Attention, Perception, & Psychophysics (2018) 80:894–912
https://doi.org/10.3758/s13414-018-1485-z

Speaker information affects false recognition of unstudied lexical-semantic associates

Sahil Luthra[1,2] · **Neal P. Fox**[1,3] · **Sheila E. Blumstein**[1,4]

Abstract

Recognition of and memory for a spoken word can be facilitated by a prior presentation of that word spoken by the same talker. However, it is less clear whether this speaker congruency advantage generalizes to facilitate recognition of unheard related words. The present investigation employed a false memory paradigm to examine whether information about a speaker's identity in items heard by listeners could influence the recognition of novel items (critical intruders) phonologically or semantically related to the studied items. In Experiment 1, false recognition of semantically associated critical intruders was sensitive to speaker information, though only when subjects attended to talker identity during encoding. Results from Experiment 2 also provide some evidence that talker information affects the false recognition of critical intruders. Taken together, the present findings indicate that indexical information is able to contact the lexical-semantic network to affect the processing of unheard words.

Keywords Attention and memory · Spoken word recognition · Phonology and semantics

In order to understand spoken language, listeners must map perceived sounds to mental representations that contact meaning. This process of spoken word recognition is complicated by the lack of a consistent one-to-one correspondence between the speech signal and cognitive representations; in particular, different instances of a given word can take on countless phonetic forms. Some factors that influence a word's acoustic realization are linguistic in nature, such as phonetic and sentential context (Daniloff & Moll, 1968; Fox, Reilly, & Blumstein, 2015; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Scarborough, 2010). However, acoustic form is also modulated by nonlinguistic factors collectively referred to as *indexical* properties of the speech signal (Abercrombie, 1967; Pisoni, 1997). Such properties include speech rate (Kessinger & Blumstein, 1998; Miller & Baer, 1983), and the identity of the speaker (Allen, Miller & DeSteno, 2003; Peterson & Barney, 1952). Importantly, although these factors do not typically inform the linguistic content of the message, listeners do appear to track such cues, exhibiting sensitivity to talker-specific phonetic variation in a variety of perceptual (e.g., Allen & Miller, 2004; Creel, Aslin, & Tanenhaus, 2008; Kraljic & Samuel, 2007; Theodore & Miller, 2010; Trude & Brown-Schmidt, 2012) and memory tasks (e.g., Goldinger, 1996; Nygaard, Sommers, & Pisoni, 1995; Palmeri, Goldinger, & Pisoni, 1993).

While it is clear that indexical properties of a spoken word can influence subsequent recognition of and memory for that word, one outstanding question regards the scope of indexical effects within the broader lexical processing system. Past work has shown that recognition of a target word (e.g., *pear*) also results in the partial activation of both phonologically related words (e.g., *bear*; Marslen-Wilson & Welsh, 1978; Goldinger, Luce, & Pisoni, 1989; Luce & Pisoni, 1998) and semantic associates (e.g., *fruit*; Collins & Loftus, 1975; McNamara, 2005; Meyer, Schvaneveldt, & Ruddy, 1975). What is not clear is how indexical effects fit into this larger theoretical framework. Are

---

✉ Sahil Luthra
sahil.luthra@uconn.edu

Neal P. Fox
neal.fox@ucsf.edu

Sheila E. Blumstein
sheila_blumstein@brown.edu

[1] Department of Cognitive, Linguistic & Psychological Sciences, Brown University, 190 Thayer St., Box 1821, Providence, RI 02912, USA

[2] Present address: Department of Psychological Sciences, University of Connecticut, 406 Babbidge Rd, Unit 1020, Storrs, CT 06269, USA

[3] Present address: Department of Neurological Surgery, University of California, San Francisco, 675 Nelson Rising Lane, Room 535, San Francisco, CA 94143, USA

[4] Brown Institute for Brain Science, Brown University, 2 Stimson Ave, Providence, RI 02912, USA

the effects of a word's indexical properties restricted to subsequent processing of that *same* word, or do they extend to modulate access to the lexical-semantic network? In particular, can indexical features of a target word (e.g., the identity of the speaker who produced *pear*) affect subsequent processing of *other* unheard words that are phonologically or semantically related to the target (e.g., *bear* or *fruit*)?

A small number of studies have attempted to answer this question and provided evidence that talker identity might influence the processing of phonologically related items, though perhaps not of semantic associates. In a visual world study by Creel and Tumlin (2009), participants learned nonword–nonword minimal pairs (e.g., *boog* and *booj*). Critically, half the pairs comprised items that were always spoken by the same talker, and half the pairs contained items that were produced by different talkers. After learning, participants were faster to recognize an item that had been part of a talker-varying minimal pair than one in a same-talker pair, suggesting that participants had maintained information about talker identity in their representations of the stimuli and were able to use this information to disambiguate the items relatively early in processing. Converging evidence comes from Dufour and Nguyen (2016), who conducted an auditory lexical decision task using French stimuli. Participants were slower to make lexical decisions on target words (e.g., *bancaire*) that were preceded by phonologically competitive primes (e.g., *banquette*) relative to targets preceded by control words. Critically, the magnitude of this inhibitory priming effect was reduced if there was a change in talker between the prime and target than if the talker was held constant; that is, participants were faster to recognize a target that had been preceded by a talker-mismatching prime than one preceded by a talker-matching prime. The authors attributed this effect to the participants having represented talker identity in memory, resulting in decreased lexical competition when the talker information was incongruent between prime and target. In contrast to studies employing phonologically related pairs, there has been little evidence that talker information can influence access to other items based on semantic associations. An auditory lexical decision study by Kittredge, Davis, and Blumstein (2006), for instance, found no evidence that varying the identity of the talker between a prime (e.g., *peace*) and semantically associated target (e.g., *war*) affected the magnitude of priming.

The current work leverages a false memory paradigm to provide additional insight into the extent to which information about a talker's identity might influence the processing of phonological and semantic associates. The Deese–Roediger–McDermott (DRM) false memory paradigm (Deese, 1959; Roediger & McDermott, 1995) is particularly well-suited for assessing how relationships between words can affect recognition memory. In a standard DRM task, subjects study items (e.g., *rye*, *jelly*, *milk*, *butter*, *sandwich*) that are all semantically associated with an unstudied word (e.g., *bread*). On a subsequent memory test, subjects are more likely to falsely remember the semantically related unstudied word (referred to as the *critical intruder*, or CI) than to remember a filler unstudied item (e.g., *chair*).

A dominant theory in the memory literature is that false memories for critical intruders result from spreading activation (Roediger, Bolota, & Watson, 2001), mirroring a proposed architecture in theories of spoken word processing (e.g., McNamara, 1992; Mirman & Magnuson, 2009). In the spreading activation framework, each time an associate is presented during encoding, the CI is partially activated. The cumulative effect of studying several associates is the robust activation of the intruder item. While the majority of DRM studies have used semantic associates to generate false memories, some studies (e.g., Sommers & Lewis, 1999) have also induced false memories (e.g., for the word *bill*) using phonological associates (e.g., *build*, *bowl*, *pill,* and *fill*). (For recent reviews of the DRM paradigm, see Gallo, 2006, 2010.)

Collectively, the literature reviewed suggests that recognition of a word after prior exposure to a list of words is bolstered (1) if the word was produced by the same speaker during exposure (i.e., congruency of indexical information) and (2) if there are strong semantic or phonological relationships between the word and those words presented during exposure. The present study employs the DRM paradigm to examine the extent to which indexical information contained in a particular token of a spoken word (e.g., information about the identity of the speaker) can contact representations in the lexical-semantic network. If talker information interacts with the lexical-semantic network, then individuals may be more likely to falsely remember a CI if it is produced by the same talker who had previously produced that CI's associates than if the CI and its related associates were produced by different talkers.

Notably, past research has implicated a host of factors that may modulate indexical effects on recognition memory, including whether indexical information was attended by the listener (Goldinger, 1996; Theodore, Blumstein, & Luthra, 2015). In a recognition memory experiment, Goldinger (1996) found that subjects showed a same-talker advantage on both hit rate and reaction time; that is, subjects were more likely to correctly remember having previously heard a word and were faster to do so when the word was produced by the same speaker at test as at encoding. This effect was largest when subjects had attended to talker identity during encoding and less pronounced when they had attended to the word's initial phoneme or made syntactic decisions at encoding. If attending to talker identity promotes the emergence of talker-specific effects on recognition memory for a previously heard word, doing so may also affect the emergence of any talker-specific effects on false recognition of a related, but novel, word (i.e., a critical intruder).

Data from an experiment by Roediger, McDermott, Pisoni and Gallo (2004, Experiment 3) are consistent with this prediction. In this study, subjects listened passively to several semantic DRM lists, each produced entirely by one of two talkers. During a subsequent auditory recognition test, CI items were either produced by the talker who had produced its associates (same talker trials) or by the other speaker (different talker trials), and subjects were asked whether the item was new or old. Although participants exhibited robust false memories for the CIs, Roediger et al. (2004) did not find an effect of speaker congruency on the rate of false recognition. However, subjects were also asked to indicate, when they judged the item (incorrectly) to be old, whether the talker was the same as at encoding. Indeed, subjects were more likely to attribute the CI to the talker who had spoken its associates at encoding than they were to attribute it to the incongruent talker. That is, while talker congruency between a CI and its studied associates did not seem to modulate the *likelihood* that a CI would be falsely remembered, subjects' *attribution decisions* for falsely remembered CIs did exhibit a talker congruency effect. These results suggest that both semantic and indexical relationships between the studied and test items influence listeners' behavior at test but also hint that their interaction may be modulated by the extent to which listeners' attention is directed to the relevant indexical cues.

