Linguistic effects on talker discrimination: The effect of semantic cohesion

Molly Babel
University of British Columbia
molly.babel@ubc.ca

Chandan Narayan
University of Toronto
chandan.narayan@utoronto.ca

As infants acquire language, they lose the ability to discriminate between non-native sounds (Werker & Tees, 1984). That is, as they acquire meaningful linguistic units, their ability to attend to low-level phonetic detail becomes somewhat attenuated. Contrast this with the fact that adult listeners are sensitive to and retain low-level phonetic detail (e.g., Palmeri et al., 1993). Goldinger (1996) found that the extent to which listeners were able to retain talker-specific information interacted with the type of processing; he found that when engaged in deep levels of linguistic processing (identifying lexical category of word), as opposed to more superficial processing (identifying speaker gender), listeners have a more difficult time recalling talkers' voices. This suggests that attention to a stimulus mediates the encoding of phonetic details of speech episodes in memory. We ask whether linguistic encoding interferes with or facilitates talker discrimination. We describe the effect of *semantic cohesion*, whereby listeners' reaction time (RT) in a talker discrimination task is affected by whether or not real word stimuli pairs form a common compound. Crucially, this effect is present only when listeners are engaged in the processing of linguistic content.

In the first set of experiments, listeners (n=15) were auditorily presented two real words with a 500 ms ISI. An ISI of this length was predicted to allow listeners to engage in more linguistically meaningful encoding (Pisoni, 1973; Pisoni & Tash, 1974). In Experiment 1a listeners' task was to determine whether the two words in each trial were spoken by the same talker or by different talkers. Stimuli pairs (recorded by 3 male and 3 female talkers) formed either semantically *cohesive,* or lexically common compounds (e.g., year-book, scare-crow, etc.) or nonsense compounds (e.g., year-crow, scare-book, etc.). There were main effects (in a linear mixed-effects model) of semantic cohesion (forms a compound or not) [β=0.02, SE=0.009, t=2.66], talker decision (same or different talker) [β=0.07, SE=0.008, t=9.45], and, crucially, an interaction between these two factors [β=-0.34, SE=0.02, t=-2.21]. These results (Fig. 1a) suggest that listeners are doubly primed with both semantic and acoustic expectations, such that the stimulus pairs are expected to be both semantically cohesive and acoustically coherent. Upon hearing the second word in a trial, when both primes are met, processing (here indicated by RT) is facilitated. When stimuli are semantically cohesive but acoustically incoherent, listeners were slower than if neither condition were met. Experiment 1a demonstrates that talker discrimination is affected, in a top-down fashion, by the higher-level linguistic structure imposed by semantic cohesion. Experiment 1b examined whether listeners' judgment of semantically cohesive or nonsense compounds was affected by variation in talker. A group of listeners (n=15) were asked to decide whether the stimuli from Experiment 1a formed semantically cohesive compounds or not. Analyses showed main effects of semantic cohesion [β=-0.13, SE=0.02, t=-8.47] and talker decision [β=-0.02, SE=0.01, t=-2.27], but no interaction (Fig. 1b). The results of Experiment 1 demonstrate that when listeners direct attention to talker differences, they are distracted by meaningful semantic content. However, when attention is directed to semantic content, listeners can successfully filter out talker differences, as evidenced by the lack of an interaction in Experiment 1b.

Our results depend on listeners' ability to detect meaningful linguistic information during the task. We predicted that when listeners are engaged in the task at a more auditory level, they are able to successfully disentangle semantic cohesion and the talker decision. Experiments 2a and 2b tested this by shortening the ISI to 100ms. In all other respects the methods of Experiments 1 and 2 are identical (n= 15 in both tasks). Experiment 2a examines how semantic cohesion influences listeners' (n=15) abilities to discriminate between talkers; there was a main effect of talker decision [β=0.09, SE=0.008, t=1.6], but, as predicted, semantic cohesion did not influence listeners' responses. Experiment 2b explored listeners' ability to determine whether stimuli pairs formed a compound with 100ms ISI. There were main effects of semantic cohesion [β=-0.14, SE=0.02, t=-8.18] and talker decision [β=0.02, SE=0.009, t=2.17], but no interaction.

Taken together, our results suggest that listeners are effectively able to ignore talker-specific phonetic details when attending to higher-level linguistic structure. Higher-level knowledge undermines listeners' abilities to selectively attend to low-level phonetic information like talker identity. This finding is important as it highlights that while phonetic memory has the potential to be highly detailed, the type of attention given to the signal determines whether the sub-phonemic detail is attended to and retained.

