Any questions?

- Practicalities?
- Any open issues from yesterday?
    - Neurobiology of rate normalization?

Lecture 2: *neural tracking*

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, 3. doi:10.3389/fpsyg.2012.00320.

Hans Rutger Bosker

Speech Perception in Audiovisual Communication [SPEAC] lab

Donders Institute, Radboud University, Nijmegen, The Netherlands

https://hrbosker.github.io

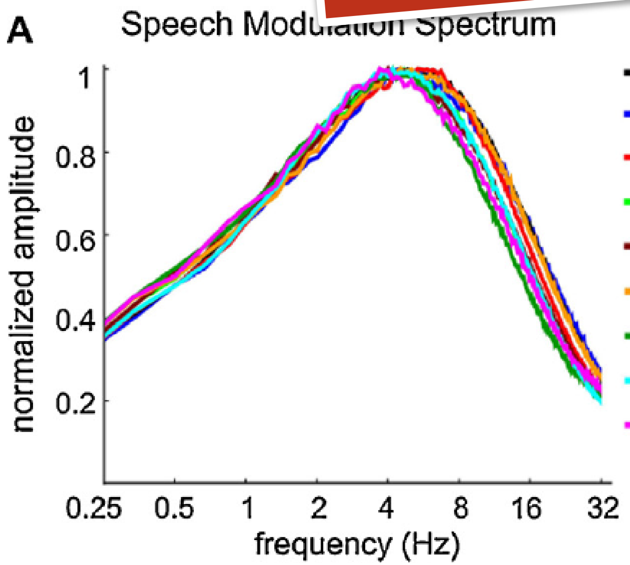
hansrutger.bosker@donders.ru.nl

European Research Council
Established by the European Commission

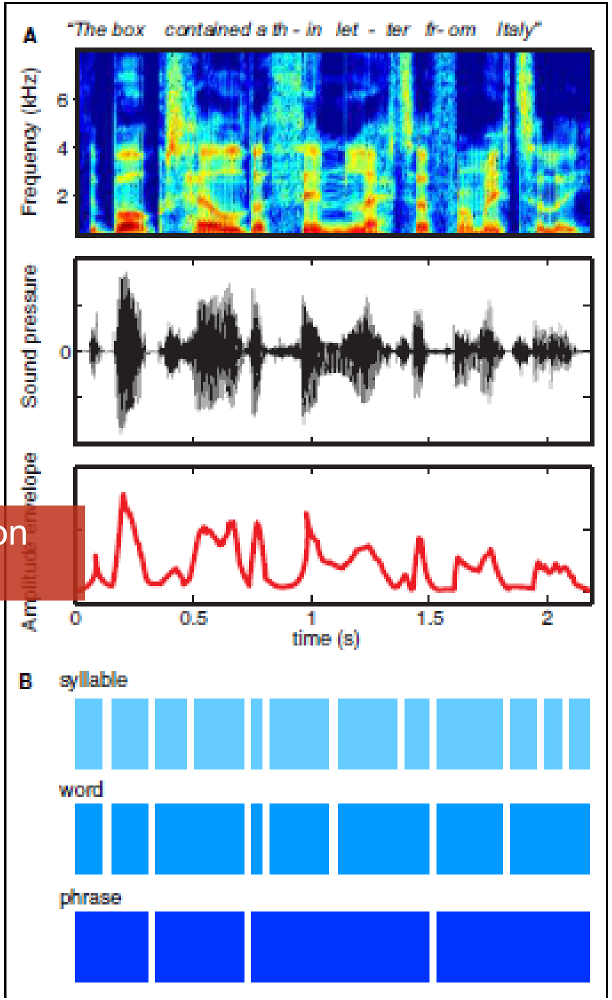
Radboud University Radboudumc

Speech is a rhythmic signal

Is this prosody?



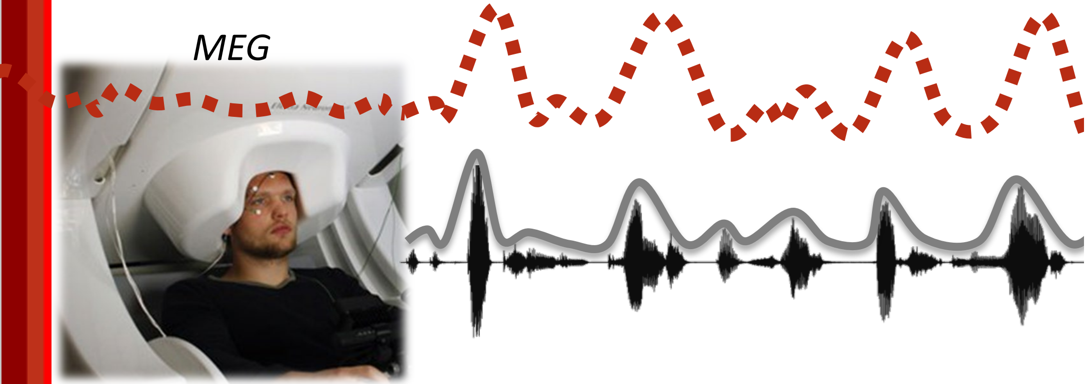
spectral decomposition
(FFT)



Peelle & Davis, 2012; Ding et al., 2016

The brain ‘tracks’ the speech rhythm

MEG



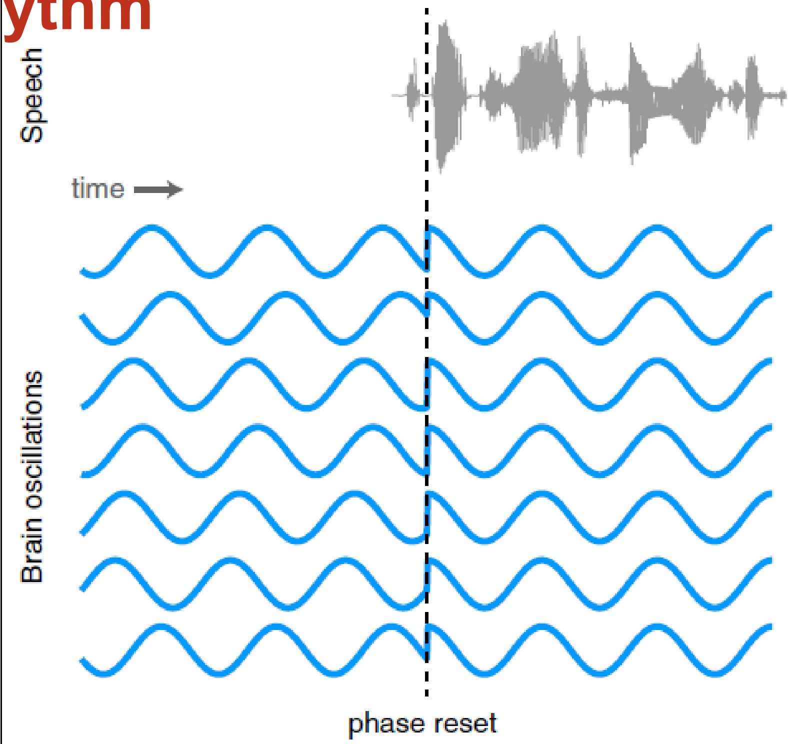
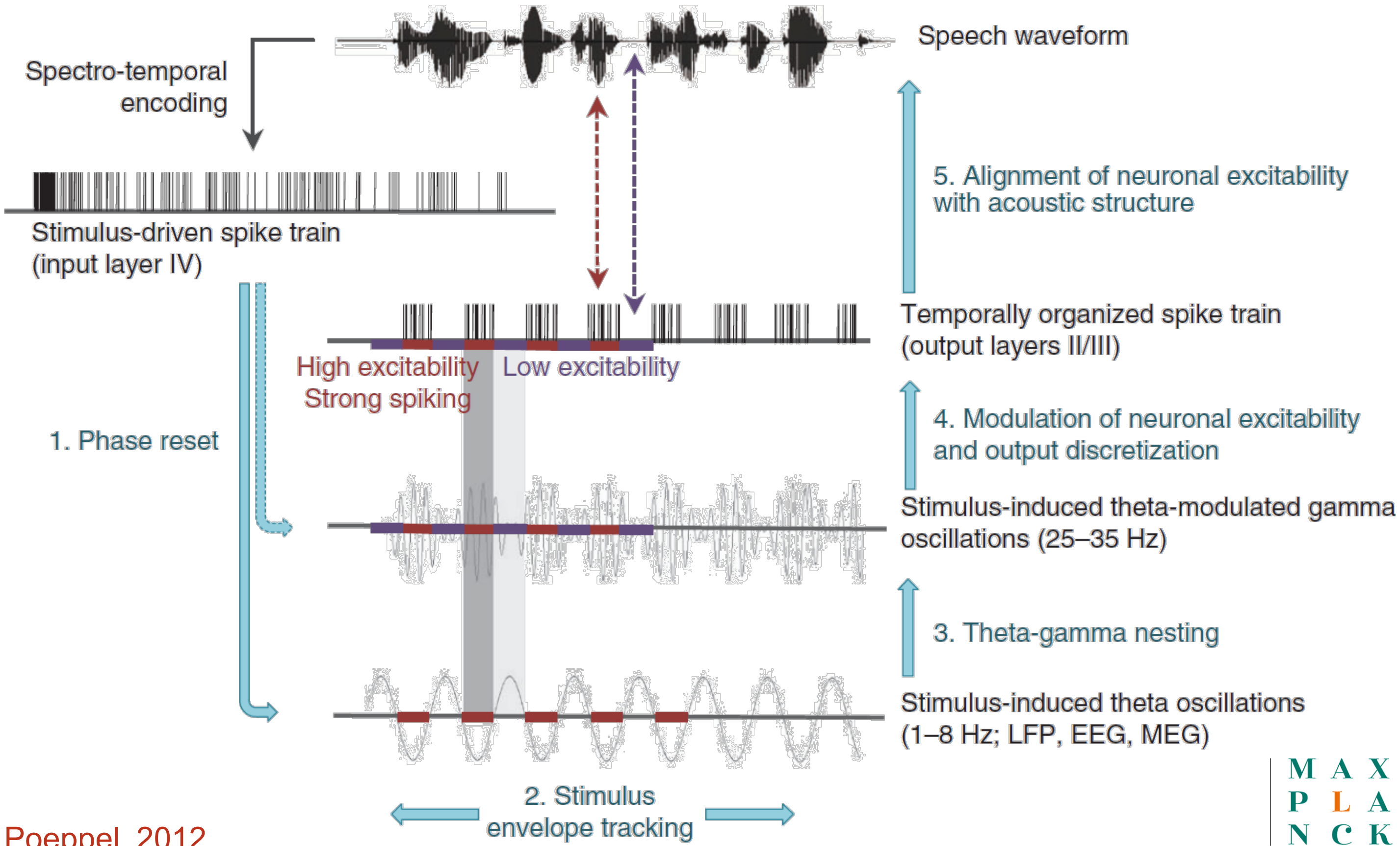


FIGURE 5 | Illustration of the phase of ongoing neural oscillations being reset by an external stimulus. Prior to the stimulus event the phases of the oscillations are random, but following the stimulus they are aligned.

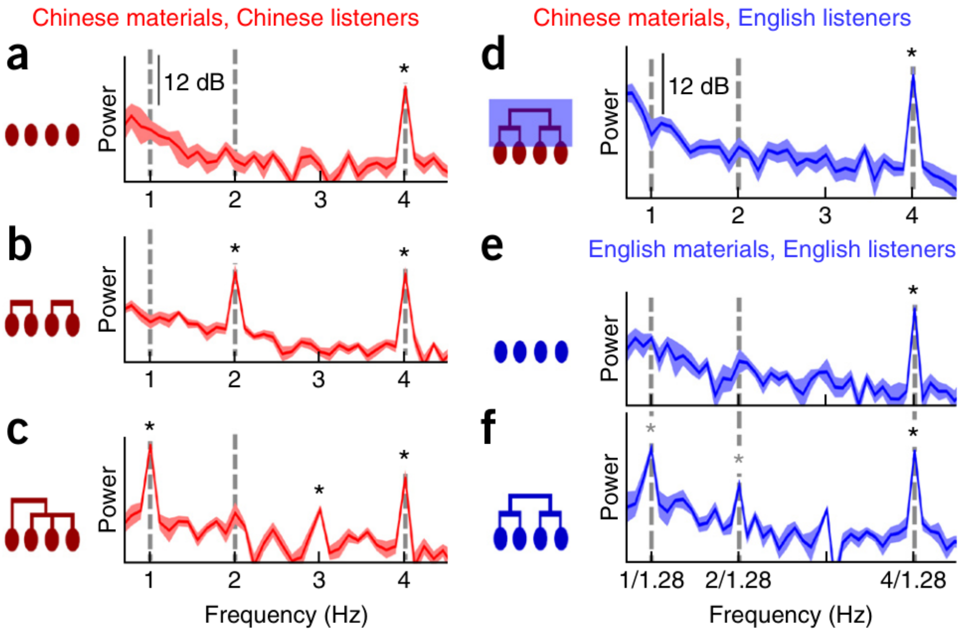
Bosker, 2017; Ghitza, 2011; Giraud & Poeppel, 2012; Peelle & Davis, 2012

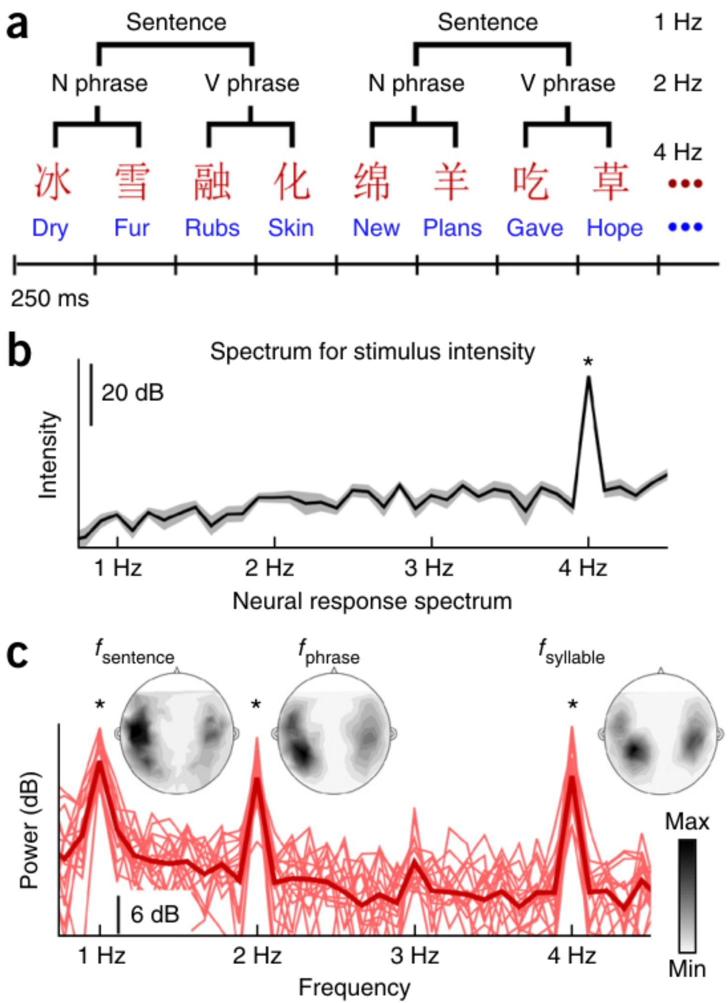
The brain 'tracks' the speech rhythm



Peelle & Davis, 2012; Giraud & Poeppel, 2012

The brain ‘tracks’ language?





