Effects of intelligibility on within- and cross-modal sentence recognition memory

TEXAS The University of Texas at Austin

5pSC15

Sandie Keerstock, Rajka Smiljanic

Department of Linguistics, The University of Texas at Austin

UTSOUNDLAB

Background

Introduction

- *Effortfulness hypothesis:* Perceptual effort during processing of degraded speech may come at the cost of attentional resources that would otherwise be available for memory encoding [1].
- *Clear speech* (CS) facilitates speech processing by increasing intelligibility for native and non-native listeners [2,3], as compared to *conversational speech* (CO).
- CS improves recognition memory for sentences in quiet [4] and in noise [5].
- Speech processing in second language is taxing and may require additional cognitive resources [6].
- Reading and hearing speech is processed separately at perceptual levels (e.g. audio, visual) and converges at higher-level stages of language processing (e.g. understanding) [7].

Research Questions

①

What underlies better sentence recognition memory in CS? Encoding of salient acoustic-phonetic cues available in hyper-articulated speech? Or encoding of higher-level linguistic information?

➔ If CS facilitates encoding of the specific acoustic-phonetic cues, the same cues should be present in the test to allow the listener to retrieve the previously heard sentence.
➔ If CS facilitates encoding of semantic information, listeners should be able to retrieve that information even when only seeing written sentence in test.

②

Does CS enhance sentence recognition memory equally for native (N) and non-native (NN) listeners?

➔ CS sentence recognition memory benefit should be smaller for non-native listeners since perceptual effort is increased when processing second language.

Methods

Procedure

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

CONDITION 1: WITHIN-MODAL

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

CONDITION 2: CROSS-MODAL

VISUAL — *"The loud noise upset the baby"*

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

Listeners

Condition 1:
- 30 native (N) monolingual American English. Mean age 19
- 30 non-native (NN). Mean age: 23; age of English acquisition: 9

Condition 2:
- 30 native (N) monolingual American English. Mean age 19
- 30 non-native (NN). Mean age: 21; age of English acquisition: 8

Material

- 80 meaningful sentences [4] (e.g. *The hot sun warmed the ground*) produced by a 26-year-old female American English speaker both in conversational and clear speaking styles.
- Intelligibility assessment: all sentences were mixed with speech-shaped noise at -5 dB signal-to-noise ratio (SNR). Clear sentences significantly more intelligible than conversational sentences.

Analyses

Signal detection framework: d' = discrimination sensitivity
= normalized probability of HIT rate – normalized probability of FALSE ALARM rate

		Sentence: OLD	Sentence: NEW
Response	OLD	HIT	FALSE ALARM
Response	NEW	MISS	CORRECT REJECTION

Statistical analysis:
Linear mixed model with d', hit, false alarm rates as the dependent variable, Speaking Style (CO vs. CS) and L1 (N vs. NN) as the independent variables and Subject as a random effect.