In summary, recognition memory for a word appears to be modulated by indexical properties of the word (e.g., whether it was produced by the same voice when it was previously heard) as well as by more abstract linguistic properties of the word (e.g., its semantic/phonological relationships with other words that were previously heard). The current study aimed to examine the relationship between indexical information and information that is encoded within the lexical-semantic network and, in particular, to investigate how these sources of information might interact to influence recognition of a spoken word. In Experiment 1, we extended the findings of Roediger et al. (2004) by investigating the conditions under which congruency of speaker information affected how likely subjects were to have false memories for unstudied CIs. Because attention to talker identity may modulate the emergence of talker-specific effects (e.g., Theodore et al., 2015), subjects either passively encoded the study lists (consistent with the approach of Roediger et al., 2004) or were asked to identify the speaker of each word during encoding. To investigate the extent to which speaker information might affect the activation of phonological neighbors and semantic associates, half the subjects received lists in which the studied items were phonologically related to the CIs and half received lists with semantic associates of the CIs.

Experiment 1

Method

Subjects Ninety-seven subjects were recruited from the Brown University community and compensated for their participation. Each was a self-reported right-handed native speaker of American English with no neurological or hearing deficits. One subject was excluded from analyses due to a computer error during the experiment, yielding an *n* of 96.

Subjects were assigned to one of two encoding conditions—half to a passive condition and half to a talker identification condition. Within each encoding condition, half the subjects received lists with phonological associates of unpresented CIs (e.g., associates: *rack*, *black*, *buck*; CI: *back*), and half received lists with semantic associates of other unpresented CIs (e.g., associates: *butter*, *jam*, *crust*; CI: *bread*). This resulted in a 2 × 2 (Encoding × List Type) between-subjects design, with 24 subjects in each condition.

Stimuli For each list type, 12 items were chosen to serve as the CIs. For each CI, a 15-item study list was created by selecting words either phonologically or semantically related to the CI. The full set of stimuli is presented in Appendix A.

Phonological lists were taken from Sommers and Lewis (1999), with two studied items changed to avoid duplicate items between lists. Each CI was a monosyllabic CVC word, and items on the study lists differed from their CIs by the addition or substitution of one phoneme. Semantic lists were taken from Roediger and McDermott (1995). Figure 1 shows an example semantic study list and the associations between the studied items and their associated CI, the latter shown as the hub of the figure. In a few instances, word association norms (Nelson, McEvoy, & Schreiber, 1998) were used to replace unusual list items. For instance, Roediger and McDermott used the word *riders* as an associate of the CI *rough*, but the experimenters felt that this historical reference would be lost on most contemporary subjects and instead used *harsh* for the CI *rough*.

Additional items were selected to serve as filler unstudied items on a postencoding recognition test. Twenty-four such items were used for each list type; six CIs and 18 associates (three per CI) from unused lists by Sommers and Lewis (1999) and Roediger and McDermott (1995) were used for the phonological and semantic conditions, respectively.

Stimuli were recorded by a male and female speaker via microphone (Sony ECMMS907) in a soundproof room using a digital recorder (Roland Edirol R-09HR). Each speaker was a native speaker of American English and produced three consecutive tokens for each item, from which the best token was selected. Stimuli were scaled for amplitude using BLISS software (Mertus, 2002).

Confirming that the speakers spoke at approximately the same rate, a one-factor ANOVA found no significant effect of speaker on stimulus duration, mean durations: 617 ± 7 ms (male), 615 ± 6 ms (female); $F(1, 570) = 0.045$, $p = .833$. For this analysis, speaker was treated as a between-item factor, since the two speakers produced different encoding lists and therefore only some items were recorded by both speakers. To confirm that there were perceptible acoustic differences between the two speakers, fundamental frequencies were computed for each speaker. For each speaker, two words with an

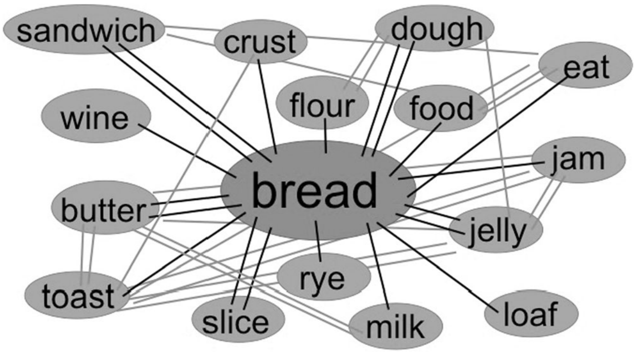

Fig. 1 Example of a semantic stimulus set for the critical intruder (CI) "bread." Lines between nodes in the network indicate semantic associations in Nelson et al. (1998), with black lines connecting the CI to its studied associates and gray lines indicating connections among studied items

/æ/ vowel, two with an /i/, and two with an /u/ were analyzed, with fundamental frequency computed over five glottal pulses from the strongest portion of the vowel. The male speaker had a fundamental frequency of 119 Hz, which was significantly different, $t(5) = 20.734$, $p < .001$, from the female speaker's fundamental frequency of 216 Hz.

Procedure As discussed above, subjects studied items that were either phonologically or semantically related to the CIs. Subjects in the passive encoding condition simply listened during encoding, while subjects in the talker ID condition were also asked to indicate via button press whether the male or female spoke each word.

Each participant completed an encoding phase and a postencoding recognition test. At test, subjects were asked to indicate whether each word was presented at encoding by pressing "yes" or "no" on a button box and were told to respond as quickly as possible without compromising accuracy.

During encoding, subjects listened to 12 study lists, six of which were always produced by the male and six by the female speaker. Importantly, CIs were not presented at encoding. Subjects in both the passive encoding or talker ID conditions were told to listen carefully, as they would later be tested on their memory. List order was pseudorandomized once such that no more than two consecutive lists were produced by a single speaker, and the same stimulus order was used for all subjects; that is, all participants received the same order of lists as well as the same order of items within those lists. In this way, stimulus exposure at encoding was as similar as possible between participants. Stimuli were presented with a 1-second ISI. In order to prevent subjects from rehearsing studied items, subjects were given a 90-second filler task of dot-to-dot puzzles (Kalvitis, 2000) between lists, a task which entailed subvocalizing each number while searching for it on the page. Subjects were told to complete as much of the puzzle as they could within the allotted time.

Following the final encoding list, subjects were presented with a 72-item auditory recognition test. The test comprised the 12 unstudied CIs (around which the studied lists were built), three studied associates of each CI (from list positions 1, 6, and 11), and 24 filler unstudied items. Thus, half of the test items were truly *old* or studied (corresponding to correct "yes" responses), and half were truly *new* or unstudied (corresponding to correct "no" responses). Talker (same/different) was manipulated within item and within subject. That is, half of the critical lists (studied items and the corresponding CI) were heard in the same voice as at encoding (*same talker* trials) and half were heard in an incongruent voice (*different talker* trials). Whether a particular item appeared in the same talker or different talker condition was counterbalanced across subjects. For a given subject, a single talker produced all associated words (e.g., *bread*, *butter*, *jam,* and *crust*) at test, and each subject received 24 same-talker trials and 24 different-talker trials. For filler unstudied items, the male and female speaker each produced half of the lists. The item order was randomized once, and this same random order was used for all subjects in a given list type group.

The full design is presented in Fig. 2. Factors are described as between-subjects/within-subjects and between-items/within-items depending on how they appeared at test. For instance, talker is listed as a within-item factor because while each item was produced by only one talker at encoding, each item was produced by both talkers (across participants) at test; talker is also described as within subject because each participant heard some items in a same-talker condition and some in a different-talker condition. Filler unstudied items are not pictured because they cannot be described as having a same/different talker.

Results

False recognition of unstudied items To first examine whether false memories were induced, we compared subject false alarms as a function of the type of unstudied item (CIs and filler items) and of list type (phonological and semantic). (Recall that filler unstudied items cannot be associated with a same/different talker, as filler lists were not presented at encoding.) Because subjects' responses to each item in the test phase were categorical (that is, "yes"/"no" indicating that they considered a given item old/new), results were analyzed using mixed-effects logistic regression (Jaeger, 2008). A detailed description of the logit mixed models used in this study is presented in Appendix B.

A two-factor analysis (item type, list type) revealed that subjects had a significantly higher false alarm rate for CIs than for filler items (mean false alarm rates: 0.67 vs. 0.22; $\beta = 1.658$, $SE = 0.141$, $|z| = 11.742$, $p < .001$). Participants also had more false alarms for the phonological lists than for semantic lists (mean false alarm rates: 0.43 vs. 0.31; $\beta = -0.485$, $SE = 0.109$, $|z| = 4.446$, $p < .001$), and there was a significant interaction between these factors ($\beta = 0.738$, $SE = 0.140$, $|z| = 5.282$, $p < .001$), indicating that the

