The human auditory cortex tracks concomitantly the syllabic and phonemic time scales via the acoustic spectral flux

Jérémy Giroud[1], Agnès Trébuchon[2,3], Manuel Mercier[2], Matthew H. Davis[1], Benjamin Morillon[2].

@JeremyGiroud

[1]MRC Cognition and Brain Sciences Unit, University of Cambridge, UK, [2]Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France, [3]APHM, Clinical Neurophysiology, Timone Hospital, Marseille, France

Introduction

How the brain parses linguistic information from a continuous speech stream is a major question in cognitive neuroscience.

Speech is thought to be sampled in parallel at both the syllabic and phonemic timescales reflected by *theta* and *gamma* neural oscillations [1,2].

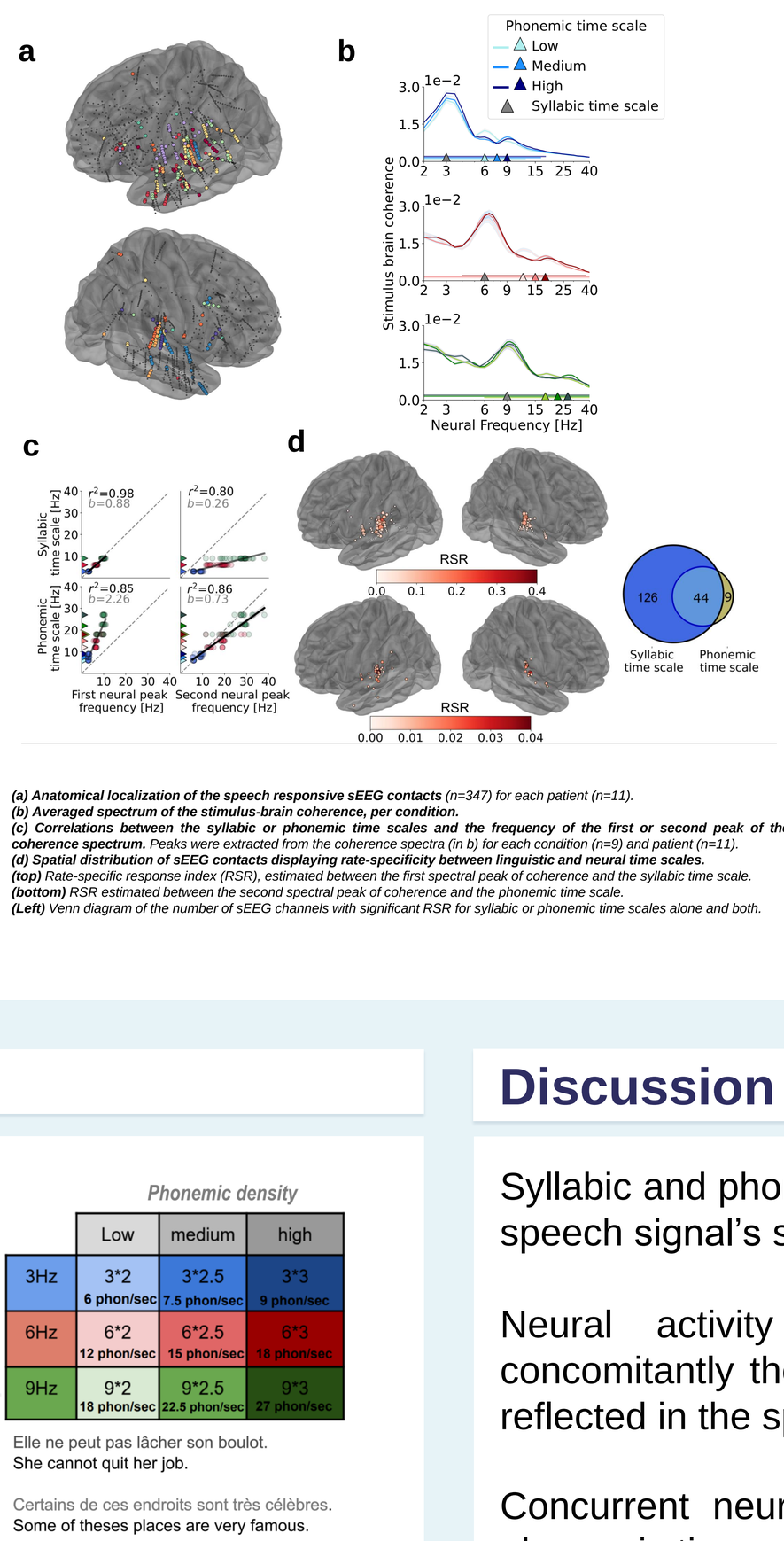
During speech perception *theta* oscillations in auditory regions track syllabic rhythm reflected in the speech envelope [3,4].

