who had moved to the United States. They found that British listeners living in the United States had less difficulty processing medially-tapped /t/s (as in "total") than British listeners with little experience with General American. They attributed this result to the fact that British listeners in the US make an assumption that their interlocutor intends to produce a real word and not a nonword (e.g., "total", not a new word "todal"). More recently, Floccia, Goslin, Girard, and Konopczynski (2006) examined processing costs associated with regional accent normalization. They examined the time course of disruption during the comprehension of targets from different French regional accents, and showed that there are initial and temporary costs associated with the comprehension of an unfamiliar accent.

Although research of this sort clearly shows that familiarity with a dialect improves processing of that dialect, there is little known about the mechanisms providing this improvement. Do listeners develop multiple representations? Do listeners become better at mapping a new sound onto an existing one? Does familiarity affect phonological representations? If multiple representations are in fact involved, additional issues must be clarified. For example, how do native dialect speakers differ from listeners who can process the dialect variants effectively, but still maintain their own dialect in production? There is growing evidence supporting both abstract and specific representations (e.g., Luce et al., 2003; McLennan et al., 2003). It remains to be seen what role these two types of representations play in cross-dialect variant processing for listeners who are (or are not) familiar with a dialect.

In the current study, we examine the processing and representation of dialect variants and the effect that prior experience with a dialect has on spoken word recognition. Specifically, we examine the processing of –er final words

[1] The term General American (GA) is used here and throughout to represent speakers who do not r-drop or exhibit any other regionally marked characteristics.

(e.g., *baker*) produced by speakers of General American[1] (GA) with a full –er (bak[ɚ]) and by speakers of the New York City dialect (NYC) who productively produce r-less variants of these forms (e.g., bak[ə]). Traditionally, both groups are analyzed as having an underlying r-ful representation, but the NYC speakers are considered to be r-droppers, as they produce r-less variants in predictable positions, as well as intrusive [*r*]s (Labov, 1972; but see Gick, 1999 for a gesture-based account of intrusive /r/). To reflect this difference at the surface/output level, we simply refer to these as r-ful or r-less variants throughout the paper, but our use of these terms should not be taken as support for the traditional view that these two subject populations with different surface variants store a single, r-ful representation.

As described in the *Preliminaries* section below, we present these forms to three different groups of listeners, who differ in their experience with the NYC dialect.

The purpose of this research is threefold. First, we are interested in how effective cross-dialect variants are at activating lexical items intended by the speakers and, more generally, how dialect variants are processed in the short-term. Second, we address issues of encoding and representation of dialect variants by speakers of different dialects by testing with a long-term priming paradigm. Finally, by controlling for the language background and dialect of our participants, we examine effects of experience with a dialect on dialect processing and representation. We are interested in resolving whether dialect variants are processed as variants of a single abstract form, or whether the dialect variants are stored as individual representations in a speaker's phonology; critically, we test whether this is dependent on language background. Although there has been progress in understanding the perception of dialects and the categorization of sounds within a dialect, relatively little is known about how listeners process and represent different dialects. Anecdotal evidence, together with studies by Floccia et al. (2006), tells us two things: first, dialectal variation can cause processing difficulties, and second, over time, these processing difficulties lessen. The current project examines the processing of dialects while examining in detail the effect experience (or lack of experience) has on spoken word recognition and representation. Ultimately, we will show that there are three distinct aspects in which a listener may have a dialect: (1) in production, (2) in representation, and (3) in perception.

Preliminaries

The central goal of this project was to understand how speakers of one dialect (NYC) process and represent forms within their own dialect, and across dialects (GA). To examine this issue, we must have knowledge of the language background of each participant. To avoid any response biases on our experimental measures, we gathered dialect information *after* the participants finished the experimental tasks. Each participant was given a post-test questionnaire and a short exit interview to assess which participants were NYC or GA speakers, and the extent to which the participant had prior experience with the other dialect.

The three groups we identified via the preliminary measures were: (1) Overt-NYC, (2) Covert-NYC, and (3) GA participants. Both Overt-NYC and Covert-NYC participants were born and raised in the New York City area. None had ever lived outside of the New York City region. Members of these two groups, although similar in residential background, differed in their production of –er final words. During the exit interviews the Overt-NYC participants productively exhibited r-less variants in viable contexts (see Labov, 1972; e.g., *dark* = [dɔək] ≠ [dɔɹk], *sister* = [sɪstə] ≠ [sɪstɚ]), whereas the Covert-NYC participants always produced r-ful variants of similar items (e.g., *dark* = [dɔɹk], *sister* = [sɪstɚ]). Thus, individuals who exhibited no r-dropping in the exit interview were put into the Covert-NYC group. Participants who produced r-less variants were rated for productivity on a four-point scale. Participants who showed no evidence of r-ful variants in viable contexts, or produced a single word with even a hint of rhoticization were assigned a rating of four. Participants rated as a three primarily exhibited productive r-less speech, but had two – five vowels which may have been rhoticized. Only those individuals with a score of three or four were placed into the Overt-NYC group; participants with ratings of one or two had more mixed productions and were not included in the study.[2] The inclusion of only those participants with productive and obvious r-less speech served to keep the two groups as clean and as distinct as possible.[3] Additional evidence gathered from the exit questionnaires further supports the separation of these subjects into two distinct groups.

Coupled with the presence or absence of r-ful variants in the speech of these two groups was a critical difference in family history. The Overt-NYC participants were third generation New York City area residents. The grandparents and parents of 46 out of 48 participants (95.8%) in this group were born and raised in the metropolitan New York City area (the parents of the remaining two participants were also born on Long Island, but no information was provided about their grandparents). In contrast, nearly all Covert-NYC participants (45/48, 93.4%) were first- or second-generation NYC residents born and raised in Long Island to parents from other states (34) or countries (11).

The labels given to the two NYC dialect groups are intended to represent two critical aspects of the dialect situation: they were all born and raised in a region with an r-less dialect (NYC) and they either do (Overt) or do not (Covert) engage in r-less productions themselves. Both NYC groups can be contrasted with the third group (GA). The GA group includes participants born and raised outside of the New York City region, who had never lived in a region with an r-less dialect, and who were first-semester Stony Brook students (Fall or Summer) at the time of test-

[2] These participants tended to be those who had lived outside the New York City area, and it is not surprising, then, that they have a more variable production of rhotic and non-rhotic vowels.

[3] We do not equate the production of a few rhoticized vowels with r-ful speech when the overwhelming majority of the speech was r-less. Even an American English speaker may not tap all [t]s intervocalically, or a GA speaker may produce an r-less variant here or there.

ing. All participants in all groups were monolingual native-English speakers.

As noted previously, the categorization of participants was based on a questionnaire and an exit interview given to each participant at the *completion* of each experiment. The questionnaire was used to assess previous experience participants had with speakers of r-less dialects and included questions about birthplace, parents' and grandparents' birthplace, awareness of r-less in speech of self, family, and friends (see Appendix A). The exit interview was used to assess actual r-ful/r-less productions in the speech of our participants. The same experimenter (an experienced phonetician) conducted the exit interview with every subject.

The only way to effectively investigate the effects of experience on processing and representation of dialectal information is to control for language background of the participants. We therefore recruited participants until the desired numbers for each identified participant group were reached. Because group assignment occurred after the experiment, a number of participants (51) were run who did not fit into any of the three groups, and were therefore not included in analyses. Reasons for exclusion included: bilingual speakers (21), incomplete questionnaires (6), participants who have lived in multiple locations (11), GA speakers who have previously lived in the New York City area or another r-dropping region (6), and GA speakers who were not first-semester students (6). One additional participant who rejected r-less speech during the experiment, questionnaire, and interview was also excluded from the analysis (this participant complained that he could not do the task since some of the speakers "could not speak English"; he also had high error rates (above 75%) for *all* targets in all real word conditions, suggesting that "Pseudoword" was the response of choice). In all cases, participants were excluded from the analyses without experimenter knowledge of their performance on the experimental tasks.

These three tightly controlled participant groups enabled us to examine the role of experience on the perception, recognition, and representation of familiar and unfamiliar dialects. Using these groups, we designed a set of experiments to determine how the processing of different dialectal variants depends on participant dialect. To examine the surface nature of the variants and how listeners handle within- and out-of-dialect variants, we use an auditory form priming task. Form priming is sensitive both to repetition of identical items and to similar items, compared to controls (Radeau, Morais, & Segui, 1995), and is therefore a useful task to examine whether multiple variants of a single lexical item are treated as identical or not by a particular subject population.

