Talker specificity vs. token specificity in recognition memory

William Clapp[a], Charlotte Vaughn[bc], Simon Todd[d], & Meghan Sumner[a]

[a]Department of Linguistics, Stanford University, [b]Language Science Center, University of Maryland, [c]Department of Linguistics, University of Oregon, [d]Department of Linguistics, UC Santa Barbara

Talker specificity vs. token specificity in recognition memory

William Clapp, Charlotte Vaughn, Simon Todd, & Meghan Sumner

Background

Talker-specificity effects: Listeners recognize words faster and more accurately when repeated in the **same voice** than in a **different voice.** (Palmeri et al., 1993)

Talker-specific information encoded in long-term memory; basis of exemplar theories of speech perception. (e.g., Goldinger, 1998; Pierrehumbert, 2001)

Replicated many times, but same-category repetitions have always used **identical** tokens, raising two issues:

Confound: Do effects reflect memory for talkers? Or memory for precise acoustic patterns? (Viswanathan et al., 2016; Winkler et al., 2002)

Ecological validity: Identical tokens never occur naturally.

Studies have treated talker voices as interchangeable, but memory for talkers may be **socially weighted** and therefore asymmetrical across talkers (e.g., Sumner et al., 2014)

Current Study

Are **talker-specificity effects** just **token-specificity effects** in disguise?
Do **talker asymmetries** emerge more clearly without the identical-token confound?

Two experiments used:

Identical tokens: Same exact sound file played twice.

Novel tokens: Different productions of the same word, recorded in the same space at the same time in the same way.

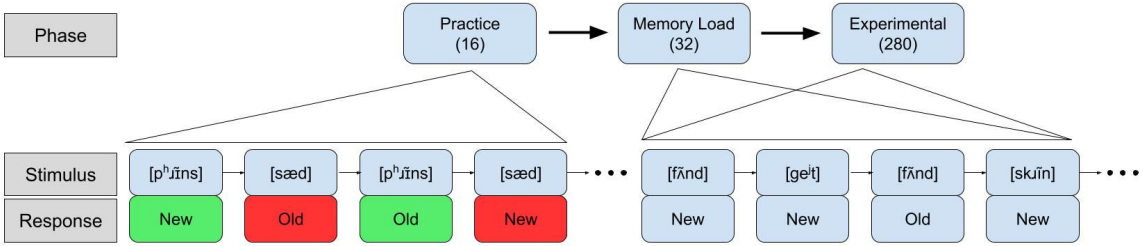
Two talkers: **M1** and **M2** (Both white male Midwesterners in their 20s)

Exp. 1: How does memory for **identical tokens** compare to memory for **novel tokens** in isolation? Is this symmetrical across talkers?

Exp. 2: Are talker-specificity effects **replicable** in the context of novel tokens? Does token type influence the encoding strength of each talker's voice?

Methods

Continuous recognition memory



Hear a string of words, respond **NEW** (first time hearing the word) or **OLD** (second time hearing the word).

Exp. 1: Hear **one** voice (M1 **or** M2, between subjects) and one repetition type (identical or novel, between subjects).

Exp. 2: Hear **both** voices (M1 **and** M2) and one repetition type (identical or novel, between subjects).

Run online via Prolific. Exp. 1: N = 355; Exp. 2: N = 375

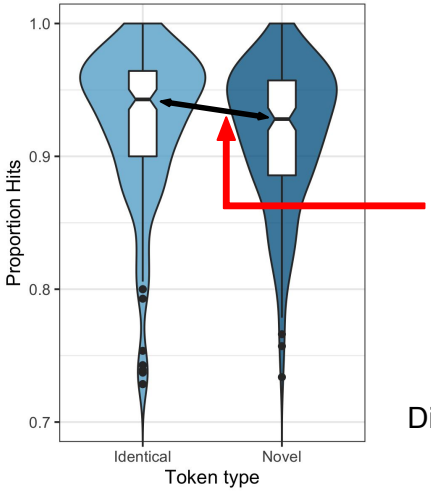
Talker specificity vs. token specificity in recognition memory

William Clapp, Charlotte Vaughn, Simon Todd, & Meghan Sumner

Exp. 1 - Results

How does memory for identical tokens compare to memory for novel tokens?