Ding et al., 2016, *Nature Neuroscience*

The brain ‘tracks’ language?

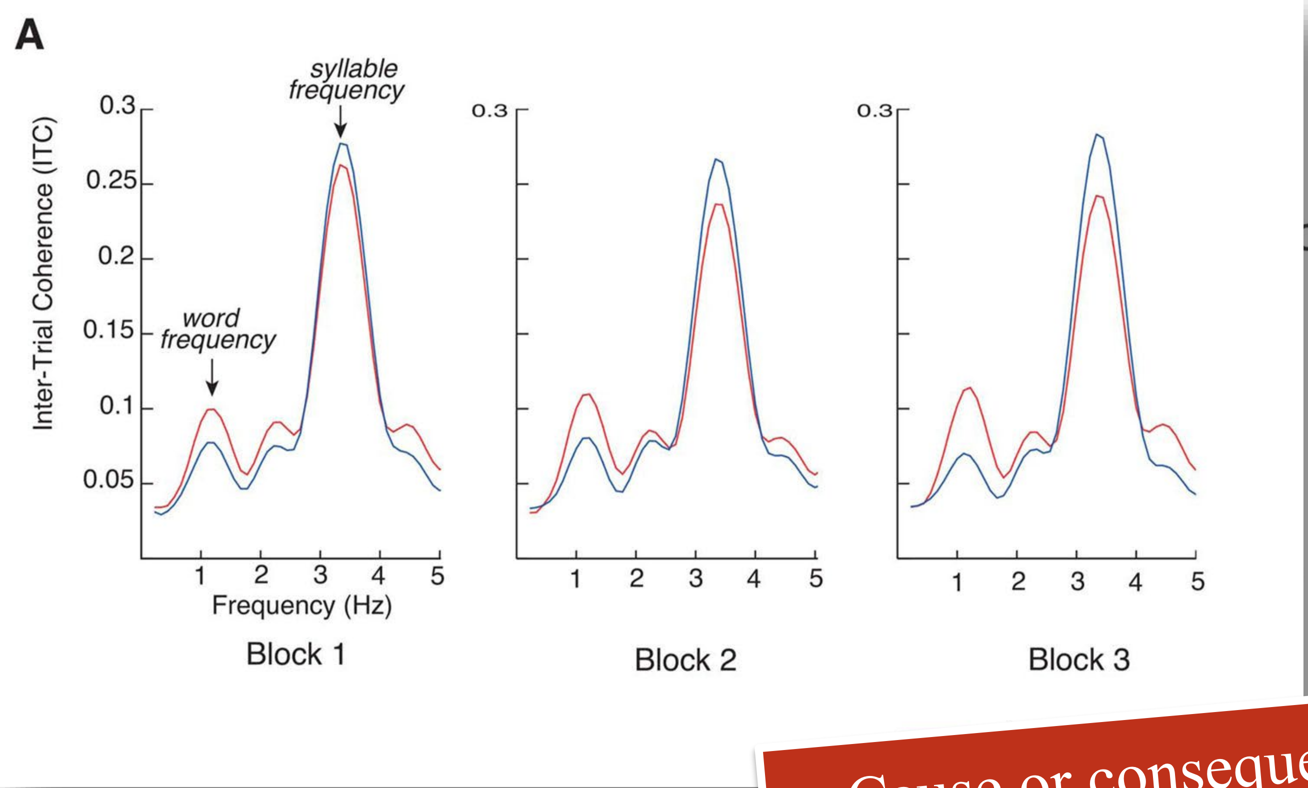
“Rand

la tu g

“Struct

tu pi r

cy

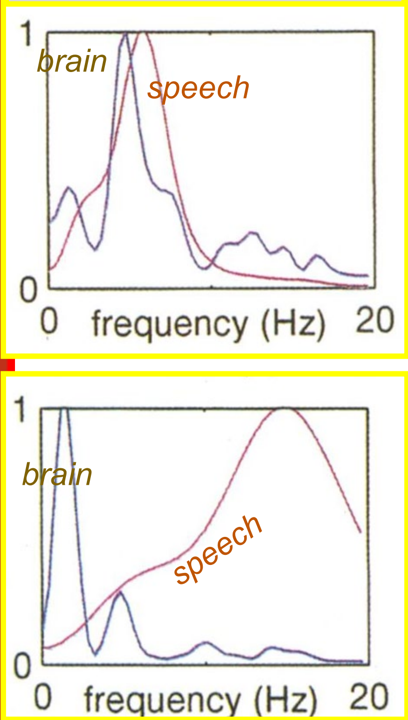


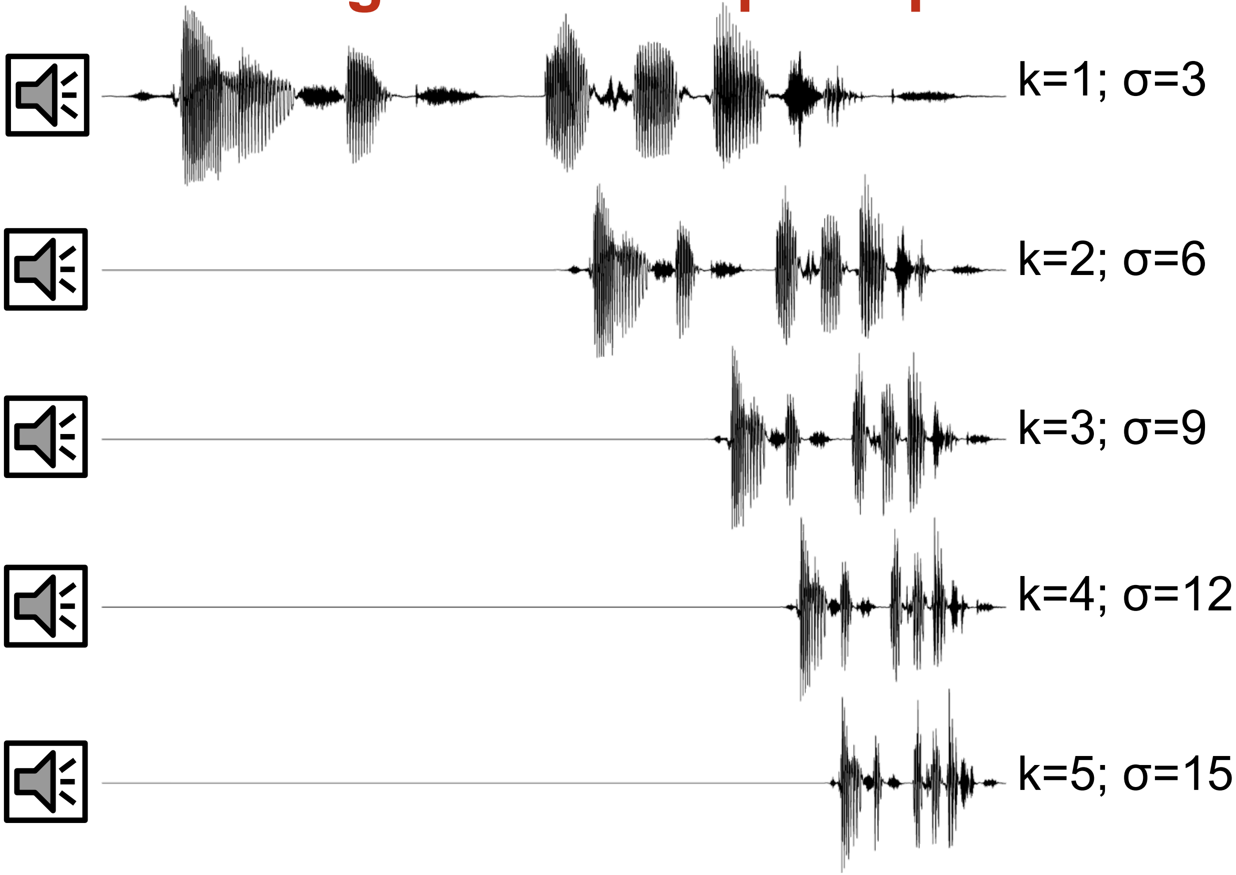
Cause or consequence?

Batterink & Paller, 2017, *Cortex*

Neural speech tracking influences perception?

