Same-Talker Benefit? Revisiting Talker-Specificity in Episodic Encoding Across Diverse Voices

William Clapp[a], Charlotte Vaughn[b], & Meghan Sumner[a]

[a]Department of Linguistics, Stanford University [b]Language Science Center, University of Maryland

How does the inclusion of diverse voices in recognition memory experiments change our understanding of talker-specificity effects?

Background

Talker-specificity effect:

- Words recognized faster/more accurately when repeated in same voice than in different voice (Palmeri et al., 1993)
- Phonetic details stored in long-term memory (Goldinger, 1998)

Limitations:

- Homogeneous sets of talkers/listeners (white, college-aged)
- Recent work motivates update to consider talker/group effects

Larger study:

- Diverse talkers (Exp. 1) & homogeneous sets of Black and white talkers/listeners (Exp. 2)
- Social weighting (Sumner et al., 2014) suggest results driven by social factors rather than asocial quantitative exposure

Specific Question for Today:

Does extension to diverse talkers change our understanding of talker-specificity effects?

Methods

Continuous Recognition Memory Paradigm
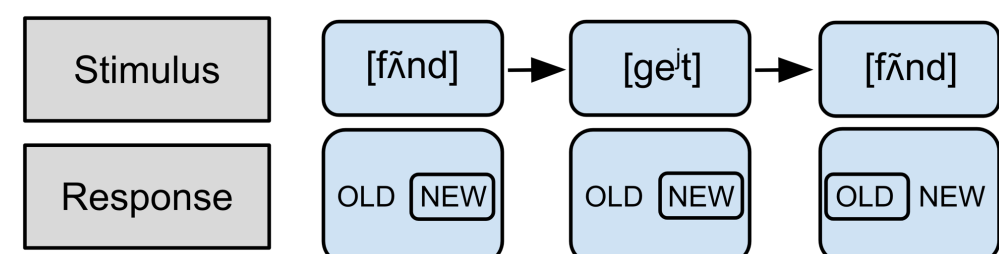
SAME (50%): OLD words repeated in same voice

DIFF (50%): OLD words repeated in different voice

Exp. 1 Stimuli	RACE	GENDER	REGION
WMC1	**White**	**Male**	**CA**
WMC2	**White**	**Male**	**CA**
WFC1	White	Female	CA
WFC2	White	Female	CA
WMA1	White	Male	AL
WMA2	White	Male	AL
WFA1	White	Female	AL
WFA2	White	Female	AL
BMC1	**Black**	**Male**	**CA**
BMC2	**Black**	**Male**	**CA**
BFC1	Black	Female	CA
BFC2	Black	Female	CA
BMA1	Black	Male	AL
BMA2	Black	Male	AL
BFA1	Black	Female	AL
BFA2	Black	Female	AL

Exp. 2 Stimuli	RACE	GENDER	REGION
WMC1	White	Male	CA
WMC2	White	Male	CA
WMC3-8	White	Male	CA
BMC1	Black	Male	CA
BMC2	Black	Male	CA
BMC3-8	Black	Male	CA

Stimuli: Monosyllabic words produced by talkers normed for perceived race, gender, region. Control talkers in **bold**

Trials: 16 practice, 32 memory load, 280 critical

Analysis: Mixed-effects regressions in lme4 for Hits, RTs, FAs and D'. Here, we focus on Hits.

Experiment 1

How do response patterns vary in the context of a diverse talker set?

Participants: 727 native English speakers recruited through Prolific. (Race not controlled; majority white)

