Voice-specific effects in semantic association

Ed King and Meghan Sumner, Stanford University

Do listeners incorporate acoustically-cued speaker characteristics when processing semantic associations?

Background

Exemplar theories of speech predict that listeners are sensitive to variation in the acoustic signal.

e.g., recognition benefits for words that match acoustic cues to speaker emotion or age
- Nygaard & Lunders, 2002; Walker & Hay, 2011

Semantic relationships in the lexicon also aid word recognition

e.g., recognition benefit for "NURSE" after "DOCTOR", compared to after "BREAD"
- Meyer & Schvaneveldt, 1971; Radeau et al, 1986

But, so far, **little exploration of the connection** between acoustically-cued speaker characteristics and semantic relationships between words.

Speaker-incongruent words in sentential contexts can lead to slowed looking time, or to N400 ERP spikes
- Creel & Tumlin, 2011; Van Berkum et al, 2008

General Question

Can speaker acoustics provide a semantic-like context for the interpretation of words, outside of a sentential semantic context?

In other words, can we find experimental evidence for these intuitive perceptions:

speaker	word	referent?
British woman	princess	Diana
American girl	princess	Cinderella
nerdy man	princess	Leia

Experiment 1

Question 1
Do listeners interpret words as having different semantic connotations, depending on speaker characteristics?

We use two speakers, **J** and **M**, who differ on a number of characteristics that can be cued by acoustic differences:

speaker	J	M
gender	male	female
age	early 80s	late 30s
race	African American	White

We conducted an **exploratory word-association task** (Battig & Montague, 1969) to determine the top semantic associates to many spoken words.

187 Amazon MT participants heard 262 randomly chosen words, read by either speaker **J** or **M**, and were asked for **the first word that came to mind**.

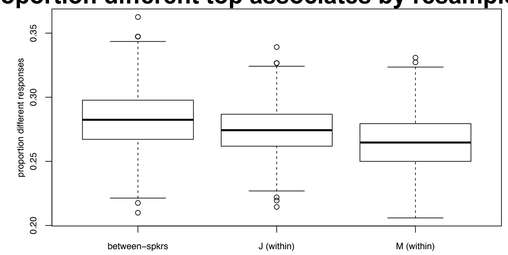
Results

We define the **top semantic associate**, for each speaker-word pair, as the most frequent response to that speaker uttering that word.

Overall, 59 prompt words (22.5%) yielded top associates that differed between speakers, e.g:
- *academy*: **J:** "school" **M:** "awards"
- *pretty:* **J:** "girl" **M:** "pink"

We verified that that across-speaker top associate differences were greater than expected by resampling.
(All comparisons were significant at $p < 0.001$)

Proportion different top associates by resample type



Experiment 2

Question 2
Do listeners interpret words as having speaker-specific interpretations in an on-line task?

Semantic associates like those gathered in Experiment 1 are typically seen as representative of association strength in the mental lexicon.

We therefore expect that **speaker-specific semantic congruence will facilitate word recognition.**

We investigate this with a cross-modal semantic priming task, crossing semantic relatedness with speaker congruence. 48 subjects participated in a lab setting.

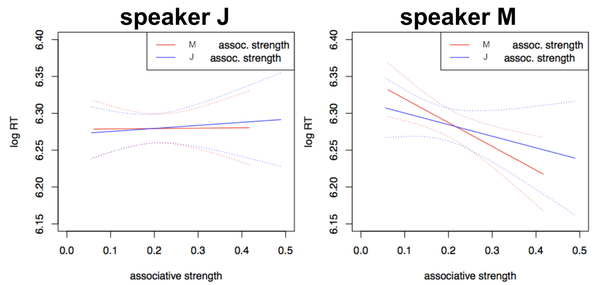
hear	see	condition
J: *academy*	"school"	congruent
J: *academy*	"awards"	incongruent
M: *academy*	"school"	incongruent
M: *academy*	"awards"	congruent

Results

The predicted interaction of speaker and associate was **not supported** by our experimental data.

We then split the data into **two groups, based on speaker voice**, with reaction time predicted by the speaker-specific response frequency from Experiment 1**:**

speaker J **speaker M**



Addition of **M**'s response frequency provides a significant increase in log-likelihood in **M**'s voice ($p = 0.02$).

Discussion

Experiment 1 found that **listeners exhibit different semantic interpretations, depending on speaker**, at a higher rate than expected.

Experiment 2 found that, for one speaker, **reaction times to target words was predictable from speaker-specific prime-target association strength**, when primes were spoken by that speaker.

In both experiments, we observe a **speaker asymmetry:**
- responses to **M** in Exp 1 showed lower within-speaker disagreement than responses to **J**
- association strength effects on reaction time in Exp 2 only appeared in **M**'s voice

We suggest that these asymmetries are related: listeners appear to have fewer unique word responses, and thus fewer semantic competitors, to words spoken by **M**.

Further work is required to determine what aspects of these speakers' vocal characteristics most robustly prompt such effects.

Conclusion

We find evidence that acoustically-cued speaker characteristics can affect the semantic interpretation – not simply the recognition – of spoken words.

Selected References

Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language*, 65(3), 264.
Nygaard, L. & Lunders, E. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition* 30(4): 583.
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.
Walker, A. & Hay, J. (2011). Congruence between 'word age' and 'voice age' facilitates lexical access. *Laboratory Phonology*. 2(1): 219.

Questions? Email Ed: etking@stanford.edu