Figure 1a. Talker-discrimination processing (RT) results (Exp.1a)

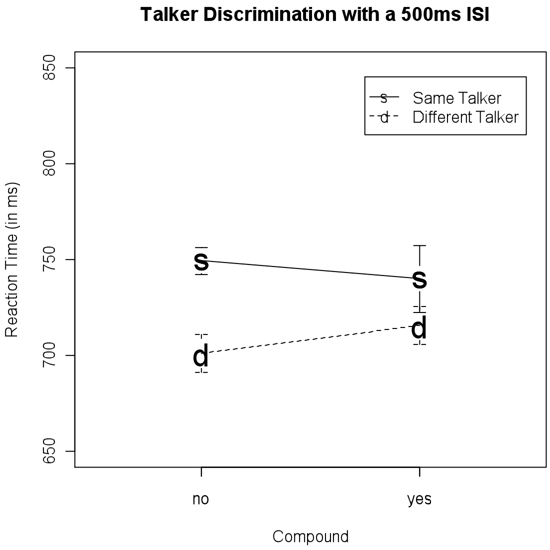


Figure 1b. Compound-decision (RT) results (Exp.1b)

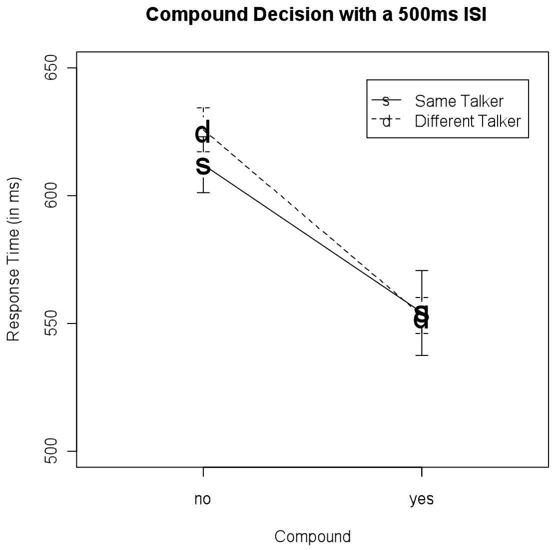


Figure 2a. Talker-discrimination processing (RT) results (Exp.2a)

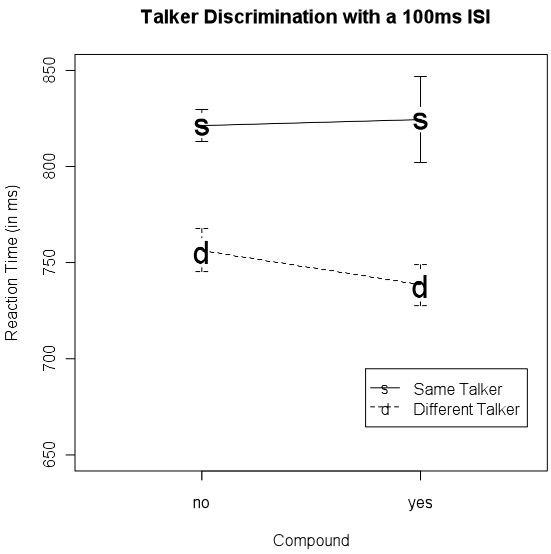
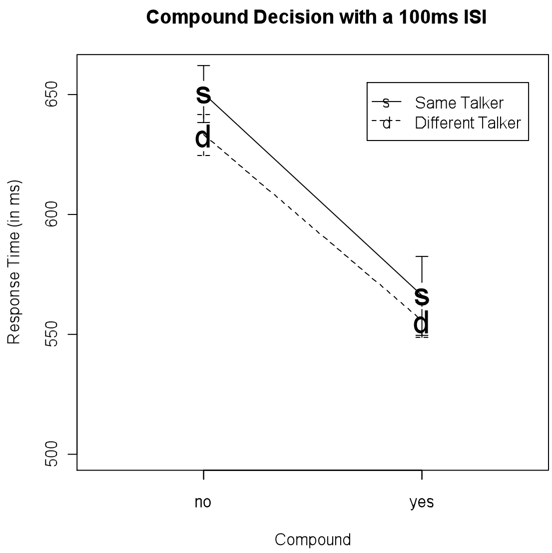


Figure 2b. Compound-decision (RT) results (Exp.2b)



References

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *JEP: Learning, Memory, and Cognition*, 22, 1166-1183.

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

Pisoni, David B. (1973) Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 132, 253–260.

Pisoni, David B. & Jeffrey Tash. (1974) Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics*, 152, 285–290.