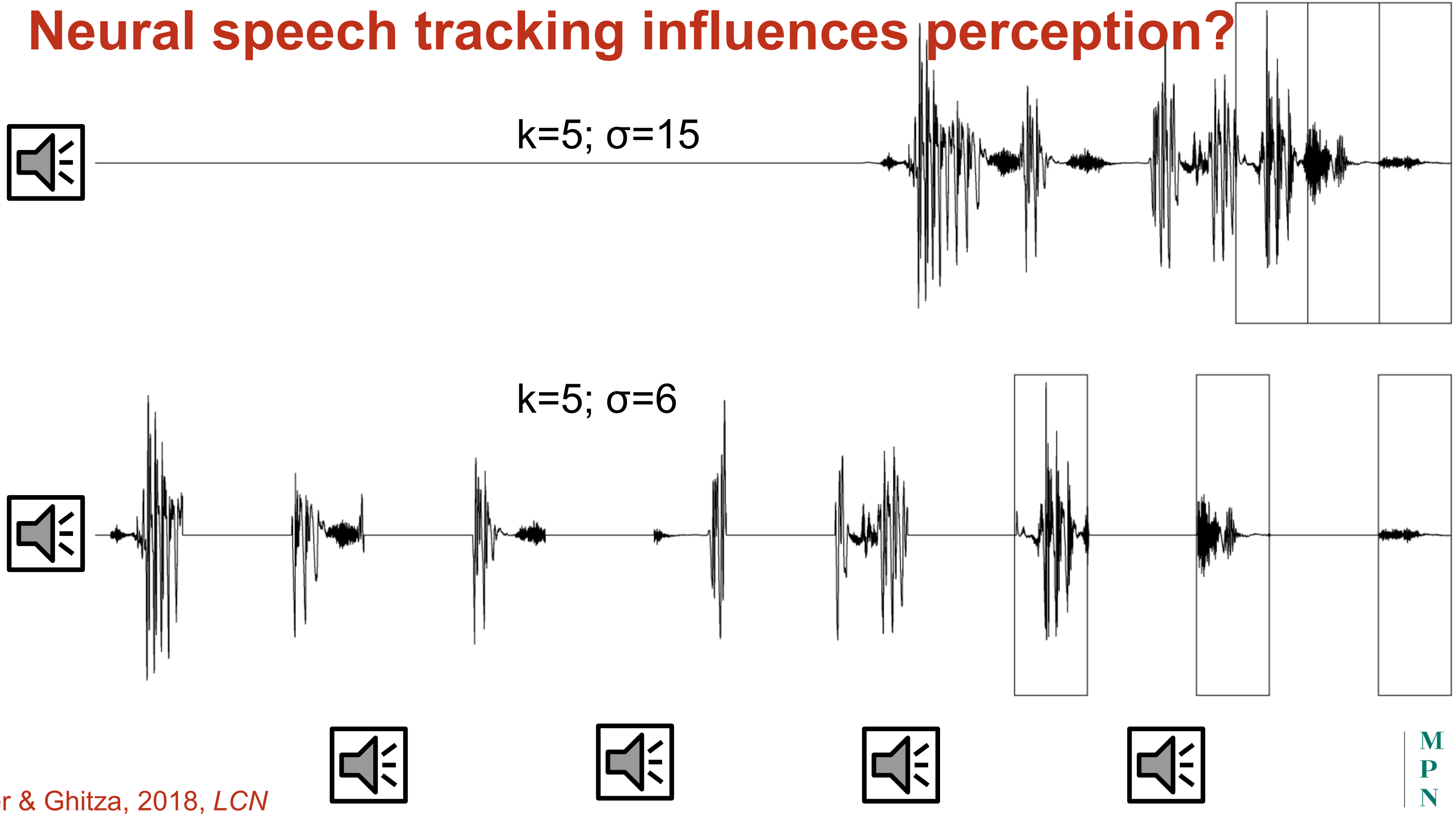
Ahissar et al., 2001





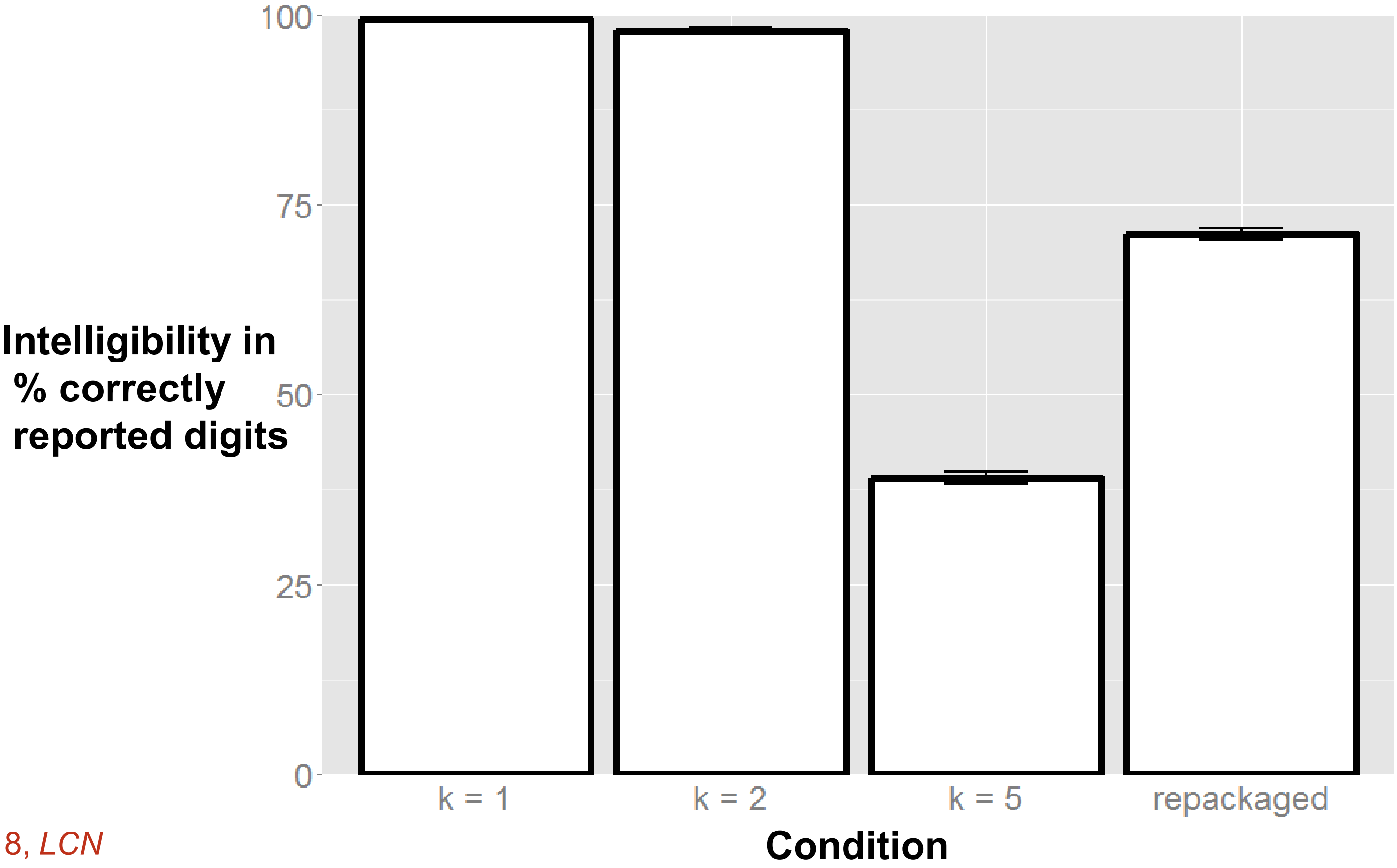
Bosker & Ghitza, 2018, *LCN*

Neural speech tracking influences perception?



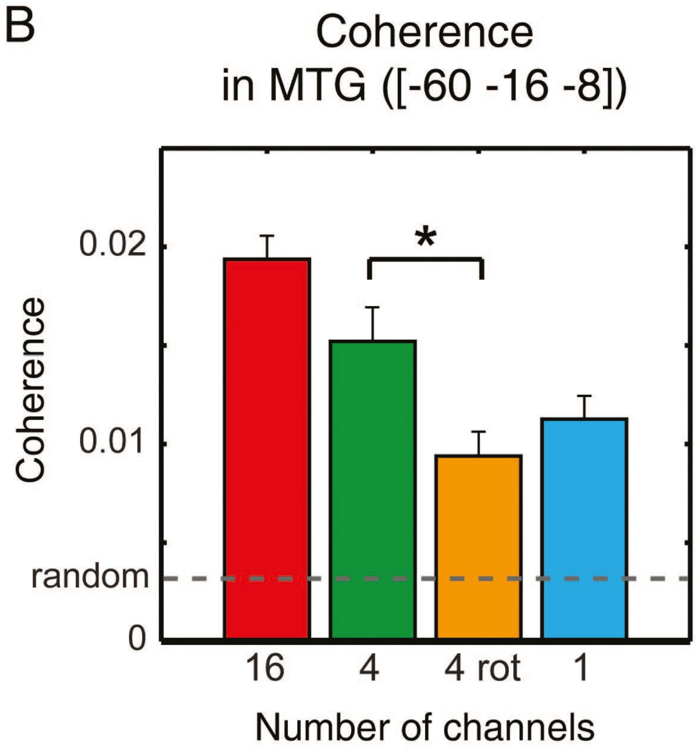
Bosker & Ghitza, 2018, *LCN*

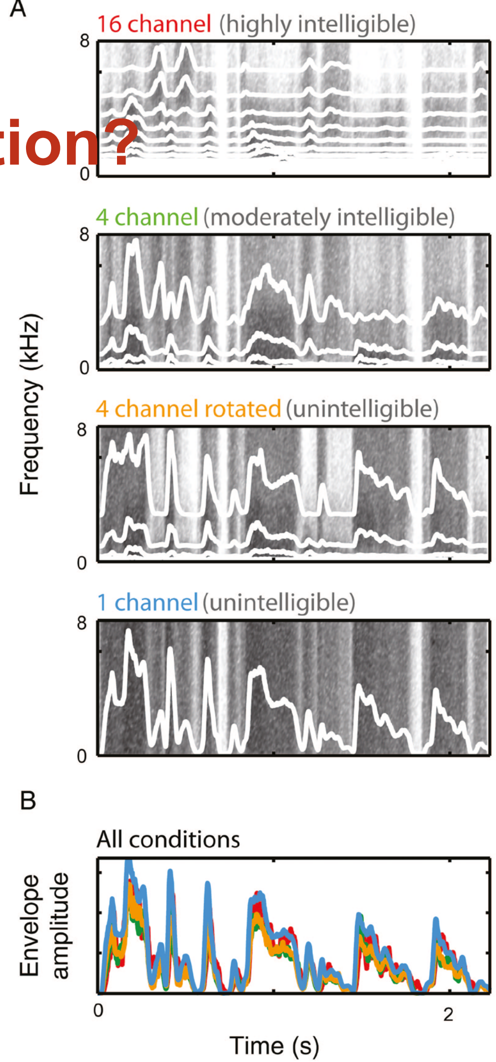
Neural speech tracking influences perception?



Bosker & Ghitza, 2018, *LCN*

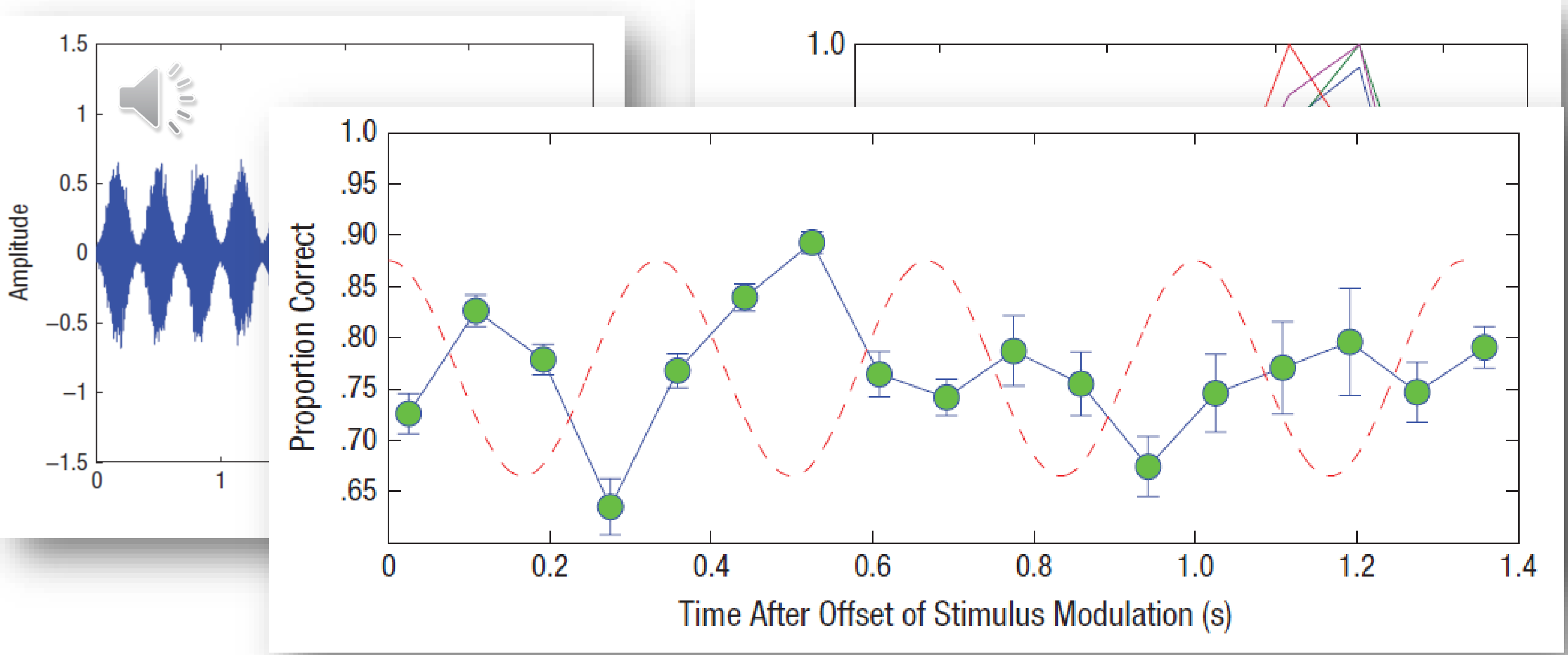
Neural speech tracking influences perception?





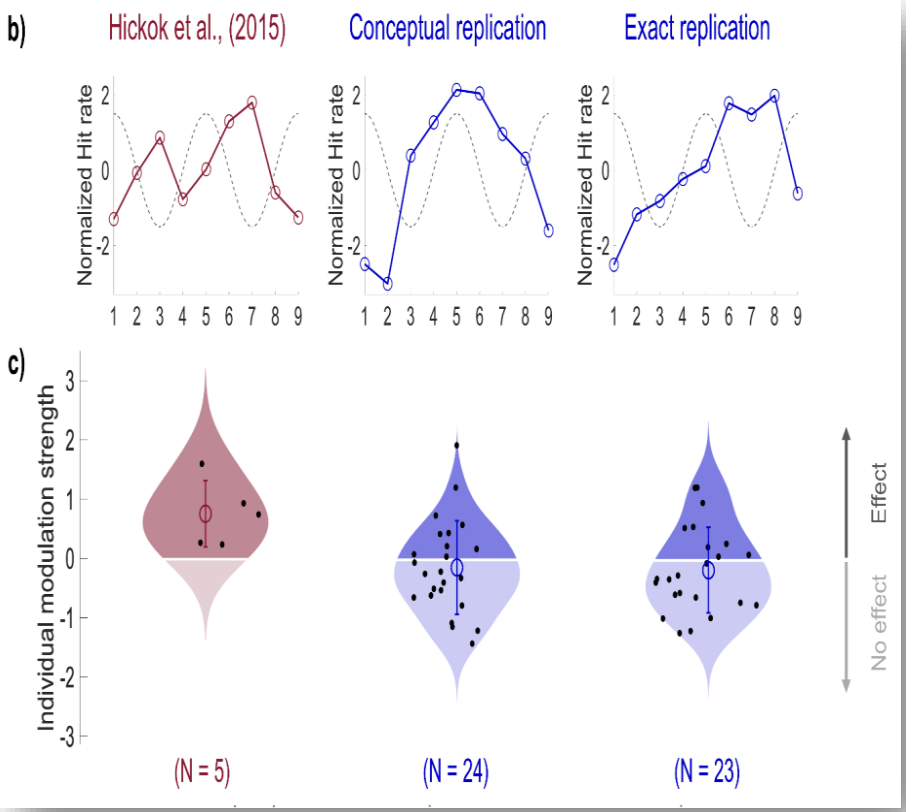
Peelle et al., 2013

Neural speech tracking influences perception?