Stimuli used: All 16 talkers

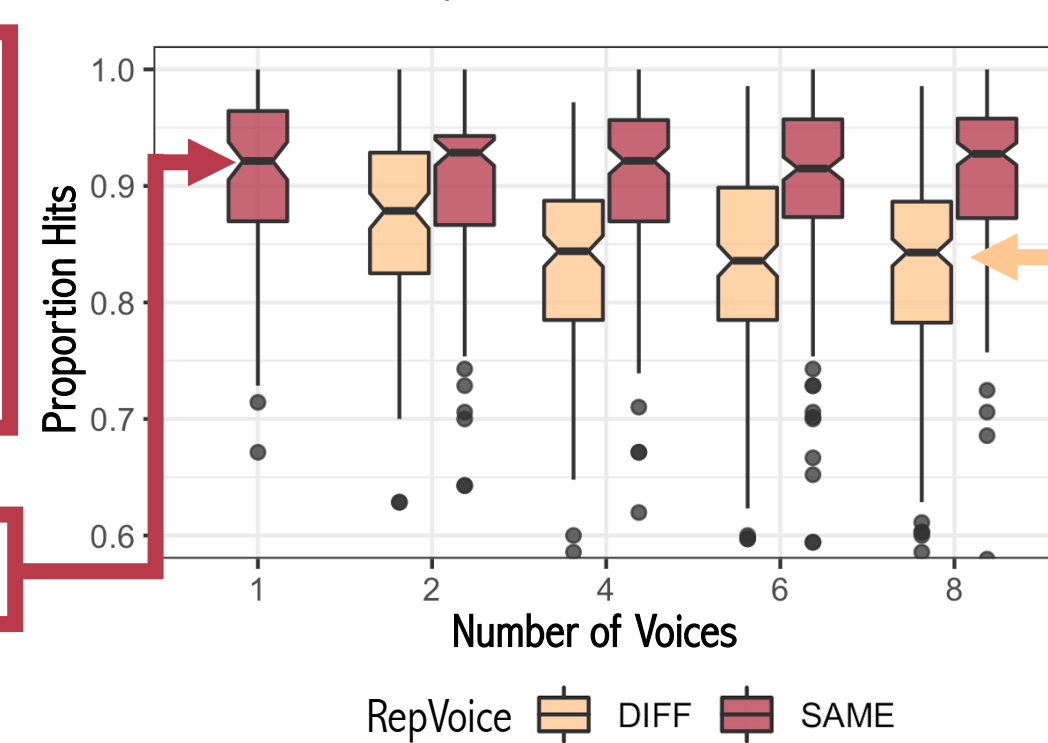**Between-subjects:** Number of voices (NOV) was 1, 2, 4, 6, or 8

Control Talkers: Both WMCs <u>or</u> both BMCs heard in all conditions with 2+ voices. Remainder selected randomly

Within-subjects: RepType (OLD vs. NEW); RepVoice (SAME vs. DIFF, among OLD), LAG (number of intervening trials; 1-65)

Results

All Talkers:
Proportion Hits by RepVoice across NOV

Replication: Listeners more accurate on SAME than DIFF

Ref. level NOV-2 – DIFF
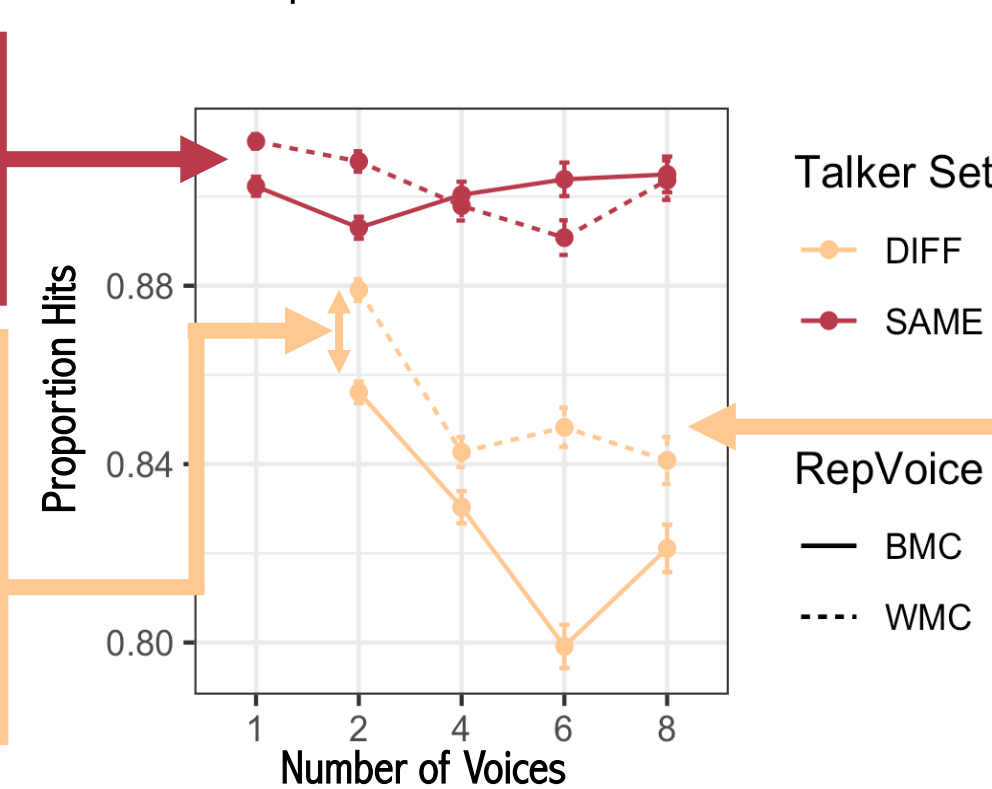SAME; $\beta = 0.58$, $p < 0.001$

