Facilitation of speech processing by both expected and unexpected talker continuity

Yaminah D. Carter, Alex M. Kapadia, Sung-Joo Lim, Tyler K. Perrachione

Department of Speech, Language, & Hearing Sciences, Boston University

BOSTON UNIVERSITY

Introduction

- Research has shown adaptation-related efficiency gains in speech perception while listening to a single continuous talker vs. mixed, unpredictable talkers [1-3].

- Expectations about the perceptual world can facilitate processing by reducing the neurocomputational demands supporting perception [4].

- Expecting to hear different talkers can reduce speech perception efficiency compared to expecting a single talker [5].

- In three speeded word identification experiments, we investigated whether efficiency *gains* come from expected vs. unexpected talker continuity and if talker adaptation can be understood as fulfilled top-down perceptual expectations.

Participants:
A new *N* = 24 sample from the same population as Expts. 1 & 2 (18-25 years).

Stimuli:
The target words "boot" and "boat" recorded by 1 female and 1 male native speaker of American English.

Methods:
Identify words from a single-talker, and
2-Talkers (uniform transition probability):
'boat' 'boot' 'boat' 'boot' 'boat' 'boot' 'boot' …
2-Talkers (alternating):
'boot' 'boat' 'boot' 'boat' 'boot' 'boat' 'boot' …

Experiment 3: Predictable repetition

Effects of expectation vs. repetition
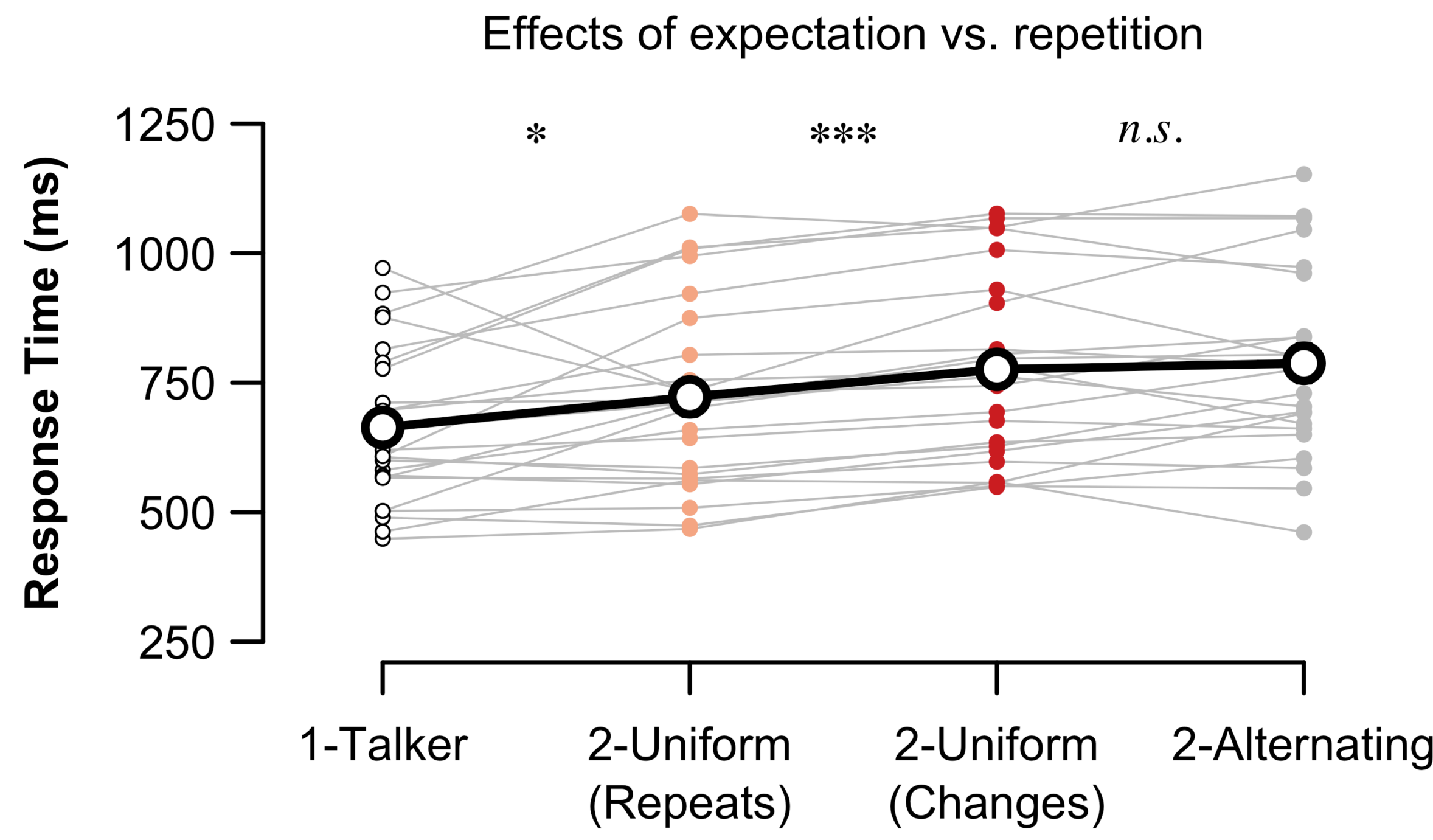
Fig. 3: Mean RT as a function of talker continuity and listeners' expectation of talker change. Expected talker continuity led to significantly faster RTs than unexpected talker continuity. RTs in the uniform condition were significantly faster on trials where the talker unexpectedly repeated vs. unexpectedly changed.