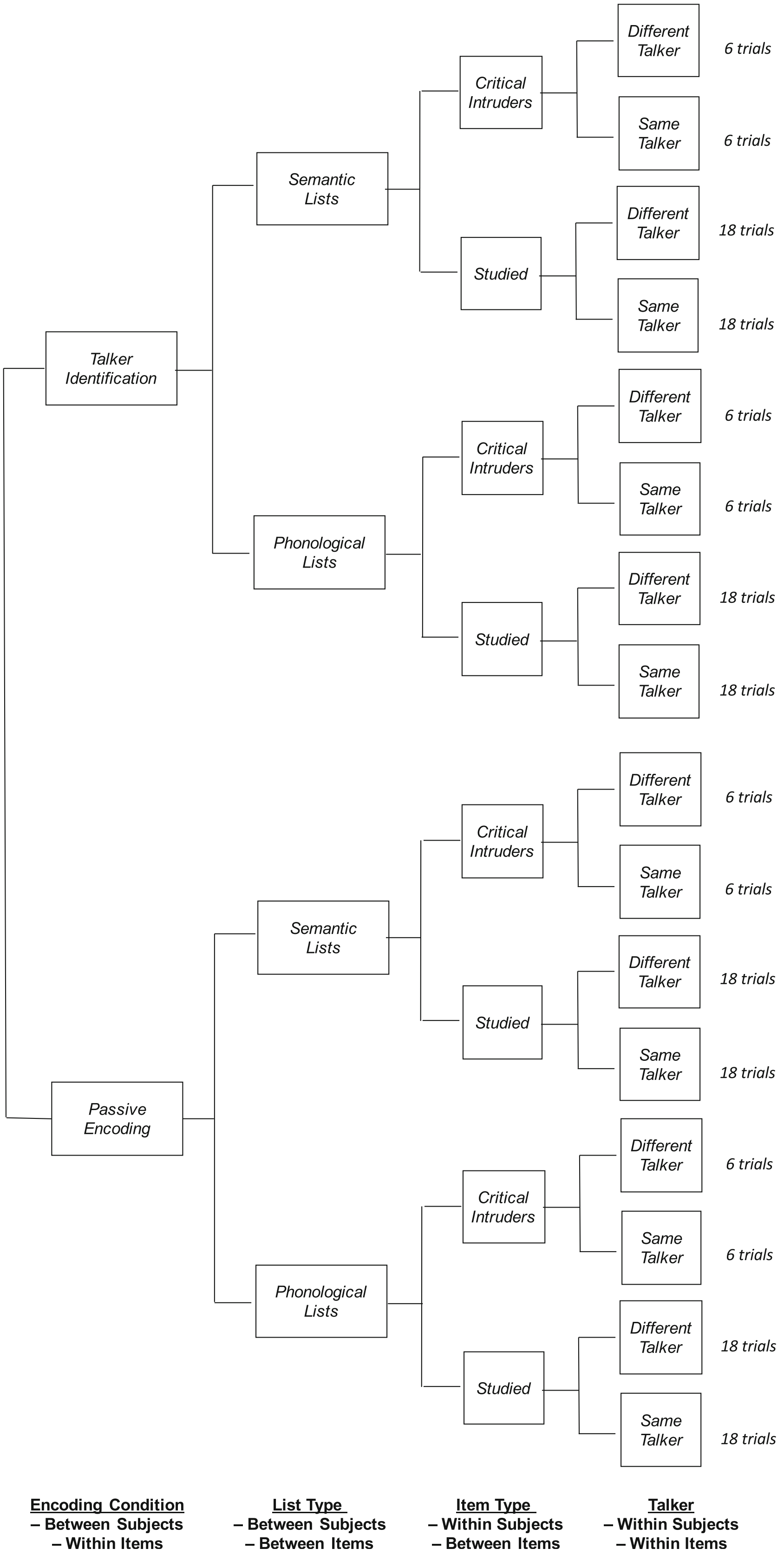

Fig. 2 The design of Experiment 1

tendency of subjects to falsely recognize CIs more often than unstudied filler items was stronger when subjects studied semantic lists than when they studied phonological lists. Critically, follow-up tests showed that false memories were reliably induced for each list type independently: Subjects falsely recognized CIs significantly more often than filler items both on phonological lists (mean false alarm rates: 0.62 vs. 0.34; β = 0.891, *SE* = 0.168, |*z*| = 5.300, *p* < .001) and semantic lists (mean false alarm rates: 0.72 vs. 0.11; β = 2.506, *SE* = 0.246, |*z*| = 10.202, *p* < .001).

Talker effects on recognition To examine the influence of talker information on recognition, results (see Fig. 3) were submitted to logit mixed models (see Appendix B) that predicted how often a subject made a "yes" response, thus identifying an item as old. The only significant effect was a four-way interaction between encoding (passive encoding vs. talker ID), list type (phonological vs. semantic), item type (studied vs. CI), and talker (same vs. different) (β = 0.006, *SE* = 0.003, |*z*| = 2.321, *p* = .020). We conducted bootstrapped simulations in order to evaluate the robustness of this four-way interaction (Efron & Tibshirani, 1986; see Appendix B for details). Results confirmed the reliability of the four-way interaction (95% CI [0.001, 0.011]) and that no other main effects or interactions were significant.

Additional analyses were conducted to investigate the basis of the four-way interaction identified in the omnibus analysis. In light of evidence that speaker effects may be modulated by attention to talker identity at encoding (e.g., Theodore et al., 2015), the data were first partitioned along encoding condition, with follow-up logit mixed models constructed for the two groups of subjects (in the passive encoding and talker ID conditions, separately). No effects were significant in the three-factor analysis of the passive condition. The model examining the talker ID condition yielded a significant three-way interaction between list type, item type and talker (β = 0.016, *SE* = 0.008, |*z*| = 2.100, *p* = .036) and no other significant effects.

To further probe the interaction from the omnibus analysis, data from the talker ID condition (bottom panel in Fig. 3) were partitioned by list type. This enabled us to examine the effects of item type and talker separately for phonological lists and for semantic lists. This model revealed no significant simple effects or interactions for phonological lists. For semantic lists, there was a trending interaction between item type and talker (β = 0.038, *SE* = 0.020, |*z*| = 1.920, *p* = .055) and no significant simple effects. A final follow-up analysis suggested that this interaction appeared to be driven by a significant simple effect of talker that emerged within critical items (β = 0.165, *SE* = 0.070, |*z*| = 2.352, *p* = .019) but not within studied items. Collectively, the analyses suggest that congruent speaker information boosted false recognition of critical intruders, but this speaker congruency effect only emerged for subjects who attended to talker identity during encoding of semantic lists.

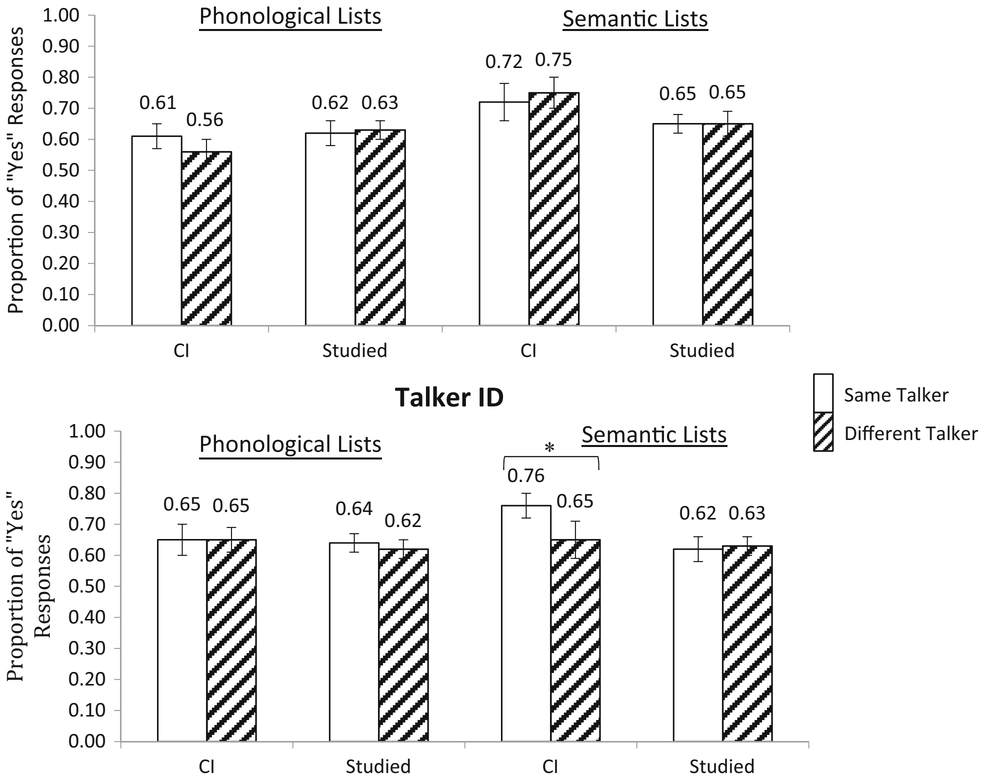

Fig. 3 Proportions of "yes" responses in Experiment 1. CI = critical intruder. Error bars indicate standard error of the mean across subjects. *Indicates a significant simple effect of talker ($p < 0.05$)

Discussion

Experiment 1 examined whether talker information affects how likely listeners are to have false memories for phonological and semantic associates of previously studied words. To the extent that it does, this would suggest that nonlinguistic, indexical information is retained during the processing of spoken words and modulates the effects of more abstract, linguistic properties on lexical activation and recognition memory for a word. Results demonstrate that talker information can, under certain conditions, influence the rate of false recognition of unstudied words that are related to studied items. In particular, when subjects actively encoded talker identity, there was a speaker congruency effect on subjects' likelihood to falsely recognize semantically related CIs. When subjects did not identify talker at encoding, no speaker congruency effect emerged on false recognition rates of these CIs, as was found by Roediger et al. (2004). Consistent with previous work on recognition memory (Goldinger, 1996; Theodore et al., 2015), the results of Experiment 1 suggest a possible role for attention to talker identity at encoding in the emergence of talker congruency effects, a point considered further in the General Discussion.

