

Introduction

Brain tracks temporal structure during continuous speech¹.

Syntax provides **top-down** temporal predictions². \longleftrightarrow Visual speech provides **bottom-up** temporal cues³.

Do syntactic and visual timing cues independently shape neural encoding of speech?

Methods

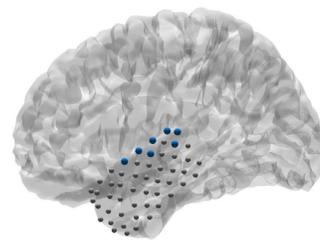
Stimuli & Design

100 English sentences⁴ in 5 conditions:

Modulation \ Modality	Temporal Degradation (AM) ⁵	Spectral Degradation (FM)
Audio Only (A)	A-AM	A-FM
Audio Visual (AV)	AV-AM	AV-FM
Visual Only (V)	V	

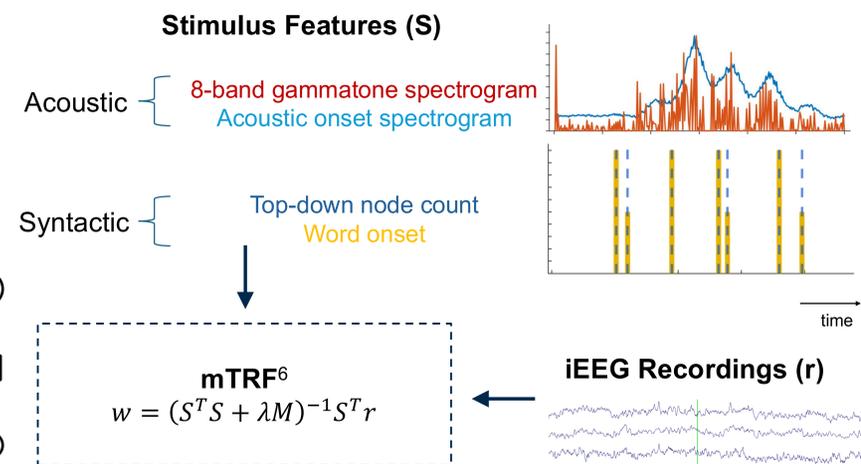
Modality	Condition	Stimulus	Response
Auditory Only	A-AM, A-FM	Waveform of "Here we go."	1. Here we go. 2. Move out of the way!
	A-AM, A-FM	Waveform of "I hate driving at night."	1. Call me a little later! 2. I hate driving at night.
Audio Visual	A-AM, A-FM	Video of speaker saying "I hate driving at night."	1. The show's over. 2. Wait just a minute!
	V	Video of speaker saying "I hate driving at night."	1. The show's over. 2. Wait just a minute!

Blank Screen + Fixation Period Sentence Stimulus Response Period



- 7 epilepsy patients in total
- Preliminary results
- 1 patient, 7 STG bipolar pairs

