TALKER VARIABILITY IN LEXICAL ACCESS: EVIDENCE FROM SEMANTIC PRIMING

Yu Zhang & Chao-Yang Lee

Ohio University, USA
yz137808@ohio.edu

ABSTRACT

It is usually assumed that lexical representations are abstract and devoid of detailed information about various sources of acoustic variability. This assumption is evaluated in this study by examining the effect of talker variability on the access to word meaning in a short-term semantic priming experiment. Prime–target pairs that were semantically related (e.g., king–queen) or unrelated (e.g., bell–queen) were produced by the same talker or different talkers. Two inter-stimulus intervals (50 ms and 250 ms) were used between the prime and target to explore the time course of semantic priming. The auditory stimuli were presented to 60 college-age listeners, who performed lexical decision tasks. It was hypothesized that a change in the talker between the prime and target would influence the magnitude of semantic facilitation. Analysis of response accuracy and reaction time showed that the magnitude of semantic priming was attenuated in the different-talker condition, although the effect was obtained only for targets produced by the female speaker. There were no effects involving different inter-stimulus intervals. These results provided partial evidence that talker variability affects lexical access.

Keywords: lexical access, talker variability, semantic priming

1. INTRODUCTION

Theories of speech perception have long assumed that representations of lexical items are highly abstract entities in the mind. Access to the mental lexicon, therefore, has been further hypothesized to be mediated by a speaker normalization process in which all information irrelevant to phoneme discrimination and word distinction is discarded early in speech processing [8]. However, recent research has suggested that detailed information from the acoustic signal, such as that related to talker or phoneme identity, can survive the initial lexical processing stage and may have an impact on spoken word processing and representation (cf. [4]). The present study examines the effect of one type of variability embedded in the acoustic details on lexical access.

Talker variability refers to speaker differences in delivering the same linguistic information. The presence of talker variability has been evidenced by previous research to have impact on speech perception performance, and therefore, speech processing. Research has also suggested a relationship between talker variability and form-based representations of words using the long-term repetition priming paradigm (see [5] for a review). That is, the magnitude of priming has been shown to be attenuated by stimuli delivered by multiple talkers in comparison with single-talker conditions. The implication is that talker-specific information may be embedded in the form-based mental word representations.

Sub-phonetic acoustic variability has been demonstrated to affect access to word meaning [1]. However, little is known about whether variability across talkers can similarly affect access to word meaning. Additionally, the time course of talker variability effect is underspecified. To bridge this gap, the current study employed the short-term semantic priming paradigm to investigate the effect of talker variability on lexical access.

2. METHOD

One short-term semantic priming experiment was conducted using prime–target pairs that were either semantically associated or unrelated.

2.1. Materials

The stimuli were adapted from Andruski, Blumstein, and Burton's study [1], which consisted of 26 words serving as targets in prime–target pairs. Each target (e.g., "queen" produced by a female speaker) was paired with four types of primes: (1) semantically related, same talker (e.g., "king" produced by the female speaker), (2) semantically related, different talker (e.g., "king"

produced by a male speaker), (3) unrelated, same talker (e.g., "bell" produced by the female speaker), and (4) unrelated, different talker (e.g., "bell" produced by the male speaker). In addition, 26 non-word targets were paired with the same set of real-word primes to serve as fillers. In sum, there were 104 word-target trials (26 targets × 4 types of primes) and 104 non-word-target trials for a total of 208 trials. Two test series were created: one with the female speaker as the voice for the target items and the other with the male speaker as the voice for the target items.

2.2. Participants

Sixty students (45 females and 15 males) at Ohio University participated in the experiment. They received partial course credit for their participation. All participants were native speakers of American English with no known hearing, speech, or language disorders. The age of the participants ranged from 18 to 25 years (M = 20, SD = 2). All the students passed a hearing screening in both ears at 20 dB HL at 1000 Hz, 2000 Hz, and 4000 Hz.

2.3. Procedure

The stimuli were recorded by a female and a male, who were native American English speakers from the same town in central Ohio, in a sound-treated booth with an Audio-technica AT825 microphone, connected to a personal computer through a USBPre microphone interface. The average fundamental frequencies for the female and male speaker were 217 Hz and 113 Hz, respectively. The recordings were sampled using the Brown Lab Interactive Speech System (BLISS) [6] at 22050 Hz with 14-bit quantization. Stimulus items were identified from the waveform display using MeV, the waveform editor of BLISS, and were saved as individual sound files after normalizing the peak amplitude of all the stimuli. The average target durations for the female and male speaker were 621 ms and 505 ms, respectively. The inter-stimulus interval (ISI) between the prime and target was specified at 50 ms or 250 ms for the experiment, motivated by considerations of auditory memory and backward masking [1].