Results

Fig.1: Average d' scores per modality and listener group

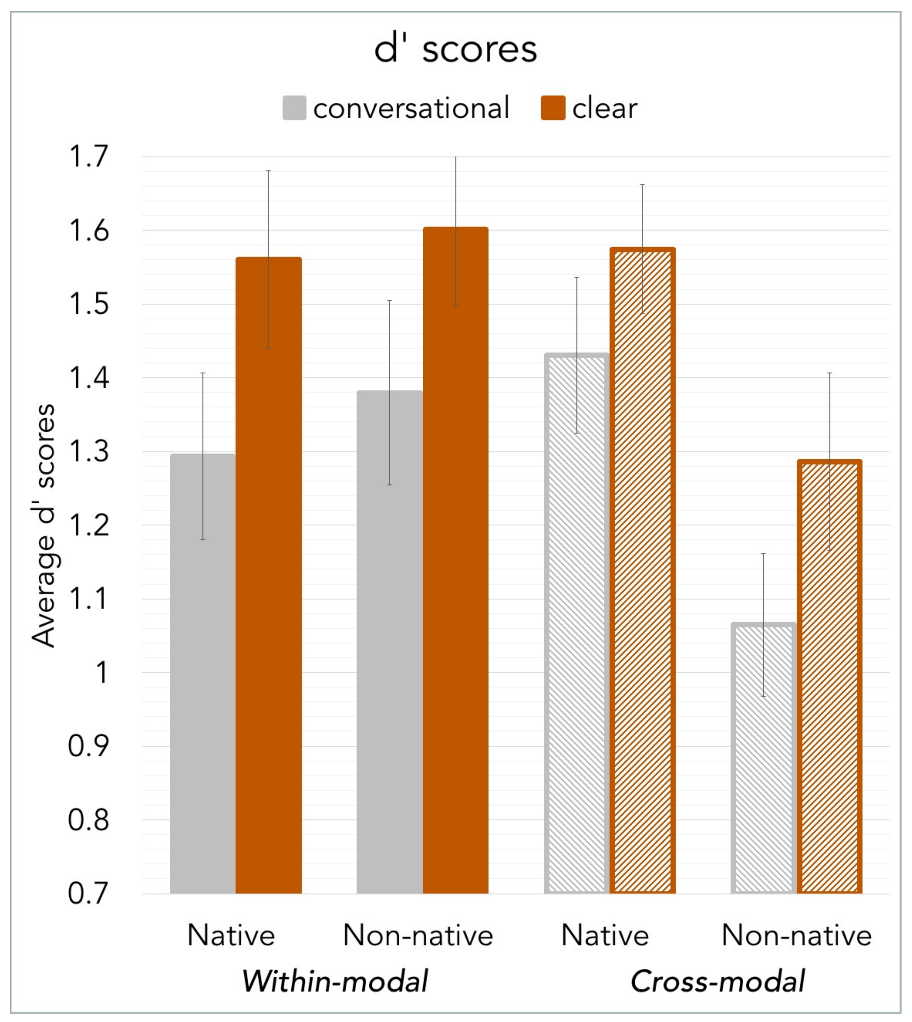


Fig.2: Average hit rate per modality and listener group

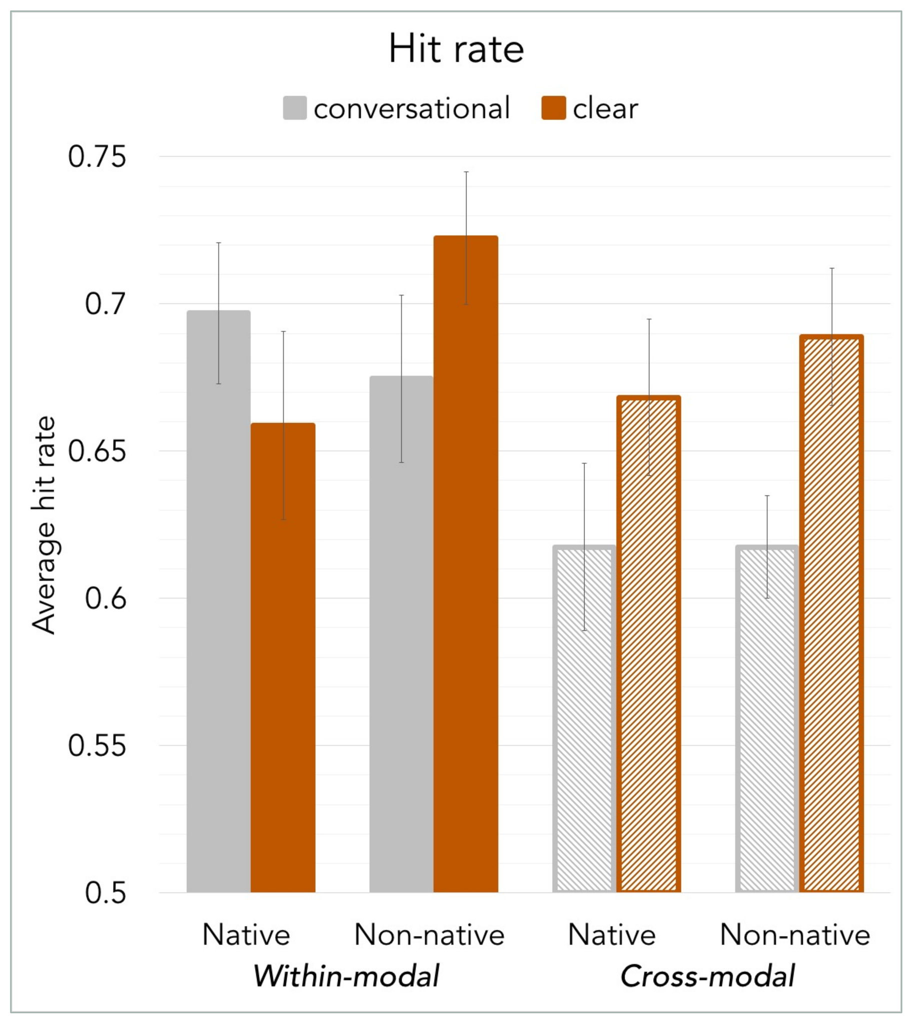


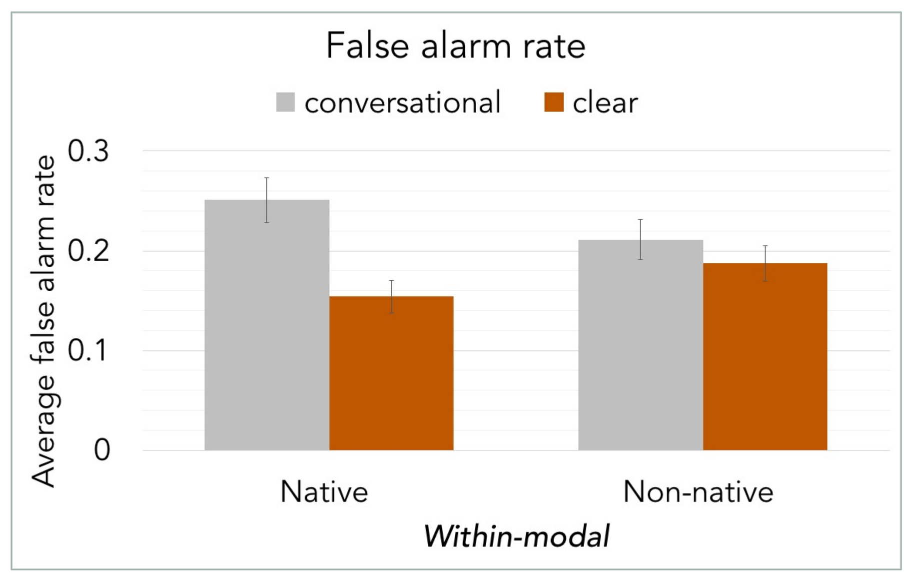
Table 1: Effect of L1 and speaking style on d', hit, false alarm

		Within	Cross
d'	Speaking style (CO vs. CS)	p=.001	p=.002
	L1 (N vs. NN)	p=.66	p=.017
	Interaction	p=.75	p=.49
hit	Speaking style (CO vs. CS)	p=.15	p=.002
	L1 (N vs. NN)	p=.56	p=.71
	Interaction	p=.02	p=.58
false alarm	Speaking style (CO vs. CS)	p=.001	n/a
	L1 (N vs. NN)	p=.15	
	Interaction	p=.009	

Table 2: Effect of modality and speaking style on d'

		N	NN
d'	Speaking style (CO vs. CS)	p=.004	p=.001
	Modality (within- vs. cross-)	p=.58	p=.03
	Interaction	p=.36	p=.99

Fig.3: Average false alarm rate per modality and listener group



- d' in CS significantly higher than in CO in both modalities and for both listener groups (Fig.1).
- Native and non-native listeners benefit equally from CS adaptations in within-modal, but natives outperform non-natives in the cross-modal condition. Sentence recognition memory in the cross-modal condition is overall more difficult for non-native listeners than in the within-modal condition. (Fig.1)
- Hit rate significantly higher for CS sentences in cross-modal, but not in within-modal (Fig.2). In within-modal condition, significantly lower false alarm rate for CS compared to CO. (Fig.3)

Summary

①

- Intelligibility enhances recognition memory for both native and non-native listeners. Supports the *Effortfulness hypothesis*[1]: greater intelligibility may free up cognitive resources for encoding speech information into memory.
