Voice-specific effects in semantic association

Ed King (etking@stanford.edu)
Department of Linguistics, Margaret Jacks Hall, Bldg. 460
Stanford, CA 94305 USA

Meghan Sumner (sumner@stanford.edu)
Department of Linguistics, Margaret Jacks Hall, Bldg. 460
Stanford, CA 94305 USA

Abstract

Benefits to lexical access are provided by acoustically-cued speaker characteristics (such as gender and age), but little work has investigated these effects in meaning-based tasks. Word recognition is affected both by a word's base-level activation and by associative spread of activation among words, and is correlated with speed of lexical access. In a free association task and a semantic priming task, we find off-line and on-line evidence of speaker-specific relationships between words. Our results suggest the need to extend existing models of spoken word recognition to include interactions between linguistic information and social information that is cued by variation in speech.

Keywords: linguistics; speech perception; spoken word recognition; semantic priming; free association

Introduction

Over the past thirty years, researchers in speech perception have established that, rather than filtering out the phonetic details of incoming speech, listeners utilize these specific phonetic cues when recognizing words. Listeners remember studied words better when they are produced by the same speaker (Goldinger, 1996), a speaker of the same gender (Schacter & Church, 1992), or at the same rate (Bradlow, Nygaard, & Pisoni, 1999) as when they were learned. Listeners shift their perception of phoneme boundaries depending on audio or visual cues to speaker sex (Johnson, Strand, & D'Imperio, 1999) or speaker dialect (Niedzielski, 1999; Hay & Drager, 2010).

In a different domain, we know that word recognition is faster following a related word than an unrelated word. For example, people are faster to recognize the word NURSE when it was immediately preceded by the related word DOCTOR than when it was immediately preceded by the unrelated word BREAD. Semantic priming effects have been established for the recognition of visual words (e.g, Meyer & Schvaneveldt, 1971) and spoken words (e.g., Radeau, 1983).

The existence of word association effects in spoken words raises the possibility of an interaction between these effects and the aforementioned phonetic specificity effects: specific phonetic cues in spoken words may be able to aid activation of other words. This was argued by, e.g., Johnson, 2006, who proposed a model of exemplar-category resonance: the acoustic signal activates exemplars based on similarity, so acoustic cues to a woman's voice give preferentially more activation to female-produced exemplars than to male-produced exemplars. The activation of all of these exemplars feed into a

social category like gender, which resonates back into all the exemplars linked to that gender. So, hearing a woman say a word eventually activates all exemplars produced by women.

The exemplar resonance model predicts associations between social categories and lexical items, but it cannot handle differences in word association given different phonetically-cued social categories. There is intuitive reason to believe that such an interaction should exist. For example, when hearing the word *princess* spoken by an adult with a British accent, people will probably think of a member of the real-life Royal Family; when hearing *princess* spoken by an American child, they may think of a fictional Disney character. Similarly, the word *clothes*, spoken by a woman, is likely to be more associated with dresses and skirts than the word *clothes* spoken by a man.

Despite the intuition that semantic association should interact with speaker characteristics, there has been relatively little empirical work done to establish whether these effects exist. Neuroscience work has shown that listeners have difficulty incorporating semantic information when a spoken message is inconsistent with perceived speaker identity (e.g., a child saying "I think I might be pregnant") (Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008; see also Creel & Tumlin, 2011).

Taken together, these studies show that listeners use voice characteristics as a context that may lead them to generate expectations about the content of an utterance given a sentential context. In other words, listeners are sensitive to the probability of a word given a specific voice in a specific sentence: as predicted by Johnson, 2006's resonance model, voice provides a context for the recognition of a word. We still do not know, however, whether voice provides a context for the *interpretation* of a word. Given the same word in two different voices, do listeners understand the word differently in voice-specific ways?

This paper tests the hypothesis that words are interpreted in speaker-specific ways. Using a free association task, we establish that, when listeners hear a spoken prompt and are asked what word first comes to mind, responses differ depending on the speaker of the word. With a subset of the prompt-response pairs from the free association task, we then show that this effect appears in on-line spoken word recognition: the speed with which listeners recognize a target (response) word, after hearing a prime (prompt) in a specific voice, improves as a function of the voice-specific associa-

tion strength derived from the free association task.