The speaker congruency effects observed in this study emerged only on critical items, not studied ones, and only with semantic lists, not phonological ones. With regard to item type, it is possible that talker effects on studied items are simply harder to detect than effects on critical intruders. In constructing the study lists, studied items were selected because of their strong associations with the CI, but they are not necessarily associated with each other (a point schematically demonstrated in Fig. 1). CIs, by their design, are therefore the *hub words* of their lists. A consequence of this may be that when each associate is activated, the CI is activated as well; after studying several associates of the CI, the CI is then highly activated. Given that studied items are not as highly associated with each other as with the CI, studied items may not have been highly activated (relative to the CI), and therefore speaker-based differences may have been more difficult to observe. The trace synthesis model proposed by McClelland (1995) offers the analogous account that remembering an item involves synthesizing all the traces that were activated at study. In a DRM task, the synthesized trace would be a composite of all the studied items, meaning that what is remembered best is the trace of the CI and not the trace of any individual word. If traces of studied items include indexical information, then the synthesized trace would also include information about speaker identity, resulting in a speaker congruency effect on the false recognition of CIs.

With regard to list type, however, it is curious that talker identity only affected semantically associated intruders and not phonological ones, particularly since the few studies investigating talker effects on lexical-semantic associates (e.g., Creel & Tumlin, 2009; Dufour & Nguyen, 2016; Kittredge et al., 2006; Lee & Zhang, 2015, 2017) might predict effects on phonologically related words but not on semantic associates. Moreover, given that talker identity directly influences the acoustic-phonetic properties of words but is not directly tied to their meaning, it is surprising that Experiment 1 failed to find an effect of talker identity on phonological intruders but found one on semantic ones when, a priori, the opposite pattern of results might have seemed more likely. Examination of the lists used in Experiment 1 reveals that the semantic lists are easily distinguishable into discrete semantic categories (see Appendix A). In contrast, the phonological lists are harder to distinguish from each other—the same sounds appear in multiple lists, and thus a critical intruder associated with one list (e.g., *peer*) may be closely related phonologically to words on several lists (e.g., *rear* and *peel*). It is possible that the similarity between phonological lists may have contributed to the failure to find a speaker effect in that condition. Specifically, if each CI is phonologically associated with studied items spoken by both speakers, then the inconsistent speaker information might wipe out speaker effects on the CIs.

Overall, Experiment 1 suggests that under some conditions, talker information can affect the false recognition of words that are unstudied but related to studied items. The goals of Experiment 2 were to further probe how indexical information affects the recognition of spoken words by testing whether methodological changes would allow for the emergence of talker congruency effects in phonological lists when listeners attend to talker information at encoding, and to replicate the talker congruency effect observed in semantic lists in Experiment 1.

Experiment 2

In Experiment 1, participants who identified talkers at encoding were sensitive to the congruency of speaker information between studied items and unstudied semantic associates (critical intruders) presented on a subsequent recognition test. Surprisingly, Experiment 1 failed to show such an effect on targets that were phonologically associated with studied words. However, while there are a large number of different semantic categories, it is relatively difficult to present lists of words which are distinguished by their phonological categories. There are roughly 24 different consonant phonemes in English, but they are not independent of each other. Rather, they group into classes of sounds that not only share sound properties (e.g., Chomsky & Halle, 1968; Halle & Stevens, 1971; Jakobson, Fant, & Halle, 1951) but are also more likely to be perceptually confused (e.g., Cutler, Weber, Smits, & Cooper, 2004; Miller & Nicely, 1955). A few methodological changes were therefore implemented for Experiment 2, most notably redesigning phonological lists to be more phonologically distinct from each other. Additionally, because effects only emerged in the talker identification condition of Experiment 1, all participants in Experiment 2 were asked to attend to talker identity. With these methodological changes in place, we sought to test for a speaker congruency effect on phonological lists as well as to replicate the speaker congruency effect on semantic lists shown in Experiment 1.

Method

Subjects Forty-nine subjects who did not participate in Experiment 1 participated in Experiment 2. Subjects were recruited using the same criteria outlined for Experiment 1. One subject was excluded from analyses, since she reported afterward having treated the recognition memory task as a lexical decision task. Thus, data from 48 participants were analyzed.

Stimuli The same talkers used in Experiment 1 recorded new stimuli for Experiment 2. The full set of stimuli for Experiment 2 is presented in Appendix A.

To minimize overlap between phonological lists, each study list employed consonants from only one of the four natural classes for manner of articulation (stops, fricatives, liquids/glides, nasals). This design choice limited the experiment to four phonological lists (in contrast to the 12 used in Experiment 1) and therefore incurred a reduction in statistical power, a point further considered in the Discussion section. Phonological lists were developed by selecting four CVC words (*but*, *seize*, *well,* and *mine*) to serve as the CIs. Of the 15 associates chosen for each CI, 32 associates differed from the CI by the substitution of only one phoneme (e.g., *beat* differs from *but* by one phoneme). The remaining 28 associates differed by the substitution of multiple phonemes (e.g., *bad* differs from *but* by two phonemes). To match the number of phonological lists, only four semantic lists were used in Experiment 2. These were constructed by replacing bisyllabic words on four Experiment 1 lists (lists for *bread*, *foot*, *rough,* and *sleep*) with monosyllabic associates using the Nelson et al. (1998) free association norms.

Additional words were selected to serve as filler unstudied items for the two postencoding recognition tests. For the phonological recognition test, the four natural classes had already been used for the studied lists, so it was necessary to repeat at least one for filler unstudied items. As such, two additional CVC words (*keep* and *dock*) were selected to serve as CIs, and three associates were determined for each CI through one-phoneme substitutions. Though this entailed employing stop consonants both on studied and unstudied lists, these filler unstudied items were selected due to their phonological dissimilarity to studied items. (For instance, the studied list that employed stop consonants contained neither word-final labial stops nor word-initial alveolar stops; by contrast, *keep* and its associates all contained word-final labial stops, and *dock* and its associates all contained word-initial alveolar stops.) For the semantic filler unstudied items, two CIs and six associates (three per CI) were taken from unused Roediger and McDermott (1995) lists.

Experiment 2 also controlled for duration differences between list types. In Experiment 1, while all the phonological words were monosyllabic, a large number of the semantic words were multisyllabic. Consequently, Experiment 1 employed semantic stimuli that were significantly longer in duration than phonological stimuli, mean durations: 628 ± 7 ms (semantic), 604 ± 6 ms (phonological); $F(1, 570) = 7.114$, $p = .008$. In Experiment 2, only monosyllabic stimuli were used. A one-factor ANOVA confirmed that, for Experiment 2's stimuli, there was no effect of list type on stimulus duration, mean durations: 583 ± 13 ms (semantic), 591 ± 13 ms (phonological); $F(1, 190) = 0.185$, $p = .667$. Similar to Experiment 1, the two speakers spoke at a comparable rate, as indicated by a one-factor analysis that found no significant effect of speaker on stimulus duration, mean durations: 571 ± 12 ms (male), 603 ± 14 ms (female); $F(1, 190) = 3.198$, $p = .075$, and the male speaker's fundamental frequency of 99 Hz significantly differed, $t(5) = 39.633$, $p < .001$, from the female speaker's fundamental frequency of 198 Hz.

Procedure The procedure for Experiment 2 was identical to that of Experiment 1, with a few exceptions. Rather than treating list type as a between-subjects factor, every subject in Experiment 2 heard both list types (phonological and semantic). Subjects were first asked to encode all the lists of one type. As in Experiment 1, half of the lists were produced by the male speaker for all subjects, and the remaining half were always produced by the female. Following each encoding list, subjects completed a dot-to-dot puzzle to prevent rehearsal of the studied items; encoding was followed by a recognition memory test for those lists. They then received the four lists of the other type and then completed the test for those lists. The order in which subjects performed the phonological and semantic tasks was counterbalanced across subjects. Because speaker effects in Experiment 1 were only observed in the talker ID condition, all subjects were required to identify the gender of the speaker of each word during encoding. Following the encoding phase for each list type, subjects were given a 24-item auditory recognition test. Each test comprised four unstudied CIs, three studied associates of each CI (from list positions 1, 6, and 11), and eight filler items. In this way, half of the test items were truly *old* or studied (corresponding to correct "yes" responses) and half were truly *new* or unstudied (corresponding to correct "no" responses). As in Experiment 1, item order was randomized once and this same order was used for all subjects. The full design of Experiment 2 is presented in Fig. 4.

Results

False recognition of unstudied items A comparison of the false-alarm rates on CIs versus unstudied filler items confirmed that Experiment 2, like Experiment 1, successfully induced false memories. A two-factor analysis of item type and list type (see Appendix B for model details) revealed that subjects were significantly more likely to falsely recognize CIs than to recognize filler unstudied items (mean false alarm rates: 0.70 (CIs) vs. 0.14 (filler), $\beta = 2.394$, $SE = 0.357$, $|z| = 6.716$, $p < .001$). There was neither a significant effect of list type nor a significant interaction between list type and item type. As before, we also analyzed phonological and semantic lists separately. Results indicated that subjects falsely recognized CIs significantly more often than filler unstudied items both on phonological lists (mean false alarm