- Speaking style adaptations promotes the encoding and storing of lower-level acoustic information (condition 1) as well as higher-level semantic information (condition 2), for both native and non-native listeners.
- Improved recognition memory in CS is mostly driven by lower false alarm rate in the within-modal condition, and by higher hit rate in the cross-modal condition.

②

- Non-native put more weight on acoustic-phonetic information when encoding second-language content in memory, based on the higher d' scores in within- than cross-modal

Future work

- Varying voice in the test phase to separate out the effects of matching of the test item to the exposure item from the effects of style on recognition memory.

References:
[1] 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.
[2] 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. DOI: 10.1121/1.2000788
[3] 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. DOI: 10.1121/1.3652882
[4] Van Engen, K., Chandrasekaran, B., and Smiljanic, R. (2012). "Effects of speech clarity on recognition memory for spoken sentences," PLoS ONE 7 (9), e43753. doi:10.1371/journal.pone.0043753
[5] Gilbert, R., Chandrasekaran, B., & Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. J Acoust Soc Am, 135, 389-399.
[6] Cutler A, et al (2008) Consonant identification in noise by native and non-native listeners: Effects of local context. J Acoust Soc Am 124, 1264–1268.
[7] Buchweitz A., Mason R.A., Tomitch, L., & Just, M.A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. Psychology & Neuroscience, 2, 111–123.

keerstock@utexas.edu - https://utsoundlab.wordpress.com