SAME words at ceiling

DIFF: Listeners less accurate on NOVs higher than 2

Ref. level NOV-2 – DIFF
NOV-4; $\beta = -0.13$, $p < 0.001$
NOV-6; $\beta = -0.16$, $p < 0.001$
NOV-8; $\beta = -0.14$, $p < 0.001$
NOV-4×SAME; $\beta = 0.32$, $p < 0.01$
NOV-6×SAME; $\beta = 0.37$, $p < 0.001$
NOV-8×SAME; $\beta = 0.42$, $p < 0.001$

RepVoice □ DIFF ■ SAME

Control Talkers: Same voices, different contexts
Proportion Hits RepVoice and Talker Set across NOV

SAME responses stable across NOV and Talker Set

(NOV, Talker Set not sig.)

DIFF: Listeners more accurate on WMC than BMC

Ref. level NOV-2 – DIFF – BMC
WMC; $\beta = 0.098$, $p < 0.01$
WMC × NOV-6 × SAME;
$\beta = 0.098$, $p < 0.001$

DIFF: Accuracy decreased as NOV increased

Ref. level NOV-2 – DIFF – BMC
NOV-4; $\beta = -0.14$, $p < 0.01$
NOV-6; $\beta = -0.19$, $p < 0.001$
NOV-8; $\beta = -0.14$, $p < 0.001$
NOV-4×SAME; $\beta = 0.24$, $p < 0.01$
NOV-6×SAME; $\beta = 0.33$, $p < 0.001$
NOV-8×SAME; $\beta = 0.14$, $p < 0.001$

Talker Set — DIFF — SAME
RepVoice — BMC ···· WMC

Increased heterogeneity in the talker set led to lower accuracy on DIFF but not SAME trials, even for talkers with high accuracy at low NOVs. Different Hit rates for control talker sets in DIFF but not SAME trials.

Experiment 2

How do responses vary when talker sets differ by (perceived) race?