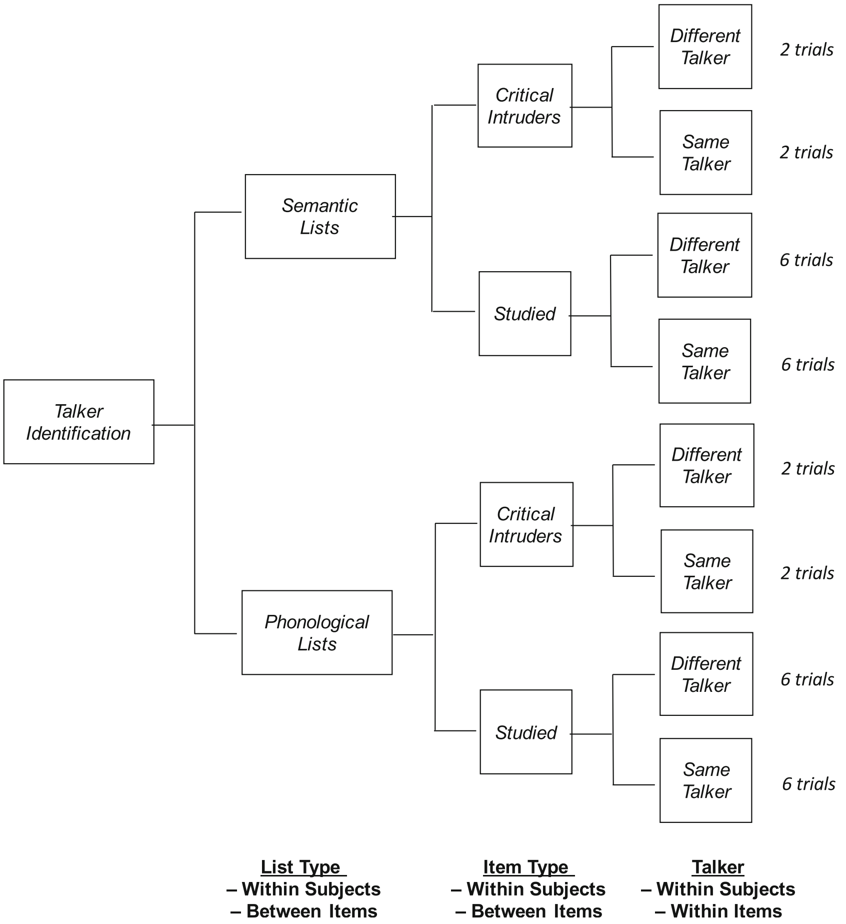

Fig. 4 The design of Experiment 2

rates: 0.59 vs. 0.12; β = 2.160, *SE* = 0.619, $|z| = 3.491$, $p < .001$) and semantic lists (mean false alarm rates: 0.81 vs. 0.15; β = 2.582, *SE* = 0.390, $|z| = 6.617$, $p < .001$).

Talker effects on recognition Logit mixed models (see Appendix B) were used to analyze speaker effects in studied and intruder items, and bootstrapped simulations were conducted to estimate confidence intervals for the effect sizes (see Appendix B for details). Mean hit rates for studied items and false alarm rates for critical intruders are shown in Fig. 5.

The omnibus three-factor (List Type × Item Type × Talker) analysis revealed a significant main effect of talker, β = 0.166 (95% CI [0.087, 0.262]), *SE* = 0.046, $|z| = 3.600$, $p < .001$, such that there were increased "yes" responses when speaker was held constant. No other main effects or interactions emerged. Of interest was whether, having made phonological lists more

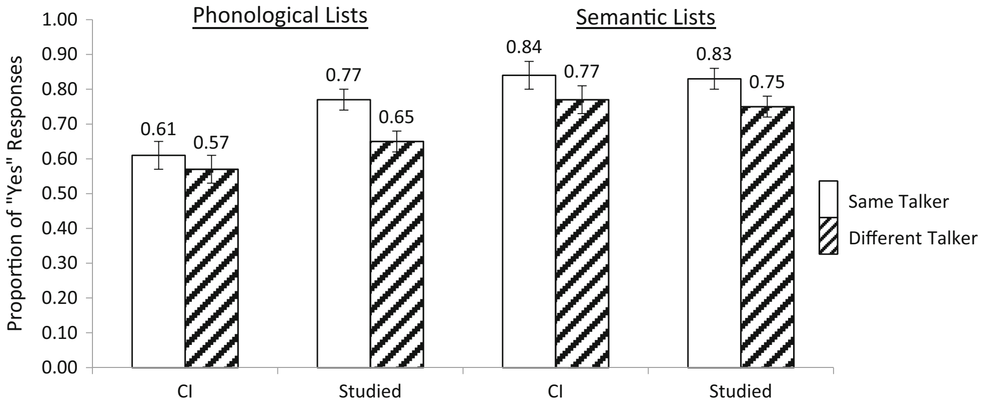

Fig. 5 Proportions of "yes" responses in Experiment 2. All subjects attended to talker identity during encoding. CI = critical intruder. Error bars indicate standard error of the mean across subjects

distinct in Experiment 2, speaker effects would emerge for both phonological and semantic lists. A planned comparison examined whether the main effect of talker held for each level of list type. This analysis indicated that the effect of talker held for phonological lists, $\beta = 0.184$ (95% CI [0.072, 0.302]), $SE = 0.059$, $|z| = 3.094$, $p = .002$) and for semantic lists, $\beta = 0.152$ (95% CI [0.014, 0.295]), $SE = 0.068$, $|z| = 2.228$, $p = .026$. No other main effects or interactions were significant for either list type. An additional comparison examined whether the talker congruency effect emerged for each level of item type, collapsing across semantic and phonological list types. This analysis indicated a significant main effect of talker for studied items, $\beta = 0.172$ (95% CI [0.083, 0.269]), $SE = 0.047$, $|z| = 3.699$, $p < .001$, and a numerical trend consistent with an effect of talker for CIs, $\beta = 0.121$ (95% CI [−0.021, 0.285]), $SE = 0.071$, $|z| = 1.693$, $p = .090$. No other main effects or interactions were significant for either item type. Overall, as can be seen in Fig. 5, the direction of the talker effect was in the predicted direction for all conditions: Words were more likely to be recognized at test when spoken by the same voice as that word's associates at encoding.

Discussion

In Experiment 2, talker congruency modulated recognition memory for both semantic and phonological word lists, with the effect of talker congruency appearing across previously studied words and unstudied associates (critical intruders). The emergence of these effects on both phonological and semantic lists suggests that the additional controls implemented in Experiment 2 were effective, although it is not immediately apparent whether this was driven by the reduction in sound overlap between phonological lists or the concomitant reduction in the number of lists from 12 to four. Since the desire to use more distinct phonological lists in Experiment 2 necessitated a reduction in statistical power, the present results cannot definitively establish whether a talker congruency effect would emerge specifically for unstudied *phonological* CIs, as was shown for *semantic* CIs in Experiment 1 and as might be predicted by other findings in the literature (Creel & Tumlin, 2009; Dufour & Nguyen, 2016). Overall, though, numerical trends in the data from Experiment 2 do suggest that talker information can modulate the false recognition of CIs, a result that is consistent with the major finding of Experiment 1 (that talker congruency between an unstudied target and its studied semantic associates can boost false recognition of the target).

Finally, the talker congruency effect on studied items in Experiment 2 was not observed in Experiment 1. A variety of factors may have contributed to this difference, including the strength of the associations between studied items and CIs (that is, how "dense" each list was around its CI), the number of lists studied (and therefore the memory demands of the experiment), and the amount of phonetic overlap between lists.

General discussion

The present study examined whether the match between talker information at encoding and retrieval can permeate the lexical-semantic network to influence the recognition of lexical items that are phonologically or semantically linked to previously heard items. The two experiments described here employed an auditory version of the DRM false memory paradigm (Deese, 1959; Roediger & McDermott, 1995) and manipulated whether the same talker producing the critical intruder at test also produced the associated studied items at encoding. In both experiments, subjects reliably had false memories for unstudied critical intruders that were phonologically or semantically associated with studied items. Of importance, in Experiment 1, we demonstrated that talker congruency can boost the false recognition of unstudied semantic associates. Unexpectedly, no such effect emerged for phonological intruders, despite previous findings (Creel & Tumlin, 2009; Dufour & Nguyen, 2016) that might have predicted such a result. Experiment 2 reconfigured phonological lists to reduce interlist phonemic overlap and showed that congruent talker information can boost the recognition of items presented at test. Consistent with the findings of Experiment 1, these results also showed a numerical trend in support of the conclusion that talker congruency affects the false recognition of unstudied CIs. Collectively, the two experiments provide evidence that talker information contacts representations of other words in the lexical-semantic network, suggesting that indexical information can have far-reaching effects on spoken word recognition.

Theoretical implications for processing indexical information

Effects of talker information have been broadly classified as arising through *talker-semantic* influences or through *acoustic-match* influences (Creel & Tumlin, 2011). Talker-semantic influences are conditioned on preexisting relationships between talker identity and the meaning of the words produced. Such relationships can be socially cued by factors like affective tone or accent (e.g., Nygaard & Queen, 2008; Sumner & Kataoka, 2013) or by pragmatic influences (e.g., Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008). Effects of talker-semantic information on lexical processing have been described in terms of a dual-route account for processing talker information (Sumner, 2015; Sumner, Kim, King, & McGowan, 2014), whereby one route is used to map acoustics to lexical representations, and one route is used to map acoustics to social representations (e.g., Freeman & Ambady, 2011). In contrast, *acoustic-match* influences comprise cases in which talker identity affects a listener's behavior despite being incidental to meaning. The present work, along with most published work demonstrating effects of nonlinguistic talker information on the perception of or memory for spoken words, falls into this latter category (e.g., Allen & Miller, 2004; Creel et al., 2008; Goldinger, 1996; Kraljic & Samuel, 2007; Luce & Lyons, 1998; Magnuson & Nusbaum, 2007; McLennan, Luce, & Charles-Luce, 2003; Nygaard et al., 1996; Palmeri et al., 1993; Theodore et al., 2015; Theodore & Miller, 2010; Trude & Brown-Schmidt, 2012). However, relatively little work has examined the extent to which this sort of indexical information can interact with the lexical-semantic network. Since acoustic-match influences are incidental to meaning, it is unclear whether such information should interact with the lexical-semantic network, and relatively few studies have investigated the extent to which it can (but see Creel & Tumlin, 2009; Dufour & Nguyen, 2016; Kittredge et al., 2006; Roediger et al., 2004).

