Effects of spectral and temporal modulation degradation on intelligibility and cortical tracking of speech signals

Ignacio Calderon De Palma[1], Laura Lopez[2], Alejandro Lopez Valdes[3]

[1]Radboud University Medical Center, Nijmegen, The Netherlands
[2]Washington State Universty, Washington, USA
[3]Trinity College Dublin, Global Brain Health Institute, Trinity College Institute of Neuroscience,Trinity Centre for Biomedical Engineering, Ireland

Background

- **Spectro-temporal modulations** are the **fundamental building blocks** of complex signals, carrying important cues for speech intelligibility.
- Consequences of **missing spectral and temporal information** are of particular i**mportance for aging populations and recipients of hearing devices** as they receive less spectro-temporal information, detrimental for speech processing.
- Cortical tracking in response to naturalistic stimuli has provided insights into speech processing and information transmission.
- Assessing cortical tracking under situations with spectral and temporal degradations can help us better **understand how stimuli are encoded in patient populations**.
- This could provide an **objective measure to study information transmission** when confronted with bottom-up degradations of speech signal.

Research questions

What is the effect of modulation filtering on speech intelligibility?

Evaluate intelligibility of speech without modulation information.
Get reference intelligibility values for EEG experiment.
Replicate previous findings (Elliot & Theunissen, 2009; Flinker et al. 2019).

What is the effect of modulation filtering on cortical tracking of spectro-temporal features?

Train model based on clean condition only → Predict responses for all conditions with that single model → Evaluate changes in prediction accuracy with respect to the clean model

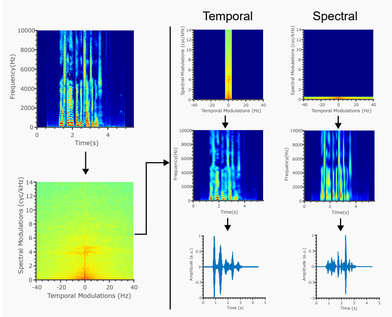
Can we recover spectro-temporal modulation degradation from neural representations?

Train an optimum model for each condition → Predict responses for each condition → Compare model weights of degraded and clean conditions

Methods

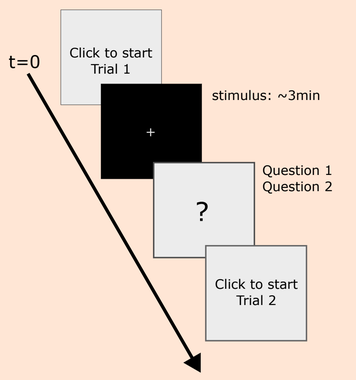
13 adult (8 female), normal-hearing, right-handed participants.
Age: 25 [20-30] years

Modulation filter

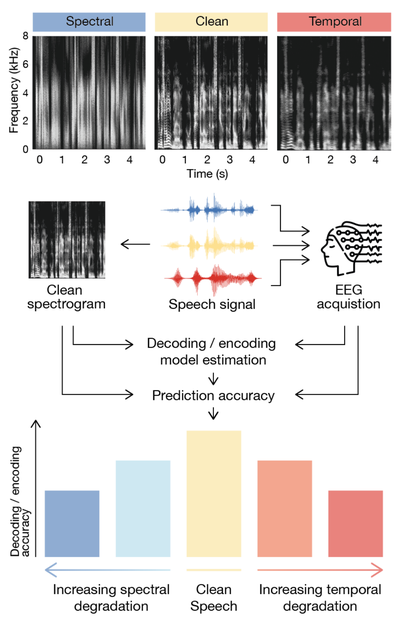


Behaviour
Stimuli: IEEE sentences (Lists 1-14)
63 dBA/ch, male speaker
Accuracy: Number of unique correct words

EEG
Stimuli: audiobook snippets (3 min)
63 dBA/ch, male speaker
64-channel Bio-Semi Active Two
mTRF toolbox, fieldtrip
5 conditions x 5 presentations per cond.