Combining phonetic detail and semantic relatedness

The idea that speaker-specific phonetic cues may affect semantic interpretation goes back at least to Geiselman and Bellezza (1976), who proposed a "voice connotation hypothesis", in which acoustic cues to speaker sex are used to link words with sex-specific connotations. Across a number of studies, they played listeners sets of sentences spoken by one of two speakers; despite having only been told to remember the sentences, the listeners later performed above chance when asked to identify the sex of the speaker of each sentence. Geiselman and colleagues suggest that this effect is due to specific semantic connotations for men and women.

Other authors have argued for voice-specific semantic associations in similar ways. Creel and Tumlin (2011) used a visual world task to track listeners' eye movements to novel word-item pairs that were previously presented in either a male or female voice. In their test session, when a sentence was spoken by the same speaker as in their learning session (and when the speaker/sentence mapping was one-to-one), listeners looked more quickly to the novel item referred to in the sentence. They argue that this effect is due to semantic encoding of speaker voice – and not due to exemplar memories for specific word/speaker associations – because listeners looked to the novel item during the frame sentence, before the actual novel word was spoken.

Evidence for an interaction between phonetically-cued social characteristics and semantic meaning has also been argued for longer-term associations, as opposed to associations that are learned within the course of an experiment. Van Berkum et al. (2008) presented listeners with a series of sentences that were either consistent or inconsistent with the speaker, such as a woman or man (respectively) saying "I always check my *make-up* before I leave"; they used event-related potential (ERP) monitoring to observe what this speaker-specific semantic consistency looks like at a neural level. They found that listeners exhibit an N400 – a negative ERP spike related to difficulty incorporating semantic information – when a spoken message was inconsistent with perceived speaker identity; this effect was similar to, but smaller than, the N400 seen when processing semantic anomalies.

These studies provide compelling evidence that speaker voice characteristics can affect the processing of the meanings of spoken words. In particular, they suggest that listeners use speaker characteristics to generate expectations about what words will appear in a sentence. They fall short, however, of completely connecting models of acoustically-cued indexical information with psycholinguistic models of semantics, because they do not consider the spread of activation between words. Since associative spread is a crucial part of semantic models, a more complete synthesis would require evidence that speaker characteristics can not only affect expectations about the presence of a word, but can additionally constrain associative interactions between different words. Evidence for this connection is particularly lacking in the current literature because work on this topic has manipulated semantic context by using different sentences; however, since listeners may store sentence-size exemplars (Bybee, 2006), we cannot assume that sentential context provides a semantic context independent of speaker-indexed exemplars, which is necessary in order to examine speaker-mediated interactions between words.

In this paper, we instead propose that the effects of speaker-specific semantic meaning can be best examined by using tasks that specifically target word interpretation: free association and semantic priming. Rather than manipulating both speaker and semantic context (the latter of which may not be independent of speaker context), we hold the baseline semantic context constant by focusing on individual words, and look for speaker-specific interpretations of those words by manipulating speaker context.

Experiment 1

Experiment 1 addresses the question of whether listeners interpret a given word as having different semantic associations depending on the voice of the speaker. We use a word association task (Battig & Montague, 1969), in which listeners hear a prompt word and provide the first word that comes to mind; the frequency with which each response word is provided for a given prompt is a strong reflection of the *associative strength* between the probe and the response (Nelson, McEvoy, & Dennis, 2000).

In our particular free-response word association task, we compare the response frequencies of prompt-response pairs across two speakers. Our hypothesis does not provide a priori predictions about what particular speaker characteristics (age, gender, race, dialect, etc.) will lead to differing semantic association; we thus chose two speakers who differ across many social categories. Speaker J is an African-American man in his early 80s, and speaker M is a White American woman in her late 30s. J was raised in the Southern United States, and M in a Northern US city, raising the possibility of dialect differences, but both produced word tokens in a Mainstream American English register.

Methods

Participants were recruited via Amazon's Mechanical Turk (MT) online survey system, and were directed to a webpage containing an in-house presentation script. After a slide of instructions, participants clicked through a series of individual pages, one for each prompt word; on each page, they clicked a button that played the prompt word, and then typed the first word that came to mind into a text box before continuing to the next slide. Each participant heard either speaker J or speaker M.

