

Conflicting Talker Information May Hinder Convergence in Synchronous Speech



☑ Orhun Uluşahin^{1,2} Hans Rutger Bosker^{3,1} Antje S. Meyer^{1,3} James M. McQueen^{3,1}

1) Max Planck Institute for Psycholinguistics, Nijmegen, NL 2) International Max Planck Research School for Language Sciences, Nijmegen, NL 3) Donders Institute for Brain, Cognition and Behaviour, Nijmegen, NL

Background

- Knowing how someone speaks shapes speech perception^[1], but research on role of talker information in production is limited
- Convergence is the tendency of interlocutors to sound more like each other over time^[2], and theories that argue for shared representations between perception and production^[3,4] predict an impact of talker information on convergence
- Synchronous speech tasks can reliably measure convergence to F0 (pitch) in laboratory settings using pitch-shifted stimuli^[5]

RQ1: Do speakers use talker information during production in a synchronous speech convergence task?

RQ2: Does knowing how a talker typically speaks help or hinder convergence?

Participants

In each experiment, 32 (different) healthy female native Dutch speakers

Randomly assigned to one of two groups (N =16):

- High F0 model talker
- Low F0 model talker

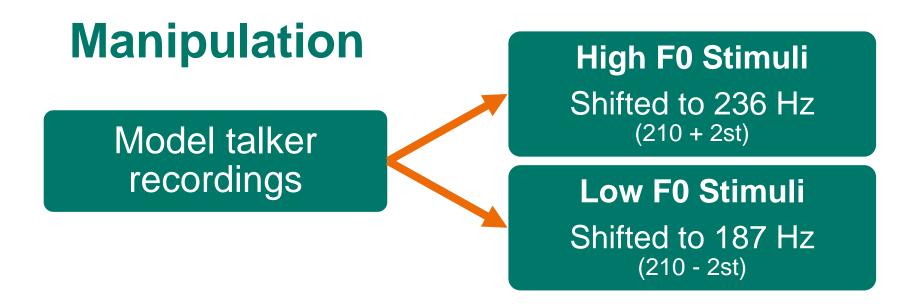
Participants read at their own pace

after a 3-second countdown

(sentence visible during countdown).

Stimuli

85 Dutch sentences in 3 sets. Sets A & B had 40 sentences each, and set C had 5 sentences.



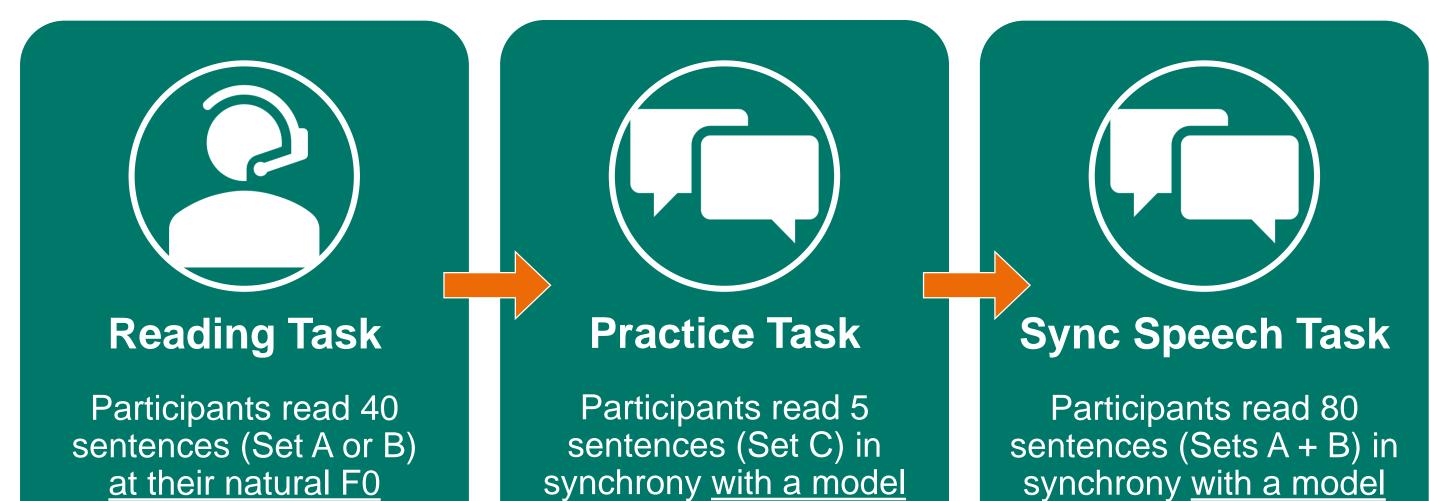
talker with high or low F0

Participants read in synchrony with

the model talker after a 3-second

countdown (sentence visible during countdown).