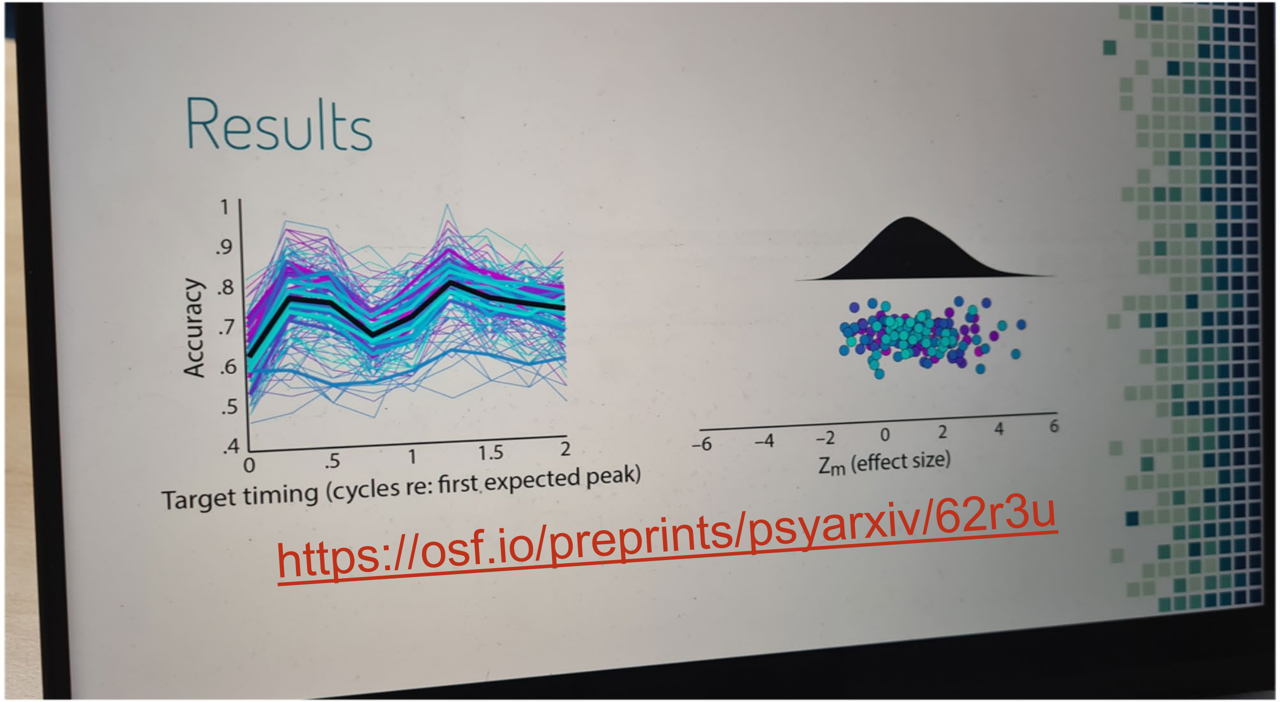
Hickok et al., 2015

Neural speech tracking influences perception?



Sun et al., 2021

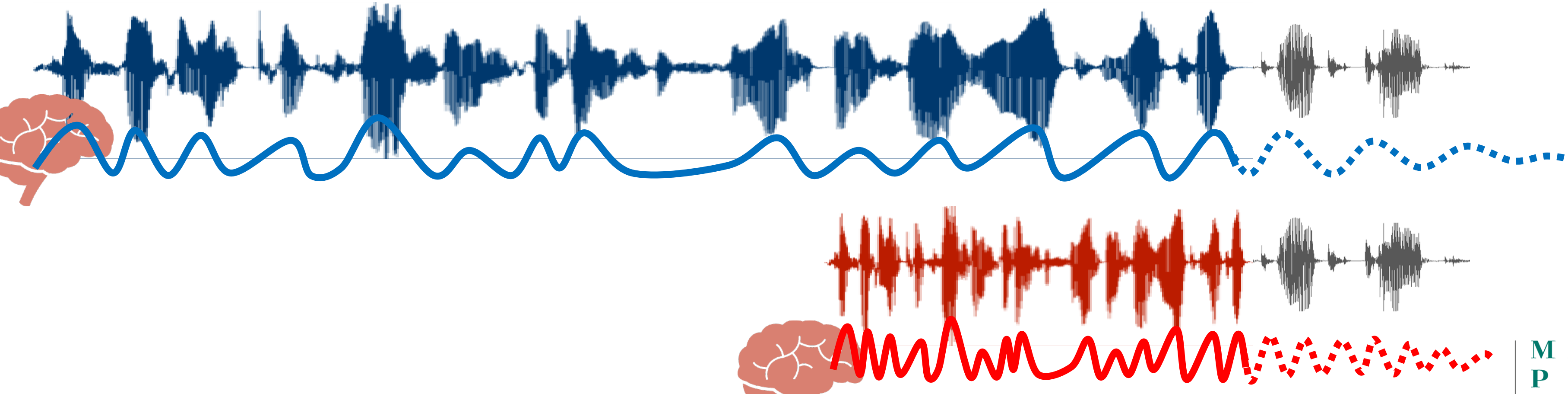
Neural speech tracking influences perception?



Multisite Overview of Oscillations [MOO] team science project, headed by Jonathan Peelle & Molly Henry

Causal effect

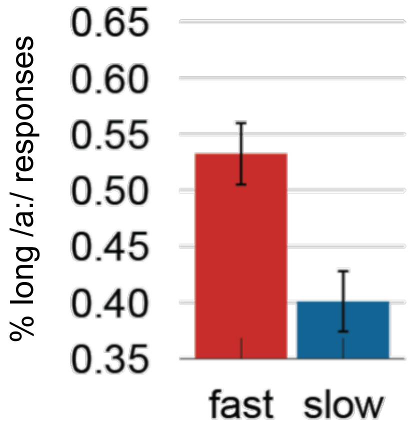
- Tracking of slow speech...
- Tracking of fast speech...
- What happens after change in speech rate?

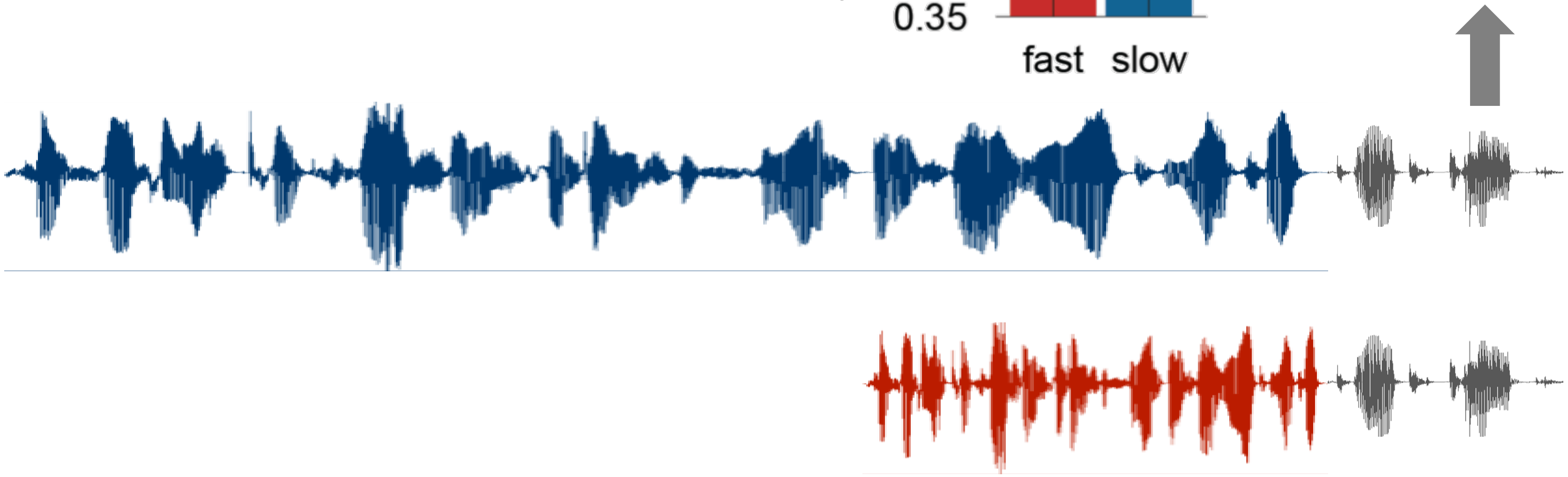


Kösem, Bosker, Takashima, Meyer, Jensen, & Hagoort (2018)

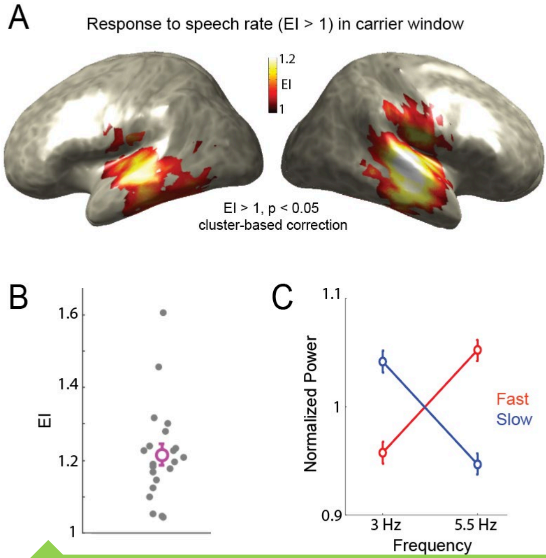
Causal effect

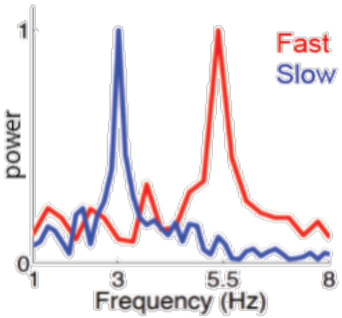
- Behavior: rate normalization

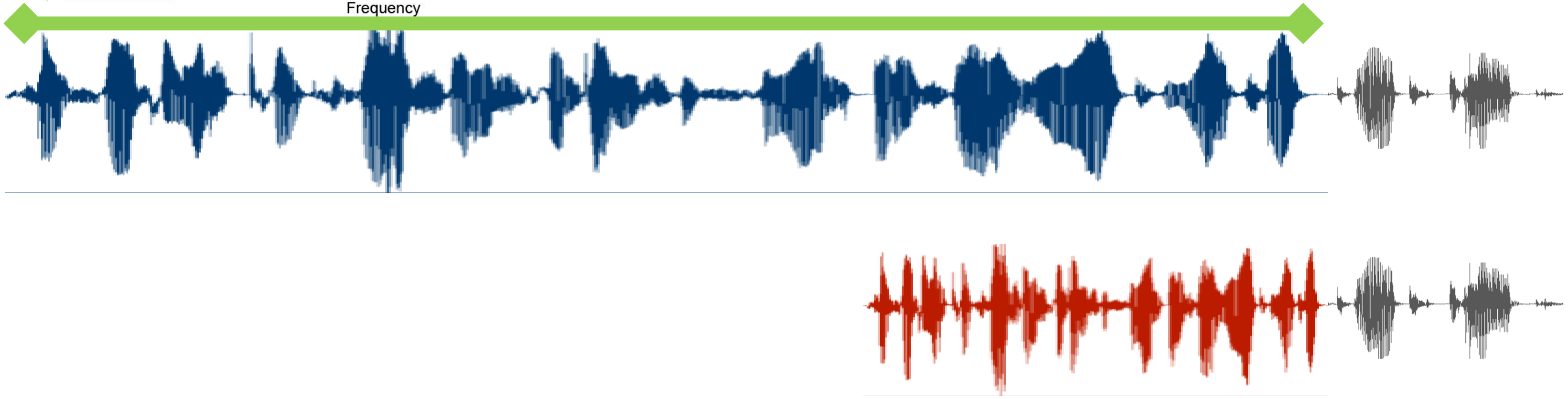




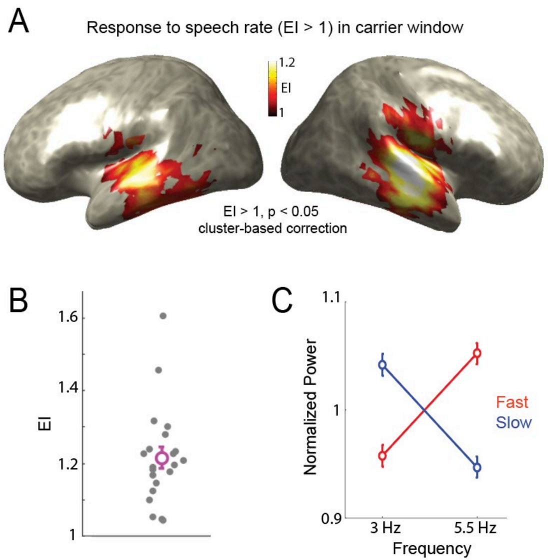
Kösem, Bosker, Takashima, Meyer, Jensen, & Hagoort (2018)

