Acoustic entrainment of speech supports comprehension under moderate noise, but degrades under more severe adversity (an OSF pre-registered study)

Benjamin Zinszer[1], Qiming Yuan[2], Zhaoqi Zhang[2], Bharath Chandrasekaran[3], & Taomei Guo[2]

1. Department of Linguistics and Cognitive Science, University of Delaware, USA & Department of Psychology, Princeton University, USA;
2. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, China; 3. Department of Communication Sciences and Disorders, University of Pittsburgh, USA

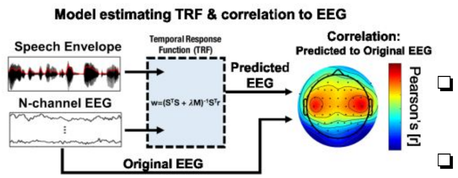
INTRODUCTION

Comprehending Speech in Noisy Conditions is central to daily life, like participating in a conversation or watching television.

- ❏ requires allocating attention to signals of interest and inhibiting others
- ❏ modulated by listener's attributes (Mattys et al., 2012)

Multivariate Temporal Response Function modeling relates a continuous model of speech (e.g., the amplitude envelope) to EEG data occuring at the same time.

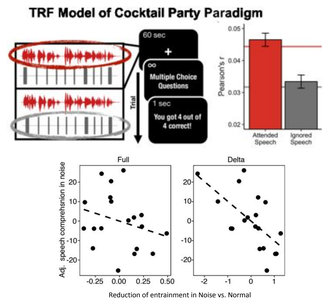
(see MTRF model: Crosse, Di Liberto, Bednar, & Lalor, 2016; https://github.com/mickcrosse/mTRF-Toolbox)



- ❏ Estimate a function for converting stimulus model to EEG data based on training set (59 tracks)
- ❏ Use the function to predict EEG data for the one left-out test track
- ❏ Correlate predicted EEG to real EEG

Listening Effort is still an ill-defined concept in speech processing *but* we have some evidence that effort is reflected in EEG activity (Francis & Love, 2019). Cortical entrainment to a speech envelope is:

1. modulated by attention to a target stimulus
2. higher in proficient non-native listeners than native listeners (Song & Iverson, 2019; Reetzke et al., 2017)
3. positively correlated with speech comprehension in older adults (Decruy et al., 2019; McHaney et al., 2017 & under review)



QUESTION & HYPOTHESES

Research Question: Do native & non-native listeners always *increase* entrainment under adverse listening, after controlling for comprehension?

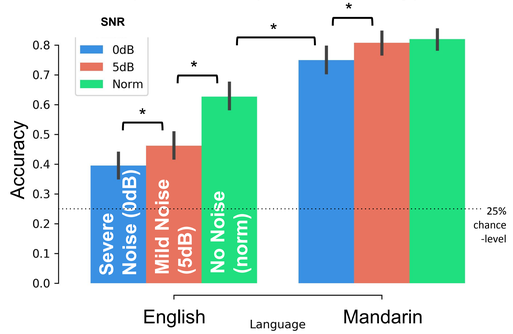
Hypothesis 1: **English > Mandarin.** Mean group-level Pearson's *r* will be greater for the English language tracks than the Mandarin language tracks.

Hypothesis 2a: Mild Noise > Normal > Severe Noise. Neural entrainment to the native language increases under mild noise and decreases under severe noise.
Hypothesis 2b: Severe Noise > Mild Noise > Normal. After controlling for accuracy entrainment consistently increases with noise.

Hypothesis 3: **Entrainment ~ -Proficiency.** Proficiency in a second language (L2, English) is inversely correlated with entrainment (controlling for accuracy).

RESULTS

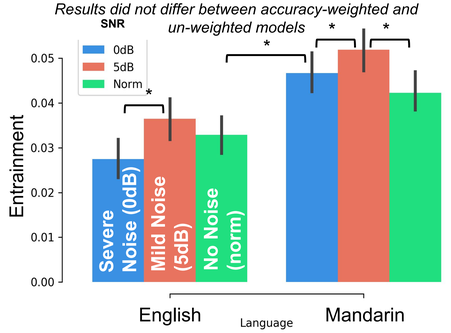
Comprehension (MCQ accuracy)



Accuracy on MCQ was averaged over 10 tracks per condition. All pairwise contrasts significantly differed (p<0.05, Tukey HSD) except between Norm & 5dB in Mandarin.