Stimuli Speakers J and M each read a list of 262 words; these words were chosen randomly, with no attempt to choose words that would specifically elicit different semantic asso-

ciates (e.g., depending on speaker gender). The stimuli words were a mix of nouns, adjectives, and verbs.

Participants 200 subjects with U.S. IP addresses participated via MT; 100 subjects heard words produced by speaker J, and 100 heard words by speaker M. 9 sets of results were excluded because subjects did not complete the task (all for speaker J), and 4 subjects were excluded for not being native speakers of English (1 from speaker J, 3 from speaker M), leaving a total of 187 sets of responses (90 to speaker J, 97 to speaker M). The remaining participants had a median age of 31 years and were 55% female (with marginally more women than men responding to speaker J).

Data clean-up Responses were spell-checked via a semi-automated process: a Python script automatically spellchecked the responses while outputting a log of changes, then a human annotator reviewed the log and manually fixed incorrect changes. Nominal and verbal morphology was removed using the WordNet stemmer in the NLTK Python package (Bird, Klein, & Loper, 2009).

Results

We define the *top associate*, for a given prompt and a given speaker, as the response that was given most frequently to that prompt spoken by that speaker. Overall, 183 prompts (69.8%) resulted in exactly the same top associate set (including ties for the top associate) for both speakers; 203 prompts (77.5%) resulted in top associates (or, including ties, sets of top associates) with at least one response that was the same across speakers. Thus, 22.5% of prompts resulted in different top associates, depending on the speaker.

These differences were difficult to attribute to any one difference in the speakers' voices: a small number of the top-associate differences might be attributable to the different speakers' sexes (e.g., "yeast" in speaker J's voice yielded the top associate "bread", but in speaker M's voice yielded "infection"), but most were relatively uninterpretable (e.g., for J and M, respectively: "question" yielded "answer"/"mark"). Further research into responses that differ across sex, age, dialect region, and other characteristics would be welcome; however, to avoid speculation about the particular differences, and to determine whether the responses were truly speaker-dependent, we analyzed the results at a more general level by randomly resampling responses.

Random resampling of responses The observation that 22.5% of prompt words resulted in a different top associate, depending on the speaker, is not meaningful without a basis for comparison: is this proportion greater than the proportion of responses that would differ within a single speaker, simply due to random variation in the response frequencies? We estimate a baseline difference proportion by randomly resampling from our observed response distributions, both within- and between-speakers. For the within-speaker baselines, we take all of the responses to a given speaker and randomly split them in half (or approximately in half; see below); we

choose two new "top associates" for each prompt based on this split, and compare them to each other to estimate the proportion of different top associates. We contrast this with the across speaker baseline, where we randomly resample from both speaker J and speaker M and compare the new "top associates" across speakers. Because agreement on the top associate increases with the number of subjects in the sample, we always sample subject groups of 45 (the largest possible number, due to the 90 responses to speaker J).

After 1000 iterations of random resampling, we find that the across-speaker differences are robustly larger than the within-speaker differences. The results are displayed in Figure 1. Across-speaker comparisons yield a difference proportion with a mean of 0.283 ($\hat{\sigma} = 0.022$), compared to within-speaker means of 0.274 for J ($\hat{\sigma} = 0.018$) and 0.266 for M ($\hat{\sigma} = 0.021$). The across-speaker difference was significantly higher than the within-speaker differences ($t(1815.0) = 16.05, p < 0.001$). There was also a significant difference in agreement rates across the within-speaker conditions ($t(1954.4) = 9.9, p < 0.001$), with speaker J yielding significantly higher within-speaker disagreement rates than speaker M. This result is corroborated by a log-log model of target frequency by rank, in which speaker M elicits a higher frequency intercept (at the most common responses to her prompts) than speaker J.
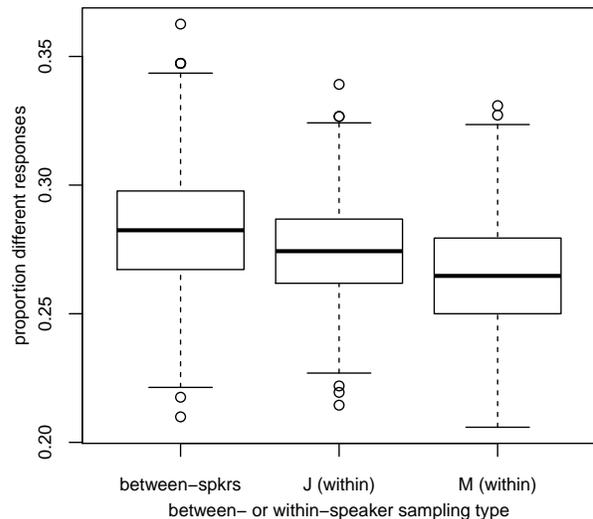

Figure 1: Proportion of different top associates in three types of random resampling: between-speakers, within-speaker (speaker J), and within-speaker (speaker M). The between-speakers condition yields higher disagreement on top associates than either within-speaker condition.