Talkers pooled:

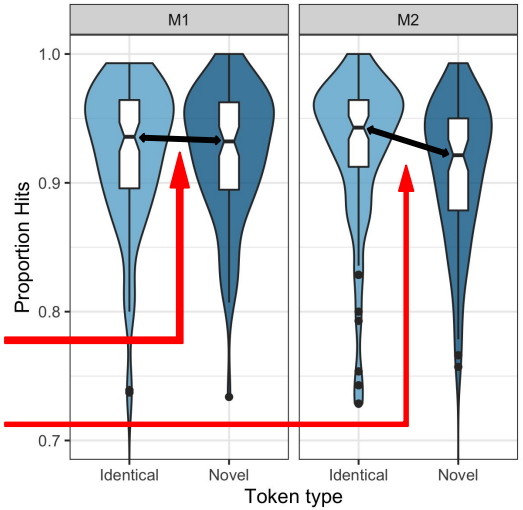


Identical tokens recognized more accurately than novel tokens

- Ref. level: Novel Token
- Main effect of token (β = 0.097, p < 0.01)

Is the influence of token-type symmetrical across talkers?

Talkers separated:



No difference for M1

Difference is driven by large effect for M2.

- Ref. levels: Novel Token, M1
- No main effect of token condition.
- Hit rates lower for M2 than M1 (β = −0.22, p < 0.001)
- Interaction between token and talker (β = 0.27, p < 0.001)

Exp. 1 - Discussion

How does memory for identical tokens compare to memory for novel tokens?

Memory for identical tokens is slightly stronger than memory for novel tokens

Are effects of token-type symmetrical across talkers?

No. Repeating the **same physical stimulus** resulted in higher hit rates than novel tokens for M2, but not for M1.

Exp. 1 was conducted with a single talker conditions to isolate and compare the memory encoding of identical and novel tokens. To investigate talker-specificity, listeners must hear multiple talkers.

Memory for identical tokens is (in some cases) better than memory for novel tokens. Exp. 2 tests whether token type influences the same-voice memory benefit in recognition memory.

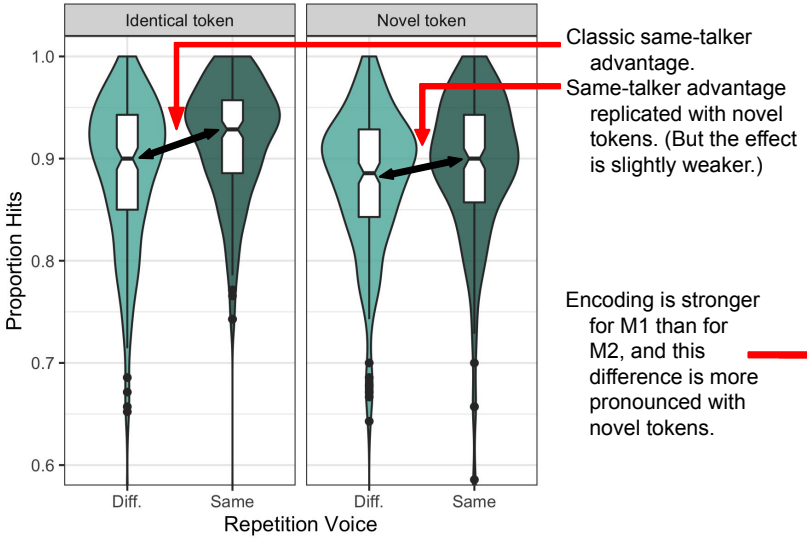
Talker specificity vs. token specificity in recognition memory

William Clapp, Charlotte Vaughn, Simon Todd, & Meghan Sumner

Exp. 2 - Results

Are talker-specificity effects replicable in the context of novel tokens?

Talkers pooled:



Classic same-talker advantage.

Same-talker advantage replicated with novel tokens. (But the effect is slightly weaker.)

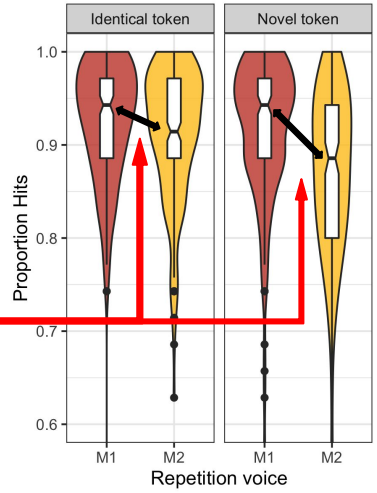
Encoding is stronger for M1 than for M2, and this difference is more pronounced with novel tokens.