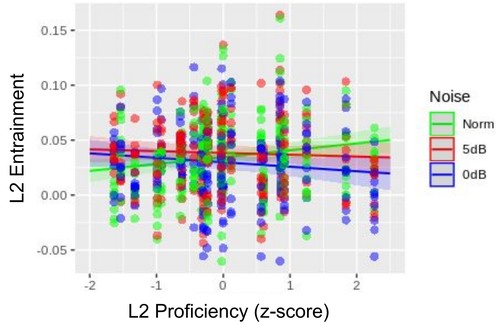
Average Entrainment (accuracy-weighted)

Results did not differ between accuracy-weighted and un-weighted models



H1: Mandarin > English
H2: Mandarin: Norm < 5dB
Mandarin: 0dB < 5dB
English: 0dB < 5dB
(all channels; p<0.05, Tukey HSD)

Differences in L2 Entrainment by Proficiency



H3: L2 Proficiency is positively related to entrainment in Norm (p<0.01). This relationship is significantly reduced in the 5dB and 0dB conditions (ps<0.01).
*(linear mixed-effects model Proficiency*Noise)*

METHOD

The study was approved by the IRB of Beijing Normal University. Each participant signed informed consent before the experiment and received debriefing forms and payment afterward.

Participants

n=24 Chinese-English bilinguals, college students from Beijing Normal University; right-handed, with normal or corrected- to-normal vision, no reported neurological disorders.

Measure	Mean	SD
CET-4 exam score	570	33
Phonemic Verbal Fluency		
English	11.60	2.87
Semantic Verbal Fluency		
English	8.59	3.00
Mandarin	17.86	3.11

Scores on CET-4 & English verbal fluency combined using PCA for composite English proficiency score

EEG Recording

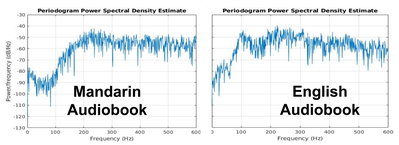
- ❏ NEUROSCAN Synamps 2
- ❏ 64-channel cap, reference at mastoids
- ❏ Bipolar HEOG & VEOG recorded
- ❏ Electrode impedances <5 kΩ

Offline EEG Processing

- ❏ Bandpass filtering at 1-15 Hz
- ❏ Down-sampling to 128 Hz

Stimulus Properties and Presentation

- ❏ Audiobook *The Old Man and the Sea*, read in English and Mandarin
- ❏ Each book cut into equivalent tracks (same content in each language)
- ❏ Two multiple choice questions (MCQ) per track, matched between languages



Measure	English	Mandarin
Tracks	60	60
Average track duration	59.5 s	64.3 s
Median pitch	103 Hz	136 Hz

Noise Conditions

- ❏ **No Masker (Norm)**
- ❏ **5 dB SNR (power of speech is 3.15 times greater than speech-shaped noise mask)**
- ❏ **0 dB SNR (speech & speech-shaped noise mask are equal in power)**

CONCLUSIONS

Hypothesis 1: In this sample, entrainment of L1 > L2

Hypothesis 2: (2a) Entrainment generally increased to mild noise, and decreased to severe noise. (2b) Adjusting for comprehension did *not* change that pattern.

Hypothesis 3: Under no-noise, L2 proficiency is positively related to entrainment (controlling for comprehension)

SUPPORT & ACKNOWLEDGMENTS

This research was funded by the Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China.

We are also grateful for the assistance of Dr. Rachel Reetzke (UC Davis), Dr. Heather Dial (Univ of Texas), Mr. Zeping Liu (Beijing Language and Culture Univ), Ms. Yingying Sun & Ms. Qianwen Chang (Beijing Normal Univ.), Ms. Jacie McHaney (Univ. Pittsburgh), Ms. Priyanka Deshmane (Tulane Univ), & Ms. Sarah Campbell (Cornell Univ) for their assistance with this research.

CONTACT

Dr. Benjamin D. Zinszer
bzinszer@gmail.com
www.benjaminz.com

Prof. Taomei Guo
gtaomei@163.com