However, the phonemic time scale is not well represented in the speech envelope and in natural speech, phonemic and syllabic time scales are highly correlated, making it challenging to study.

To overcome these challenges, we developed speech materials with partially decorrelated syllabic and phonemic time scales and focused on the acoustic spectral flux to reveal neural processing at both the syllabic and phonemic time scales.

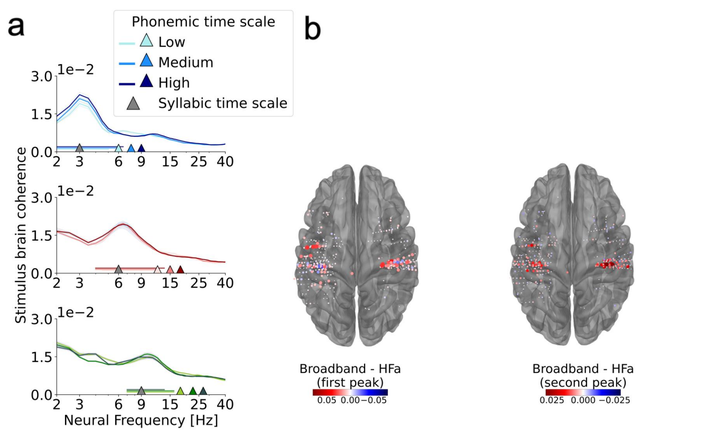
Results

Neural correlates of the acoustic spectral flux



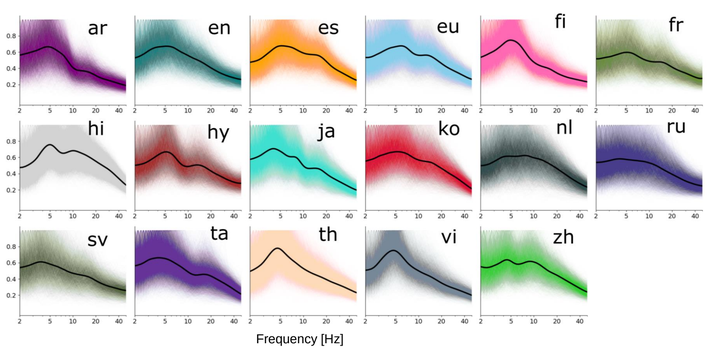
***(a) Anatomical localization of the speech responsive sEEG contacts** (n=347) for each patient (n=11).*
(b) Averaged spectrum of the stimulus-brain coherence, per condition.
***(c) Correlations between the syllabic or phonemic time scales and the frequency of the first or second peak of the coherence spectrum.** Peaks were extracted from the coherence spectra (in b) for each condition (n=9) and patient (n=11).*
(d) Spatial distribution of sEEG contacts displaying rate-specificity between linguistic and neural time scales.
***(top)** Rate-specific response index (RSR), estimated between the first spectral peak of coherence and the syllabic time scale.*
***(bottom)** RSR estimated between the second spectral peak of coherence and the phonemic time scale.*
***(Left)** Venn diagram of the number of sEEG channels with significant RSR for syllabic or phonemic time scales alone and both.*

Broadband vs High frequency activity



(a) Averaged spectrum of the stimulus-brain HFa coherence, per condition.
(b) Spatial distribution of the difference in coherence strength estimated with broadband and HFa

Spectral flux across languages



***Modulation spectrum of the spectral flux in 17 languages**.*
***ar**: Arabic, **en**: English, **es**: Spanish, **eu**: Basque, **fi**: Finnish, **fr**: French, **hi**: Hindi, **hy**: Armenian, **ja**: Japanese, **ko**: Korean, **nl**: Dutch, **ru**: Russian, **n**: Swedish, **ta**: Tamil, **th**: Thai, **vi**: Vietnamese, **zh**: Chinese Mandarin.*

Methods

Stimuli

105 French sentences extracted from *TedTalks*.

Different ***phonemic densities*** (low, medium, high) corresponding to 2, 2.5 or 3 phonemes per syllable.