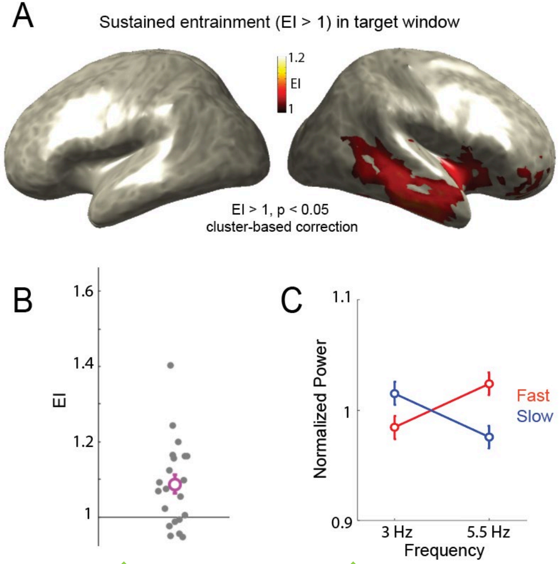


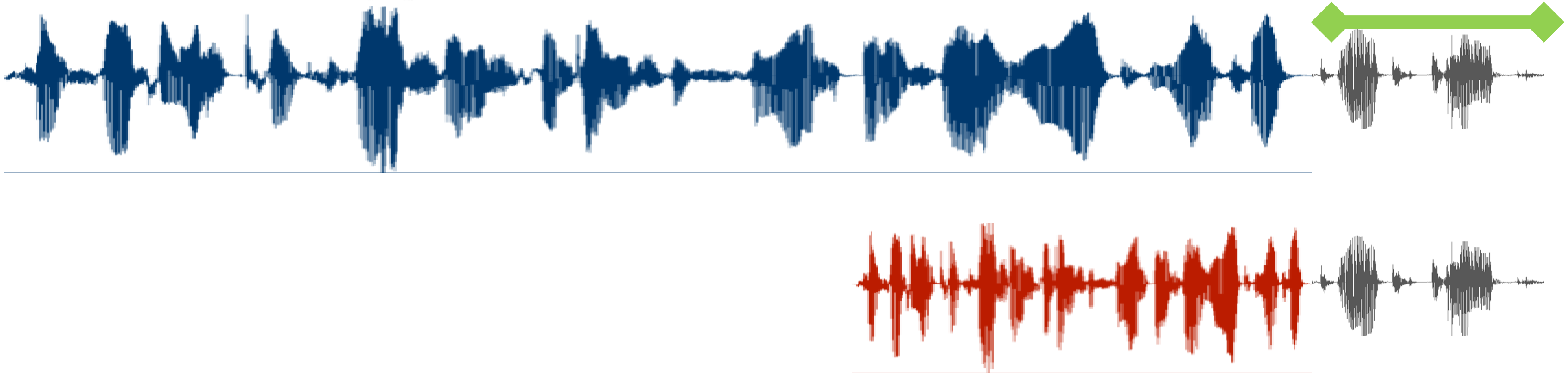




Kösem, Bosker, Takashima, Meyer, Jensen, & Hagoort (2018)



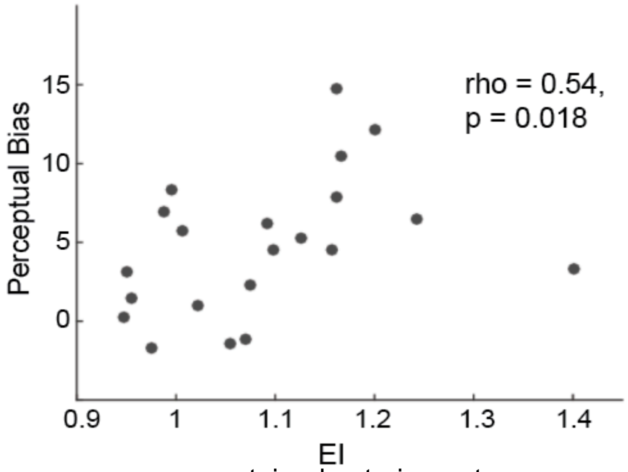


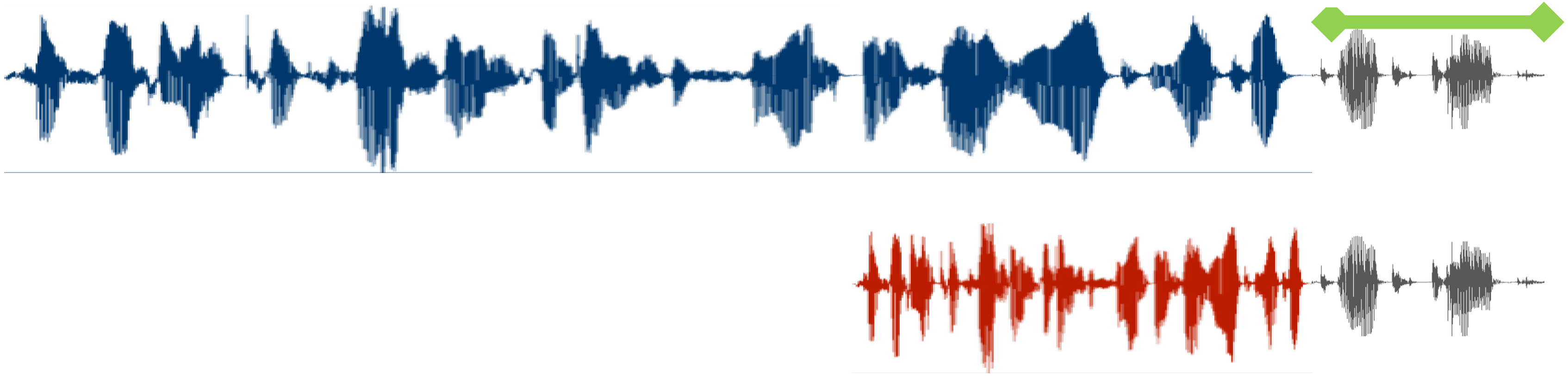


Kösem, Bosker, Takashima, Meyer, Jensen, & Hagoort (2018)

Causal effect

- Sustained entrainment correlates with behavior

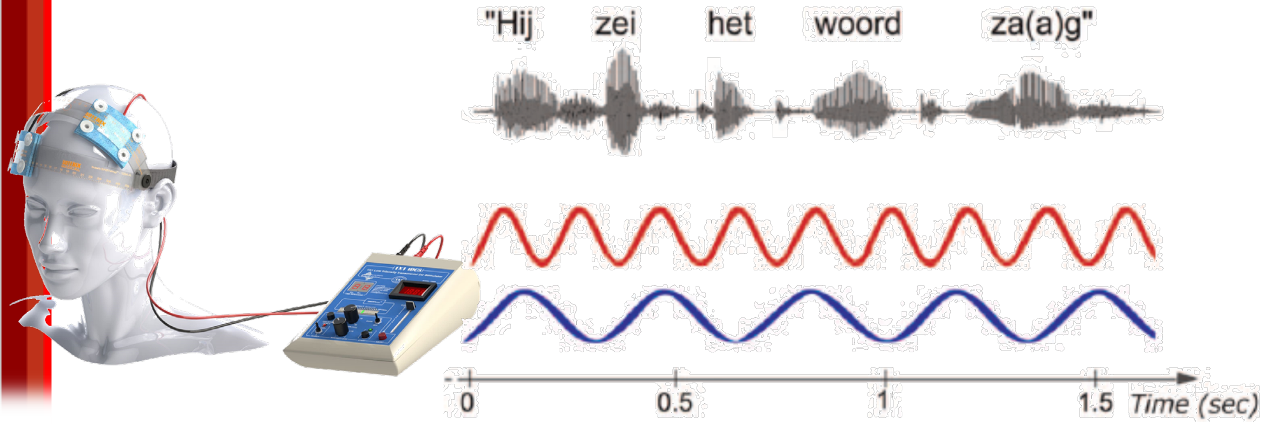


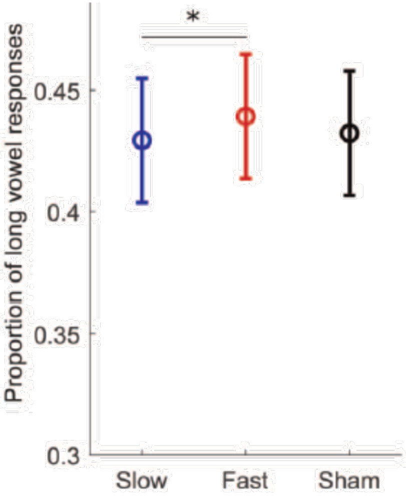


Kösem, Bosker, Takashima, Meyer, Jensen, & Hagoort (2018)

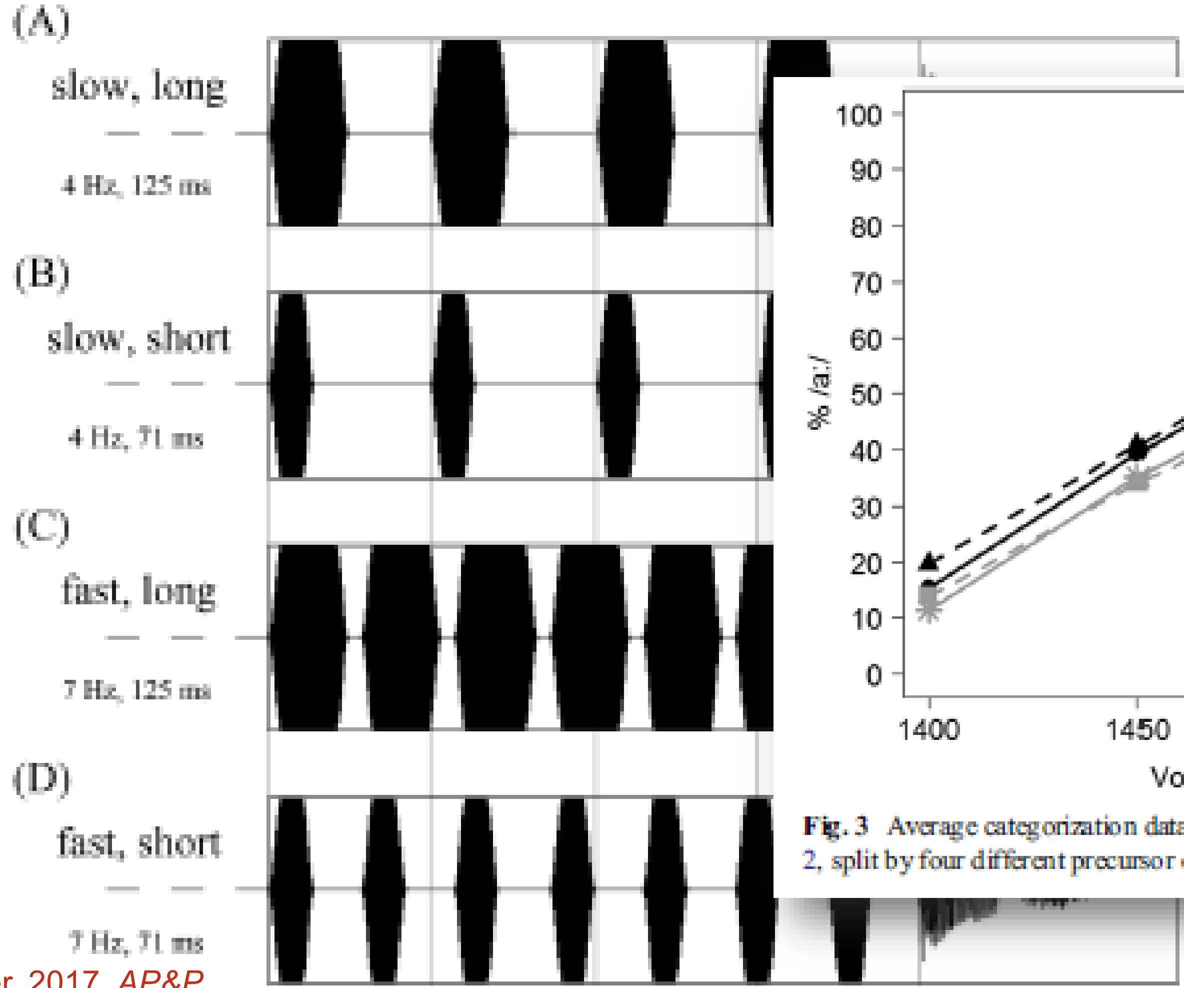
…well, that predicts that:

- …‘zapping the brain’ (tACS) at fast vs. slow rhythms influences vowel length perception





Kösem, Bosker et al., 2020, *JoCN*



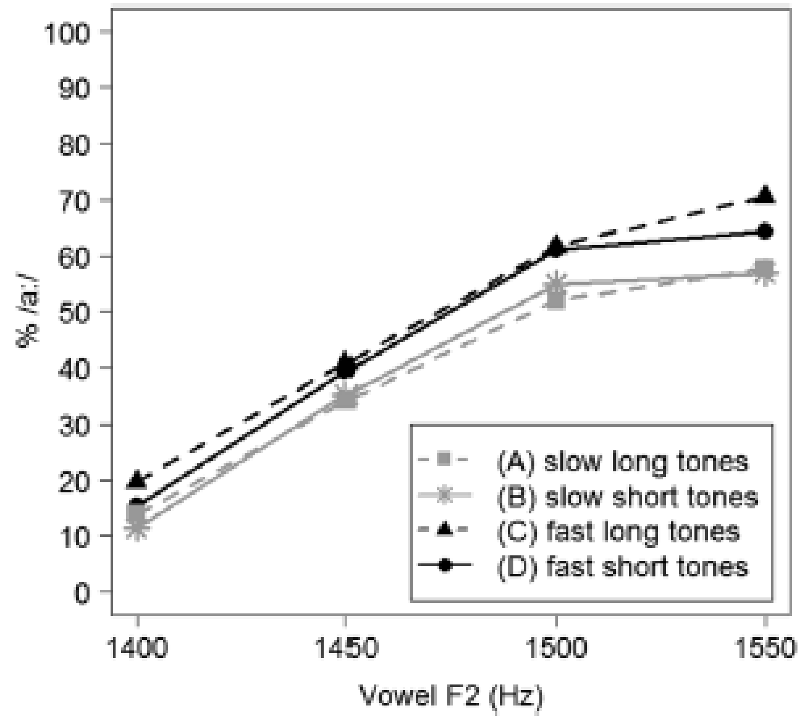


Fig. 3 Average categorization data (in % /a:/ responses) for Experiment 2, split by four different precursor conditions