Experiment 1 – Baseline Convergence



talker with high or low F0

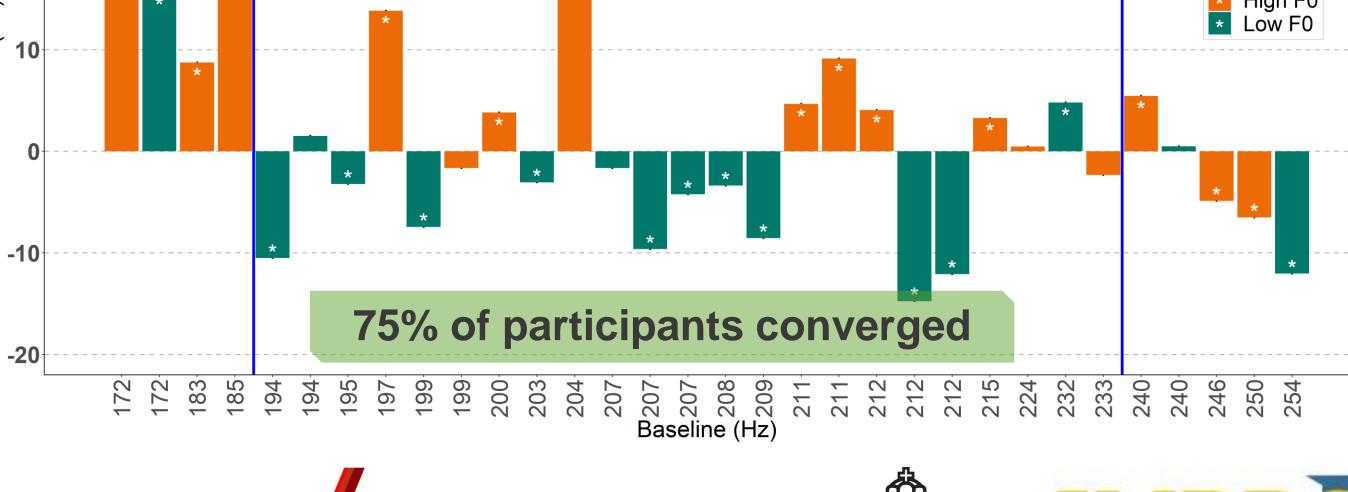
Participants read in synchrony with

the model talker after a 3-second

countdown (sentence visible during

countdown).

Mean F0 distance from baseline per group Mean reading baselines per F0 group + 5.92 Hz Group Group **Comparable baselines** - 4.25 Hz High F0 Low F0 High F0 Low F0 20 Trial (sentence) number Trial (sentence) number Per-participant distances from baseline Group High F0



