Effect of spectral resolution on neural entrainment of the speech envelope

Jonas Vanthornhout[1] Lien Decruy[1] Tom Francart[1]

[1]KU Leuven - University of Leuven, Department of Neurosciences, ExpORL

1. Introduction

Neural entrainment of the speech envelope

Speech envelope is a primary cue for speech understanding (Shannon et al., 1995)
Cortical activity tracks the envelope of running speech (Peelle and Davis, 2012)
Reconstruction of speech envelope from cortical activity is possible (Ding and Simon, 2011)
Reconstruction quality correlates with behaviourally measured speech understanding (Vanthornhout et al., 2017)

Research question

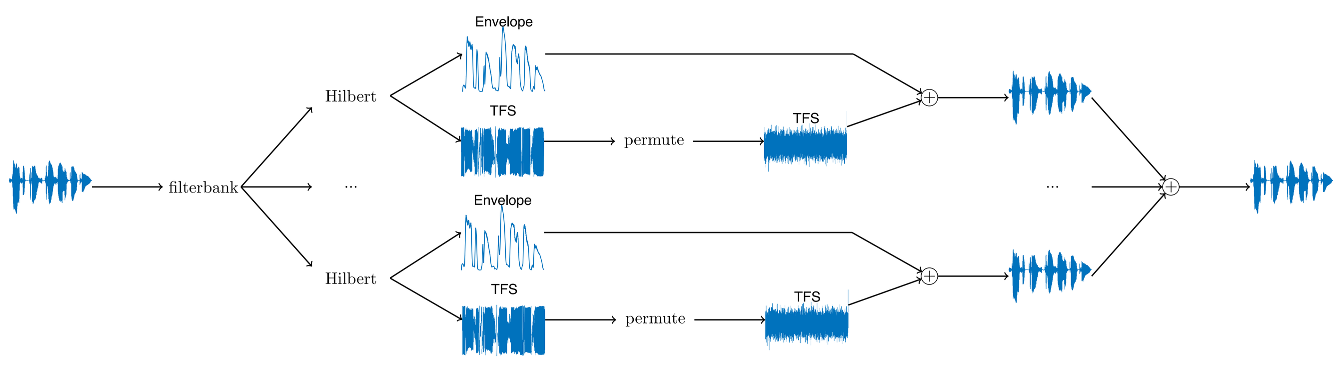
Can speech understanding alone influence envelope entrainment?
In previous experiments noise was added to the speech stimuli
⇒ SNR(speech) ↓ ⇒ SNR(envelope) ↓ ⇒ envelope entrainment ↓
⇒ Is neural entrainment a measure of stimulus acoustics or speech understanding?
⇒ Reduce speech understanding without distorting the envelope
⇒ Use vocoded & chimaera speech (Ding et al., 2014; Kong et al., 2015; Obleser and Weisz, 2011)

2. Methods

Stimuli

Vocoder
Retain envelope while replacing temporal fine structure with noise
2, 4, 6, 8 channels & clean speech

Chimaera
4 channel vocoder but retain a fixed amount of temporal fine structure
0%, 25%, 50%, 75% & 100% TFS



`Permute` a given percentage of TFS samples. If no permutation takes place (100% correct TFS) the original speech is obtained, if full permutation takes place (0% correct TFS) a noise vocoder is obtained.

Experiments

Vocoder: 7 young normal hearing subjects, aged 22-26 years
Chimeara: 9 different young normal hearing subjects, aged 22-27 years

- speech understanding: word scores
- EEG: BioSemi system with 64 electrodes
- binaural stimulation at 62 dBA
- speech in silence and with 3 dB SNR
- stationary speech weighted-noise (SWN)
- 2 s Flemish Matrix sentences: 20 sentences/condition
- 15 minutes Flemish story to train linear decoder