The BLISS AVRunner subject-testing program was used for stimulus delivery and response data acquisition. The participants were tested individually in a sound-treated booth. They were instructed to listen to the stimulus pairs and make a decision on whether or not the second item of a pair is a real English word or not. The stimuli were presented through a pair of KOSS R/80 earphones connected to a desktop computer. The participants made their responses by pressing buttons labeled WORD and NONWORD on a computer keyboard. They were also instructed to use their dominant hand to respond as quickly and accurately as possible. Ten practice trials were provided prior to the actual experiment to familiarize the participants with the response format. Each test session lasted approximately 20 minutes. The order of the stimuli was randomized for each participant.

2.4. Data analysis and predictions

Information on the response accuracy and reaction time was acquired by BLISS automatically. Reaction time was measured from the onset of targets. Only correct responses were included in the reaction time analyses. Reaction times exceeding two standard deviations from the means were excluded from the reaction time analyses. Non-word target responses were not analyzed.

Mixed model ANOVAs were performed on the accuracy and reaction time data. The within-subjects factors include prime–target semantic relationship (related and unrelated) and prime–target talker relationship (same and different). The between-subjects factors include ISI (50 ms and 250 ms) and target voice (female and male). Both analyses by subject (F_1) and by item (F_2) were performed on the data.

Based on previous research, we predicted that there would be a semantic priming effect, meaning that responses to the semantically related trials would be faster and more accurate than those to the unrelated trials. It was also predicted that there would be a main effect of talker variability. That is, participants would respond to the same-talker trials faster and more accurately than the different-talker trials. Further, it was expected that there would be an interaction effect between semantic relationship and talker relationship. Compared with the different-talker trials, the amount of semantic facilitation reflected by the difference between the two semantic relationship conditions would be greater for the same-talker trials. Finally, the interaction effect was expected to appear at the ISI of 250 ms.

3. RESULTS

The results of the experiment are displayed in Table 1 and Table 2.

Table 1: Mean accuracy of lexical decision responses (in % with SD, n = 15) in the semantic priming experiment.

ISI (ms)	Target voice	Same speaker		Different speakers	
		Semantic	Unrelated	Semantic	Unrelated
50	Female	96 (4)	95 (5)	97 (4)	96 (4)
	Male	98 (3)	95 (3)	98 (2)	96 (4)
250	Female	99 (2)	94 (5)	97 (3)	96 (4)
	Male	98 (2)	95 (3)	97 (3)	95 (3)

Table 2: Mean reaction time of lexical decision responses (in ms with SD, n = 15) and magnitude of facilitation (MOF) in the semantic priming experiment.

ISI (ms)	Target voice	Same speaker			Different speakers		
		Related	Unrelated	MOF	Related	Unrelated	MOF
50	Female	857 (90)	942 (110)	85	848 (113)	915 (101)	67
	Male	775 (104)	834 (115)	59	774 (106)	856 (88)	82
250	Female	892 (144)	989 (139)	97	881 (141)	945 (132)	64
	Male	823 (171)	869 (135)	46	816 (156)	899 (156)	83

3.1. Accuracy

The ANOVA on the arcsine transformed response accuracy data revealed one significant effect. The main effect of semantic relationship between the prime and target was significant [$F_1(1, 56) = 21.02$, $p < .001$, $\eta_p^2 = .27$; $F_2(1, 100) = 15.73$, $p < .001$, $\eta_p^2 = .14$]. Participants were more accurate in their responses to the target words that were preceded by the semantically related primes than those preceded by the unrelated primes (97.4% and 95.3% correct, respectively). This is consistent with the semantic priming prediction. However, the other predictions were not evidenced by the accuracy analysis.

3.2. Reaction time

The ANOVA on the reaction time data revealed several significant effects. The main effect of semantic relationship between the prime and target was significant [$F_1(1, 56) = 239.68$, $p < .001$, $\eta_p^2 = .81$; $F_2(1, 100) = 118.34$, $p < .001$, $\eta_p^2 = .54$]. This is shown in Fig. 1. As expected, the reaction time was shorter for participants in their responses to the semantically related trials. However, the main effect of talker relationship was not significant. The participants responded equally fast to both the same-talker trials and the different-talker trials.

Figure 1: Mean reaction time (± 1 SE) as a function of semantic relationship condition.

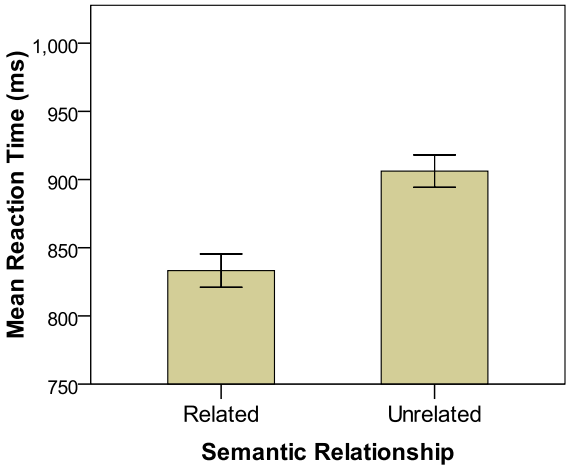


Figure 2: Mean reaction time (± 1 SE) as a function of semantic relationship and talker condition, separated by target voice.