Discussion

The results of this experiment suggest that there are semantic associations that are differentially cued by speaker-specific phonetics. When responses to a set of prompt words are compared across speakers, there is significantly more disagree-

ment (in terms of the most common response) than when responses are compared within each individual speaker.

An unexpected result is that there is more agreement on "what first comes to mind" when a prompt word is spoken by speaker M, relative to when it is spoken by speaker J: The random resampling analysis indicates that, across prompt words, responses to speaker M exhibit fewer differences in what constitutes the most frequent response; listeners are more likely to give the same response to speaker M's prompts, while responses to speaker J's prompts are more varied.

Experiment 2

The results from Experiment 1 support the hypothesis that speaker-specific information influences semantic interpretation, at least in self-reports of what words first came to listeners' minds. The response frequencies derived from this type of free association task are typically thought to be (or, at least, to be related to) the strength of association between words in the mental lexicon. If this is the case, we would expect to find evidence for speaker-specific associations in an online task sensitive to meaning.

In this experiment, we investigate listener reactions to targets when preceded by primes, based on the top associate results found in Experiment 1. We augment this standard cross-modal semantic priming task with our factor of interest: the spoken primes are produced by both of the speakers, and the targets are the speaker-specific responses that were observed in Experiment 1. In other words, we compare listeners' reaction times, in a lexical decision task, to the target "infection" when preceded by the prime "yeast" produced either by speaker M (speaker match) or speaker J (speaker mismatch); we similarly compare reaction times to the target "bread" when primed by "yeast" spoken by M (mismatch) or J (match).

It is important to note that, unlike typical semantic priming studies which compare related and unrelated primes, we are comparing two related primes across speakers. We expect relatedness priming as a baseline, but additionally predict that priming is affected by the association strength (operationalized as the response frequency from Experiment 1) that is specific to the speaker of the prime.

Methods

Participants 48 monolingual speakers of American English participated in this study for pay. The participants were all undergraduate students. None reported hearing-related issues.

Stimuli We chose our prime-target stimuli from the results of Experiment 1, using two criteria: (1) the prime (prompt) yielded different top associate responses, depending on the speaker, and (2) the top associate was given as a response to the prompt by at least 20% of the participants in Experiment 1.

Design We used a cross-modal auditory-visual semantic priming paradigm. Twenty-four critical prime-target triplets

(prime; J target; M target) were created based on the criteria above. The design was within-subject with two experimental conditions (VoiceMatch and VoiceMismatch) and two speakers (J and M); Depending on the trial, listeners heard a prime spoken by J and responded to a target that was a top response to J (J-VoiceMatch) or to M (J-VoiceMismatch); or they heard a prime spoken by M and responded to a target that was a top response to J (M-VoiceMismatch) or M (M-VoiceMatch). Four counterbalanced lists were created to ensure that each target was preceded by a prime in each voice, with no subject responding to any prime or target more than once. Each list of twenty-four critical items was augmented with twenty-four unrelated (control) pairs, and forty-eight non-word targets preceded by a real-word prime.

Procedure Participants were run individually or in groups of 2-3 in a sound-attenuated booth. Each trial consisted of an auditory prime, a 100ms ISI, and a visual target. Listeners were instructed to decide whether the visual target was a word or pseudoword by pressing the correspondingly labeled response button.

Results

Reaction times below 300ms and above 1101 milliseconds (the latter equal to two standard deviations above mean log reaction time) were excluded from all analyses. Initially, log-transformed reaction times were subjected to mixed-effects linear regression with main effects of condition (VoiceMatch v. VoiceMismatch) and speaker (J v. M) and the interaction of condition and speaker; we included a random intercept of prime word, and a random slope of condition. The results of this model were inconclusive: with the exception of the intercept, all t-values were less than 1.0; we therefore cannot reject the null hypothesis that the VoiceMatch condition and the VoiceMismatch condition produce categorically different priming effects.