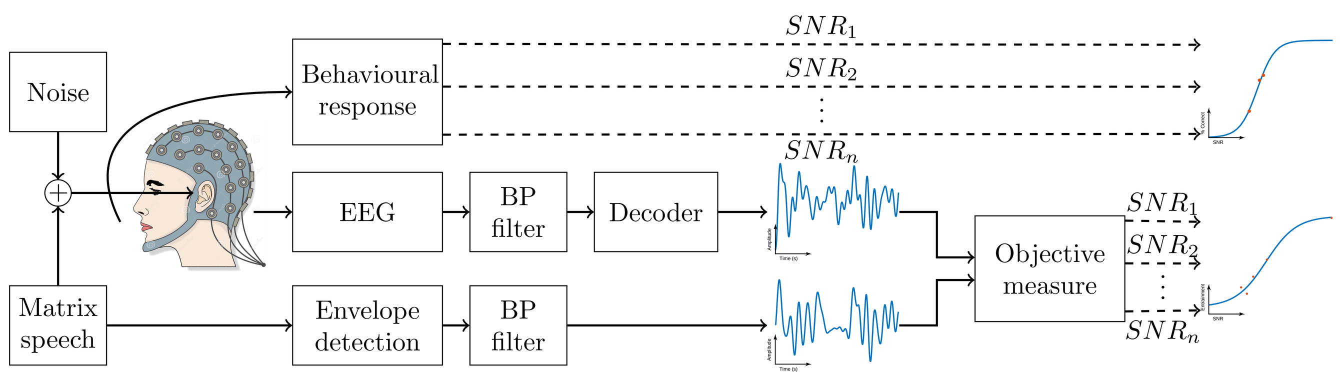


Figure: Behaviourally and objectively measured speech understanding using envelope entrainment

Signal processing

$$decoder = (RR^T)^{-1}(RS^T)$$

R time-lagged neural data   time lags 0-75 ms
S stimulus envelope of story
decoder minimises MSE between actual and reconstructed envelope

$$\hat{s}(t) = \sum_n \sum_\tau decoder(n,\tau) R(t+\tau, n)$$

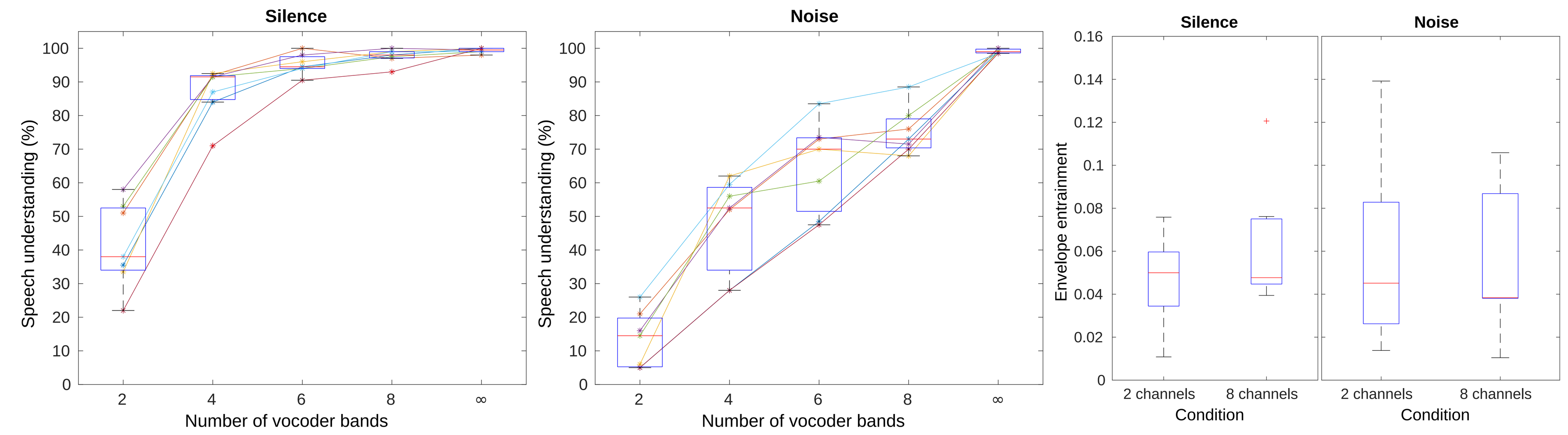
$\hat{s}$ reconstructed envelope
t time ranging from 0 to T
n recording electrodes ranging from 1 to N
τ post-stimulus samples used to reconstruct the envelope: integration window

