Sources of enhanced sentence recognition memory for native and non-native listeners

Sandie Keerstock, Rajka Smiljanic
Department of Linguistics, The University of Texas at Austin

Introduction

Background

Intelligibility and memory benefit of Clear Speech (**CS**)

- Word recognition in noise higher for clear speech than conversational speech (**CO**) for native and non-native listeners [1,2].
- Sentence recognition memory (**SRM**) higher for CS than CO in quiet [3] and in noise [4] for native English speakers.
- By enhancing intelligibility, CS frees up cognitive resources which are then available for memory encoding: Effortfulness Hypothesis [5].

Experiment 1: Within-modal SRM

❶ Does the CS benefit for sentence recognition memory extend to non-native listeners? ➜ Effortful speech processing in second language (L2) may require additional cognitive resources to the detriment of storing speech content in memory [6].
Prediction: Non-native listeners may benefit less from CS in SRM

Experiment 2: Cross-modal SRM

❷ Does the CS benefit persist if the stimuli in the test phase are written instead of presented auditorily (cross-modal testing)? ➜ Cross-modal SRM allows us to tease apart linguistically encoded information from surface features. *Prediction*: If the CS benefit persists in cross-modal SRM, available cognitive resources are used for deeper processing of the speech signal and storing in memory.

Experiment 3: SRM for Compressed Sentences

❸ What is the contribution of the slower CS speaking rate for sentence recognition memory? ➜ Longer stimulus duration can increase memory retention [7]. *Prediction*: If SRM is improved for CS sentences even when CS and CO sentences are compressed to the same syllabic rate, the CS benefit cannot be attributed to slower speaking rate only.

Methods

Recognition memory procedure

AUDIO — **EXPOSURE**
"Commit each sentence to memory"
40 sentences (20 CO; 20 CS)

EXPERIMENT 1 & 3: WITHIN-MODAL

AUDIO — **TEST**
"Identify each one as new or old"
80 sentences (40 CO; 40 CS)
40 from exposure + 40 new

EXPERIMENT 2: CROSS-MODAL

VISUAL — **TEST**
"The hot sun warmed the ground"
"Identify each one as new or old"
80 sentences (20 CO; 20 CS)
40 from exposure + 40 new

Material

80 meaningful sentences produced by a 26-year-old female American English speaker both in CO and CS. CS intelligibility higher for both native [2] and non-native listeners [8]. In Experiment 3, sentences temporally compressed into syllable rates based on the average syllable rates in CO and CS (using PSOLA in Praat) [9]. For native listeners, at rate 06, similar intelligibility in CO and CS; at rate 11, CS intelligibility > CO [9].

Listeners information

		Experiment 1 (N=60)	Experiment 2 (N=60)	Experiment 3 (N=81)
Native monolingual English listeners	Age	19 (18-23)	20 (18-32)	20 (18-37)
Non-native English listeners	Age	23 (18-31)	22 (18-31)	25.5 (18-38)
	Age of first exposure to English	9 (6-17)	8 (6-13)	10 (4-30)
	Age of arrival to USA	18.5 (6-30)	16 (1-28)	20 (3-38)
	L1	Mandarin (10), Korean (7), Spanish (5), French (2), Farsi, Turkish, Cantonese, Dutch, Portuguese, Amharic (1).	Spanish (11), Mandarin (10), Korean (4), French (2), Turkish, Hindi, Indonesian (1).	Portuguese (12), Korean (6), Mandarin (5), Spanish, Hindi, Marathi (2), Danish, Serbian, Nepali, Vietnamese (1)

Analyses

Signal detection framework: computing hit rates (identifying old item as old) and false alarm (identifying new item as old). Accuracy measured in d'. Statistical analyses: Linear mixed models lme4 package in R [10]. DV: d', hit, false alarm, reaction time (RT). IV: Style (CO vs. CS) and Group (native vs. non-native). Random effect: Subject.