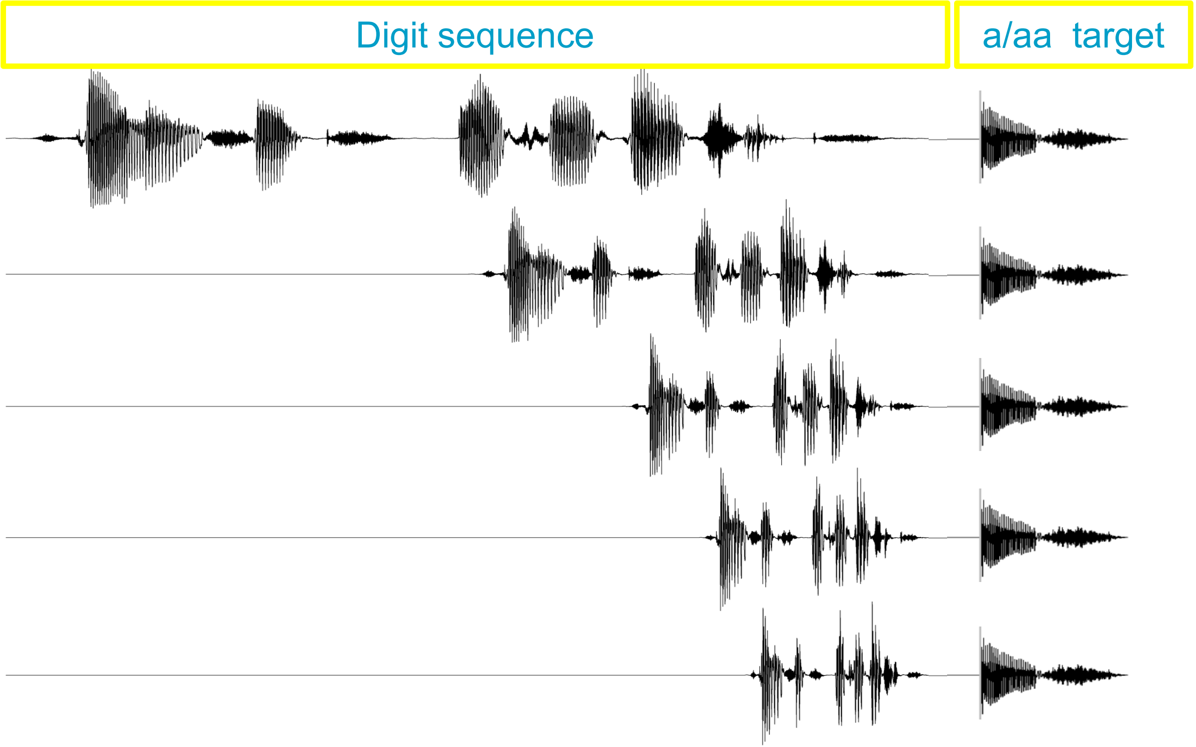
Bosker, 2017, *AP&P*

…well, that predicts that:

- …‘zapping the brain’ (tACS) at fast vs. slow rhythms influences vowel length perception
- …rate normalization depends on the context’s rhythm, not the preceding unit’s duration
- …there’s an upper limit to the speech rates that induce rate normalization

Bosker & Ghitza, 2018, *LCN*

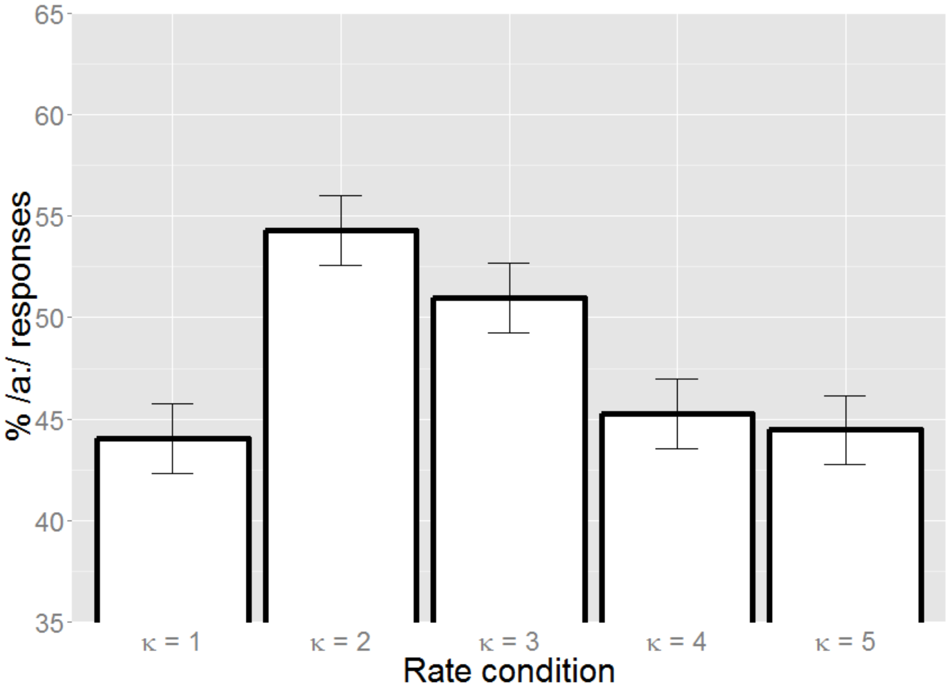
UPPER LIMIT



Bosker & Ghitza, 2018, *LCN*

UPPER LIMIT

- Behavior



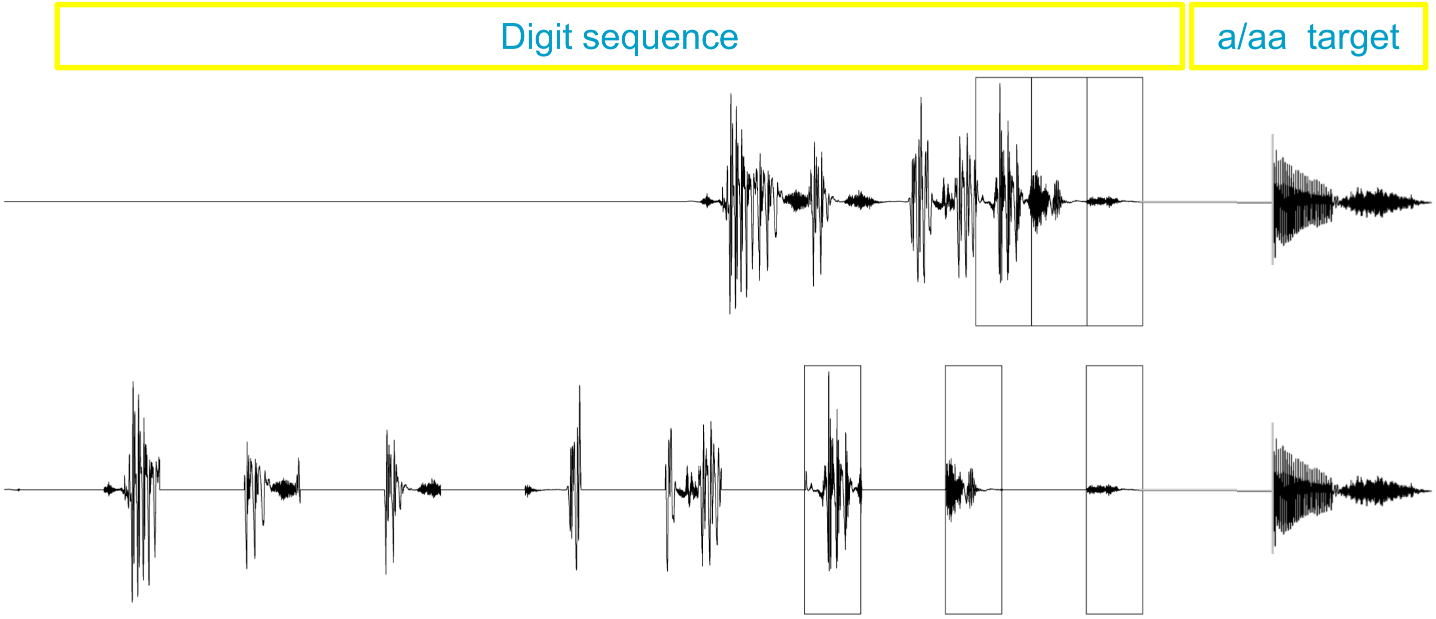
Bosker & Ghitza, 2018, *LCN*

UPPER LIMIT

- Only rhythms inside *theta* range induce 'sustained rhythm' effects (i.e., rate normalization)

Bosker & Ghitza, 2018, *LCN*

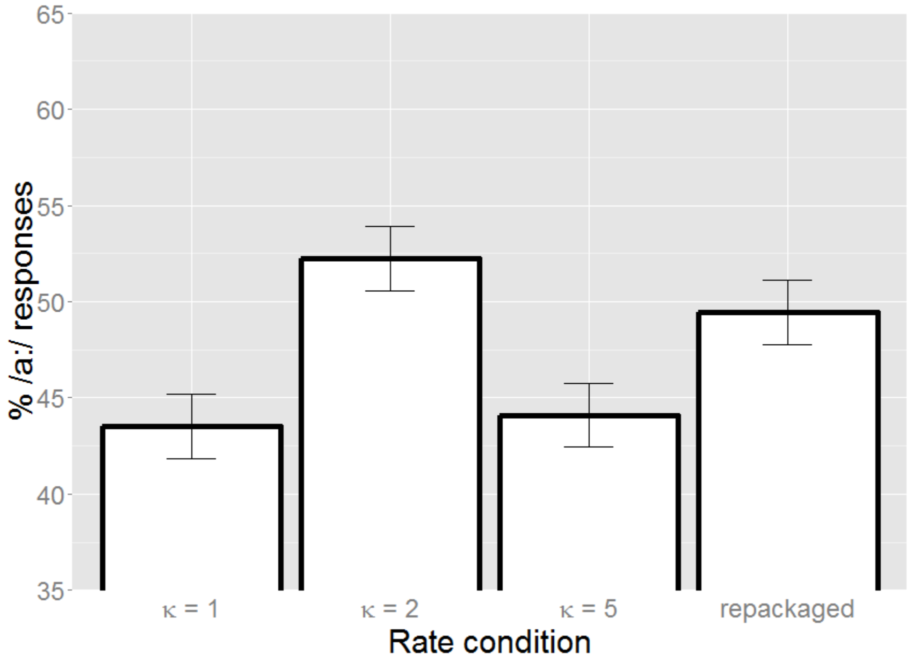
UPPER LIMIT



Bosker & Ghitza, 2018, *LCN*

UPPER LIMIT

- Behavior



Bosker & Ghitza, 2018, *LCN*

UPPER LIMIT

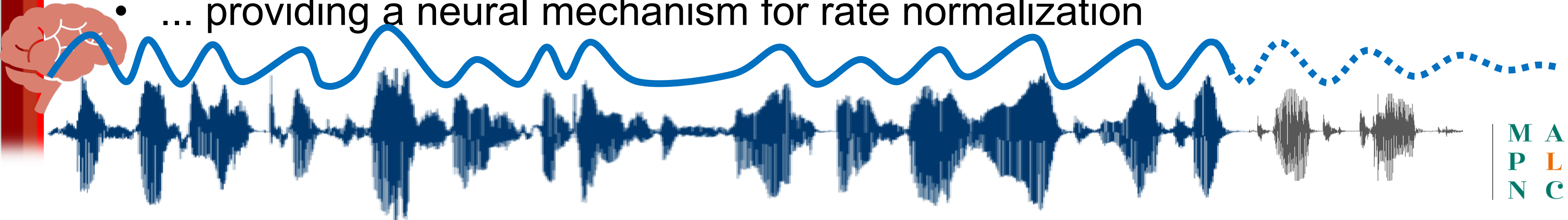
- Only rhythms inside *theta* range induce 'sustained rhythm' effects (i.e., rate normalization)

- If *theta* rhythm is imposed onto heavily compressed speech signal (through 'repackaging'), behavioral rate normalization is restored.