$$envelope\ entrainment = correlation(s(t), \hat{s}(t))$$

s actual envelope

3. Results & Discussion

Vocoder

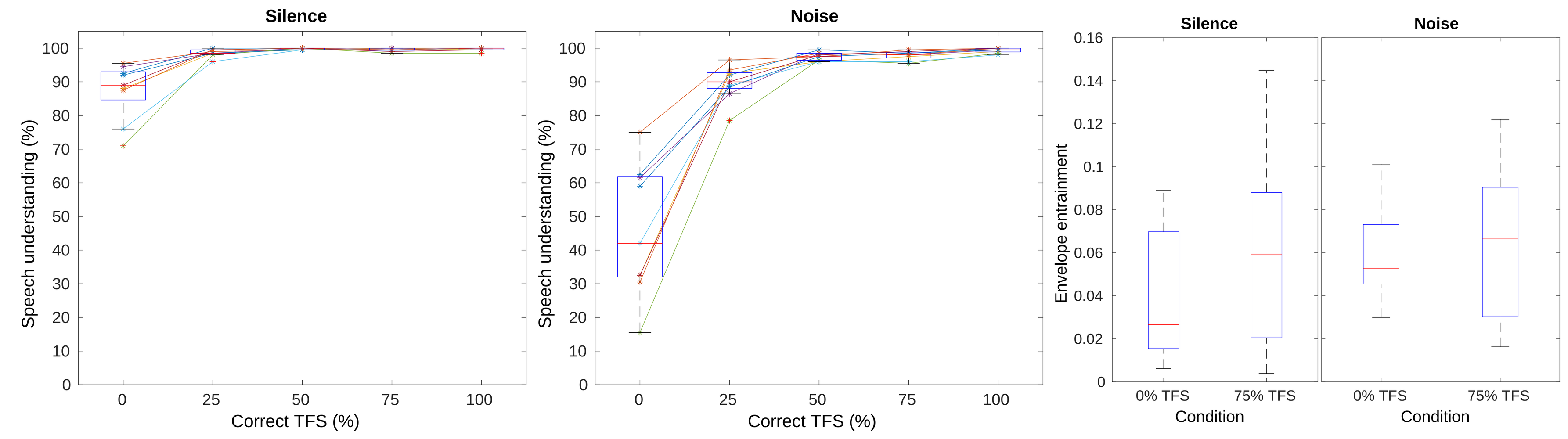


- speech understanding ↑ when numbers channels ↑
- speech in noise more difficult than speech in quiet
- vocoded speech in noise: largest range of speech understanding scores

Comparing 2 channel vocoder with 8 channel vocoder
- entrainment increases as spectral resolution increases
- not significantly ($p = 0.33$ & $p = 0.94$)

⇒ envelope entrainment ↑ when speech understanding ↑?

Chimaera



- speech understanding ↑ when TFS is less distorted
- speech in noise more difficult than speech in quiet

Comparing 0% TFS with 75% TFS
- entrainment increases as spectral resolution increases
- not significantly ($p = 1.00$ & $p = 0.82$)

⇒ envelope entrainment ↑ when speech understanding ↑?