Results & Discussion

Figure 1. Discrimination sensitivity (d′)

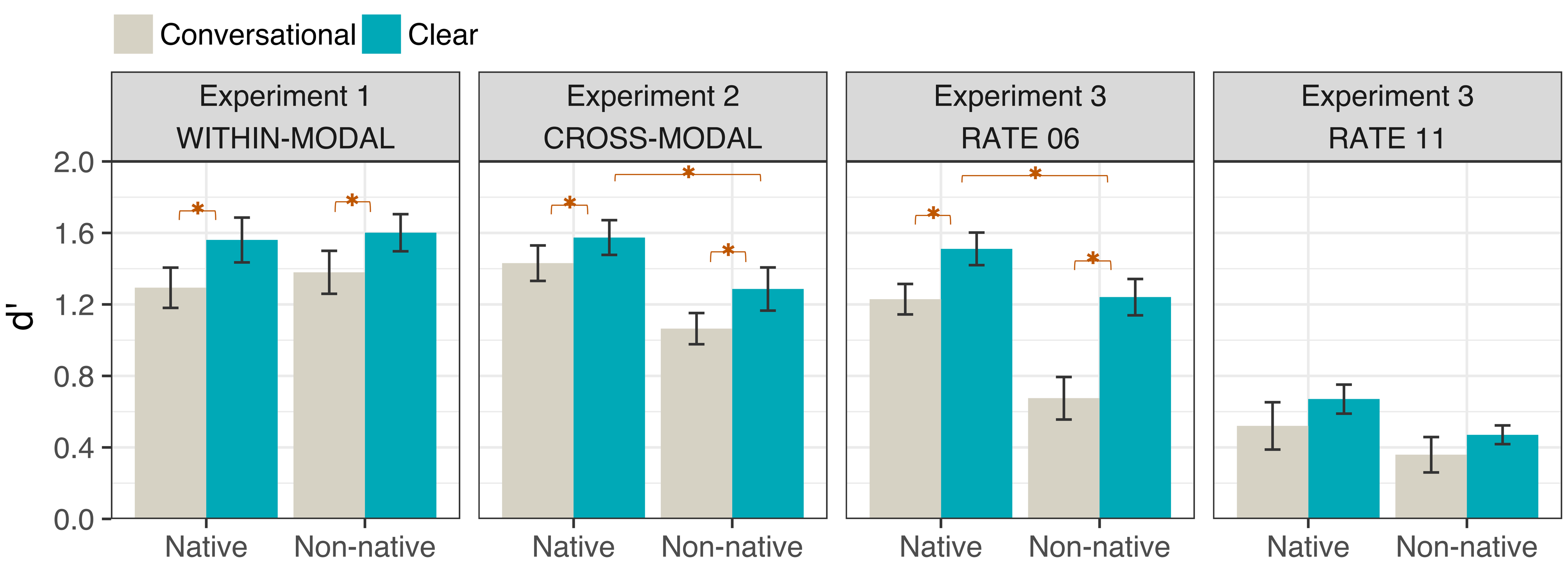


Figure 2. Hit rates

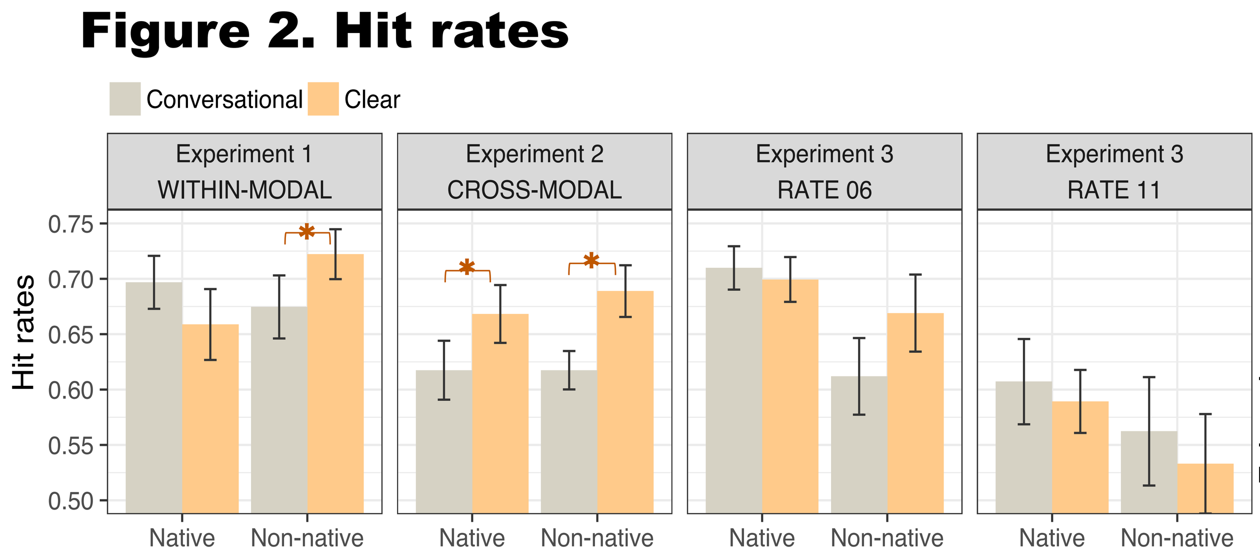


Figure 3. False alarm rates

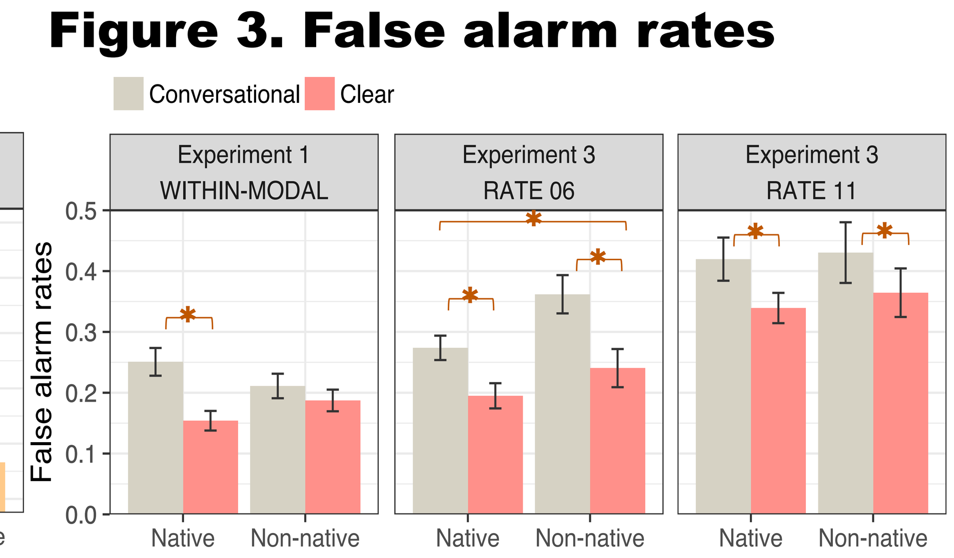
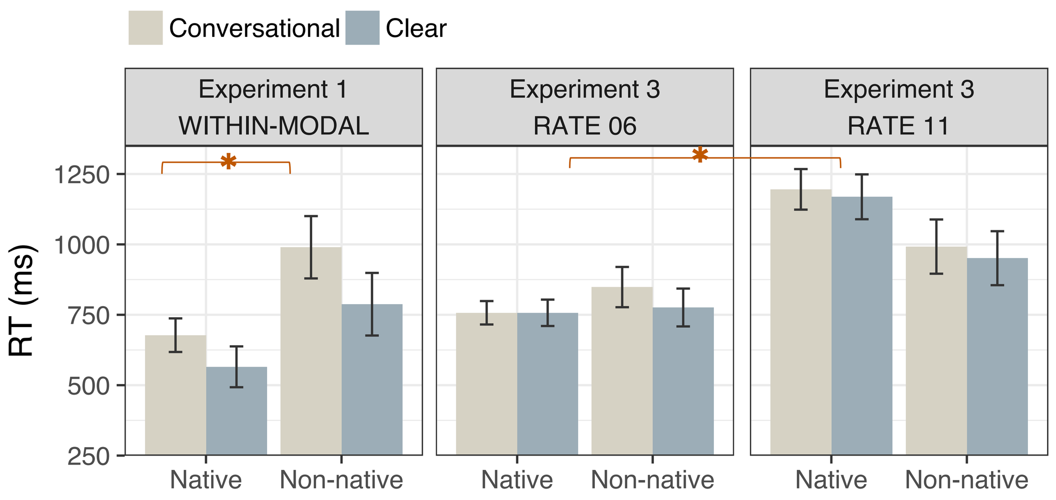


Figure 4. Reaction times



TAKEAWAY POINTS

- Clearly spoken sentences are better remembered by L1 and L2 listeners.
- Clear speech benefit remains even when test items are written only, suggesting higher-level linguistic encoding. L2 listeners may rely more on lower-level cues in SRM than L1 listeners.
- The CS benefit on SRM remains when equating syllable rate between clear and conversational sentences, suggesting that the memory benefit of CS is not due to the longer duration characteristic of that speaking style modification.