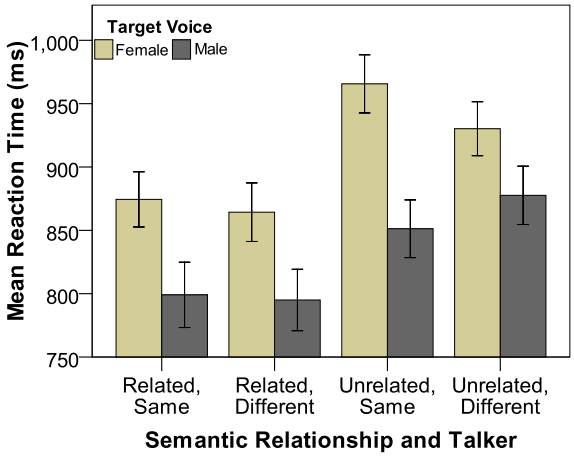
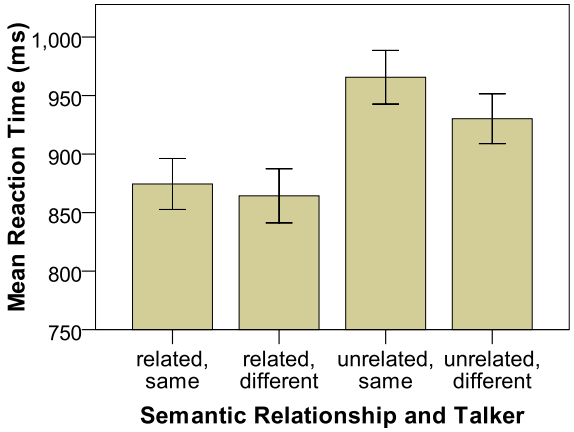


Figure 3: Mean reaction time (± 1 SE) as a function of semantic relationship and talker condition, displayed for female-voice targets.



The interaction between semantic relationship and talker relationship was not significant. Nevertheless, based on subject analysis, there was a significant three-way interaction involving semantic relationship, talker relationship, and target voice [$F_1(1, 56) = 6.65$, $p = .013$, $\eta_p^2 = .11$], shown in Fig. 2. The magnitude of semantic priming was attenuated by the different-talker condition, but only for female-voice target trials, shown in Fig. 3. There were no significant effects involving ISI.

4. DISCUSSION

This study examined the effect of talker variability on lexical access. In particular, how listeners process talker information was investigated, as was the time course of the talker variability effect in semantic priming. Short-term semantic priming was evaluated in an auditory lexical decision task. The experimental data lent partial support for the predictions made earlier.

First, the results exhibit a robust semantic priming effect in terms of both accuracy and reaction time, indicating that semantically associated lexical items activate each other to a greater extent compared to unrelated items. This was the case regardless of the talker condition imposed on the trials. Second, the semantic facilitation in terms of reaction time was attenuated by the introduction of a different talker, although it was true only for the female-target group. This finding provides partial support to the hypothesis that talker-specific information can influence lexical access. In addition, this is consistent with the exemplar-based theory of mental lexicon in which all instances of spoken words form a cloud of exemplars represented in the mind. Third, inconsistent with previous findings [2, 3, 7, 9], the main effect of talker was not significant. Finally, the time course of talker variability was not verified by specifying ISI at 50 ms or 250 ms.

To ensure that the durations of target words were not a confound in the study, another set of analyses was conducted on the reaction time data measured from the offset of targets. The results remained unchanged.

5. CONCLUSION

The effect of processing talker variability, along with other types of variability, in speech perception is not fully understood. Our findings indicate that talker variability is not discarded early in spoken word recognition and can affect access to word meaning. Results from this semantic priming experiment partially support the presence of talker-specific information in the access to word meaning. However, the time course of processing talker variability could not be confirmed because none of the effects containing ISI was significant.

6. REFERENCES

[1] Andruski, J.E., Blumstein, S.E., Burton, M. 1994. The effects of subphonetic differences on lexical access. *Cognition* 52, 163-187.

[2] 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.

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

[4] Luce, P.A., McLennan, C.T. 2005. Spoken word recognition: The challenge of variation. In Pisoni, D.B., Remez, R.E. (eds.), *The Handbook of Speech Perception*. Malden, MA: Blackwell Publishing, 591-609.

[5] Luce, P.A.., McLennan, C.T., Charles-Luce, J. 2003. Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. In Bowers, J.S., Marsolek, C.J. (eds.), *Rethinking Implicit Memory*. New York, NY: Oxford University Press, 197-214.

[6] Mertus, J.A. 2000. *The Brown Lab Interactive Speech System*. Brown University.

[7] 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.

[8] Pisoni, D.B. 1997. Some thoughts on "normalization" in speech perception. In Johnson K., Mullennix, J.W. (eds.), *Talker Variability in Speech Processing*. San Diego, CA: Academic Press, 9-32.

[9] 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, 915-930.