Participants: 680 native English speakers recruited through Prolific. 340 self-identified Black, 340 self-identified white
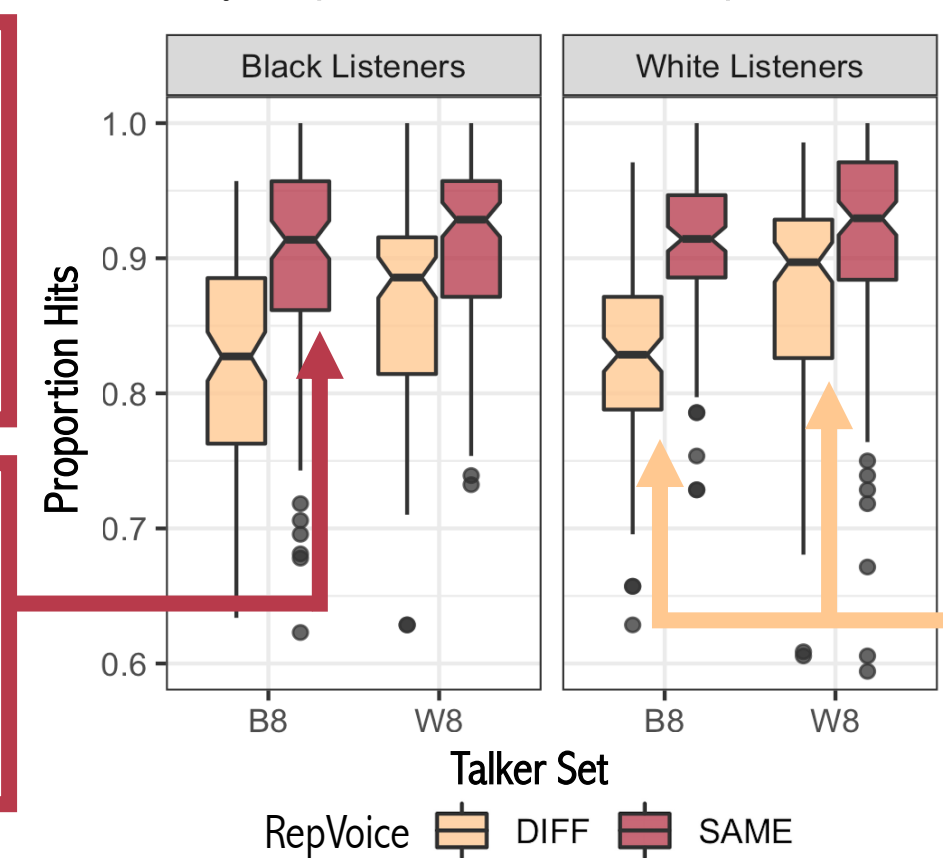
Stimuli used: 8 BMCs and 8 WMCs

Between-subjects: Talker Set: 8 BMC (B8), 8 WMC (W8), & 4 BMC/4 WMC (B4W4)

Within-subjects: RepType (OLD vs. NEW); RepVoice (SAME vs. DIFF), Lag (number of intervening trials; 1-65)

Results

B8 and W8:
Proportion Hits by RepVoice, Listener Population, and Talker Set

Replication: Listeners more accurate on SAME than DIFF

Ref. level B8 – BL – DIFF
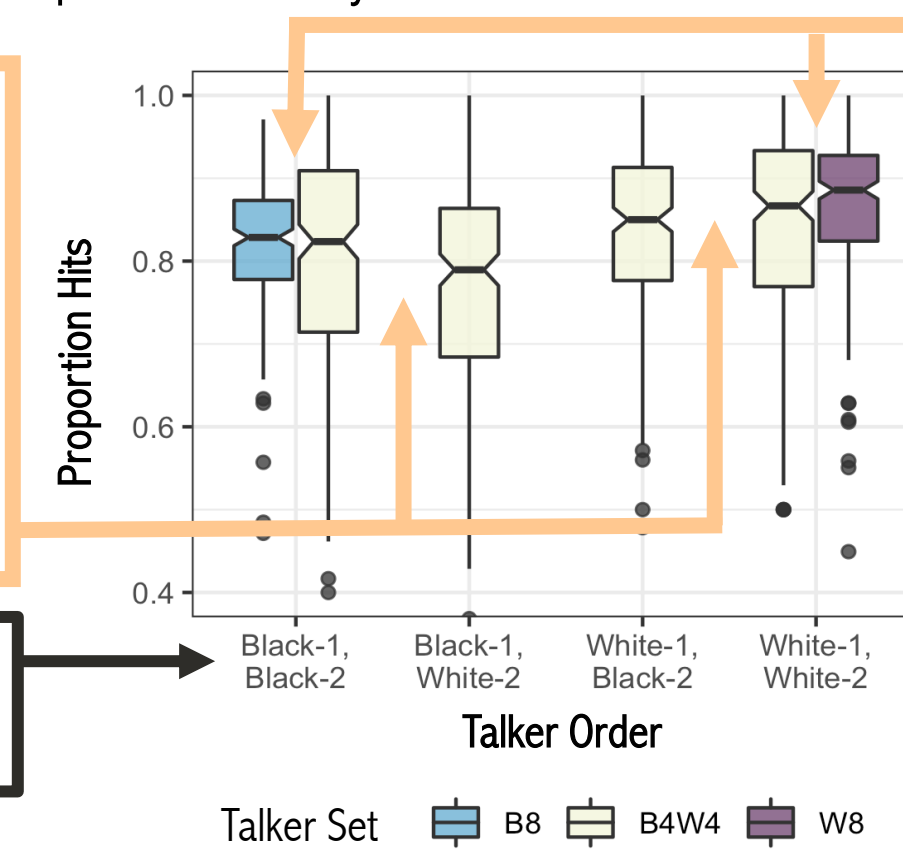SAME; $\beta = 0.67$, $p < 0.001$