In some word recognition models, there is a distinction between the phonological (or orthographic) access codes for a lexical item, and the semantic information associated with that lexical entry. In our effort to clarify exactly what it means to have a dialect, we are also interested in how well different dialect variants activate semantic information in the lexicon. It is well known that a particular lexical item is retrieved more easily following the presentation of a semantically-related item than following an unrelated item (Radeau, 1983). In addition to the immediate form priming task, then, we also use an auditory semantic priming task to examine whether different variants (r-ful or r-less) are equally able to facilitate activation of a semantically-related target depending on participant dialect.

These two priming paradigms provide a nice contrast of surface and lexical effects of dialect variation within a paradigm that traditionally provides robust effects. With these tasks, we can probe different types of processing, and in so doing, gain new insight into the costs or benefits of dialect familiarity. While these two tasks provide information about a listener's ability to accommodate different variants immediately, they do not necessarily provide information regarding the long-term phonological representations for each subject population. To probe potential differences in representation for participants, we use a long-term priming task. This task has been used to examine the specificity of phonological representations (Luce et al., 2003; McLennan et al., 2003; Sumner & Samuel, 2005). By combining this task with the immediate priming tasks, we can directly compare immediate and long-term form effects as a function of a participant's experience with a dialect. Given the accumulating evidence that listeners maintain very rich and detailed representations (e.g., Goldinger, 1996), it would be very useful to know whether these representations vary depending on listener experience with a variant. Collectively, the three experiments can provide a rich picture of the effect of dialect experience on perception, recognition, and representation – spoken word processing. To foreshadow the results, the priming effects across the three experiments demonstrate that a dialect cannot be defined by production alone.

Experiment 1: form priming

One fundamental comparison to make across phonologically-variant dialectal forms is how well the surface features of each form are processed by speakers familiar versus unfamiliar with both dialects. The form priming paradigm has often been used to examine effects of surface features on immediate processing. In this paradigm, listeners are presented with a prime followed by a phonologically-related target (Radeau et al., 1995). A target is generally responded to more quickly when preceded by an identical prime than when preceded by a phonologically unrelated prime. What this paradigm allows us to see is whether phonologically deviating from a dialect variant has an effect on processing that variant.

Methods

Participants

Forty-eight Stony Brook University students participated for course credit or for pay. Sixteen participants from each participant group (Overt-NYC, Covert-NYC, GA) were run in this experiment. No participant reported any hearing problems.

Stimuli and design

The stimuli were recorded by four speakers in a sound-attenuated booth. Two speakers of General American (GA) and two native speakers from the New York City area

(NYC) with obvious r-less accents were recorded. One male and one female from each dialect were recorded. Throughout the experiments, primes were always produced by female speakers, and targets were always produced by male speakers. Therefore, even when primes and targets matched in phonological composition, the word tokens were physically different. All speakers were naïve with respect to the purpose of the experiment.

The stimuli for this experiment included 320 two- and three-syllable English words ending with –er (e.g., baker, soccer) (See Appendix B for a representative sample of stimuli). The primes and targets were chosen based on the results of two large-scale mass-testing sessions (roughly 500 participants per session). The selection process was designed to identify stimuli that could be used both in the current experiment and in a semantic priming task in Experiment 2. A total of 500 –er-final probes were given to participants. Each participant filled out a questionnaire with 100 –er-final (printed) words and was asked to write down the first word that came to mind for each item. Each item received approximately 200 responses. The 320 items with the highest number of shared responses were used as primes in this experiment, and as critical items in Experiments 2 and 3. The average of shared responses across items was 53% with a minimum requirement of 40% shared responses in order for a target to be used in the experiment. All four speakers recorded all items.

Experiment 1 was designed to test whether r-less forms are as effective as r-ful forms in priming an identical word. Four conditions were used to examine this issue: (1) GA prime – GA target (e.g. baker – baker); (2) NYC prime – GA target (e.g., bakə – baker); (3) GA prime – NYC target (e.g., baker – bakə); (4) NYC prime – NYC target (e.g., bakə – bakə). For each critical prime-target pair, we also constructed an unrelated control prime-target pair, which provided a within-item baseline for any priming effects. The control primes were selected by randomly re-pairing prime-target pairs, and as such, would generally have no association at all with their targets. Examples of stimuli in the four conditions, with unrelated controls, are provided in Table 1.

The control primes ending in –er (e.g., *filter*) were used to eliminate any repetition effect based solely on the sharing of the final rime of each word. For all conditions, the tokens produced by the female speakers were used as the primes, and those produced by the male speakers were used as the targets. As noted above, four speakers were used to ensure that any effects in the within-dialect identity conditions were free of voice repetition effects, i.e., the prime and the target were always produced by different speakers.

Sixteen counterbalanced lists were created so that each item served as a critical target for an identical prime (baker – baker) and a control target for an unrelated prime (filter – baker). Each list included 80 critical pair combinations (GA – GA, NYC – GA, GA – NYC, NYC – NYC) and 80 unrelated control pairs. No prime or target item (or variant) was repeated for any participant. In addition to the 160 critical and control pairs, 640 filler pairs were included, none of which ended in –er; speakers were balanced in the control pairs, as in the critical pairs. Half of the filler pairs had real word targets and half had pseudoword targets. For the 320 fillers with real word targets, 80 had pseudoword primes, 80 had identical primes, and 160 had unrelated real word primes. For the 320 fillers with pseudoword targets, 80 had real word primes, 80 had identical pseudoword primes, and 160 had unrelated pseudoword primes. The fillers were used to avoid biasing participants toward processing r-less variants as either words or pseudowords.

Table 1
Sample stimuli for Experiment 1.

Conditions	Sample prime	Sample target
GA prime		
Related	'baker' [beɪkɚ]	
Unrelated	'filter' [fɪltɚ]	
NYC prime		*GA Target:* [beɪkɚ]
Related	'baker' [beɪkə]	
Unrelated	'filter' [fɪltə]	
GA prime		
Related	'baker' [beɪkɚ]	
Unrelated	'filter' [fɪltɚ]	*NYC Target:* [beɪkə]
NYC prime		
Related	'baker' [beɪkə]	
Unrelated	'filter' [fɪltə]	

Note. The IPA symbol [ɚ] represents the final –er sound in the GA forms, and the symbol [ə] represents the –uh-like sound in the NYC forms. An asterisk indicates a within-dialect identity condition (same item, same dialect, different voice).

The form priming paradigm typically produces strong priming for identical items and moderate priming for similar items compared to unrelated prime-target pairs (see Radeau et al., 1995 for overview). In this experiment, we would expect to find full priming in the GA – GA condition independent of the background of the participants. Whether full priming occurs for the NYC – NYC identity condition, however, may differ across participant groups if the variants are handled differently depending on prior experience with r-dropping. We would expect the GA speakers to have more difficulty than the other two participant groups in processing NYC targets if experience with a variant is critical to accommodating multiple surface variants during perceptual processing. Any perceptual differences between Overt- and Covert-NYC participants would demonstrate the impact of the parental environment and/or consequences of the production difference itself.

Procedure

Participants were tested in a sound-attenuated booth individually or in groups of two or three. Participants were presented with pairs of utterances over headphones and were asked to make a lexical decision to the second item in each pair. Participants were instructed to respond as quickly and accurately as possible. For each trial, an auditory prime was presented first, followed by a 500ms ISI, followed by an auditory target. If participants did not respond within three seconds, a new pair was presented. A new trial began one second after the responses had been made. As discussed in the *Preliminaries*, all participants filled out a questionnaire and were given a short exit interview following the form priming task.

Results

Reaction times more than 2.5 standard deviations from the mean were discarded (<2%). We found significant effects on error rates for critical items, so participant exclusion due to high error rates was based on responses to non-critical items. Two participants with error rates greater than 10% were replaced by new participants from the same participant population. Mean reaction times for the eight experimental conditions and for the three participant groups are provided in Table 2. A corresponding breakdown of the error rates is given in Table 3. Priming effects, based on the reaction times, are illustrated in Fig. 1.