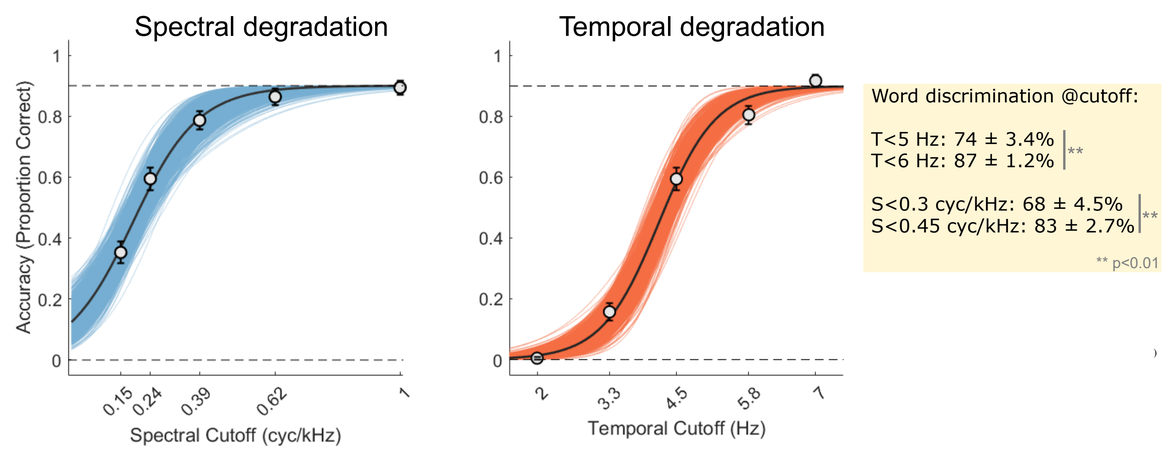
Experiment progression



Behaviour

What is the effect of modulation filtering on speech intelligibility?

Intelligibility decreases with increasing degradation (comparable to previous studies)

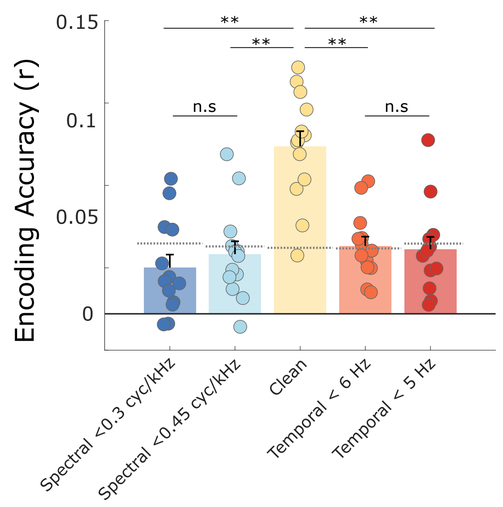


Percent correct words using IEEE sentences (male speaker) for different spectral and temporal modulation filters

EEG

What is the effect of modulation filtering on cortical tracking of spectro-temporal features?

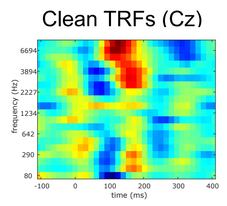
The clean model cannot predict degraded conditions for our dataset.

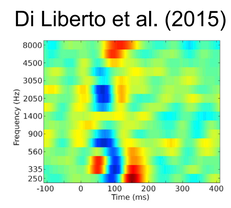


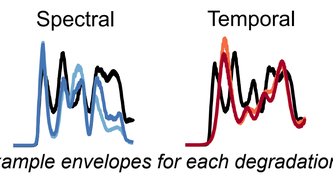
Prediction accuracy of clean model for all conditions

t-statistics for significance level:

Clean: t(12)=6.47, p<0.01
Temporal<5 Hz: t(12)=-0.28, p=0.78
Temporal<6 Hz: t(12)=0.11, p=0.91
Spectral<0.3 cyc/kHz: t(12)=-1.66, p=0.12
Spectral<0.45 cyc/kHz: t(12)=-0.76, p=0.46