Experiment 1: Within-modal SRM

FIG. 1 • Higher d' for CS than CO (p <.001)
FIG. 2, 3 • Lower false alarm rate in CS than CO for native (p<.001); higher hit rate in CS for non-native (p<.05)
FIG. 4 • Longer RT for non-native than native (p<.05)

❶ Non-native listeners benefited from CS enhancements in SRM. Despite increased cognitive load in L2 perception (slower RTs), memory accuracy did not significantly differ between listener groups. Longer RTs for non-native listeners suggest more effortful processing.

Experiment 2: Cross-modal SRM

FIG. 1 • Higher d' for CS than CO (p<.001). Higher d' for native than non-native (p<.05). Lower d' in cross than within-modal task for non-native (p<.05).
FIG. 2 • Higher hit rates for CS than CO (p<.001)

❷ CS sentences better remembered than CO, even in a cross-modal task. Encoding in the exposure phase is not only at the perceptual lower-level but involves higher-level holistic processing. Increased accuracy in old item retrieval in CS. Cross-modal integration of information costlier in L2 than L1.

Experiment 3: SRM for Compressed Sentences

FIG. 1 • At rate 06, higher d' for CS than CO (p<.001), higher d' for native than non-native (p<.001). At rate 11, no significant effect of Style or Group.
FIG. 3 • Rate 06 & 11: Lower false alarm rate in CS than CO (p<.001). Rate 06: Lower false alarm rate for native than non-native (p<.05).
FIG. 4 • Longer RTs for rate 11 than 06 (p<.001).

❸ Slower CS speaking rate alone does not explain benefit on SRM. This suggests an important role for the exaggerated acoustic cues for memory encoding.

Future Directions

Explore the effect of clear speech on sentence recall, a more complex and effortful type of declarative memory that requires processing at phonological, lexical-semantic, morphosyntactic, and syntactic levels (in progress).

Acknowledgments

Thank you to the URAs and Kirsten Meemann at the UT SoundLab for assistance with data collection. Special thank you to Dr. Steven Alcorn for support with recruiting non-native participants for these experiments.

UTSOUNDLAB
TEXAS
The University of Texas at Austin

References

[1] Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. J Acoust Soc Am, 118(3 I), 1677-1688.
[2] Smiljanić, R., & Bradlow, A. R. (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness. J Acoust Soc Am, 130(6), 4020-4031.
[3] Van Engen, K., Chandrasekaran, B., and Smiljanic, R. (2012). "Effects of speech clarity on recognition memory for spoken sentences," PLoS ONE 7 (9), e43753.
[4] Gilbert, R., Chandrasekaran, B., & Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. J Acoust Soc Am, 135, 389-399.
[5] McCoy, S. L., Tun, P. A., Cod, L. C., Colangelo, M., Stewart, R.A., et al. (2005). "Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech," The Quarterly Journal of Experimental Psychology, 58A, 22-33.
[6] Cutler, A., Garcia Lecumberri, M. L., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context, J Acoust Soc Am, 124, 1264–1268.
[7] McDermott, K. B., & Watson, J. M. (2001). The Rise and Fall of False Recall: The Impact of Presentation Duration. Journal of Memory and Language, 45(1), 160–176.
[8] Keerstock, S., & Smiljanic, R. (2018). Effects of intelligibility on within- and cross-modal sentence recognition memory for native and non-native listeners. Manuscript submitted for publication.
[9] Shafiro, V., Smiljanic, R., & Keerstock, S. (2018). Semantic context modulates intelligibility advantage of clear speech in temporally compressed sentences. J Acoust Soc Am, 143(3), 1921–1921.
[10] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67, 1–48.

Contact: keerstock@utexas.edu | https://utsoundlab.wordpress.com