Reaction times

A three-factor ANOVA (participant group × priming condition × related/unrelated) on the reaction times revealed a main effect of related items ($F1(1,45) = 22.758$, $p < .01$; $F2(1,79) = 11.861$, $p < .01$; $MinF'(1,124) = 7.797$, $p < .01$) and a main effect of condition ($F1(3,135) = 7.823$, $p < .01$; $F2(3,237) = 9.874$, $p < .01$, $MinF'(3,319) = 4.385$, $p < .01$). Additionally, an interaction between participant group and condition was found ($F1(6,135) = 3.832$, $p < .01$, $F2(6,474) = 2.674$, $p < .05$; $MinF'(6,504) = 0.157$, $p = .152$). The main effect of participant group was marginal ($F(2,45) = 2.786$, $p = .079$, $F2(2,158) = 2.484$, $p = .087$; $MinF'(2,149) = 1.313$, $p = .272$) and all other results were insignificant. From the results, we can see that related items are, in general, recognized more quickly than unrelated items. The main effect of condition appears to be due to the lack of facilitation for conditions with NYC targets by the GA participants (i.e., the significant participant by condition interaction mentioned above).

The pattern of responses reveals a striking asymmetry between those participants with prior exposure to the r-less dialect and the GA speakers, who have had little experience with the variant. This asymmetry is reflected in the interaction between participant group and priming condition and is illustrated by the difference scores shown in Fig. 1.

Table 2
Mean reaction times for correct responses to targets in Experiment 1.

	Participant dialect		
	Overt-NYC	Covert-NYC	GA
GA prime	GA target		
Related[a] (beɪkɚ – beɪkɚ)	917	909	911
Unrelated (fɪltɚ –beɪkɚ)	975	977	973
NYC prime			
Related (beɪkə – beɪkɚ)	938	944	939
Unrelated (fɪltə – beɪkɚ)	977	974	972
GA prime	NYC target		
Related (beɪkɚ – beɪkə)	954	959	994
Unrelated (fɪltɚ – beɪkə)	989	992	996
NYC prime			
Related[a] (beɪkə – beɪkə)	930	929	989
Unrelated (fɪltə – beɪkə)	991	994	993

[a] Denotes an identity condition (same dialect, same word, though different speakers).

Table 3
Mean error rates for targets in Experiment 1.

	Participant dialect		
	Overt-NYC	Covert-NYC	GA
GA prime	GA target		
Related[a] (beɪkɚ – beɪkɚ)	2.5	2.6	2.5
Unrelated (fɪltɚ – beɪkɚ)	3.1	1.9	3.1
Difference	+.6	−.7	+.6
NYC prime			
Related (beɪkə – beɪkɚ)	2.6	2.4	2.5
Unrelated (fɪltə – beɪkɚ)	2.7	2.3	2.2
Difference	+.1	−.1	−.3
GA prime	NYC target		
Related (beɪkɚ – beɪkə)	3.1	3.4	11.9
Unrelated (fɪltɚ – beɪkə)	9.1	10.3	21.6
Difference	+ 6.0	+6.9	+9.7
NYC prime			
Related[a] (beɪkə – beɪkə)	9.9	10.4	21.7
Unrelated (fɪltə – beɪkə)	9.5	10.1	22.0
Difference	−.4	−.3	+.3

[a] Denotes an identity condition (same dialect, same word, though different speakers).

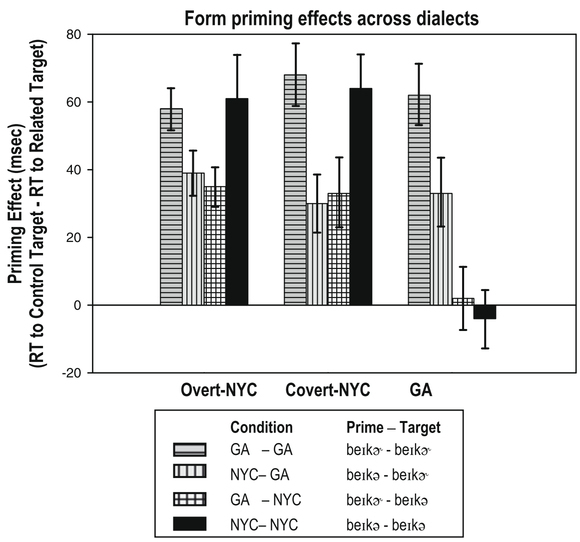

Fig. 1. Priming effects for the three participant groups across the four experimental conditions. The error bars correspond to the 95% confidence interval for each condition mean.

The difference scores plotted in Fig. 1 illustrate two important patterns in the data: (1) the Overt-NYC and Covert-NYC participants exhibit similar priming effects across priming conditions that differ from those of the GA participants, and (2) priming effects are present for GA participants *only* when the target is a within-dialect item. This asymmetry is supported by pair-wise comparisons of difference scores showing that there is no difference in effect between Overt-NYC and Covert-NYC participants across conditions ($F1(1,30) < 1$; $F2(1,79) < 1$), but both Overt-NYC and Covert-NYC participants behave differently from GA participants (Overt-NYC – GA ($F1(1,30) = 11.67$, $p < .01$; $F2(1,79) = 8.553$, $p < .01$; $MinF'(1,98) = 4.936$, $p <$

.05); Covert-NYC – GA ($F1(1,30) = 15.782$, $p < .01$; $F2(1,79) = 4.898$, $p < .05$; $MinF'(1,108) = 3.738$, $p = .056$).

Errors

In this experiment participants were required to respond directly to both r-less forms and r-ful forms. A three-way ANOVA on error rates resulted in a number of interesting effects. Main effects were found for Group ($F1(2,45) = 8.616$, $p < .01$; $F2(1,79) = 6.354$, $p < .01$; $MinF'(2,164) = 3.657$, $p < .05$), and Condition ($F1(3,135) = 4.117$, $p < .05$; $F2(3,237) = 7.971$, $p < .01$; $MinF'(3,270) = 2.715$, $p < .05$). Interactions between Condition and related items ($F1(3,135) = 4.645$, $p < .05$; $F2(3,237) = 4.199$, $p < .01$; $MinF'(3,353) = 2.205$, $p = .087$), and between Group and Condition were found ($F1(6,135) = 4.641$, $p < .01$; $F2(6,474) = 3.792$, $p < .05$, $MinF'(6,468) = 2.087$, $p = .053$). No other results were significant. The main effect of group is expected, as GA participants had significantly higher error rates when responding to NYC targets than the other two participants groups ($F1(2,45) = 3.518$, $p < .05$, $F2(2,77) = 6.301$, $p < .01$; $MinF'(2,92) = 2.257$, $p = .110$). The main effect of condition reflects the higher overall error rates to NYC targets than to GA targets. The more interesting result is the interaction between related items and condition. This interaction was driven by the error rates in the GA – NYC condition. Pairs in this condition included an –er-final prime followed by an r-dropped target. When responding to a target following an unrelated prime, Overt-NYC and Covert-NYC participants made incorrect responses approximately 10% of the time. This number decreased significantly to approximately 3% when the target was preceded by a related GA prime (Overt: $F1(1,15) = 11.998$, $p < .01$; $F2\ (1,79) = 4.663$, $p < .05$; $MinF'(1,85) = 3.358$, $p = .07$; Covert: $F1(1,15) = 9.730$, $p < .01$; $F2\ (1,79) = 5.190$, $p < .05$; $MinF'(1,74) = 3.384$, $p = .07$). This is in direct contrast to the NYC – NYC condition where the error rates for both critical and control items hovers around 10%. The same benefit is found in GA participants (22% vs. 12%) ($F1(1,15) = 7.689$, $p < .05$; $F2\ (1,79) = 7.031$, $p < .01$; $MinF'(1,54) = 3.673$, $p = .061$). Thus, *for all listeners*, hearing a GA form improves the later processing of a variant form. Even for listeners raised in the NYC dialect, hearing the General American form of a word is not only easier, it makes recognition of a following variant easier, as well. This result is consistent with recent work that has found a benefit for a less frequently produced form that is supported by orthography (Ranbom & Connine, 2007; Sumner & Samuel, 2005).

Experiment 2: semantic priming

The results of Experiment 1 suggest that listeners who have experience with both dialects are more *flexible* in form processing; these listeners are flexible in that they do not have a rigid criterion for a particular form when it is a regular variant. As discussed previously, within-language regular variation has been shown not to disrupt processing. Many studies on assimilation, for example, have shown that listeners use subtle acoustic cues to distinguish potentially ambiguous words (Coenen, Zwitserlood, & Bölte, 2001; Gaskell & Marslen-Wilson, 1996; Gow, 2002, 2003). In Experiment 1, we suggest that the Overt and Covert listeners were more flexible because over time, the consistent exposure to multiple variants makes them perceptually equivalent. These listeners show greater perceptual adaptability, with both phonetic forms supporting the activation needed for priming to occur. Flexibility in processing multiple surface variants is apparently *not* a given; experience with a surface variation seems to be critical in enabling listeners to process multiple forms equally.