Vocoder in noise

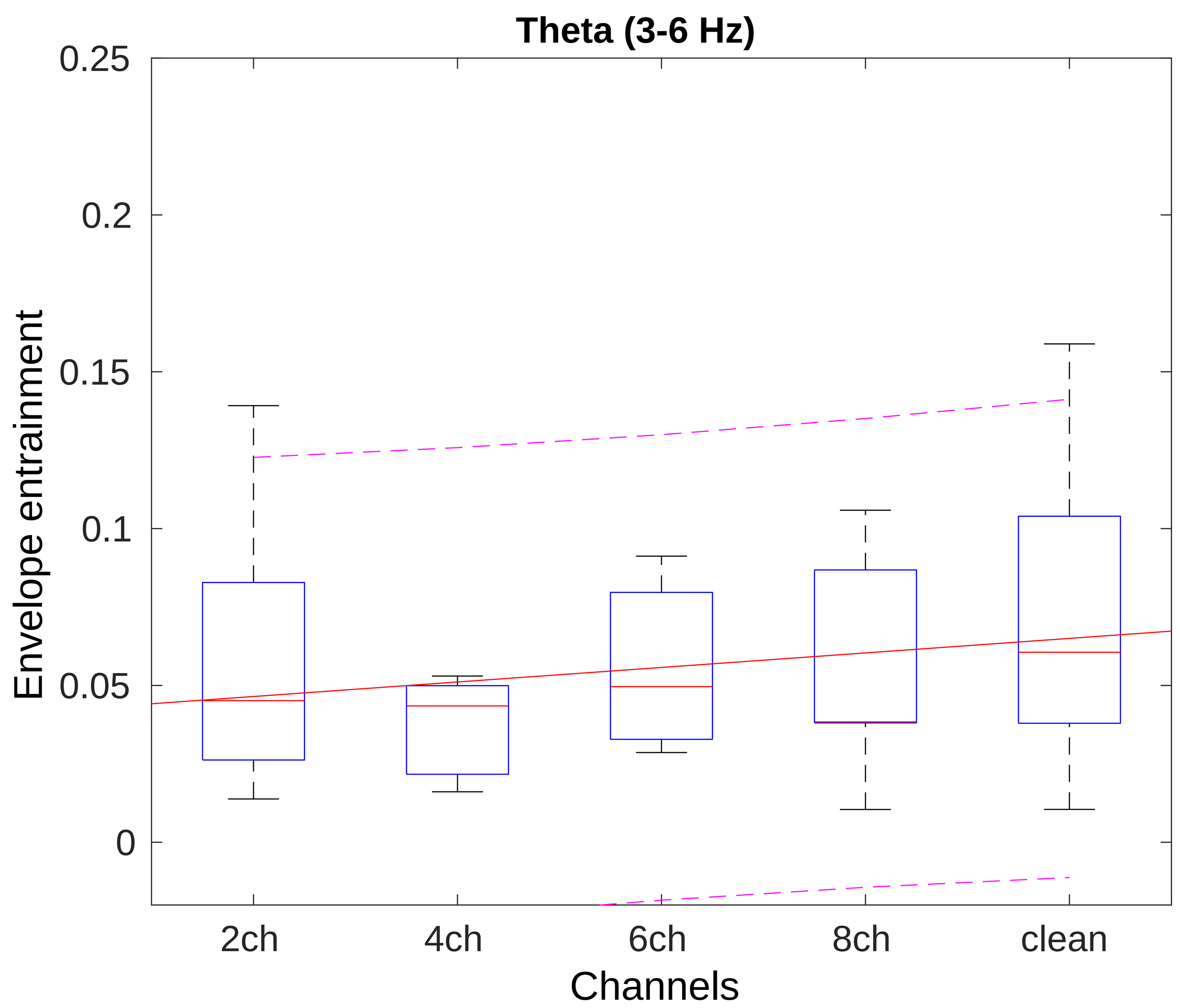
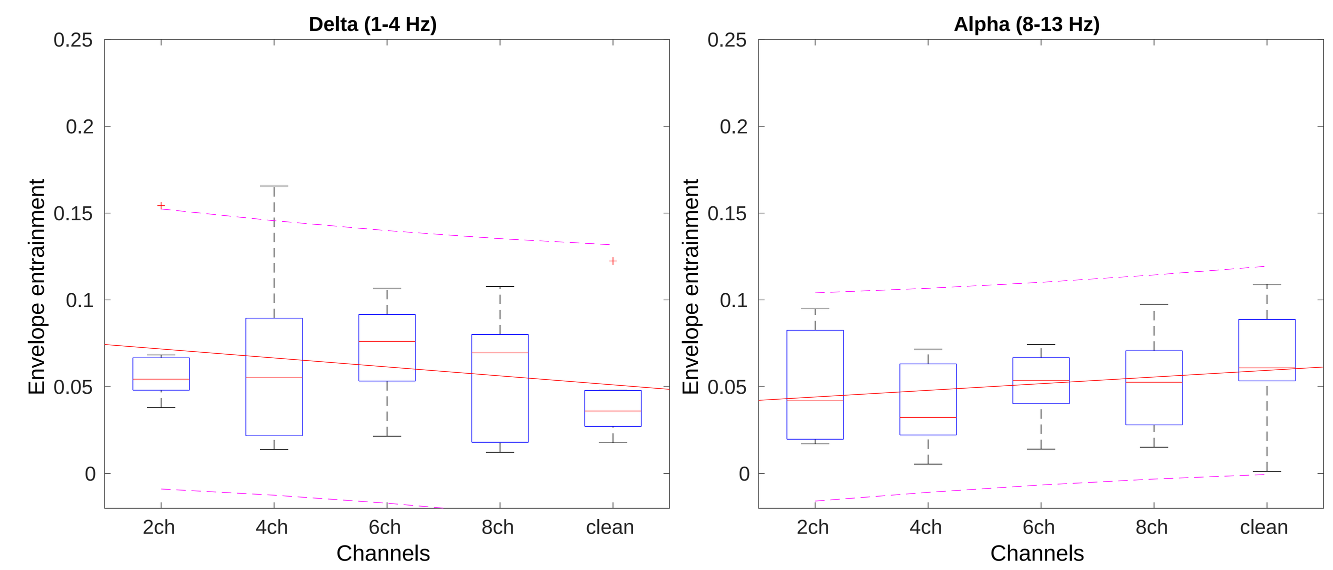