Different ***syllabic timescales*** (temporally altered) corresponding to 3, 6 or 9 syllables per second.

Syllabic time scale \ Phonemic density	Low	medium	high
3Hz	3*2 **6 phon/sec**	3*2.5 **7.5 phon/sec**	3*3 **9 phon/sec**
6Hz	6*2 **12 phon/sec**	6*2.5 **15 phon/sec**	6*3 **18 phon/sec**
9Hz	9*2 **18 phon/sec**	9*2.5 **22.5 phon/sec**	9*3 **27 phon/sec**

Elle ne peut pas lâcher son boulot.
She cannot quit her job.

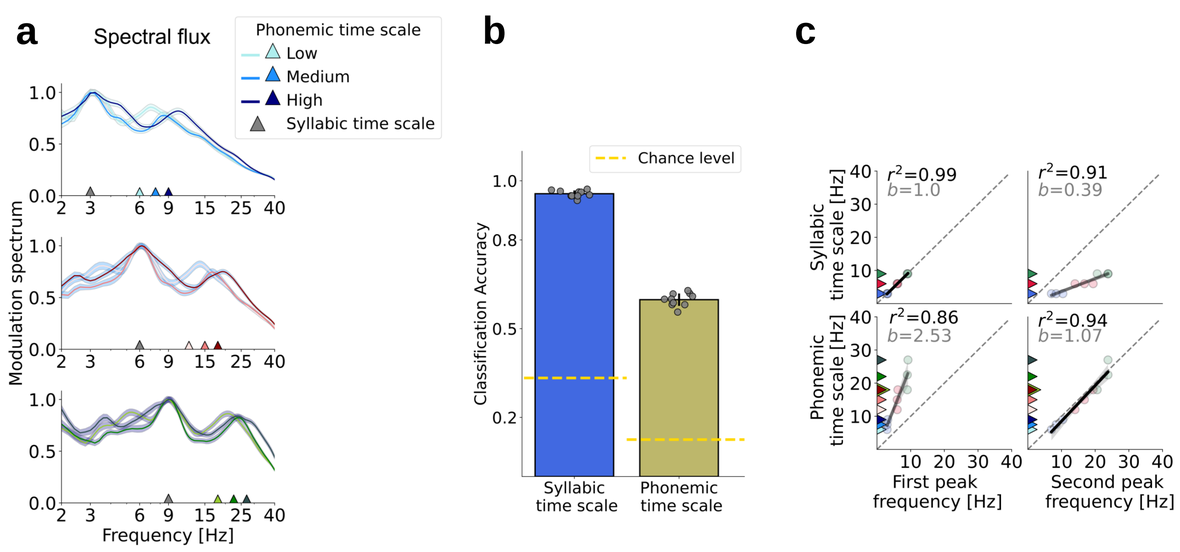
Certains de ces endroits sont très célèbres.
Some of theses places are very famous.

Cette première traduction semble être vraiment parfaite.
This first translation seems really perfect.

Participants, neural recordings & paradigm

11 French epileptic patients were recorded with iEEG while they performed a speech comprehension task with repetitions.

Acoustic representation of the stimuli



(a) Averaged power spectrum of the spectral flux, per condition.
(b) Classification accuracy of the linguistic time scales, based on the power spectrum of the spectral flux.
(c) Correlations between the syllabic or phonemic time scales and the frequency of the first or second peak of the power spectra.

Discussion

Syllabic and phonemic time scales are reflected in the speech signal's spectral flux.

Neural activity in the auditory cortex tracks concomitantly the syllabic and phonemic time scales reflected in the spectral flux.

Concurrent neural dynamics reflecting syllabic and phonemic timescales are co-localized.

Speech syllabic and phonemic time scales are more prominently reflected in the raw signal (broadband activity) as compared to the high frequency activity (HFa).

The speech spectral flux modulation spectrum is quite consistent across languages.

References

[1] Giraud & Poeppel (2012). *Nature Neuroscience*.
[2] Giroud et al. (2020). *PLoS biology*.
[3] Schmidt et al. (2021). *Psychophysiology.*
[4] Luo & Poeppel (2007). *Neuron*.