Two factors led us to consider a second analysis. First, our Experiment 1 found a difference in the associative strength of top targets between speakers J and M; we may see a similar speaker-specific response effect in the lexical decision task. Second, and more importantly, each target word was associated with its prime to some degree, even in the VoiceMismatch condition where the target was not the most highly associated word given that prime and that speaker. In a meta-review of semantic priming experiments, Lucas (2000) suggests that strength and type of word association can affect priming; we therefore want to consider association strength as a continuous measure, and determine whether it has a speaker-specific effect on reaction time.

To account for these two factors, we split our data based on speaker: one data set (553 trials) contained responses to targets preceded by speaker J, and the other (548 trials) contained responses to targets preceded by speaker M. We fit two separate models to each data set: one in which log reaction time is predicted by the strength of the prime/target association in speaker J's voice, and one in which log reaction time

is predicted by the strength of the prime/target association in speaker M's voice. We predict a gradient effect of speaker-specific association strength: J's association strength should linearly improve reaction times to targets following primes spoken by J, and M's association strength should linearly improve reaction times to targets following primes spoken by M. We crucially predict that, despite the correlation between prime/target association strengths across speakers (as calculated by the response frequencies from Experiment 1), we should not observe M's association strength affecting responses to J, or J's association strength affected responses to M.

All models include the maximal random effects, including a random intercept of target and random slopes of association strength (of either one or both speakers, depending on model) by subject; random slopes of target are not justified because each target has only one strength value per speaker. Due to the moderate correlation of association strength across speakers (Pearson's $r = 0.54, T(1099) = 21.5, p < 0.001$) model comparison was conducted using R's `anova()` function: models containing only effects of one speaker's association strength were compared to a full interactive model of both speakers' association strength (with the interaction justified by that model's better fit compared to a non-interactive model, $\chi^2(1) = 4.71, p = 0.03$).

For the speaker J dataset, the model fitting log reaction time to J's prime/target association strength resulted in a log-likelihood of 101.15, and the model fitting log reaction time to M's association strength resulted in a log-likelihood of 101.97; the full model resulted in a log-likelihood of 101.55. When compared to the partial models, the full model did not perform any better (full v. J: $\chi^2(2) = 0.78, p = 0.67$; full v. M: $\chi^2(2) = 0, p = 1$), indicating that neither speaker's association strength contributed anything more than the other's.

For the speaker M dataset, the model fitting log reaction time to J's prime/target association strength resulted in a log-likelihood of 97.157, and the model fitting log reaction time to M's association strength resulted in a log-likelihood of 98.753; the full model resulted in a log-likelihood of 101.108. When compared to the partial model of M's association strength, the full model did not perform significantly better (full v. M: $\chi^2(2) = 4.9, p < 0.1$); however, when compared to the partial model of J's association strength, the full model provided a significant increase in log-likelihood (full v. J: $\chi^2(2) = 7.9, p = 0.02$), indicating that adding the partial effects of M's association strength improves the model containing only the effects of J's association strength. The partial effects of J's and M's association strengths on reaction times to M's voice are displayed in Figure 2.

Discussion

This experiment tested whether listeners responded more quickly to target words when the targets were preceded by a spoken prime when the prime/target pair was the most strongly associated pair for that particular speaker, as compared to when the prime was spoken by a different speaker.
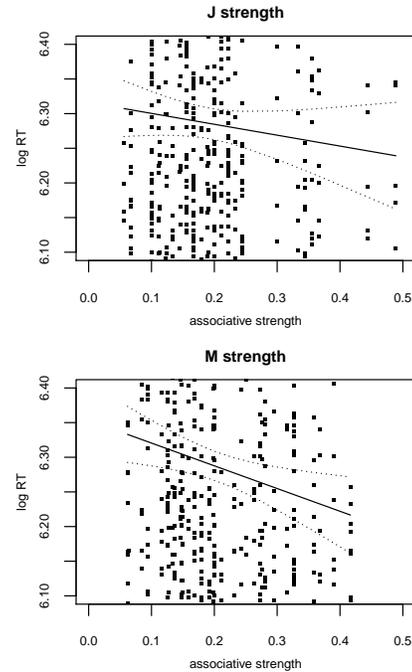

Figure 2: Partial effects of J's association strength (top) and M's association strength (bottom) on log reaction times when primes were spoken in M's voice.