Bosker & Ghitza, 2018, *LCN*

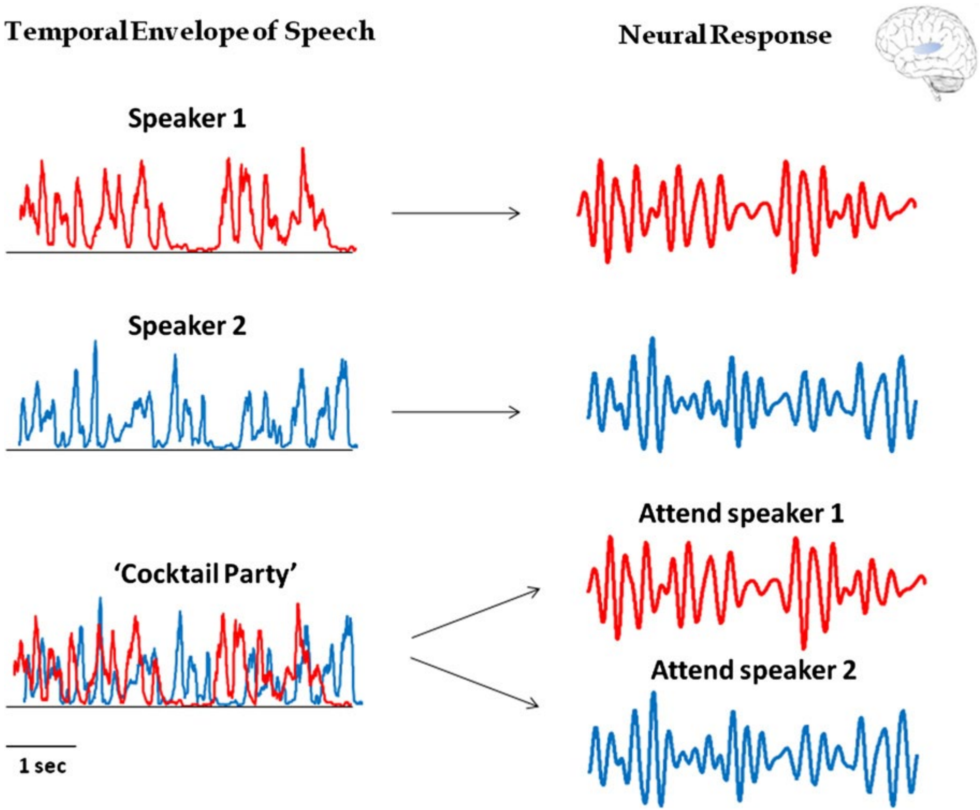
Interim summary

- Speech is highly variable: between and also within individual talkers
    - For instance, speech can be produced at a vast range of speech rates
- Still, listeners manage to understand fast and slow speech
- Neural oscillations contribute to speech comprehension by...
- ... entraining to the rhythm of speech,
- ... imposing an appropriate temporal sampling regime,
- ... influencing word recognition by over-/undersampling of segmental durations
- ... providing a neural mechanism for rate normalization



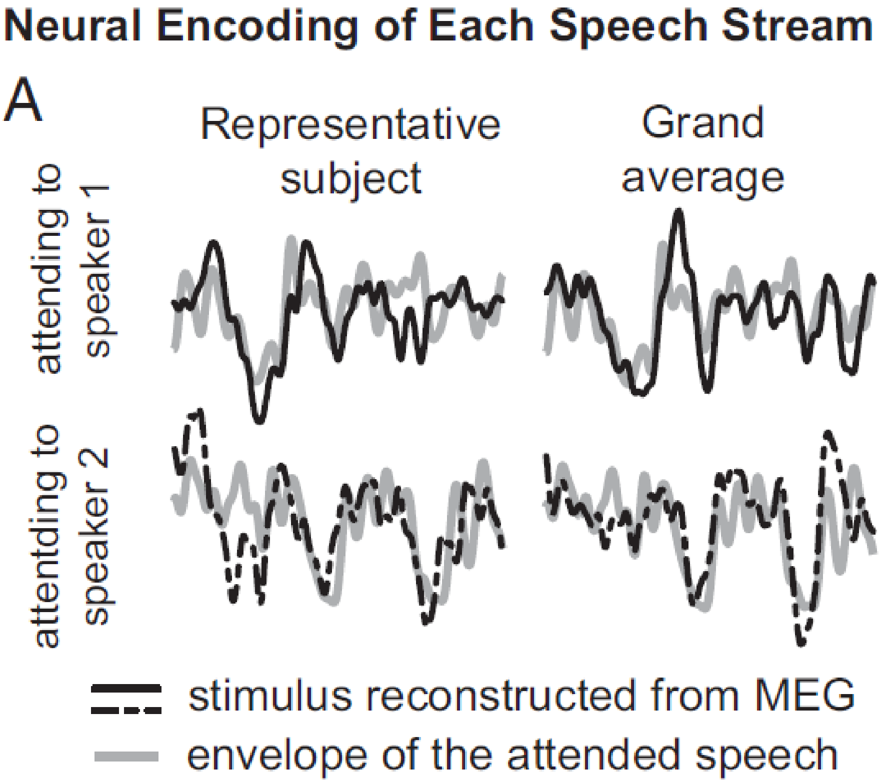
‘Cocktail party’ listening

Neural ‘speech tracking’ in noise



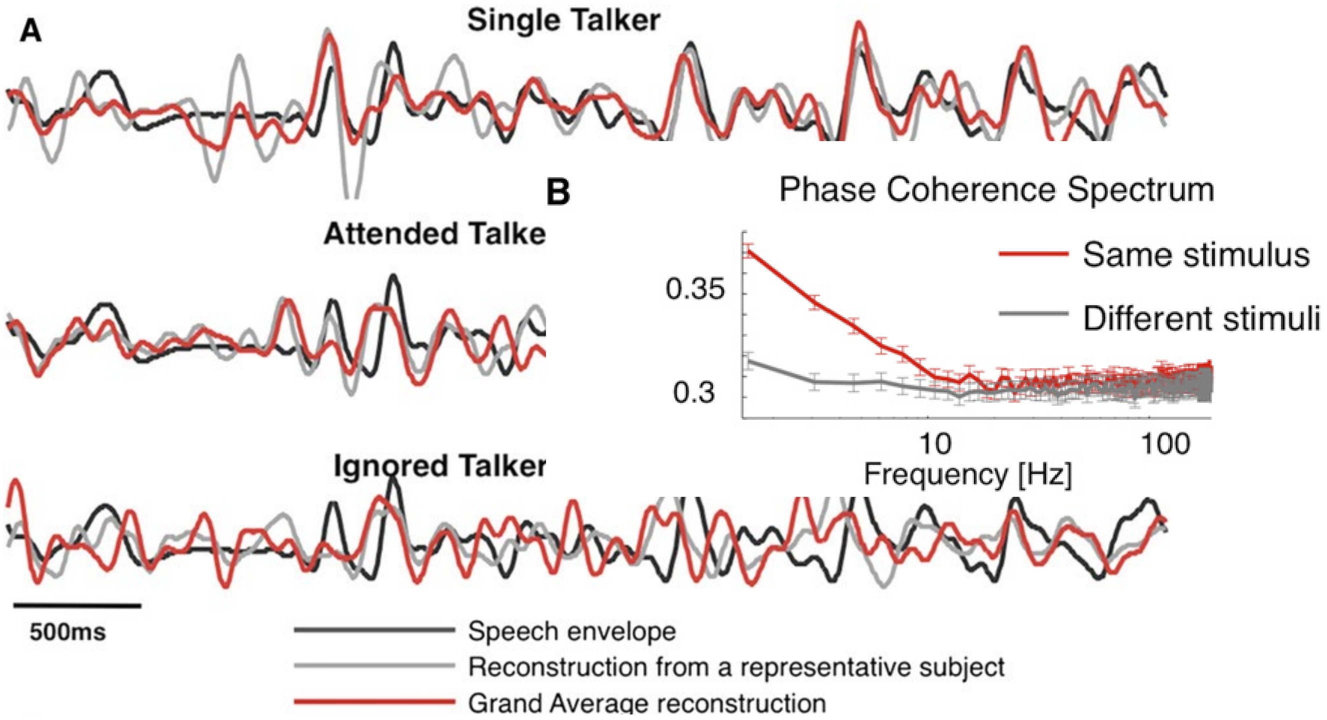
Zion Golumbic, Poeppel, & Schroeder, 2012

Neural ‘speech tracking’ in noise



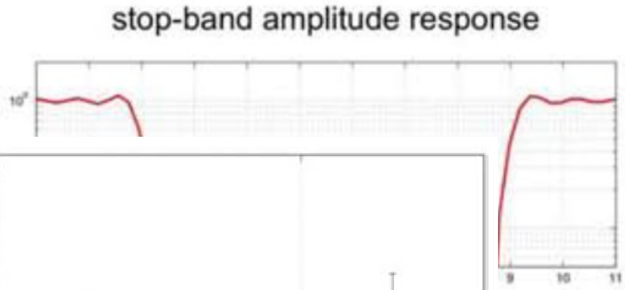
Ding & Simon, 2012

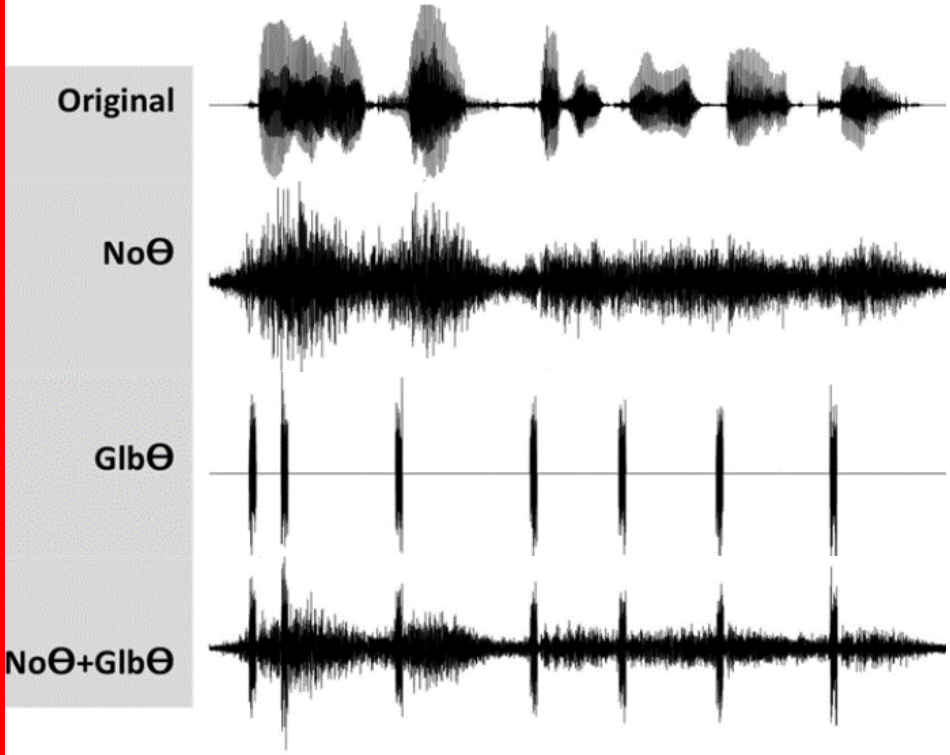
Neural ‘speech tracking’ in noise

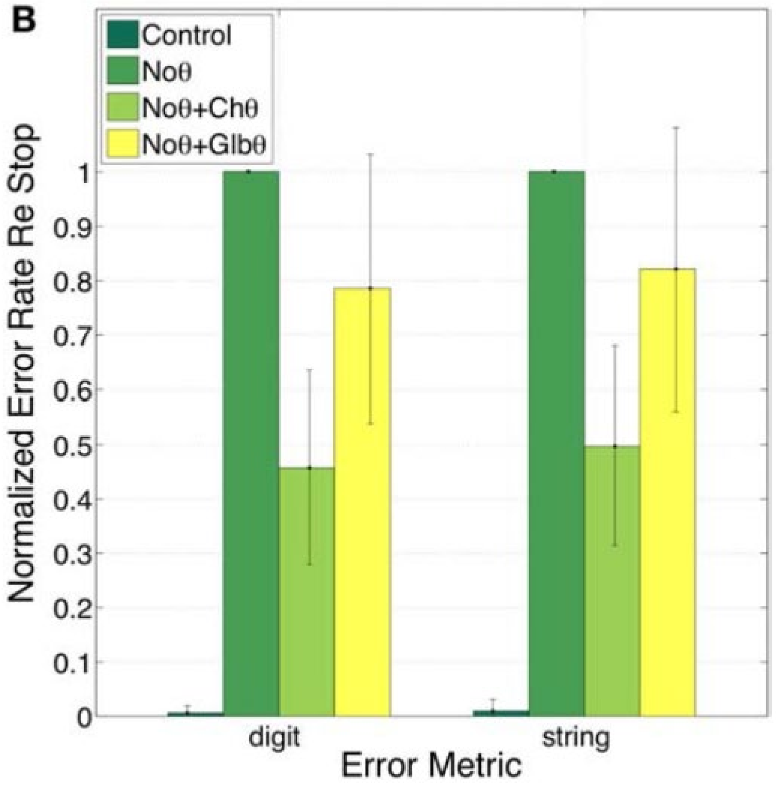


Zion Golumbic et al., 2013

Amplitude modulations are the key







Ghitza, 2012; Doelling et al., 2014

Speech production in noise

- If the neural tracking of the rhythm of the attended speech stream is essential for speech-in-noise comprehension...
- ... then does the talker adjust his/her voice to facilitate this neural tracking, for instance by producing more ‘rhythmic’ speech?

Speech production in noise

Is this prosody?

- 'Lombard spee
  - the speech we produce when we're talking in noise
  - vs. plain speech (speech-in-quiet)

  - increased intensity
  - slower speech rate
  - raised *f*0
  - flatter spectral tilt
  - ...

Cooke et al., 2014

Speech production in noise

- Lombard speech is more intelligible in noise than plain speech, even after matching overall intensity.
- However, exactly what makes Lombard speech more intelligible in noise (i.e., which acoustic adjustments) is unclear.
  - Slowing down plain speech doesn't boost intelligibility
  - Increasing the *f*0 of plain speech doesn't help neither
- Acoustic investigation of amplitude fluctuations ('rhythmicity') in Lombard speech.

Lu & Cooke, 2009; Cooke et al., 2014, *JASA*

Speech production in noise

Table 1. Characteristics of the four speech corpora [M = male; F = female; BMN = 9-talker babble-modulated ICRA noise from Dreschler *et al.* (2001); SSN = speech-shaped noise; SMN = speech-modulated noise].