The present findings demonstrate that indexical information can permeate the structure of the lexical-semantic network and affect processing when speaker information matches across stimuli. Such results place important constraints on accounts of how talker information is treated during spoken word recognition. Early theories of spoken word recognition took a so-called *abstractionist* approach, a term derived from the view that recognition entailed a mapping from the acoustic signal to abstract representations devoid of indexical information (Shankweiler, Strange, & Verbrugge, 1975; Syrdal & Gopal, 1986). More recent abstractionist accounts argue that there may be some representations that contain talker information and others that do not (Magnuson & Nusbaum, 2007). The present results clearly demonstrate that there must be some level of representation that can incorporate both indexical information and lexical-semantic relationships. Furthermore, if it is the case that talker information is indeed separated from linguistic information prelexically, as suggested by strict abstractionist accounts, there must be a way for it to be reintegrated with the lexical representation so that it can contact the lexical-semantic network. Consistent with this view, neuroimaging studies suggest the existence of two distinct pathways for processing of talker identity and linguistic content but propose that these pathways may interact (e.g., Belin, Fecteau, & Bédard, 2004; Myers & Theodore, 2017; Zhang et al., 2016).

An alternative to the abstractionist view comes via *episodic* accounts of representation (e.g., Goldinger, 1998; Goldinger & Azuma, 2003; Johnson, 2006), so named because they hold that representations are a composite of detailed speech episodes and therefore comprise both indexical and linguistic information. Such accounts are motivated by a variety of studies showing that listeners demonstrate sensitivity to indexical information in speech (in adults: Church & Schacter, 1994; Goldinger, 1996; Palmeri et al., 1993; in infants: Houston & Jusczyk, 2000; Singh, Morgan, & White, 2004). The present results might be explained in such an episodic framework if indexical information were allowed to permeate the lexical-semantic network. Under such a view, an activated lexical item (e.g., a studied item) would include a trace of information about the speaker who produced it. Linguistic information is able to contact the lexical-semantic network using extant lexical-semantic connections, and talker information would permeate the network along these same lexical-semantic connections. The cumulative effect of linguistic and talker information contacting other words would be the enhanced activation of unstudied but lexically/semantically related associates that match in speaker, resulting in talker congruency effects on false recognition. The trace synthesis model proposed by McClelland (1995) represents one potential mechanism through which this could occur. Given a potential role for attention in promoting talker congruency effects (discussed in Factors Modulating the Emergence of Talker Effects section), episodic frameworks that incorporate attention (e.g., Goldinger & Azuma, 2003) may provide an appropriate framework for characterizing speaker effects in accessing words and their lexical-semantic representations.

One other framework consistent with the present results is articulated by Pufahl and Samuel (2014), who found that recognition memory for spoken words was not only affected by congruency of indexical cues, like speaker information, but also by the congruency of co-occurring nonlinguistic, nonindexical environmental sounds (e.g., a dog barking). The authors argue that because episodic accounts incorporate indexical features but not other environmental sounds, such accounts cannot explain the congruency effects they observed. They thus posit that speech-specific lexical representations may not exist as they are traditionally described, but that representations for spoken words may exist in a multidimensional space in auditory memory, where any auditory cue—linguistic, indexical, or environmental—may be represented. The present results might be explained under this sort of nonlexical account; congruency of indexical information between intruders and studied associates would enhance the similarity between representations that already share linguistic features, potentially allowing for the emergence of the speaker congruency effects demonstrated here.

In considering the possibility that indexical information may be maintained in lexical representations, it is important to note that the present experiments utilized recognition memory tasks and therefore did not probe lexical access directly. Though recognition memory tasks are often used to investigate the components of lexical representations (e.g., Mattys & Liss, 2008; Palmeri et al., 1993; Theodore et al., 2015), their appropriateness for such questions is also a subject of debate. Magnuson and Nusbaum (2007), for instance, argue that while recognition memory paradigms can be used to show that listeners represent indexical information, such representations may not be the ones accessed in the course of online spoken word recognition. Nonetheless, studies more directly probing lexical access have provided some evidence that talker information may be able to contact the lexical-semantic network (e.g., Creel & Tumlin, 2009). To the extent that spoken word recognition and tasks that probe recognition memory for words engage the lexical-semantic network similarly, the current experiments place key constraints on accounts of spoken word recognition. In particular, any theory of spoken word recognition must allow indexical information to influence the lexical-semantic network. Consequently, talker information cannot be viewed simply as "noise" that is discarded prelexically (see Pisoni, 1997).

Factors modulating the emergence of talker effects

Experiment 1 suggested that attention to talker identity at encoding influences the emergence of talker-specific effects. In particular, there were no speaker congruency effects when subjects passively encoded study items (replicating the earlier results of Roediger et al., 2004), though such effects did emerge when subjects actively identified talkers at encoding. Experiment 2 replicated the finding that speaker congruency effects can emerge when listeners attend to talker identity at encoding. The present results therefore add to a growing body of literature positing an interaction between attention and the emergence of indexical effects, whether in word identification (Magnuson & Nusbaum, 2007; Mullenix & Howe, 1999) or recognition memory tasks (e.g., Bradlow, Nygaard, & Pisoni, 1999). Importantly, speaker congruency effects on recognition memory tasks do not seem to be the result of deeper processing at encoding, as such effects do not emerge when subjects perform syntactic judgments at encoding (Goldinger, 1996; Theodore et al., 2015). Rather, it may be the act of directing attention toward the relevant indexical variation that increases the likelihood of indexical congruency effects emerging.

Though identifying talker at encoding is a relatively simple task (particularly in the present experiment, where encoding lists were blocked by talker), doing so may increase the salience of talker-specific information, making such information more likely to influence later recognition. This may be achieved if attention boosts the activation of a low-level layer in processing, which could be realized under several architectures, including adaptive resonance frameworks (e.g., Goldinger & Azuma, 2003) and interactive models (e.g., Mirman, McClelland, Holt, & Magnuson, 2008). Alternatively, attending to talker at encoding may cue the subject to the importance of dimensions in auditory memory that include talker information (Magnuson & Nusbaum, 2007; Nosofsky, 1989; Pufahl & Samuel, 2014).

Importantly, none of these accounts posit that attention to talker identity is *required* for talker-specific effects to emerge, and talker specificity effects often emerge in the absence of an explicit attentional manipulation (e.g., Palmeri et al., 1993). In many studies, it may be the case that attention is implicitly drawn to talker information, as when a particular talker's productions are relatively unusual, as in foreign-accented (McLennan & González, 2012) or dysarthric speech (Mattys & Liss, 2008). However, several other factors may also modulate the emergence of talker-specific effects, including whether the test is implicit or explicit in nature (Church & Schacter, 1994; Schacter & Church, 1992) and how familiar a listener is with a particular talker (Dufour & Nguyen, 2016).

The amount of processing time at retrieval has also emerged as a key determinant of whether indexical information can affect processing. McLennan et al. (2003; see also McLennan & Luce, 2005) have advanced a time-course hypothesis, whereby abstract codes dominate processing in situations when response times are fast, but indexical information influences processing when response times are relatively long. Indeed, a number of studies (e.g., Mattys & Liss, 2008; Newman, 2016) are consistent with this hypothesis (but see also Theodore et al., 2015). The current work adds to this body of literature by suggesting that future studies on indexical effects should consider not only the role of factors such as processing time but also the possible role of attention to indexical information at encoding.

Conclusions

The present experiments demonstrate that indexical information can contact the lexical-semantic system, affecting recognition memory for items that have themselves not been studied but that are linguistically related to studied items. Speaker congruency effects emerged only when indexical information was actively encoded, suggesting a possible role for attention in increasing the salience of talker-specific details. Most informative for theories of spoken word recognition, results indicate that information about a speaker's identity is not restricted to its source word; rather, talker information can affect recognition memory for words based on their inherent structural properties (i.e., their lexical-semantic relationships).

Acknowledgements This research was supported by NIH Grant DC000314. S.L. was supported by the Society of Royce Fellows at Brown University. N.P.F. was supported by NSF Graduate Research Fellowship DGE 0228243, a Graduate Fellowship from the Brown Institute for Brain Science at Brown University, and NIH Grant F32 DC015966. These findings were presented at the Society for Neurobiology of Language in October 2015. Raw data and analysis scripts are available online at osf.io/5b7ct. We thank Rachel M. Theodore and members of the Blumstein lab for valuable input throughout this project. We also extend thanks to Meghan Clayards, Gareth Gaskell, James Magnuson, Arthur Samuel, Meghan Sumner, and anonymous reviewers for feedback on earlier versions of this manuscript.