We did not observe the expected categorical effect of voice matching: listeners responded just as quickly to associated prime/target pairs regardless of the specific speaker.

Because of the qualitatively different responses to speakers J and M that we found in Experiment 1, and because of the gradient differences in prime/target association strength across speakers, we fit two sets of gradient models: one in which each speaker's association strength was used to predict reaction times following primes spoken by J, and one in which each speaker's association strength predicted reaction times following primes spoken by M. We observed that speaker-specific association strength significantly improved within-speaker reaction times, but only for speaker M; no model suggested a gradient effect of association strength to speaker J's voice.

General Discussion

The goal of this paper was to determine whether speaker-specific phonetic cues affect the interpretation of spoken words. In two different experiments, we establish that listeners respond to spoken words in speaker-specific ways: in the first experiment, the most common responses to spoken words differed across-speakers to a greater extent than expected; in the second experiment, listeners responded to one of our speakers in a way that depended only on that speaker's specific association strengths from the first experiment. We thus found robust effects of speaker-specific word associations in both off-line (free association) and on-line (semantic

priming) tasks.

Our two speakers differ along many dimensions that are cued by phonetic details in speech – including age, race, gender, and dialect background – making it difficult to interpret the variety of speaker-specific semantic differences we found. One particularly odd effect, consistent across our experiments, is the asymmetry between our two speakers. In the first experiment, speaker M's voice prompted significantly more agreement in the composition of top responses than did speaker J's voice; listeners were more likely to give the same response to prompts spoken by M, and gave more varied responses to prompts spoken by J. We suggest that this difference – particularly the possibility that listeners have fewer unique word associations, and thus fewer semantic competitors, to words spoken by M – explains why association strength played a role in Experiment 2 only for words spoken by M.

A potential explanation for this asymmetry is that our subjects may have more experience with the voice characteristics of M – a younger, white, woman – than with those of J – an older, African-American, man; this additional experience with voices like M's would lead to more robust activation of lexical items and thus to greater priming in M's voice. This interpretation, however, cannot be verified without additional research into how characteristics such as age, race, and gender affect listeners' reactions to these speakers, and a much closer look at how these social characteristics relate to the free responses to M and J. Future work will investigate these characteristics and their effects on word associations in order to better understand the free response results and the cross-task asymmetry between our two speakers.

Our experiments provide evidence for a role of speaker-specific phonetic information in semantic interpretation. Across two experiments, single words robustly prompt different word associations depending on speaker; this interaction cannot be accounted for by standard accounts of semantic priming (which could handle word associations) or standard exemplar-based accounts (which could handle speaker specific effects for individual words). These results require a model of spoken word recognition which explicitly incorporates social information (as encoded by speaker-specific acoustic cues) and linguistic information (including semantic relatedness); a model like that of Sumner, Kim, King, and McGowan (2014), for example, provides a framework for understanding how these two sources of information can interact in spoken word recognition.

References

Battig, W., & Montague, W. (1969). Category norms of verbal items in 56 categories A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80*(3).

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python: Analyzing text with the natural language toolkit*. O'Reilly Media.

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), 206–19.

Bybee, J. L. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, *82*(4), 711–733.

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.

Geiselman, R. E., & Bellezza, F. S. (1976). Long-term memory for speaker's voice and source location. *Memory & Cognition*, *4*(5), 483–9.

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*(5), 1166–83.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4).

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

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359–384.

Lucas, M. (2000). Semantic priming without association: a meta-analytic review. *Psychonomic Bulletin & Review*, *7*(4), 618–30.

Meyer, D., & Schvaneveldt, R. (1971). Facilitation in Recognizing Pairs of Words : Evidence of a Dependence Between Retrieval Operations. *Journal of experimental psychology*, *90*(2), 227–234.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & cognition*, *28*(6), 887–99.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*(1), 62–85.

Radeau, M. (1983). Semantic priming between spoken words in adults and children. *Canadian Journal of Psychology/Revue canadienne . . .*, *37*(4), 547–556.

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

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

Van Berkum, J., van den Brink, D., Tesink, C., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–91.