SAME near ceiling

Only B8 – BL – SAME sig.
Rel. to W8 – BL – SAME
B8: $\beta = 0.26$, $p < 0.001$
W8-BL, B8-WL, and W8-WL N.S.

DIFF: Listeners more accurate on W8 than B8; WL than BL in B8

Ref. level B8 – BL – DIFF
W8; $\beta = 0.22$, $p < 0.001$,
WL; $\beta = 0.069$, $p < 0.01$
W8 × WL; $\beta = -0.20$, $p < 0.001$

| Black Listeners | White Listeners |

RepVoice □ DIFF ■ SAME

DIFF trials only:
Proportion Hits by Talker Set and Talker Order

Talker Order: Listeners more accurate when first talker was white than Black

Ref. level B4W4 – Black-1, Black-2
White-1; $\beta = 0.24$, $p < 0.01$
White-2; $\beta = -0.19$, $p = 0.05$
White-1, White-2; $\beta = 0.31$, $p < 0.001$

Talker Set: Listeners more accurate on same-race talkers in B8/W8 than B4W4

Ref. level B4W4 – Black-1, Black-2
B8; $\beta = 0.15$, $p < 0.01$
Rel. to B4W4 – White-1, White-2 (post-hoc)
W8; $\beta = 0.16$, $p < 0.01$

Talker Order: "Black-1, Black-2" means word was repeated by a **different** Black talker, etc.

Talker Set □ B8 □ B4W4 □ W8

When talker sets differed by race, listeners were more accurate responding to OLD words presented in different voices for white talkers than for Black talkers, while accuracy on words presented in the same voice were near ceiling.

Discussion

Classic effect replicated: Words repeated in the same voice recognized more accurately than those in a different voice.

High-level findings:

- Regardless of talker, listeners recognize words repeated in the same voice equally well.
- Including diverse voices illustrated new patterns in data, concentrated in different-voice trials:
  - The magnitude of talker-specificity effects is greater than previously reported.
  - Listeners biased toward voices perceived as standard.
  - Similar patterns were observed across listener groups.
- Listener behavior depends on voices and context.

What drives these effects?

- Memory for surface form vs. memory for lexical item.
- **Standard language ideology:** bias towards the language of those in positions of power (Lippi-Green, 2012).
- Similar patterns among Black/white listeners: asymmetries **do not result** from **different amounts of experience** or **perceptions of typicality.**
- Near-ceiling SAME may reflect confound between repetition of **voice** and of **token** (Clapp, et al., *under review*).

Conclusions

Same-talker benefit is better framed as **different-talker cost.**

Homogeneous talker populations in past recognition memory work have **obscured nuance** in talker-specificity effects.

Talkers are not interchangeable: We can't swap one voice for another and expect the same results.

References

Clapp, W., Vaughn, C., Todd., S., & Sumner, M. (*under review*). Talker-Specificity and Token-Specificity in Recognition Memory.
Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychol. Rev.* 105, 251–279.
Lippi-Green, R. (2012). English with an Accent: Language, Ideology, and Discrimination. London; New York: Routledge.
Palmeri, T. I., 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.
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.