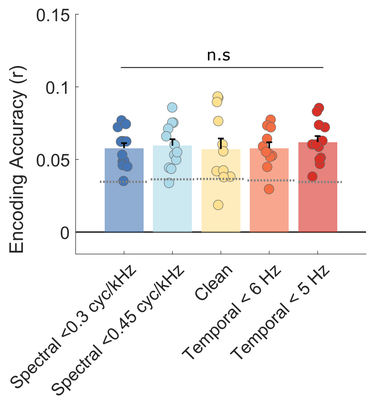




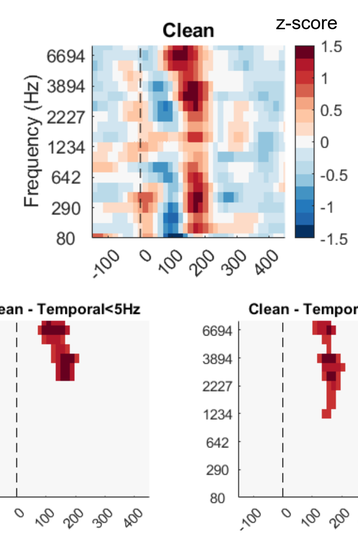
Example envelopes for each degradation

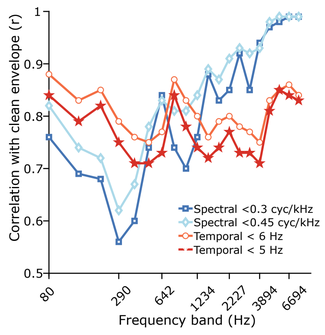
Can we recover spectro-temporal modulation degradation from neural representations?

Differences in weights reveal the most degraded regions in the spectrogram.

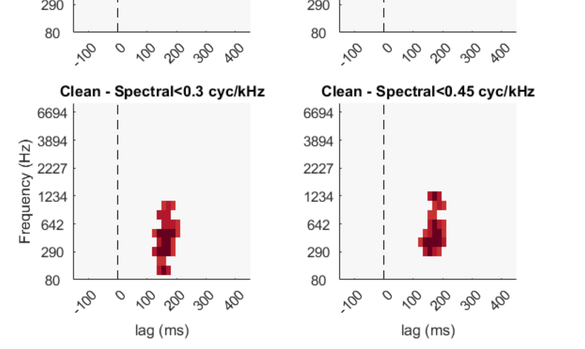


Prediction accuracy of models for each condition





Envelope correlation w/clean for a single audio file



Cluster based analysis (N=5000), corrected for multiple comparisons

Take-home messages

- Temporal and spectral modulations are fundamental for speech understanding, with temporal modulations being the most important (as shown in previous research).
- Systematic removal of spectral and temporal modulations leads to a decrease in encoding accuracy of EEG data.
- Comparing the model weights between clean and degraded conditions is a potential tool to capture spectral and temporal degradations from neural signals.

MOSAICS is a European Industrial Doctorate project funded by the European Union's Horizon 2020 framework programme for research and innovation under the Marie Sklodowska-Curie grant agreement No 860718.

References

J.E. Peelle, M.H. Davis. "Neural Oscillations Carry Speech Rhythm through to Comprehension". Frontiers in Psychology. vol 3, no. 320, pp 1-17, Sep. 2012.
G.M. Di Liberto, J.A. O'Sullivan, E.C. Lalor. "Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing". Current Biology, vol. 25, no. 19, pp. 2457-2465. Sep. 2015.
T.M. Elliott, F.E. Theunissen. "The modulation transfer function for speech intelligibility". PLoS Computational Biology, vol. 5, no. 3, e1000302, Mar 2009.
A. Flinker, W.K. Doyle, A.D. Mehta, O. Devinsky, D. Poeppel. "Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries". Nature Human Behaviour, vol. 3, no. 4, pp. 393-405. Apr. 2019.
E. Sohoglu, M. H. Davis. "Rapid computations of spectrotemporal prediction error support perception of degraded speech". Elife, vol. 4, no. 9, e58077, Nov. 2020.
C.R. Holdgraf, W. de Heer, B. Pasley, J. Rieger, N. Crone, J.J. Lin, R.T. Knight, F.E. Theunissen. "Rapid tuning shifts in human auditory cortex enhance speech intelligibility". Nature Communications, vol 7 (13654), Dec. 2016.
G.M. Di Liberto, M.J. Crosse, E.C. Lalor. "Cortical Measures of Phoneme-Level Speech Encoding Correlate with the Perceived Clarity of Natural Speech". eNeuro, vol. 5, no. 2, Apr. 2018.
K.D. Prinsloo, E.C. Lalor. "General Auditory and Speech-Specific Contributions to Cortical Envelope Tracking Revealed Using Auditory Chimeras". J Neuroscience, vol 42, no. 41, pp. 7782-7798. Oct. 2022.

ignacio.calderondepalma@radboudumc.nl
Neural Engineering Group: Lopez Valdes Lab
@LoVaLab_TCD
www.lovalab.net