- Ref. levels: Different voice, novel tokens
- Hit rates higher for identical than novel tokens ($\beta = 0.17$, $p < 0.05$)
- Higher for same than different voice ($\beta = 0.16$, $p < 0.001$)
- Interaction between rep. voice and token condition ($\beta = 0.21$, $p < 0.001$)

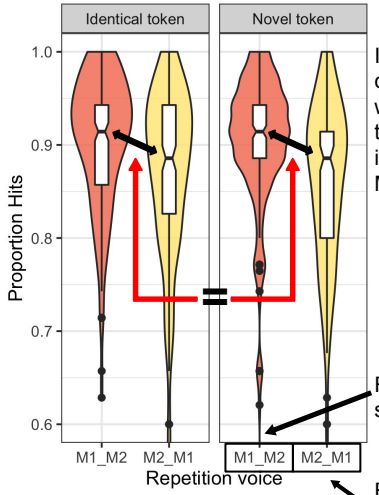
Does token type influence the relative encoding strength of each talker's voice? Are the talkers encoded equally well?

Talkers separated:

Just **Same-Talker** repetitions



Just **Different-Talker** repetitions



In both token conditions, encoding was stronger when the talker heard at a word's initial presentation was M1 rather than M2.

First presentation M1, second presentation M2.

First presentation M2, second presentation M1.

- Ref. levels: Novel Token, First Voice-M1, Second Voice-M1
- Hit rates higher in identical token condition ($\beta = 0.20$, $p < 0.05$).
- Higher when first talker was M1 ($\beta = -0.61$, $p < 0.001$)
- Higher when second talker was M1 ($\beta = -0.20$, $p < 0.01$)
- Interaction between first and second talker ($\beta = 0.32$, $p < 0.001$)
- Interaction between first and second talker and token condition ($\beta = 0.32$, $p < 0.001$)

Talker specificity vs. token specificity in recognition memory

William Clapp, Charlotte Vaughn, Simon Todd, & Meghan Sumner

Exp. 2 - Discussion

Are talker-specificity effects replicable in the context of novel tokens?

Yes, hit rates were higher on same-talker than different-talker trials, even when repetitions were novel tokens. With talkers pooled, effect was slightly weaker with novel tokens than with identical tokens.

Are the talkers encoded equally well?

No, results were variable across talkers. Responses on same and different-voice trials point to better encoding of M1.

Does token type influence the relative encoding strength of each talker's voice?

Yes and No, recognition was equal in identical and novel token conditions for M1, but better in identical than in novel token condition for M2. Listeners were more dependent on low-level acoustic characteristics for M2 than for M1.

General Discussion

Previously observed talker-specificity effects are **not solely** the result of recognition of precise acoustic matches, but talker-specificity and token-specificity have **distinct memory patterns**.

Although the classic talker-specificity effect was replicated with novel tokens in Exp. 2, the pattern of results in both experiments differed between novel and identical tokens.

The use of novel tokens provides a more sensitive measure of memory, and talker-level asymmetries are more likely to emerge when novel tokens are heard.

The range of response patterns was greater for novel tokens than for identical tokens, and differences between M1 and M2 were observed more clearly in the novel token condition. For these reasons, and because of their ecological validity, we suggest that novel tokens are better suited to investigating talker-specificity effects.

Counter to typical assumptions in perception and memory research, all talkers are **not equivalent**. Swapping one voice for another, even within a macro-demographic category, is likely not to yield the same results.

Many studies use stimuli from a single talker to make general claims about human cognition. If Exp. 1 had used only M1 or only M2, results would have been quite different. While the source of this variability is a subject for future research, this study demonstrated that this variability exists. Talkers can be a rich source of variation in data.

References

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

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (1993). Episodic Encoding of Voice Attributes and Recognition Memory for Spoken Words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 309–328.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). John Benjamins Publishing Company.

Sumner, M., Kim, S.K., King, E., & McGowan, K.B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech. *Frontiers in Psychology, 4,* 1015.

Viswanathan, J., Rémy, F., Bacon-Macé, N., & Thorpe, S. J. (2016). Long Term Memory for Noise: Evidence of Robust Encoding of Very Short Temporal Acoustic Patterns. *Frontiers in Neuroscience, 10*, 490.

Winkler, I., Korzykov, O., Gumenyuk, V., Cowan, N., Linkenkaer-Hansen, K., Alho, K., Ilmoniemi, R. J., & Näätänen, R. (2002). Temporary and longer retention of acoustic information. *Psychophysiology, 39*, 530–534.