Multivariate Temporal Response Function



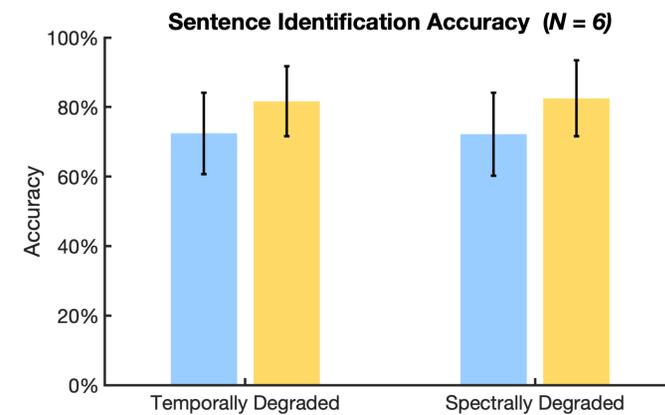
R²: explained variance TRF coefficients

2 model structures

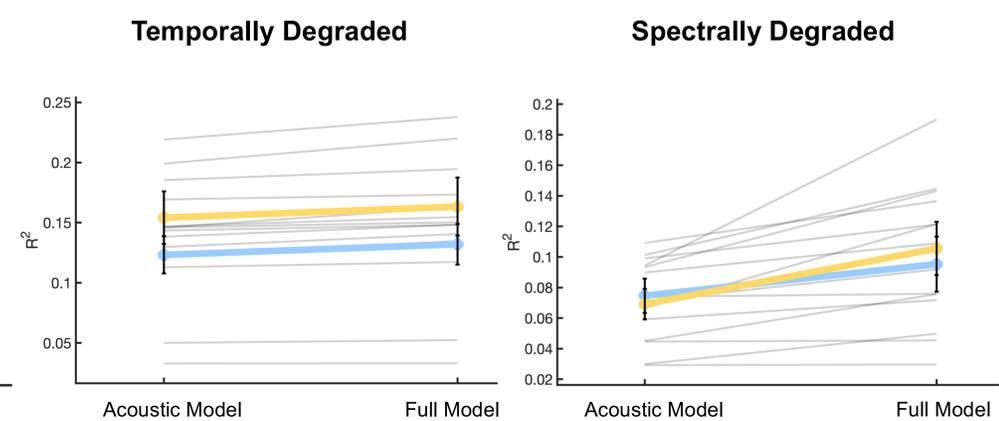
1. Acoustic model
 - acoustic predictors only
2. Full model:
 - acoustic + syntactic predictors

Results

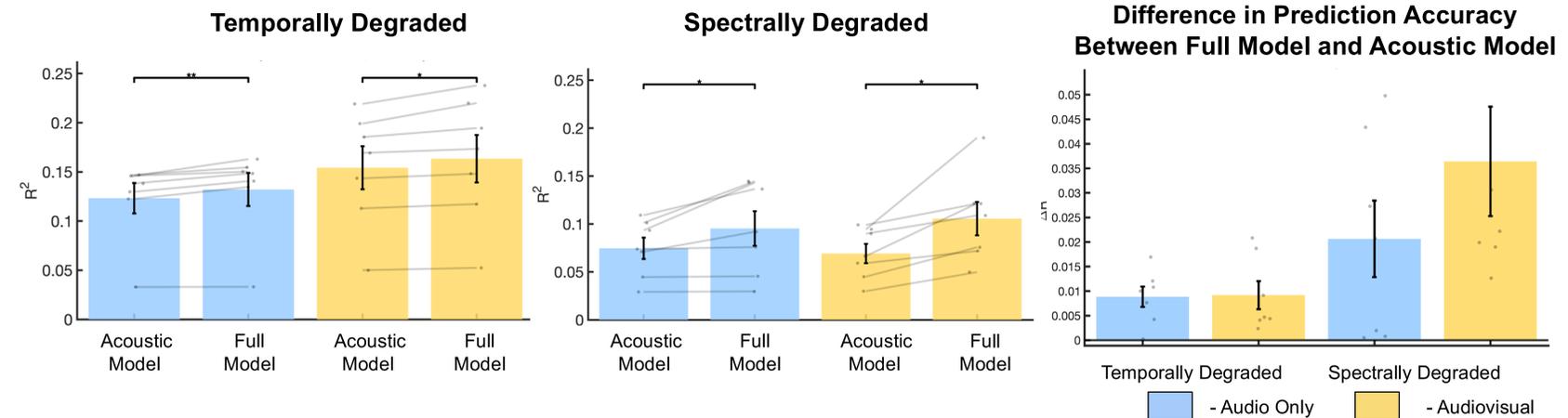
Behavioral Benefits of Visual Speech



Interaction Between Visual Speech and Syntax



Effects of Visual Speech on Syntactic Encoding



Conclusion & Discussion

- Syntax and visual timing cues interact during speech comprehension.
- Future work will investigate how visual speech contributes to syntactic temporal prediction across different acoustic degradation levels.

References

[1] N. Ding, L. Melloni, H. Zhang, X. Tian, and D. Poeppel, "Cortical tracking of hierarchical linguistic structures in connected speech," *Nat Neurosci*, vol. 19, no. 1, pp. 158–164, Jan. 2016.
 [2] C. W. Coopmans, H. de Hoop, F. Tezcan, P. Hagoort, and A. E. Martin, "Language-specific neural dynamics extend syntax into the time domain," *PLoS Biology*, vol. 23, no. 1, p. e3002968, Jan. 2025.
 [3] H. Park, C. Kayser, G. Thut, and J. Gross, "Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility," *eLife*, vol. 5, p. e14521.
 [4] Richie, Carolyn, Warburton, Sarah, and Carter, Megan, "Audiovisual Database of Spoken American English," Linguistic Data Consortium, p. 7759462 KB, Feb. 16, 2009.
 [5] J. Plass, D. Brang, S. Suzuki, and M. Grabowczyk, "Vision perceptually restores auditory spectral dynamics in speech," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 117, no. 29, pp. 16920–16927, Jul. 2020.