Appendix A: Stimuli

Experiment 1: Phonological lists

Item	LIST 1	LIST 2	LIST 3	LIST 4	LIST 5	LIST 6	LIST 7	LIST 8	LIST 9	LIST 10	LIST 11	LIST 12
1	bag	build	sneak	stale	knight	gain	rot	year	loge	forth	dine	keel
2	rack	bowl	beak	cell	cite	bane	knot	hear	hone	ford	dung	wheel
3	pack	kill	cede	fail	rice	raid	put	veer	lobe	phone	den	feel
4	book	will	scene	scale	tight	rave	got	dear	cone	fort	dune	veal
5	jack	built	see	shale	height	cane	pop	spear	loam	far	none	meal
6	bake	pill	meek	say	rhyme	pain	tot	leer	moan	soar	dawn	eel
7	bass	ill	seep	sill	riot	rake	hot	fear	flown	bore	does	peel
8	bat	till	cheek	hail	might	main	palm	par	line	force	don	rile
9	black	been	chic	mail	white	ray	plot	rear	loaf	tore	bun	deal
10	bath	hill	seethe	wail	writ	rail	lot	pair	known	forge	dunce	teal
11	sack	bid	seat	seal	trite	race	posh	tear	loath	oar	din	heal
12	buck	fill	sick	snail	ripe	ran	dot	cheer	bone	lore	dumb	role
13	lack	ball	cease	dale	rife	raise	cot	beer	blown	core	dutch	wreath
14	bank	bit	leak	save	rise	rate	pat	pierce	load	poor	ton	reap
15	bad	big	siege	soil	light	lane	pit	mere	roan	fair	one	read
CI	**BACK**	**BILL**	**SEEK**	**SAIL**	**RIGHT**	**RAIN**	**POT**	**PEER**	**LOAN**	**FOR**	**DONE**	**REAL**

Filler unstudied items

LIST 1	LIST 2	LIST 3	LIST 4	LIST 5	LIST 6
that	heat	dwell	kin	bide	tip
cut	beef	wool	shin	side	rib
mat	cheat	bell	gin	tide	rid
CAT	**BEAT**	**WELL**	**SIN**	**RIDE**	**RIP**

Experiment 1: Semantic lists

Item	LIST 1	LIST 2	LIST 3	LIST 4	LIST 5	LIST 6	LIST 7	LIST 8	LIST 9	LIST 10	LIST 11	LIST 12
1	butter	shoe	hill	note	thread	water	smooth	bed	fast	sour	steal	door
2	food	hand	valley	sound	pin	stream	bumpy	rest	lethargic	candy	robber	glass
3	eat	toe	climb	piano	eye	lake	road	awake	stop	sugar	crook	pane
4	sandwich	kick	summit	sing	sewing	Mississippi	tough	tired	listless	bitter	burglar	shade
5	rye	sandals	top	radio	sharp	boat	sandpaper	dream	snail	good	money	ledge
6	jam	soccer	hike	band	point	tide	jagged	wake	cautious	taste	cop	sill
7	milk	yard	peak	melody	prick	swim	ready	snooze	delay	tooth	bad	house
8	flour	walk	plain	horn	thimble	flow	coarse	blanket	traffic	nice	rob	open
9	jelly	ankle	glacier	concert	haystack	run	uneven	doze	turtle	honey	jail	curtain
10	dough	arm	goat	instrument	thorn	barge	harsh	slumber	hesitant	soda	gun	frame
11	crust	boot	bike	symphony	hurt	creek	rugged	snore	speed	chocolate	villain	view
12	slice	inch	climber	jazz	injection	brook	sand	nap	quick	heart	crime	breeze
13	wine	sock	range	orchestra	syringe	fish	boards	peace	sluggish	cake	bank	sash
14	loaf	knee	steep	art	cloth	bridge	ground	yawn	wait	tart	bandit	screen
15	toast	mouth	ski	rhythm	knitting	winding	gravel	drowsy	molasses	pie	criminal	shutter
CI	**BREAD**	**FOOT**	**MOUNTAIN**	**MUSIC**	**NEEDLE**	**RIVER**	**ROUGH**	**SLEEP**	**SLOW**	**SWEET**	**THIEF**	**WINDOW**

Filler unstudied items

LIST 1	LIST 2	LIST 3	LIST 4	LIST 5	LIST 6
mad	white	table	hot	apple	web
fury	funeral	desk	wet	ripe	arachnid
hatred	ink	swivel	freeze	basket	creepy
ANGER	**BLACK**	**CHAIR**	**COLD**	**FRUIT**	**SPIDER**

Experiment 2: Phonological lists

Item	LIST 1	LIST 2	LIST 3	LIST 4
1	bug	size	yell	moon
2	bite	siege	wool	nine
3	boot	cease	wear	mean
4	bat	these	wall	man
5	bet	sage	will	mime
6	buck	save	wail	moan
7	bought	sauce	wheel	main
8	beat	shave	rare	noon
9	boat	shove	lair	noun
10	bait	sass	roar	none
11	cut	sieve	rear	known
12	putt	thus	leer	maim
13	bud	fuss	lure	numb
14	gut	fish	role	name
15	bad	five	lore	gnome
CI	**BUT**	**SEIZE**	**WELL**	**MINE**

Filler unstudied items

LIST 1	LIST 2
peep	duke
cup	talk
cape	dog
KEEP	**DOCK**

Experiment 2: Semantic lists

Item	LIST 1	LIST 2	LIST 3	LIST 4
1	stale	shoe	smooth	bed
2	food	hand	crude	rest
3	eat	toe	road	lay
4	crumb	kick	tough	trance
5	rye	heel	grit	dream
6	jam	leg	rogue	wake
7	milk	yard	raw	snooze
8	bun	walk	coarse	inn
9	wheat	paw	draft	doze
10	dough	boot	harsh	dorm
11	crust	arm	rock	snore
12	slice	inch	sand	nap
13	wine	sock	board	peace
14	loaf	knee	ground	yawn
15	toast	mouth	flat	deep
CI	**BREAD**	**FOOT**	**ROUGH**	**SLEEP**

Filler unstudied items

LIST 1	LIST 2
sour	steal
taste	cop
cake	bank
SWEET	**THIEF**

Appendix B: Logit model implementation details

Logit mixed models were used in this paper to avoid the pitfalls of traditional ANOVAs on categorical data (Jaeger, 2008) as well as to allow for the testing of fixed effects and interactions while simultaneously accounting for multiple crossed random factors (e.g., subjects and items; Baayen, Davidson, & Bates, 2008). Models were constructed in R (R Core Team, 2016) using the glmer function of the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). To assess whether beta estimates were robust to changes in sample, 95% confidence intervals were bootstrapped for four critical statistical analyses: the omnibus tests for Experiments 1 and 2, and the two planned comparisons for Experiment 2 (see main text for details). One thousand simulations were performed for each model using the confint function of the stats package (R Core Team, 2016).

The full dataset and analysis stream can be found online at osf.io/5b7ct.

Presence of false memories

Prior to examining potential speaker effects, a two-factor analysis was conducted to verify whether false memories were reliably induced in each experiment. Similar models were constructed and applied to the data sets from each experiment; in particular, the two models had identical fixed effect structures and slightly different random effect structures, as necessitated by design changes between the two experiments. Both models had fixed effects of item type (critical/filler unstudied) and list type (phonological/semantic), with the positive levels corresponding to the critical intruders and the semantic lists. The item type factor was centered and sum-coded with a contrast of (−0.5, 1), as there were twice as many unstudied items as critical intruders on each recognition test, and the list type factor was centered and sum-coded with a contrast of (−1, 1), since there were an equal number of items from each list. Following the recommendation of Barr, Levy, Scheepers, and Tily (2013), the maximal random effects structure was used for these models. In Experiment 1, random by-subject slopes for item type (since item type was manipulated within subjects) and random by-subject and by-item intercepts were included. Experiment 2 differed from Experiment 1 in that list type was also manipulated within subject. Consequently, the random effect structure for Experiment 2 included random by-subject slopes for item type and list type and random by-subject interactions between these factors. As in the model used for Experiment 1, random by-subject and by-item intercepts were also included.

In addition to these two-factor analyses, we conducted one-factor analyses of item type, considering each level of list type separately. Models for these analyses maintained the structure of the two-factor models but dropped fixed effects of list type, random slopes for list type, and random interactions that included list type.

Effects of speaker information

Additional models were created to investigate potential speaker effects in each experiment. Fixed effects for encoding (passive/talker ID), list type (phonological/semantic), and talker (different/same) were centered and sum-coded on contrasts of (−2, 2). Item type (studied/critical) was centered on a contrast of (−1, 3), as there were three times as many studied items as critical items. The talker ID, semantic, critical and same talker conditions served as the positive levels. The use of these contrasts resulted in more consistent model convergence than did the use of comparable normalized contrasts of (−0.5, 0.5) and (−0.25, 0.75), but both models yielded the same pattern of results.

Random effect structures included every random intercept, slope and interaction for subjects and items that was motivated by the design (Barr et al., 2013). In particular, the omnibus analysis for Experiment 1 included random intercepts for subjects and items, random by-subject slopes and interactions for item type and talker (both of which were manipulated within subject), and random by-item slopes and interactions for encoding and talker (both of which were manipulated within item). For Experiment 2, the omnibus analysis included random intercepts for subjects and items, random by-subject slopes and interactions for list type, item type and talker (as all were manipulated within subjects) and random by-item slopes for talker (manipulated within item). Follow-up models maintained these structures and dropped random slopes and interactions appropriately when two levels of a factor were analyzed separately.

Occasionally, the models encountered convergence failures and/or problematic fixed effects correlations (correlations > 0.30; see Agresti, 2002). Following the recommendation of Barr et al. (2013), this was resolved by constraining random effect structures such that they could not co-vary by-subject or by-item. Though such models are not maximal in the strictest sense, they do include all of the appropriate random intercepts, slopes and interactions; the inference ability of these models does not appear to be hindered relative to that of truly maximal models (Barr et al., 2013; see Eager & Roy, 2017, for a detailed discussion of an alternative analysis approach).

In Experiment 2, partitioning the data by item type resulted in only two observations per cell for the analysis of critical intruders; recall that for design reasons, each subject heard only two critical intruders in each condition. With this number of data points, it would be inappropriate to model random interactions, and so this term was dropped from the model. This decision is in line with the recommendation of Barr et al. (2013) to employ the "maximal random effect structure justified by the design."

The full set of best-fit parameters for the omnibus model in each experiment is presented in Table 1 and Table 2.