Figure: Envelope entrainment as a function of the number of vocoder bands.

Author	Measure	|*stimulus*|	noise	band	understanding	neural response
Ding et al. (2014)	entrainment	long	no	delta	↑	↓
Ding et al. (2014)			no	theta	↑	↑
Ding et al. (2014)			SSWN	delta	↑	↑(1)
Ding et al. (2014)			SSWN	theta	↑	↑
Kong et al. (2015)			speaker	∞	↑	↑
Obleser and Weisz (2011)	power	short	no	theta	↑	↑
Obleser and Weisz (2011)			no	alpha	↑	↓
ExpORL	entrainment	short	no	delta	↑	↑ ($p = 0.20$)?
ExpORL			no	theta	↑	↑ ($p = p = 0.76$)?
ExpORL			no	alpha	↑	↑ ($p = p = 0.83$)?
ExpORL			SSWN	delta	↑	↓ ($p = 0.26$)? (1)
ExpORL			SSWN	theta	↑	↑ ($p = 0.29$)?
ExpORL			SSWN	alpha	↑	↑ ($p = 0.26$)?

(1) difference can be explained by the stimulus length, related to top-down attention & listening effort



(a) (b)

Open question:

$$entrainment = frequency * (\alpha \cdot speech\ understanding + \beta \cdot envelope\ coding + \gamma \cdot age + \delta \cdot effort + \varepsilon \cdot attention)$$

References

Nai Ding and Jonathan Z. Simon. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, 2011.

Nai Ding, Monita Chatterjee, and Jonathan Z Simon. Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *Neuroimage*, 88:41–46, 2014.

Ying-Yee Kong, Ala Somarowthu, and Nai Ding. Effects of spectral degradation on attentional modulation of cortical auditory responses to continuous speech. *Journal of the Association for Research in Otolaryngology*, 16(6):783–796, 2015.

Jonas Obleser and Nathan Weisz. Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral cortex*, 22(11):2466–2477, 2011.

J. E. Peelle and M. H. Davis. Neural Oscillations Carry Speech Rhythm through to Comprehension. *Front Psychol*, 2012.

R V Shannon, F G Zeng, V Kamath, J Wygonski, and M Ekelid. Speech recognition with primarily temporal cues. *Science*, 1995.

Jonas Vanthornhout, Lien Decruy, Jan Wouters, Jonathan Z. Simon, and Tom Francart. Speech intelligibility predicted from neural entrainment of the speech envelope. *Submitted*, 2017.

Acknowledgements

Research funded by a PhD grant of the Research Foundation Flanders (FWO). This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 637424). The authors thank Karlien Smolders and Cleme Roels for helping collecting data.

QR code

A digital version of this poster can be found on `https://exporl.med.kuleuven.be/web/index.php/Public:ICAC2017` or by scanning the QR code.