1-talker vs. 2-uniform (repeats)
$\beta = 0.039$, $t = 2.92$, $p = 0.008$
2-uniform (repeats) vs. 2-uniform (changes)
$\beta = 0.033$, $t = 6.11$, $p < 0.0001$
2-uniform (changes) vs. 2-alternating
$\beta = 0.003$, $t = 0.35$ $p = 0.73$

Experiment 1: Expectation of talker continuity

Participants:
Native speakers of American English (*N* = 20; 18-26 years) without speech, hearing, or language disorders.

Stimuli:
The target words "boot" and "boat" were recorded by 2 female and 2 male native speakers of American English.

Methods:
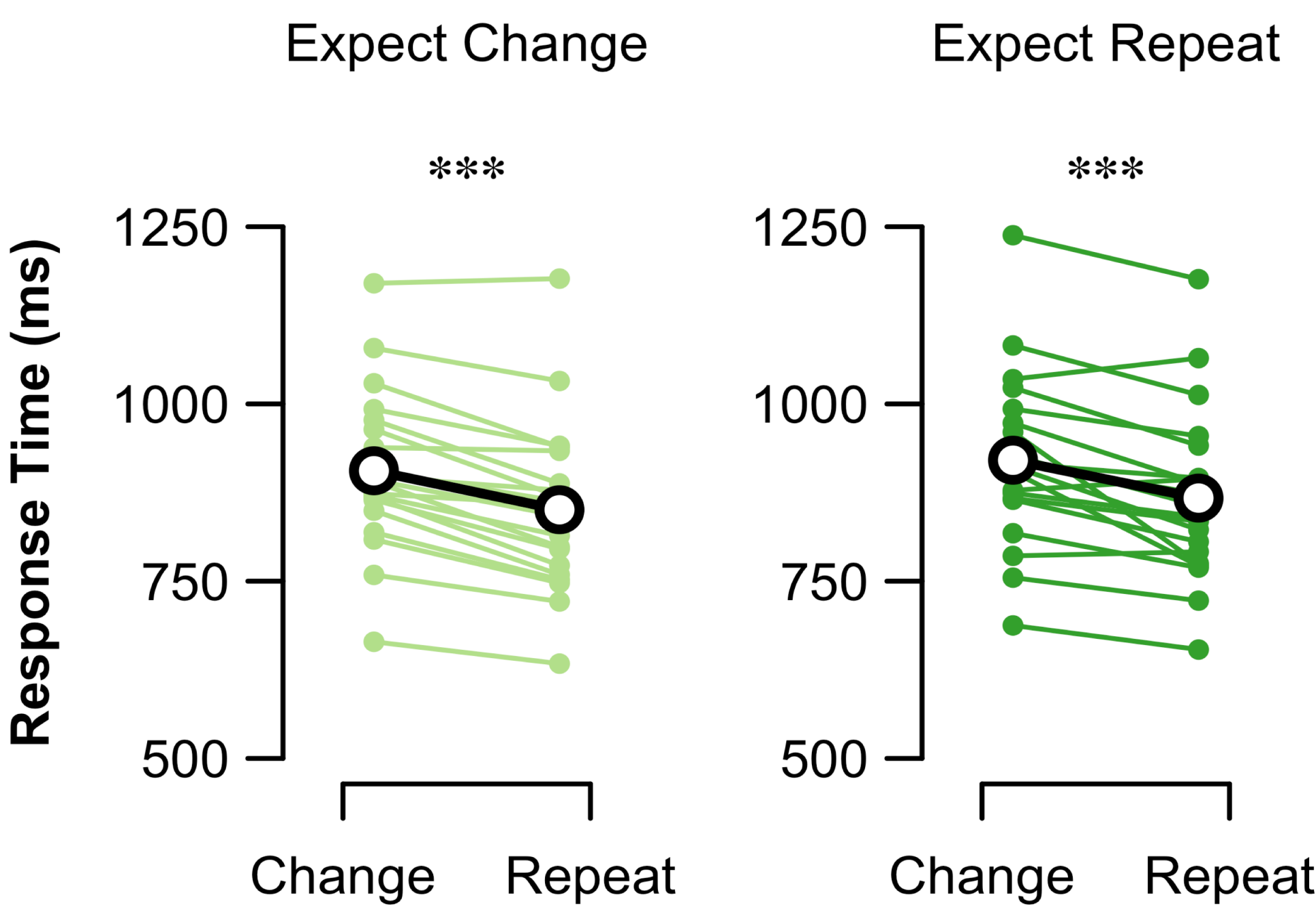Identify words quickly and accurately in two conditions:
High probability (80%) of talker change:
'boat' 'boat' | 'boot' 'boat' | 'boat' **'boot'** | 'boot' 'boot' | …
High probability (80%) of talker repetition:
'boat' 'boot' | 'boot' 'boot' | 'boat' **'boot'** | 'boot' 'boot' | …
2-s ISI between words in a pair, 2-s ITI between each pair