The goal of Experiment 2 is to examine the effect that different dialect variants have on lexical activation as a function of listener experience. Studies examining the effect of mismatching information on lexical access have shown that arbitrary variation from a canonical form slows or precludes lexical access (Connine, Blasko, & Titone, 1993; Gaskell & Marslen-Wilson, 1993; Marslen-Wilson, 1987; Soto-Faraco, Sebastián-Gallés, & Cutler, 2001; Sumner & Samuel, 2005; Warren & Marslen-Wilson, 1988). The results from Experiment 1 suggest that even for listeners raised in an r-less environment, the influence of the GA dialect of American English (potentially through literacy, education, and media, for example) is strong. Therefore, we use the canonical –er final primes and semantically-related targets as a baseline for all participants. We assess how listeners process the phonological variants by comparing the effectiveness of r-less primes in facilitating the processing of semantically-related targets to the canonical baseline priming. If experience with a particular variant is a factor, then we would expect to find differences among the participant groups.

Specifically, an r-less NYC word should not be an effective prime for a semantically-related target (e.g., slend[ə] – *thin*) unless the –er ∼ [ə] variation is a predictable and productive alternation in a listener's phonology. In this analysis, the GA participants should not show semantic priming in this situation because slend [ə] is not a legal variant for them. In contrast, the Overt-NYC group should show such an effect, as the schwa-variant is clearly a predictable and productive part of their phonology. The Covert-NYC group presents a very interesting case: these participants have *heard* the alternation predictably and productively since birth, but the alternation is not apparent in their own productions.

Method

Participants

Forty-eight participants were included in this study. As in Experiment 1, 16 individuals from each dialect group participated. None of the participants reported any hearing problems.

Stimuli and design

In this experiment, the 320 two- and three-syllable –er-final words from Experiment 1 were used as primes for semantically-related targets. Primes were always in a female voice in either the GA or NYC dialect, and semantically-related targets were always in a male voice.

Four critical conditions were formed: (1) GA prime – GA target, (2) NYC prime – GA target, (3) GA prime – NYC target, (4) NYC prime – NYC target. The critical pairs were split in half to form two lists of 160 critical pairs. From each list, eight separate counterbalanced lists were created to include 80 semantically-related targets and 80 control targets. The eight counterbalanced lists were created to ensure that no target or prime was repeated for any participant, and that each target (e.g., thin) was paired with a related prime (e.g., slender) in all critical conditions and with an unrelated prime from a different list in all control conditions. Examples are provided in Table 4. Only the primes in this experiment end in –er; the targets do not. Nonetheless, the same speaker combinations used in Experiment 1 (female-based primes, male-based targets) were used to control for any more subtle dialectal differences. Targets were matched in duration, with mean target durations across speakers differing by only 3.1 ms.

In addition to the 160 critical items and controls, 480 filler pairs were included to avoid the development of response strategies. Of the 480 filler pairs, 160 included real word targets and 320 included pseudoword targets. No final –er words or nonwords were used as fillers.

Procedure

Participants were tested in a sound-attenuated booth individually or in groups of two or three. Participants were presented with pairs of words over headphones and were asked to make a lexical decision to the second word in each pair. Participants were instructed to respond as quickly and accurately as possible. For each trial, an auditory prime was presented first, followed by a 500ms ISI, followed by an auditory target. If participants did not respond within three seconds, a new pair was presented. A new trial began one second after the responses had been made.

Results

Response times greater than 2.5 standard deviations from the mean reaction time were discarded (2.7% of all responses). Participants were excluded based on incorrect responses to non-critical stimuli. Three participants with mean error rates above 10% were replaced with participants from the same group. Table 5 provides the lexical decision reaction time means and error rates for this experiment.

A three-way ANOVA (participant group × condition × related/unrelated) revealed a main effect of semantic-relatedness ($F1(1,47) = 12.341$, $p = .01$, $F2(1,79) = 6.829$, $p < .05$; $MinF'(1,126) = 4.396$, $p < .05$). Additionally, the main effect of Group was significant ($F1(2,45) = 6.36$, $p < .01$; $F2(2,158) = 4.17$, $p < .01$; $MinF'(2,173) = 2.519$, $p = .084$), as was the main effect of Condition ($F1(3,135) = 4.823$, $p < .01$; $F2(3,237) = 2.996$; $MinF'(3,371) = 1.849$, $p = .138$). A significant interaction between Group and Condition was also found ($F1(6,135) = 3.859$, $p < .05$; $F2(6,474) = 3.135$, $p < .05$; $MinF'(6,469) = 1.73$, $p = .112$). No other interactions were significant. In addition, no significant effects of error rates were found, reflecting the low and consistent error rates in all conditions, likely due to the fact that targets were similar in production across dialects.

The priming effects presented in Fig. 2 are similar to those found in Experiment 1. Overt-NYC and Covert-NYC participants again behave in a unified manner and facilitation effects for GA participants are dependent on dialect. The reliable main effect of Group (which was marginal in Experiment 1), may reflect the strong priming across conditions for NYC participants compared to GA participants (i.e., the Condition by Group interaction is likely to be driving the main effect of Group). Both GA r-ful and NYC r-less forms facilitated responses to a semantically-related target (e.g., slend[ɚ] primes *thin* as does slend[ə]) for NYC participants. One condition, NYC Prime – GA target, showed a reduced but still significant priming effect of 38 ms ($F1(1,30) = 4.671$, $p < .05$; $F2(1,79) = 3.962$, $p = .05$; $MinF'(1,93) = 2.144$, $p = .146$). This reduced effect can likely be attributed to mismatching dialects and not to the ineffectiveness of an r-less prime since strong priming was found when an r-less prime preceded a target in the same dialect. A similar reduction was not found in the GA Prime – NYC Target condition, possibly because the strength of the GA prime alleviates the cost of mismatching prime-target dialects in this task.

Table 4
Sample stimuli for Experiment 2.

	Prime	Target
GA prime		*GA target*: thin
Related[a]	slender	
Unrelated[a]	shower	
NYC prime		
Related	slend[ə]	
Unrelated	show[ə]	
GA prime		*NYC target*: thin
Related	slender	
Unrelated	shower	
NYC prime		
Related[a]	slend[ə]	
Unrelated[a]	show[ə]	

[a] Indicates that prime and target are presented within-dialect (but across speaker).

Table 5
Mean reaction times for correct responses to targets in Experiment 2.

	Participant dialect		
	Overt-NYC	Covert-NYC	GA
GA prime	GA target		
Related[a]	913 (2.1)	913 (2.3)	909 (2.6)
Unrelated[a]	974 (2.3)	970 (2.6)	975 (2.3)
NYC prime			
Related	933 (2.4)	937 (2.1)	968 (2.4)
Unrelated	971 (2.0)	973 (2.3)	972 (2.2)
GA prime	NYC target		
Related	919 (2.8)	915 (2.2)	921 (2.3)
Unrelated	987 (2.5)	980 (2.4)	989 (2.0)
NYC prime			
Related[a]	920 (2.2)	916 (2.2)	988 (2.3)
Unrelated[a]	984 (2.1)	985 (2.4)	985 (2.6)

Note. Error rates are shown in parentheses.
[a] Indicates a within-dialect condition.

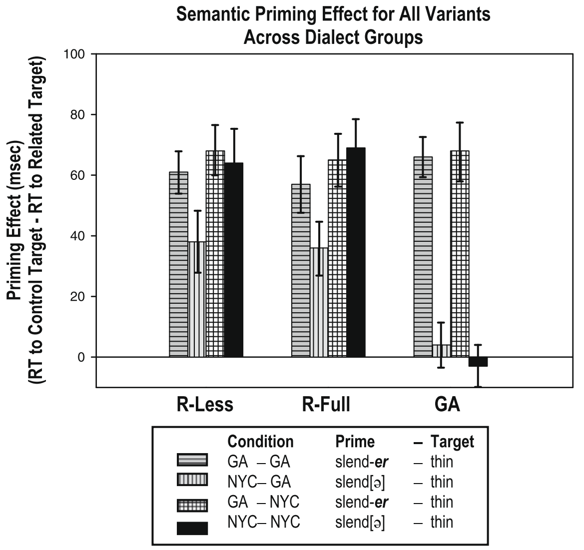

Fig. 2. Semantic priming effects for the three participant groups across conditions. Error bars correspond to the 95% confidence interval for each condition mean.