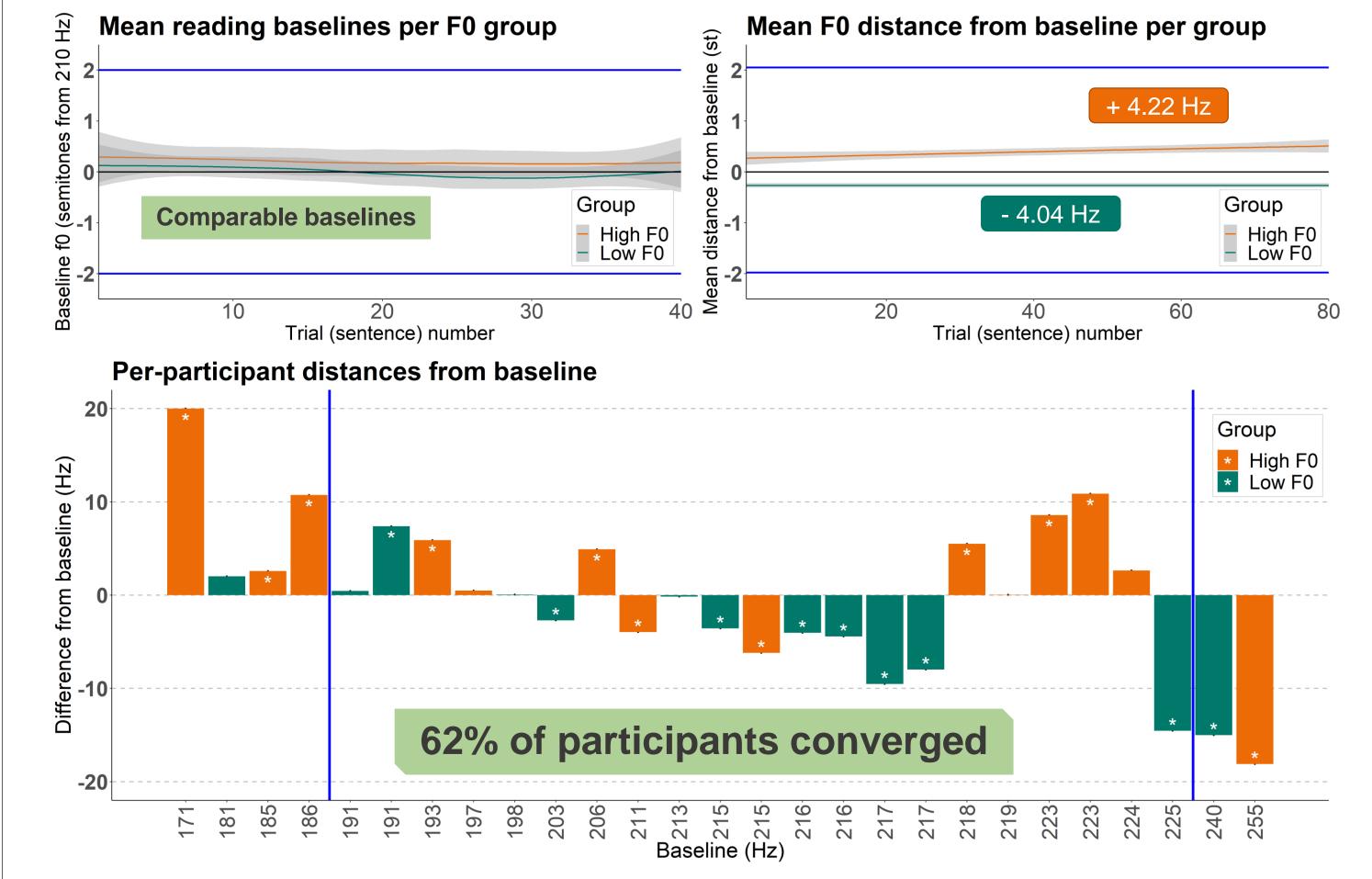
References [1] Holmes, E., To, G., & Johnsrude, I. S. (2021). How long does it take for a voice to become familiar? Speech intelligibility and voice recognition are differentially sensitive to voice training. Psychological Science, 32(6), 903-915. [2] Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. Attention, Perception, & Psychophysics, 79, 637-659. [3] McQueen, J. M., & Meyer, A. S. (2019). Key issues and future directions: Towards a comprehensive cognitive architecture for language use. In Human language: From genes and brain to behavior (pp. 85-96). MIT Press

[4] Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and brain sciences*, 36(4), 329-347.

Experiment 2 – Convergence to Conflicting Talker Listening Task Practice & Sync Reading Task Identical to Exp 1 Participants listen to the Identical to Exp 1 model talker for 20 Model talker at reversed F0 minutes at group F0 Mean reading baselines per F0 group Mean F0 distance from baseline per group + 5.09 Hz Group - 2.39 Hz **Comparable baselines** High F0 Low F0 High F0 Low F0 Trial (sentence) number Trial (sentence) number Per-participant distances from baseline High F0 * Low F0 56% of participants converged

Experiment 3 – Convergence to Congruent Talker





Conclusion

- Participants consistently converged to F0 across three synchronous speech experiments
- Conflicting talker information led to a **numerically smaller** proportion of convergers than congruent talker information

Talker information may be used in speech production, suggesting involvement of shared representations or mechanisms across production and perception







[5] Bradshaw, A. R., & McGettigan, C. (2021). Convergence in voice fundamental frequency during synchronous speech. PloS one, 16(10), e0258747.