	Talkers	Sentences	Noise	Source
Corpus 1	$N = 1$ (M)	"Normal"; $N = 25$	BMN; intense	Mayo *et al.* (2012)
Corpus 2	$N = 8$ (4 M/4 F)	Matrix; $N = 400$	SSN; 96 dB (L)	Lu and Cooke (2008)
Corpus 3 (Hurricane)	$N = 1$ (M)	"Normal"; $N = 720$	SMN; 84 dB (A)	Cooke *et al.* (2013)
Corpus 4 (MRT)	$N = 1$ (M)	Frame; $N = 300$	SMN; 84 dB (A)	Collected by Valentino-Botinhao (2013)

Bosker & Cooke, 2018; 2020

Speech production in noise

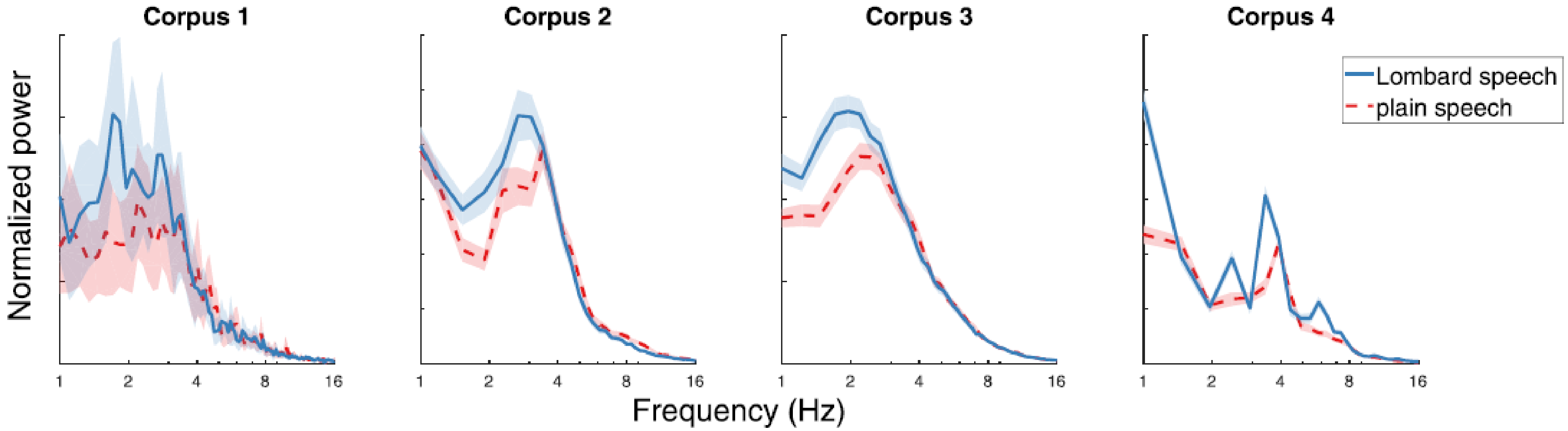
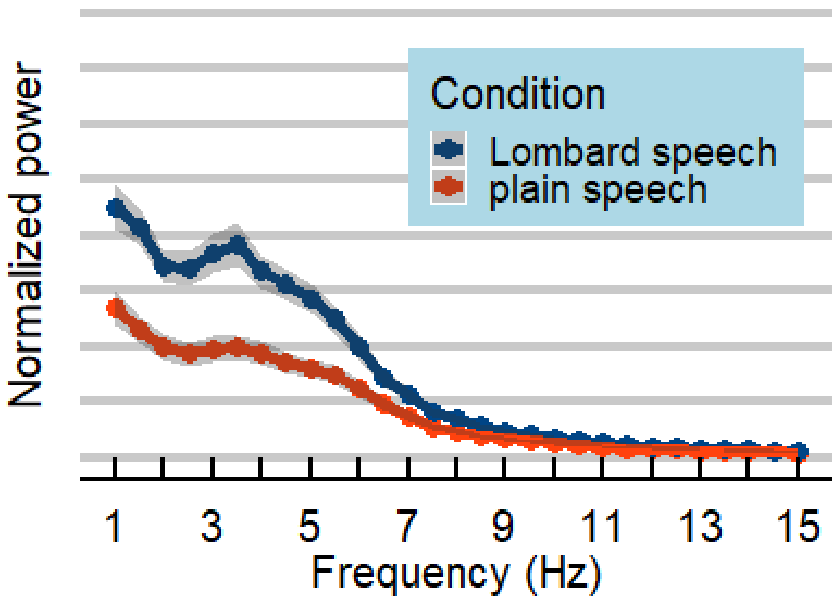


Fig. 1. (Color online) Average modulation spectra, calculated from the broadband analysis (250–4000 Hz), of Lombard speech (solid line) and matching plain speech (dashed line), for each corpus. Shaded areas indicate 95% CIs.

Bosker & Cooke, 2018; 2020

Speech production in noise



... replicates in Dutch (NiCLS corpus)

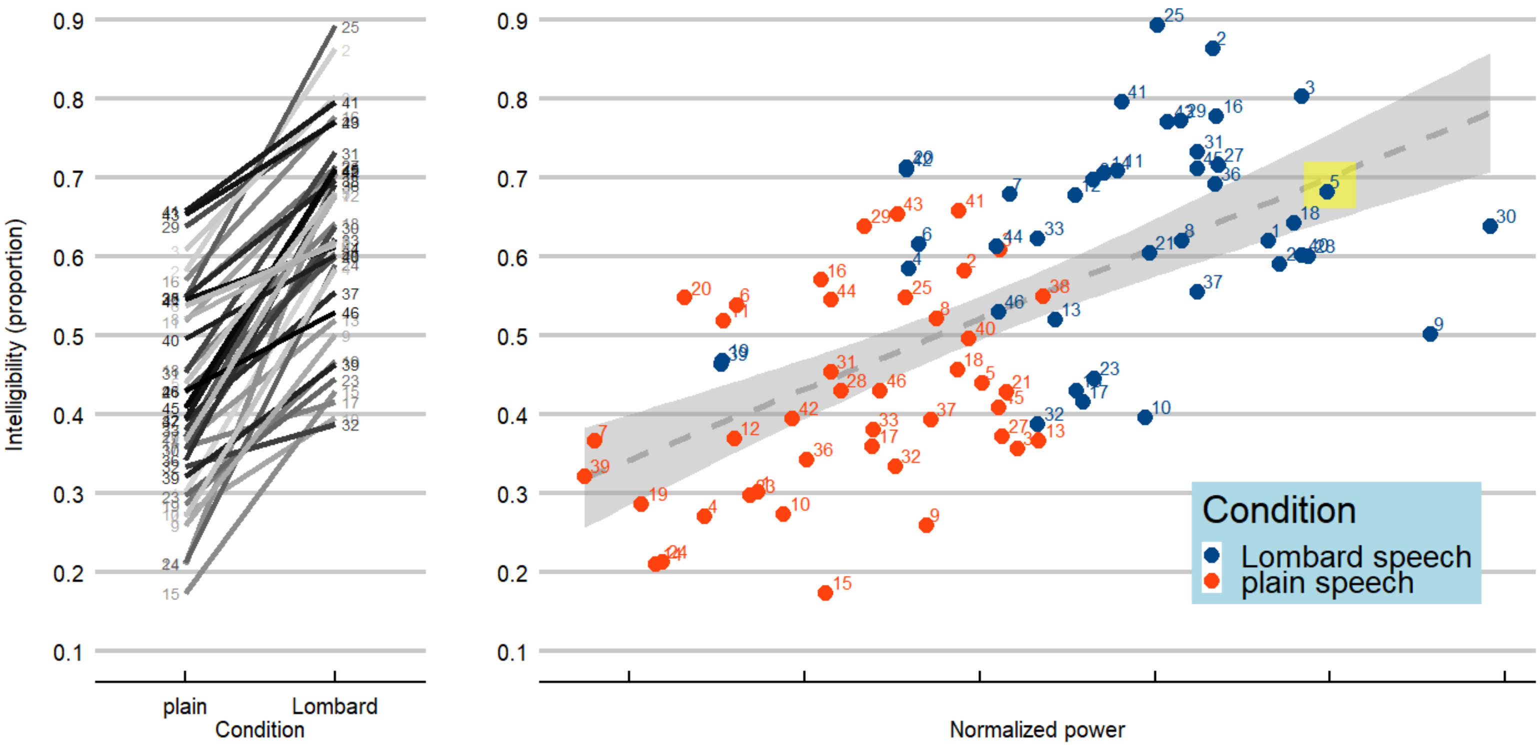
Bosker & Cooke, 2018; 2020

Speech production in noise

- More pronounced amplitude modulations in Lombard speech compared to plain speech, potentially facilitating neural 'speech tracking'.
- Do these more pronounced temporal modulations help perception?

Bosker & Cooke, 2018; 2020

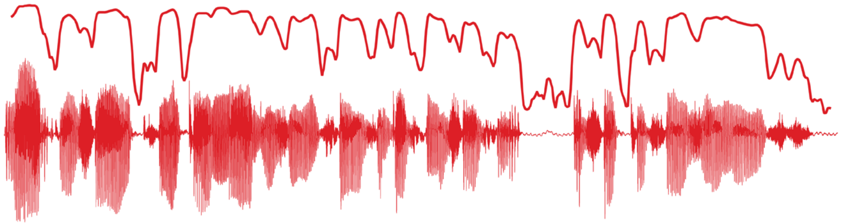
Speech perception in noise



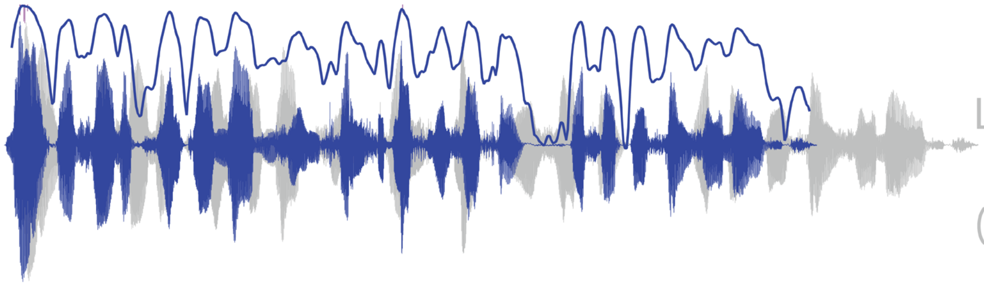
Bosker & Cooke, 2018; 2020

Speech perception in noise

Plain speech

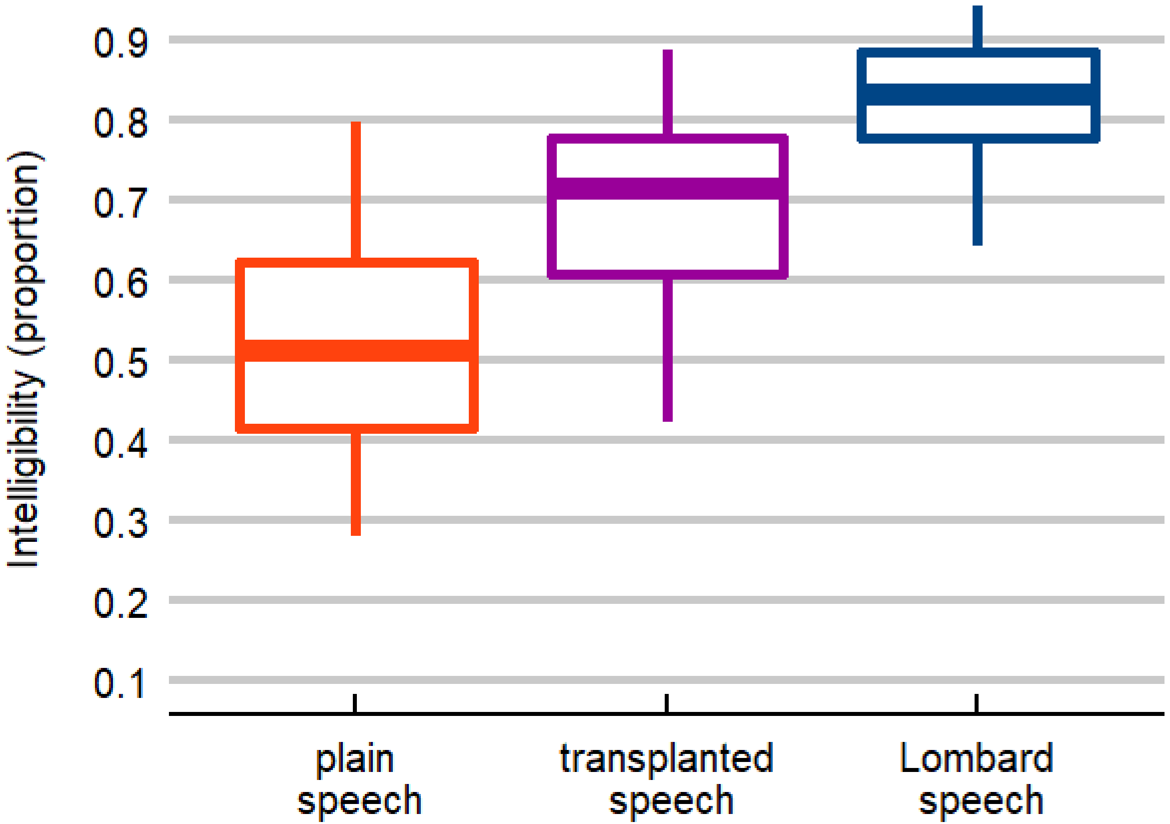


Lombard speech (DTW)



Lombard speech (original)

Speech perception in noise



Bosker & Cooke, 2018; 2020

Speech comprehension is hard...

- Neural 'speech tracking' provides a neural mechanism that may explain how listeners manage to understand 'noisy' speech.
- ***Signal-intrinsic*** noise: variation in speech rate
  - neural oscillations impose appropriate 'sampling frequency' on signal
  - ... normalize the spoken input for the rate at which it is produced
- ***Signal-extrinsic*** noise: speech-in-noise
  - talkers produce more pronounced temporal modulations in noise
  - facilitates comprehension by allowing more opportunity for neural oscillations to 'latch onto' the attended speech stream

Next up:

- Lecture 3: *Prosody-guided prediction*

Hans Rutger Bosker

Speech Perception in Audiovisual Communication [SPEAC] lab

Donders Institute, Radboud University, Nijmegen, The Netherlands

https://hrbosker.github.io

hansrutger.bosker@donders.ru.nl