As in Experiment 1, the interaction between participant group and Condition reflects the difference in the results for participants with experience with the NYC dialect (Overt-NYC and Covert-NYC) and those who lack this experience (GA). The GA participants exhibit strong semantic priming when the prime is a within-dialect item, but the out-of-dialect NYC primes are ineffective in priming semantically-related items. As in Experiment 1, the priming effects for Overt-NYC and Covert-NYC participants are not statistically different ($F1(1,30) < 1$; $F2(1,79) < 1$), but there is a significant difference between the priming effects for Overt-NYC and GA participants ($F1(1,30) = 9.924$, $p < .01$; $F2(1,79) = 7.01$, $p < .01$; $MinF'(1,99) = 4.108$, $p < .05$) and for Covert-NYC and GA participants ($F1(1,30) = 5.627$, $p < .05$; $F2(1,79) = 4.381$, $p < .05$; $MinF'(1,96) = 2.463$, $p = .12$).

These results demonstrate that GA final r-ful primes facilitate responses to semantically-related targets, regardless of the listener's dialect. They also suggest that NYC final r-less primes are able to facilitate the recognition of semantically-related targets, but the effect is only facilitative for participants with prior exposure to the NYC dialect. Without such experience, NYC primes are completely ineffective at activating semantically-related targets. This pattern of results mirrors the *arbitrary mismatch* condition examined in Sumner and Samuel (2005). In that paper, we examined the lexical activation of words with the regular phonetic variants of final /t/ (*flute*, [flut], [fluʔt˺], [fluʔ]) by measuring their ability to prime a semantically-related word (e.g., *music*). We also included an arbitrarily mismatched variant that deviated from the base word by a single feature (e.g., *floose*), just as one of the variants ([fluʔ]) differed from the canonical form by a single feature. All three legal variants facilitated recognition of related words, but arbitrary variants were ineffective primes, having no effect on target recognition. In the current study, the NYC primes behaved like arbitrary variants for the GA speakers. It appears, then, that r-less variants are treated more like nonwords by GA participants, and that there are significant processing costs associated with out-of-dialect, or unfamiliar, phonological variants. This result is consistent with several lines of research that examined phonological mismatch (e.g., Connine et al., 1993; Frauenfelder, Scholten, & Content, 2001; Marslen-Wilson, Moss, & Van Halen, 1996).

Both the form priming results of Experiment 1 and the semantic priming results of Experiment 2 argue for a critical role of dialect experience: the two NYC-raised groups produced identical priming patterns, and these patterns contrasted sharply with the results for the GA-raised participants. Experience facilitates the activation of lexical items, as seen in the NYC listeners processing NYC variants. A lack of experience has obvious costs, as well, since the GA listeners treat the NYC variants as arbitrary (not part of their sound system). These variants preclude lexical activation in the GA listeners. At minimum, the results suggest that word recognition suffers from new or infrequent variants, consistent with variant frequency effects found by Connine and colleagues (Connine, 2004; Connine, Ranbom, & Patterson, 2008).

Thus far, the pattern of results is similar to that found by Sumner and Samuel (2005) for the immediate processing of phonetic variants. An interesting dichotomy in that study was found when comparing the results from an immediate processing task (semantic priming) to a long-term priming task. Immediately, all regular variants facilitated recognition of a related target. In the long-term, however, the canonical form was the only effective prime for an identical target (even though the canonical form is not the most frequent form); the canonical form seemed to dominate representationally. Experiments 1 and 2 have shown a consistent pattern for perception as a function of dialect. We explore the issue of representation in Experiment 3.

Experiment 3: long-term repetition priming

Experiments 1 and 2 demonstrated clear dialect-based differences in the immediate perceptual processing of speech. These experiments do not, however, inform us as to how variants are represented in the long-term. For example, even though we see differences in the immediate processing of other-dialect forms, it might still be the case that variants are effectively stored as the intended items independent of the listener's language background. Alternatively, other-dialect variants might simply not be encoded effectively, with both immediate perceptual consequences and representational differences as well. We use the long-term priming paradigm in Experiment 3 to examine these issues. In this paradigm, participants hear critical items and fillers presented in two blocks, with critical items in the first block ultimately acting as primes for targets presented in the second block. With two long blocks of items (e.g., 560 items), the time between prime and target is 20–30 min. Despite this long lag, primed words are identified more quickly than unprimed ones in the second block of trials (Church & Schacter, 1994; Goldinger, 1996; Kempley & Morton, 1982; Luce & Lyons, 1998). We can use this technique to assess whether a variant (e.g., *bakə*) is stored in a lexical form that is capable of producing priming (e.g., for *bakɚ*) over the long-term, and whether such representational forms vary with dialect of the listeners.

Methods

Participants

Forty-eight Stony Brook University students participated in this experiment for course credit or for pay. Sixteen participants from each listener group were included in this study. No participant reported any hearing problems.

Stimuli and design

The critical stimuli used in this experiment were the same as in Experiment 1. The design was similar in that we are examining form repetition, but instead of receiving one block of items with prime-target pairs separated by 500 ms, participants in this experiment received two blocks of items with each item presented separately. The first block of items included primes and fillers, and the second block of items consisted of targets and fillers. The initial block of items contained 80 final –er forms – 40 produced by a female GA speaker and 40 produced by a female NYC speaker.

The second block of items had 80 corresponding items (i.e., the same 80 lexical items) produced by male speakers, and thus produced four experimental conditions: GA – GA, NYC – GA, GA – NYC, and NYC – NYC. Of the 40 items produced by the female GA speaker in Block 1, 20 were produced by a male GA speaker in Block 2 (GA – GA) and 20 were produced by a male NYC speaker in Block 2 (GA – NYC). Of the 40 items produced by a female NYC speaker in Block 1, 20 were produced by a male GA speaker (NYC – GA) and 20 were produced by a male NYC speaker (NYC – NYC). An additional 80 final –er forms (40 from each male speaker) were included in Block 2 and served as new controls; these 80 words were unrelated to the 80 prime/target words. Sixteen lists were created to ensure that every item appeared in every condition (critical and control) and that no items or variants of an item were repeated for any participant.

In addition to the 80 primes in Block 1, 160 real word fillers and 320 pseudoword fillers were added, resulting in 560 trials. Block 2 contained the same number and type of fillers, resulting in 640 Block 2 trials. Half of the real word fillers from Block 1 were repeated in Block 2 and half were new; this was also true for the pseudoword fillers. This ensured that the critical items were not the only repeated items. The entire experiment, then, consisted of an initial block of 560 items (primes) presented one at a time, followed by a second block of 640 items (targets) presented one at a time. In both blocks, subjects made lexical decision judgments for all items.

Results

Responses with reaction times more than 2.5 standard deviations from the mean were excluded from analysis (<2%). Two participants with mean error rates above 10% to non-critical stimuli were excluded and replaced by new participants from the same participant population. Table 6 presents the reaction time results, and Table 7 shows the error data.

Reaction times. A three-way ANOVA (participant group × condition × repetition) performed on the reaction times revealed main effects of related items ($F1(1,45) = 13.646$, $p < .01$; $F2(1,79) = 14.889$, $p < .01$; $MinF'(1,109) = 6.858$, $p < .05$), Group ($F1(2,45) = 10.09$, $p < .01$; $F2(2,158) = 7.736$, $p < .01$; $MinF'(2,161) = 4.379$, $p < .05$), and Condition ($F1(3,135) = 4.127$, $p < .01$; $F2(3,237) = 4.803$, $p < .01$; $MinF'(3,329) = 2.22$, $p = .086$). Despite long delays, repeated items are recognized more quickly than new items. Unlike Experiment 1, the facilitation that occurs reflects the impact of having heard a word after about a half-hour delay, requiring long term representation of the lexical and/or sublexical information. The amount of priming depended on both group and condition. The main effect of participant group was driven by GA participants responding more slowly overall because of trials including out-of-dialect items. The main effect of that condition reflected the fact that NYC primes were poorer at inducing facilitation than the GA primes. An interaction between participant group and Condition ($F1(6,135) = 3.288$, $p < .01$; $F2(6,474) = 2.905$, $p < .01$; $MinF'(6,450) = 1.542$, $p = .163$) shows that the lack of facilitation (for NYC primes, for GA listeners) is experience-based. The interaction of Condition and Repeated items was significant as well ($F1(3,135) = 4.441$, $p < .01$; $F2(3,237) = 2.728$, $p < .05$; $MinF'(3,372) = 0.169$), driven by the lack of priming for conditions with NYC primes for Covert-NYC and GA participants.