Fig. 1: Mean RTs to the second target words in the trial pair. Both expected and unexpected repetitions of a talker led to faster RTs compared to trials where talker changed. There was no difference in the magnitude of this facilitation based on listeners' expectations.

Talker repetition
$\chi^2(1) = 62.70$, $p < 0.001$
Expectation
$\chi^2(1) = 1.14$, $p = 0.29$
Talker repetition × Expectation
$\chi^2(1) = 0.12$, $p = 0.73$

Discussion

- Listeners were significantly faster identifying words spoken by the same talker, even if they expected to hear a different talker or had no certain expectation about which talker they would hear.

- Only fulfilled expectations of word repetition were facilitatory for speech processing; expected changes or unexpected repetitions had no effect compared to having no expectations about word repetition at all.

- These results suggest that talker adaptation-related efficiency gains in speech processing may reflect bottom-up mechanisms that interpret speech in the context of preceding speech [6].

- Speech processing efficiency appears to primarily be driven by continuity in the speech source, not top-down expectations [7], suggesting talker adaptation in speech perception is an obligatory, feedforward process [2].

Experiment 2: Expectation of word repetition

Participants:
A new *N* = 20 from the same population as Experiment 1 (18-26 years).

Stimuli:
The target words "boot" and "boat" from one male speaker from Exp. 1.
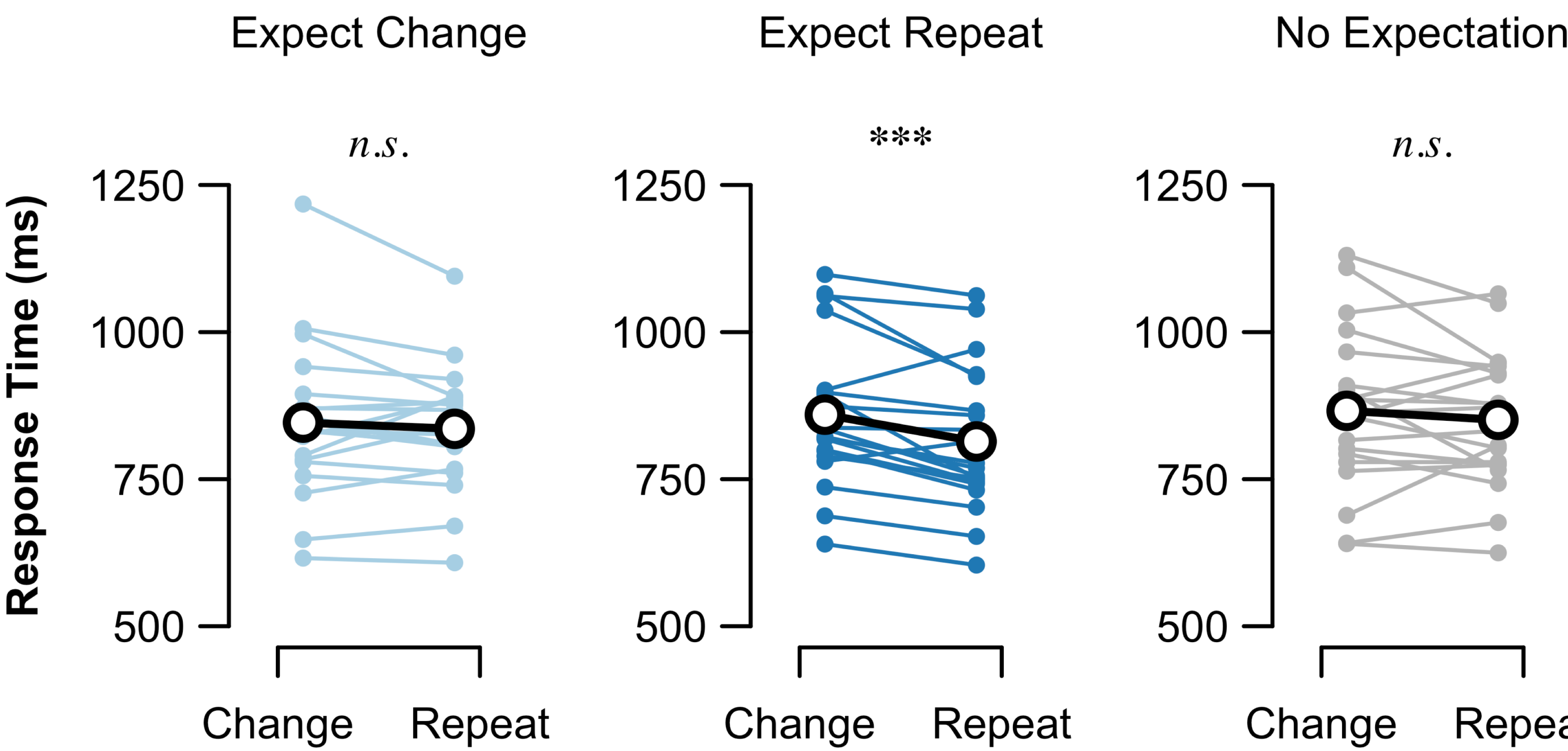
Methods:
High probability (80%) of word change:
'boat' 'boot' | 'boot' 'boat' | 'boot' **'boat'**…
High probability (80%) of word repeat:
'boat' 'boat' | 'boot' 'boot' | 'boot' **'boat'**…
and Equally probable change / repeat
2-s ISI within pair, 2-s ITI between pairs

Fig. 2: Mean RTs to the second target words in the trial pair. Only expected repetition of a word led to faster RTs compared to trials in which the word changed. When listeners had no expectations about word repetition, or when they expected the word to change, repetition was not facilitatory.

Word repetition
$\chi^2(1) = 4.70$, $p = 0.030$
Expectation
$\chi^2(2) = 7.94$, $p = 0.019$
Word repetition × Expectation
$\chi^2(2) = 16.14$, $p < 0.001$

References

[1] Mullennix & Pisoni (1990). *Perception & Psychophysics, 47*, 379–390.
[2] Choi, Hu, & Perrachione (2018). *Attention, Perception, & Psychophysics, 80*, 784–797.
[3] Perrachione, et al. (2016). *Neuron, 92*, 1383–1397.
[4] Summerfield & de Lange (2014). *Nature Reviews Neuroscience, 15*, 745–756.
[5] Magnuson & Nusbaum (2007). *Journal of Experimental Psychology - HPP, 33*, 391–409.
[6] Johnson (1990). *Journal of the Acoustical Society of America, 88*, 642–654.
[7] Kleinschmidt & Jaeger (2015). *Psychological Review, 122*, 148–203.

Communication Neuroscience Research Laboratory

tkp@bu.edu
http://sites.bu.edu/cnrlab/