Table 1 The full set of best-fit parameters for the omnibus model in Experiment 1.

Fixed Effects	Estimate	*SE*	*Z* value	*p* value	
Intercept	0.729	0.099	7.360	.000	***
Item Type	0.053	0.052	1.010	.312	
Encoding	0.006	0.032	0.184	.854	
List Type	0.034	0.049	0.695	.487	
Talker	0.019	0.020	0.957	.339	
Item Type × Encoding	0.006	0.014	0.390	.697	
Item Type × List Type	0.032	0.026	1.219	.223	
Encoding × List Type	−0.019	0.016	−1.187	.235	
Item Type × Talker	0.007	0.011	0.633	.527	
Encoding × Talker	0.010	0.010	1.001	.317	
List Type × Talker	0.001	0.010	0.152	.879	
Item Type × Encoding × List Type	−0.004	0.007	−0.598	.550	
Item Type × Encoding × Talker	0.001	0.005	0.162	.872	
Item Type × List Type × Talker	0.005	0.005	0.890	.373	
Encoding × List Type × Talker	0.002	0.005	0.478	.633	
Item Type × Encoding × List Type × Talker	0.006	0.003	2.322	.020	*

* indicates significance at $p < 0.05$, ** indicates significance at $p < 0.01$, and *** indicates significance at $p < 0.001$

Table 2 The full set of best-fit parameters for the omnibus model in Experiment 2.

Fixed Effects	Estimate	*SE*	*Z* value	*p* value	
Intercept	1.507	0.254	5.932	0.000	***
Item Type	−0.073	0.130	−0.565	0.572	
List Type	0.154	0.113	1.362	0.173	
Talker	0.171	0.046	3.726	0.000	***
Item Type × List Type	0.055	0.064	0.857	0.391	
Item Type × Talker	−0.014	0.020	−0.693	0.488	
List Type × Talker	−0.009	0.021	−0.448	0.654	
Item Type × List Type × Talker	0.008	0.010	0.771	0.441	

* indicates significance at $p < 0.05$, ** indicates significance at $p < 0.01$, and *** indicates significance at $p < 0.001$

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, UK: Edinburgh University Press.

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). New York, NY: John Wiley & Sons.

Allen, J. S., & Miller, J. L. (2004). Listener sensitivity to individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America, 115*, 3171–3183.

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, *113*(1), 544–552.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B, & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–135.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, *61*(2), 206–219.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York, NY: Harper & Row.

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.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition, 108*, 633–664.

Creel, S. C., & Tumlin, M. A. (2009). *Talker information is not normalized in fluent speech: Evidence from on-line processing of spoken words*. Paper presented at the 31st Annual Meeting of the Cognitive Science Society, Amsterdam.

Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language, 65*(3), 264–285.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6), 3668–3678.

Daniloff, R., & Moll, K. (1968). Coarticulation of lip rounding. *Journal of Speech, Language, and Hearing Research*, *11*(4), 707–721.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17–22.

Dufour, S., & Nguyen, N. (2016). Does talker-specific information influence lexical competition? Evidence from phonological priming. *Cognitive Science*. Advance online publication. doi:https://doi.org/10.1111/cogs.12472

Eager, C., & Roy, J. (2017). *Mixed effects models are sometimes terrible*. Retrieved from Cornell University Library: https://arxiv.org/abs/1701.04858

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science, 1*(1), 54–75.

Fox, N. P., Reilly, M., & Blumstein, S. E. (2015). Phonological neighborhood competition affects spoken word production irrespective of sentential context. *Journal of Memory and Language*, *83*, 97–117.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of personal construal. *Psychological Review, 118*, 247–279.

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks.* New York, NY: Psychology Press.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38*(7), 833–848.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Language, Memory, and Cognition, 22*(5), 1166–1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Goldinger, S. D., & Azuma, T. (2003). Puzzle-solving science: The quixotic quest for units in speech perception. *Journal of Phonetics, 31*, 305–320.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28*, 501–518.

Halle, M., & Stevens, K. (1971). A note on laryngeal features. *Quarterly Progress Report of the Research Laboratory of Electronics, 101*, 198–213.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570–1582.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446.

Jakobson, R., Fant, C. G., & Halle, M. (1951). *Preliminaries to speech analysis: The distinctive features and their correlates.* Cambridge, MA: MIT Press.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*(4), 485–499.

Kalvitis, D. (2000). *The greatest dot-to-dot book in the world* (Book 1). New York, NY: Monkeying Around.

Kessinger, R. H., & Blumstein, S. E. (1998). Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics*, *26*(2), 117–128.

Kittredge, A., Davis, L., & Blumstein, S. E. (2006). Effects of nonlinguistic auditory variations on lexical processing in Broca's aphasics. *Brain and Language, 97*, 25–40.

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language, 56*, 1–15.

Lee, C. Y., & Zhang, Y. (2015). Processing speaker variability in repetition and semantic/associative priming. *Journal of Psycholinguistic Research*, *44*(3), 237–250.

Lee, C. Y., & Zhang, Y. (2017). Processing lexical and speaker information in repetition and semantic/associative priming. *Journal of Psycholinguistic Research*, 1–14. Advance online publication. doi: https://doi.org/10.1007/s10936-017-9514-

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431.

Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory and Cognition, 26*(4), 708–715.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology, 33*(2), 391–409.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63.

Mattys, S. L., & Liss, J. M. (2008). On building models of spoken-word recognition: When there is as much to learn from natural "oddities" as artificial normality. *Perception & Psychophysics, 70,* 1235–1242.

McClelland, J. L. (1995). Constructive memory and memory distortions: A parallel-distributed processing approach. In D. Schacter (Ed.), *Memory distortion: How minds, brains, and societies reconstruct the past* (pp. 69–90). Cambridge, MA: Harvard University Press.

McLennan, C. T., & González, J. (2012). Examining talker effects in the perception of native- and foreign-accented speech. *Attention, Perception, & Psychophysics*, *74,* 824–830.

McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 306–321.

McLennan, C., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 306–321.

McNamara, T. P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review, 99*(4), 650–662.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York, NY: Psychology Press.

Mertus, J. A. (2002). *BLISS: The Brown lab interactive speech system* [Computer software]. Providence, RI:Brown University. Retrieved from http://mertus.org

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word-recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 98–118). London, UK: Academic Press.

Miller, J. L., & Baer, T. (1983). Some effects of speaking rate on the production of /b/ and /w/. *The Journal of the Acoustical Society of America*, *73*(5), 1751–1755.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, *27*(2), 338–352.

Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & Cognition*, *37*(7), 1026–1039.

Mirman, D., McClelland, J. L., Holt, L. C., & Magnuson, J. S. (2008). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science, 32*, 398–417.

Mullenix, J. W., & Howe, J. N. (1999). Selective attention in perceptual adjustments to voice. *Perceptual and Motor Skills, 89*, 447–457.

Myers, E. B., & Theodore, R. M. (2017). Voice-sensitive brain networks encode talker-specific phonetic detail. *Brain and Language*, *165*, 33–44.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms.* Retrieved from http://w3.usf.edu/FreeAssociation/

Newman, R. S. (2016). Lexical access across talkers. *Language, Cognition, and Neuroscience, 31*, 709–727.

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics*, *45*(4), 279–290.

Nygaard, L. C., & Queen, J. S. (2008). Communicating emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 1017–1030.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1995). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception & Psychophysics*, *57*(7), 989–1001.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 309–328.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). San Diego, CA: Academic Press.

Pufahl, A., & Samuel, A. G. (2014). How lexical is the lexicon? Evidence for integrated auditory memory representations. *Cognitive Psychology*, *70,* 1–30.

R Core Team. (2016). *R: A language and environment for statistical computing* [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Roediger, H. L., Bolota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington DC: American Psychological Association.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814.

Roediger, H. L., McDermott, K. B., Pisoni, D. B. & Gallo, D. A. (2004). Illusory recollection of voices. *Memory*, *12*(5), 586–602.

Scarborough, R. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. *Laboratory Phonology*, *10*, 557–586.

Schacter, D. L., & Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 915–930.

Shankweiler, D., Strange, W., & Verbrugge, R. (1975). Speech and the problem of perceptual constancy. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 315–345). Hillsdale, NJ: Erlbaum.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*(2), 173–189.

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language*, *40*, 83–108.

Sumner, M. (2015). The social weight of spoken words. *Trends in Cognitive Sciences, 19*(5), 238–239.

Sumner, M., & Kataoka, R. (2013). Effects of phonetically-cued talker variation on semantic encoding. *Journal of the Acoustical Society of America, 134*, EL485–EL491.

Sumner, M., Kim, S. K., King, K. & McGowan, K. (2014). The socially-weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology, 4*, 1–13.

Syrdal, A. K. & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America, 79*, 1086–1100.

Theodore, R. M., Blumstein, S. E., & Luthra, S. (2015). Attention modulates specificity effects in spoken word recognition: Challenges to the time-course hypothesis. *Attention, Perception, & Psychophysics*, *77*(5), 1674–1684.

Theodore, R. M., & Miller, J. L. (2010). Characteristics of listener sensitivity to talker-specific phonetic detail. *Journal of the Acoustical Society of America, 128,* 2090–2099.

Trude, A. M., & Brown-Schmidt, S. (2012). Talker-specific perceptual adaptation during online speech perception. *Language & Cognitive Processes, 27*(7/8), 979–1001.

Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–591.

Zhang, C., Pugh, K. R., Mencl, W. E., Molfese, P. J., Frost, S. J., Magnuson, J. S., … Wang, W. S. (2016). Functionally integrated neural processing of linguistic and talker information: An event-related fMRI and ERP study. *NeuroImage*, *124*, 536–549.