Fig. 3 shows the priming effects across participant groups and conditions.

The figure reveals a striking distinction between two groups of participants who had patterned as one in the immediate perceptual tasks. Introducing a lag between prime and target produced a distinction between Overt-NYC and Covert-NYC participants ($F1(1,30) = 14.198$,

Table 6
Mean reaction times to targets in Experiment 3.

	Participant Dialect		
	Overt-NYC	Covert-NYC	GA
GA prime	GA target		
Related[a] (beɪkɚ – beɪkɚ)	917	913	911
Unrelated (fɪltɚ – beɪkɚ)	975	977	973
NYC prime			
Related (beɪkə – beɪkɚ)	938	973	980
Unrelated (fɪltə – beɪkɚ)	977	976	972
GA prime	NYC target		
Related (beɪkɚ – beɪkə)	934	959	992
Unrelated (fɪltɚ – beɪkə)	989	992	996
NYC prime			
Related[a] (beɪkə – beɪkə)	930	994	989
Unrelated (fɪltə – beɪkə)	991	991	992

[a] Denotes an identity condition (same dialect, same item, different voice).

Table 7
Mean error rates to targets in Experiment 3.

	Participant dialect		
	Overt-NYC	Covert-NYC	GA
GA prime	GA targets		
Related[a] (beɪkɚ – beɪkɚ)	2.5	3.1	3.0
Unrelated (fɪltɚ – beɪkɚ)	2.8	2.8	2.8
Difference	+.3	−.3	−.2
NYC prime			
Related (beɪkə – beɪkɚ)	2.6	2.8	3.0
Unrelated (fɪltə – beɪkɚ)	2.7	2.9	2.9
Difference	+.1	+.1	−.1
GA prime	NYC targets		
Related (beɪkɚ – beɪkə)	2.3	2.1	10.6
Unrelated (fɪltɚ – beɪkə)	10.4	10.6	24.3
Difference	+8.1	+8.5	+13.7
NYC prime			
Related[a] (beɪkə – beɪkə)	10.3	10.5	24.1
Unrelated (fɪltə – beɪkə)	10.1	10.4	24.4
Difference	−.2	−.1	+.3

[a] Denotes an identity condition (same dialect, same item, different voice).

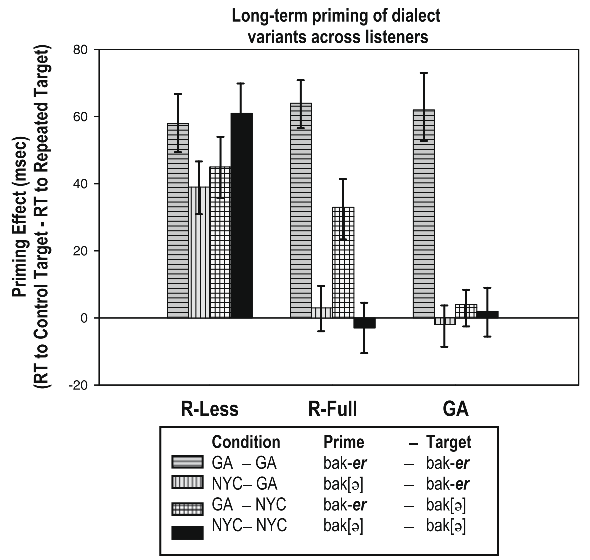

Fig. 3. Priming effects for all participants across experimental conditions. Error bars correspond to the 95% confidence interval for each condition mean.

$p < .001$; $F2(1,79) = 5.639$, $p < .05$; $MinF'(1,109) = 4.036$, $p < .05$). These two listener populations, who behaved uniformly in the immediate priming task, diverge when a long lag is added between prime and target.

To make sure that this divergence was not some artifact of using single-item presentation, as opposed to the pairs used in the first two experiments, we examined lexical decision times and error rates in the first Block of the long-term priming task. As expected, the reaction times to Block 1 primes for Overt-NYC participants (981 ms) and Covert-NYC (987 ms) participants did not significantly differ ($F1(1,30) < 1$; $F2(1,77) < 1$). Moreover, as expected, response times to the GA tokens in Block 1 were comparable for both Overt (971 ms) and Covert (974 ms) participants to the range of times to NYC targets in Block 2 that did not have a Block 1 prime (972–977 ms; see Table 6); similarly, the Overt (989 ms) and Covert (1003) groups' responses to the Block 1 NYC tokens were comparable to the Block 2 NYC targets without Block 1 primes (989–996 ms; Table 6). The error rates for all participant groups for GA items in Block 1 were similar (Overt: 2.2%, Covert: 2.7%, GA: 2.6%). The error rates to NYC targets for each participant group resemble those for unrepeated controls in Block 2; critically, there is no difference between Overt- and Covert-NYC participants (Overt: 10.2%, Covert: 9.9%, GA: 25.2%). These results show that the difference in reaction times and error rates to repeated items in Block 2 is due to encoding differences, not due to the single-item presentation of regionally-marked dialect items.

The pattern of results for GA speakers is different than that found for the other two participant groups, but is quite similar to the pattern found by Sumner and Samuel (2005): the only condition that resulted in a strong priming effect was the GA – GA condition. A GA prime only facilitated target recognition when the target and prime were both produced in the GA dialect. Unlike the Covert participants who are able to accommodate the variation present in an r-dropped target, GA participants only benefit from an identity condition that is purely GA. In Sumner and Samuel (2005), despite the efficacy of all three final /t/ variants in immediate priming, only the canonical /t/ (prime) – canonical /t/ (target) case was effective in the long-term paradigm.

The only difference between the Covert-NYC and GA participants in the long-term is the performance of GA – NYC prime – target pairs. This difference can be attributed to the GA listeners' inability to process NYC variants as targets (as in Experiment 1), or to the Covert-NYC participants' flexibility in processing previously encountered variants.

Error rates

The error rate results for Experiment 3 are similar to those for Experiment 1. Main effects were found for Group ($F1(1,45) = 10.23$, $p < .01$; $F2(1,79) = 9.843$, $p < .01$; $MinF'(1,116) = 5.106$, $p < .05$), and Condition ($F1(1,45) = 4.621$, $p < .05$; $F2(1,79) = 5.518$, $p < .05$; $MinF'(1,109) = 2.515$, $p = .116$). The main effect of group reflects the higher error rates of GA participants to NYC items, while the main effect of condition reflects the higher overall error rates to NYC targets than to GA targets. Interactions between Condition and Repeated items ($F1(3,135) = 3.274$, $p < .05$; $F2(3,237) = 4.962$, $p < .01$; $MinF'(3,298) = 1.972$, $p = .118$), and Group and Condition were found ($F1(6,135) = 3.99$, $p < .01$; $F2(6,474) = 3.316$, $p < .01$; $MinF'(6,464) = 1.81$, $p = .095$). No other results were significant. The interactions between Group and Condition and Group and Related may again be driven by two factors: (1) the general benefit of having a GA prime preceding a NYC target, and (2) the robust benefit for GA participants. Pairwise comparisons show that there is a strong benefit to hearing a GA prime before a NYC target, *even in the long term*, for all three participant groups: error rates are significantly lower for related targets than for unrelated targets in the GA – NY condition (Overt: $F1(1,15) = 5.851$, $p < .05$, $F2(1,79) = 7.064$, $p < .01$; $MinF'(1,33) = 2.990$, $p = .093$; Covert: $F1(1,15) = 8.603$, $p < .01$, $F2(1,79) = 7.29$, $p < .05$; $MinF'(1,34) = 4.176$, $p < .05$; GA: $F1(1,15) = 9.955$, $p < .01$, $F2(1,79) = 9.771$, $p < .01$; $MinF'(1,34) = 4.716$, $p < .05$).

General discussion

The goal of this project was to examine the processing and encoding of dialect variants by individuals who have had different personal histories with respect to their exposure to dialects of American English. Perhaps not surprisingly, given the occurrence of r-ful forms in orthography, education, the media and even among many speakers in the NYC-area, there appears to be a strong overall benefit for r-ful forms independent of participant dialect (in terms of the error rates for the form and long-term priming experiments). Despite this generalization, the results suggest that experience with a dialect has a major impact on the ability of listeners to process variants and that geographic origin alone is not a good predictor of this ability. Perhaps more surprisingly, dialect production and/or home environment can lead to the establishment of independent phonological representations for each dialect, even as listeners show an ability to access lexical entries regardless of the dialectal variant that is encountered.

The case of the fluent listener

The results associated with the NYC stimuli are worth discussing in detail. As one would expect, the NYC dialect causes difficulties for GA speakers. Also expected, Overt-NYC participants produce, represent, and perceive words in the NYC dialect with no problems. The most unexpected result is the behavior of the Covert-NYC participants when responding to NYC variants, and the dichotomy between Covert- and Overt-NYC participants. Unlike Overt-NYC listeners, Covert-NYC participants consistently produced final [ɚ] during the exit interviews and show long-term representational effects for only the GA forms consistent with production, yet they exhibit the same perceptual flexibility as the Overt group. We believe that this pattern of performance is evidence for what we will call "fluent listening".

Consider, as an illustration, a marriage between a native American English speaker born and raised in the United States and a native British English speaker born and raised in England. Both adults, these speakers have established representations of their respective languages. Over time, each develops the ability to listen to the other fluently. When asked how to pronounce a particular word in the non-native dialect, neither may know, or may exhibit predictable stress or phonotactic errors due to native dialect. When *listening* to the non-native dialect, however, there are no (or few) issues. We could say in this case, and in the case of the Covert-NYC participants, that these are fluent listeners. They exhibit the immediate perceptual abilities of a native speaker, but differ from a native speaker in both production and representation.

Mechanistically, what *is* fluent listening? We would like to propose here that fluent listening involves flexibility at the surface level in that these listeners are perceptually able to handle multiple *regular* variants of a particular word across dialects, just as they are perceptually able to handle multiple regular variants of a particular word within their dialect. Hence, they are fluently able to immediately process these variants as a single word. The reason they are fluent listeners and not bi-dialectal, however, is that over time (as in Experiment 3), these variants are encoded as abstracted to a single, r-ful form. The mechanism for fluent listenering stems from two ideas in the literature. First, enabling fluent listening, consistent experience with variants results in listener accommodation of variation (Johnson, 1997). Second, promoting the accommodation of both surface variants as variants of a single form, frequent variants are encoded more efficiently than infrequent variants. This immediate fluency is likely manifested via a mediated representation to which both variants map in immediate processing (which, in this paradigm is indistinguishable from mapping to two distinct mediated-representations as expected in the Overt-NYC participants). It may also be that it is exactly this mapping that results in processing costs in the long-term.

Experience as a factor in the immediate processing of variants

In terms of immediate lexical activation, there were substantial differences between listeners experienced with r-less variants and those with little prior experience with r-dropping. Both GA and NYC primes facilitated target recognition in form and semantic priming tasks for listeners in the Overt- and Covert-NYC groups. In contrast, NYC tokens were ineffective primes to related targets for the GA listener group in both experiments. This suggests that variable experience with a single form enhances flexibility in processing.

This result echoes those of within-language phonetic variation in which regular variants of a language are processed more efficiently than irregular, or arbitrary, variants. For example, Sumner and Samuel (2005) showed that all three legal phonetic variants of final /t/ in English (e.g., *flute* [flut], [fluʔt˺], [fluʔ]) were equally effective (immediate) primes for semantically-related targets; in contrast, an arbitrary variant (e.g., *floose*) that like [fluʔ] differs from the canonical form in place of articulation, was unable to improve target recognition. The results taken together suggest that once a variant is internalized, it can provide the cues needed for immediate perception. In both the final /t/ case, and the case of individuals raised with two legitimate variants of final –er, lexical access is triggered by all/both legal variants, but not by unlicensed variation. Thus, from an immediate *processing* perspective, the variants are functionally equivalent. However, this does not imply that they are *representationally* equivalent.

Evidence for dialect-specific representations

In the long-term paradigm, very different results were observed for Overt- and Covert-NYC participants. Specifically, Overt-NYC participants appear to encode both variants of final –er equally well; either form is able to facilitate the recognition of either form even after a 20–30 min period. This is not so for the Covert-NYC and the

GA participants. Those individuals appear to encode only the GA final –er form; r-less NYC variants are not encoded by these participants, leading to no long-term priming by the NYC variants. How can this be so?

This result at first appears to be at odds with the results of Experiments 1 and 2. The immediate priming data were virtually identical for the Overt-NYC and Covert-NYC participants, with these participants yielding results quite different than those for the GA participants. The two NYC groups treated the r-less forms as productive variants, while the GA participants did not. However, the long-term priming data reveal that although the results for the two NYC groups converge on the measures of immediate recognition, there is actually a fundamental representational difference between the two: Overt-NYC participants store both variants in memory, whereas Covert-NYC participants store only the GA final –er form.

This implies that the Covert-NYC individuals encode the r-less form as a variant of an underlying r-ful form during word recognition. This process can be seen as the ability to map a wider set of inputs onto the single underlying representation. Note that this is exactly the same implication forced by Sumner and Samuel's (2005) results for regular variants of final /t/, and the growing body of literature suggesting some special status for canonical forms (McLennan et al., 2003; Ranbom & Connine, 2007): they can all be used in word recognition, but only the canonical form seems to be stored for later use and thus produces long-term priming.

Benefit for general American –er

A number of results suggest that a GA –er form benefits a participant more than a NYC –ə form, even for Overt-NYC listeners. This effect surfaces in the error rates from Experiments 1 and 3. In both experiments, listeners were required to make lexical decisions to –er final words from both dialects. In both experiments, listeners did this reliably for GA forms (with error rates ranging from 1.9% to 3.1% across experiments). All listeners, independent of dialect or language background, faltered somewhat when responding to target items in the NYC dialect ending in –ə. In these instances, error rates ranged from 9.9% to 10.6% for Overt- and Covert-NYC listeners and 21.6–24.4% for GA listeners.

An important caveat is needed here: these high error rates were substantially reduced when the r-dropped NYC target was primed by the same lexical item in the GA dialect (e.g., [beɪkɚ] – [beɪkə]). In this case, the error rates dropped to 2.1–3.4% for Overt- and Covert-NYC listeners and to 10.6–11.6% for GA listeners. Thus, although it is true that GA listeners had more difficulty responding to NYC targets overall, the benefit for NYC forms primed by GA forms was universal. Moreover, this benefit even obtained when the GA form had been presented a half-hour (and several hundred trials) before the NYC target.

Finally, to return to our original questions regarding what it means to have a dialect and the extent to which experience influences dialect perception and representation, we can make the following suggestions: one can *have* a dialect in one domain (word recognition) and not another (production); or at one time (immediate processing), but not another (long-term processing). The Overt-NYC participants provide an interesting case contrasting perception and production. On the one hand, they pattern as we would expect with facilitative processing of all variants in the short-term and equivalent encoding of all variants in the long-term. On the other hand, they tend to produce a single form. If the traditional definition of dialect is used to assess speakers (e.g., a dialect is what is produced), their perceptual accommodation of multiple dialects is unconsidered. These results demonstrate that language use and language representation are not equivalent, at least for listeners exposed to two dialects. This result is reminiscent of the common observation in language acquisition that production often lags behind perception and comprehension.

As for the role of experience in the recognition and representation of dialects – it is critical. The most obvious influence of experience (or, the lack of it) is the inability of General American participants to process out-of-dialect forms. There is a clear and consistent processing cost for speakers new to a non-native dialect region. These costs are evident both in recognition and lexical activation, and give credence to the numerous anecdotes involving miscommunication across dialects. The more interesting effect of experience, however, surfaces in a comparison of the linguistic environments of the Covert- and Overt-NYC participants. As previously discussed, the two groups behave similarly in the short-term, but contrast sharply in their ability to encode the schwa-final variants.

In comparing the linguistic background of the two groups of NYC participants, we find that the type and amount of experience a listener has are critical to the processing of variants. Members of both groups are lifelong NYC-area residents. However, Covert-NYC participants receive input containing –ɚ final words at home, and are likely to have exposure to schwa-final forms through peers. The Overt-NYC participants receive schwa-final words at home, but are likely to have more formal and more direct experiences with the –er final variant in more environments (school, media, peers) than the Covert-NYC participants have to schwa-final forms. Potentially critical to the asymmetry in the representations between Overt-NYC and Covert-NYC participants is the age of acquisition: both groups of participants have been exposed to r-less variants, but only the Overt-NYC participants may have been exposed to these variants at a young age.[4] Perhaps, then, the Overt-NYC participants have exposure to both variants in similar quantities early on, motivating individual phonological representations, but the Covert-NYC participants have a more robust bias towards a single form. At this point, we can only suggest this as a possible explanation, and propose this as an empirical question open for future research.

These differences in exposure to dialectal variants produce intriguing dissociations in phonetic production, perception, and representation. The results presented here suggest a dissociation between production and representation on the one hand, and the perceptual process and rep-

[4] We would like to thank an anonymous reviewer for this suggestion.

resentation, on the other. The dissociation between production and representation can be seen in the Overt-NYC group. The results support the suggestion that Overt-NYC dialect members typically produce a single form ([ə]), but use both forms in perception, and store representations for both variants. Comparing the Covert-NYC participants to the GA participants brings to light the dissociation between perception and representation: we have evidence that both groups of speakers store representations only for a single dialectal variant, ending in –er. Yet, the two groups behave differently in the immediate processing tasks. GA participants, with little exposure to r-dropping, have significant costs associated with the out-of-dialect variant. Covert-NYC participants, though, are perfectly capable of processing the NYC dialect variant even though they have only a single representation that conflicts with the surface variant. What they appear to have gained by living in a dialect region is perceptual flexibility, and this might be what is acquired when processing difficulties eventually dissipate for speakers moving to a region with an unfamiliar dialect. They become fluent listeners.

From these patterns, we claim that there is not a simple and unitary answer to the question of what it means to *have* a dialect. We propose that a dialect must be characterized by how it is produced, perceived, and represented. Contrary to the typical assumption that a person's production reveals his or her dialect, our results indicate that aspects of a dialect may differ within an individual, just as they differ between individuals.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 0325188 and by NIMH under Grants No. F32MH68204-01A and R0151663. We greatly appreciate helpful comments and suggestions from William Labov and two anonymous reviewers. We are especially grateful to Herb Clark for discussion of fluent listening. We also thank Donna Kat for help with this project.

Appendix A

A.1. Post-experiment questionnaire

1. In what city, state, country were you born?

   a. Your parents?
   b. Your grandparents?

2. Where do you currently live?
3. Have you ever lived in a different town? If yes, where and for how long?

   a. Your parents?
   b. Your grandparents?

4. What language(s) do you speak at home?

   a. Your parents?
   b. Your grandparents?

5. Do you have friends or family that r-drop? If yes, who (mother, best friend, acquaintance)?

(Term *r-drop* introduced in post-experiment interview)

Appendix B. Critical items used in all experiments, along with corresponding semantically-related targets used in Experiment 2

armor – KNIGHT, *bagger* – GROCERY, *baker* – CAKE, *banker* – MONEY, *banner* – FLAG, *blubber* – FAT, *blunder* – MISTAKE, *bother* – ANNOY, *bouncer* – CLUB, *calendar* – DATE, *camper* – TENT, *chapter* – BOOK, *charger* – PHONE, *cheater* – TEST, *chowder* – CLAM, *chrysler* – CAR, *cider* – APPLE, *clever* – SMART, *climber* – ROCK, *clipper* – NAIL, *clutter* – MESS, *copier* – XEROX, *creamer* – COFFEE, *cucumber* – SALAD, *dealer* – DRUG, *easter* – BUNNY, *feather* – LIGHT, *fever* – SICK, *fibber* – LIAR, *fiddler* – ROOF, *finger* – HAND, *flounder* – FISH, *flower* – DAISY, *gender* – MALE, *glacier* – ICE, *grammar* – ENGLISH, *hanger* – CLOTHES, *hunger* – FOOD, *hunter* – ANIMALS, *jeweler* – DIAMONDS, *joker* – WILD, *kosher* – JEWISH, *laser* – BEAM, *minister* – PRIEST, *murder* – DEATH, *number* – ONE, *officer* – POLICE, *pacifier* – BABY, *panther* – CAT, *paper* – PENCIL, *pepper* – SALT, *quiver* – SHAKE, *razor* – SHAVE, *register* – CASH, *rider* – BICYCLE, *robber* – THIEF, *roller* – BLADES, *seltzer* – SODA, *silver* – GOLD, *simmer* – BOIL, *skimmer* – POOL, *slayer* – FUNNY, *slender* – THIN, *slither* – SNAKE, *sober* – DRUNK, *soccer* – BALL, *sour* – SWEET, *spider* – WEB, *summer* – HOT, *teacher* – SCHOOL, *thunder* – LIGHTNING, *timber* – TREE, *toaster* – BREAD, *weather* – RAIN, *whisper* – TALK, *winter* – COLD, *choker* – NECKLACE, *monster* – CLOSET, *shower* – CLEAN, *slugger* – BASEBALL.

References

Bowie, D. (2000). *The effect of geographic mobility on the retention of a local dialect*. Unpublished doctoral dissertation, Linguistics, University of Pennsylvania, Philadelphia.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 521–533.

Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics, 32*, 111–140.

Clopper, C. G., & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics, 35*, 421–438.

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences of assimilation for word recognition and lexical representation. *Language and Cognitive Processes, 16*, 535–564.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review, 11*, 1084–1089.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*, 193–210.

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics, 7BI0*, 403–411.

Deelman, T., & Connine, C. M. (2001). Missing information in spoken word recognition: Nonreleased stop consonants. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 656–663.

Evans, B. G., & Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America, 115*, 352–361.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? A lexical decision study in French listeners. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1276–1293.

Frauenfelder, U. H., Scholten, M., & Content, A. (2001). Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes, 16*, 583–607.

Gaskell, G., & Marslen-Wilson, W. D. (1993). Match and mismatch in phonological context. In *Proceedings of the fifteenth annual meeting of the cognitive science society* (pp. 446–451). Hillsdale, NJ: Erlbaum.

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 144–158.

Gick, B. (1999). A gesture-based account of intrusive consonants in English. *Phonology, 16*, 29–54.

Giles, H. (1970). Evaluative reactions to accents. *Educational Review, 22*, 211–227.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166–1183.

Gow, D. W. Jr., (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language, 45*, 133–159.

Gow, D. W. Jr., (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance, 28*, 163–179.

Gow, D. W. Jr., (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics, 65*, 575–590.

Janson, T., & Schulman, R. (1983). Non-distinctive features and their use. *Journal of Linguistics, 19*, 321–336.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–166). San Diego: Academic Press.

Kempley, M., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology, 73*, 441–454.

Kurath, H. (1939). *Handbook of the linguistic geography of New England*. Providence, RI: Brown University Press.

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

Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English: Phonetics phonology and sound change*. Berlin: Mouton de Gruyter.

Labov, W., Karan, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change, 3*, 33–74.

LoCasto, R. C., & Connine, C. M. (2002). Rule-governed missing information in spoken word recognition: Schwa vowel deletion. *Perception & Psychophysics, 64*, 208–219.

Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory & Cognition, 26*, 708–715.

Luce, P. A., McLennan, C. T., & Charles-Luce, J. (2003). Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. In J. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory* (pp. 197–214). Oxford, England: Oxford University Press.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.

Marslen-Wilson, W. D., Moss, H. E., & Van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1376–1392.

McLennan, C., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 529–553.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics, 65*, 956–969.

Preston, D. R. (1989). *Perceptual dialectology: Non-linguists' view of aerial linguistics*. Dordrecht: Foris.

Radeau, M. (1983). Semantic priming between spoken words in adults and children. *Canadian Journal of Psycholinguistics, 37*, 547–556.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1297–1311.

Ranbom, L., & Connine, C. M. (2007). Lexical representation of phonological variation. *Journal of Memory and Language, 57*, 273–298.

Scott, D. R., & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior, 23*, 450–466.

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*, 412–432.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final-/t/. *Journal of Memory and Language, 52*, 322–338.

Warren, P., & Marslen-Wilson, W. (1988). Cues to lexical choice: Discriminating place and voice. *Perception & Psychophysics, 43*, 21–30.