Modelling Demonstrates Phase-Reset is an Important Feature of Neural Speech Processing

Andrew Shannon (andrew.shannon@bristol.ac.uk)
School of Computer Science, University of Bristol, MVB, 75 Woodland Rd, Bristol, BS8 1UB

Martin Homer (Martin.Homer@bristol.ac.uk)
School of Engineering Mathematics and Technology, University of Bristol, Ada Lovelace Building, Tankard's Cl, Bristol, BS8 1TW

David A. W. Barton (David.Barton@bristol.ac.uk)
School of Engineering Mathematics and Technology, University of Bristol, Ada Lovelace Building, Tankard's Cl, Bristol, BS8 1TW

Conor Houghton (Conor.Houghton@bristol.ac.uk)
School of Engineering Mathematics and Technology, University of Bristol, MVB, 75 Woodland Rd, Bristol, BS8 1UB

Abstract

Syllable segregation and source separation are foundational components of neural speech processing, yet consensus on their underlying mechanism remains elusive. Several hypotheses have been proposed, suggesting that the brain may align its activity to incoming linguistic stimuli via evoked responses, entrainment of endogenous oscillations, or some combination of the two. We investigate the origin of oscillatory behaviour in syllable segregation by modelling the dynamical response to periodic linguistic stimuli. We compare a biophysically accurate neural mass model and a phase-resetting oscillator with prior experimental EEG data. We find that a correlation between neural activity entrainment strength and the sharpness of incoming phonemes, identified in the EEG experiment, is readily reproduced by both the neural mass model and the oscillator. However, when the phase-resetting dynamics are removed, the oscillator fails to reproduce the correlation. This demonstrates that phase-resetting is required for sharpness specific tuning of neural entrainment to speech. Identifying the neural correlates of this phenomenon may be possible through interrogation of the biophysical features of the neural mass model.

Keywords: Neural Entrainment; Phase-Resetting; Speech Processing; Syllable Segregation; Neural Mass Modelling.

Introduction

Two key components of speech processing are source separation and syllable segregation. These allow us first to attend to a particular voice amongst background noise, including potentially many other voices, and subsequently identify key sounds comprising words. Neural activity thus aligns with the syllabic rhythm of speech, and in particular, to the rhythm of the voice being attended to. Multiple underlying mechanisms have been proposed for this apparent entrainment of neural activity to speech, with debate around whether it results from activity evoked by acoustic edge features or through phase-resetting of endogenous oscillations (Schroeder & Lakatos, 2009; Oganian et al., 2023; Obleser & Kayser, 2019). Identifying the signatures of each and probing the neural response to speech through careful experiment is key to settling this debate.

In this direction, (Cucu, Kazanina, & Houghton, 2022) investigated the mechanisms behind the neural response to low-frequency speech components by analysing the entrainment of electroencephalography (EEG) activity to periodic sequences of consonant-vowel (CV) phonemes. They identified a robust correlation $(-0.91,\ p<0.001)$ between entrainment strength and phoneme envelope sharpness, quantified using a range of acoustic features. The dynamics that produces such a correlation remains unknown: is it the result of particular biological features of neural populations, or a fundamental feature of driven oscillator systems?

A promising mathematical model of macroscopic neural behaviour that has the potential to address this question is the next generation neural mass model (NGNMM) (A. Byrne, Brookes, & Coombes, 2017). This is an exact mean field reduction of microscopic spiking network dynamics that incorporates biophysical features to reproduce complex oscillatory neural dynamics (A. Byrne et al., 2017; Pietras, Devalle, Roxin, Daffertshofer, & Montbrió, 2019; Á. Byrne, O'Dea, Forrester, Ross, & Coombes, 2020; Á. Byrne, Ross, Nicks, & Coombes, 2022). As such, it may also capture the oscillatory dynamics of the speech processing phenomenon we are considering here.

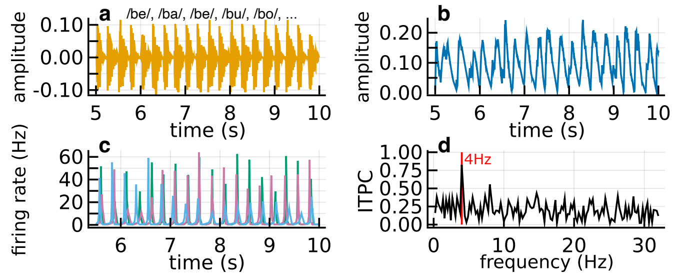

Figure 1: Experimental audio stimuli (**a**) were processed into envelopes (**b**), with added noise, and used to drive the models. The mean ITPC (**d**) was calculated over the activity (**c**) of 60 trials for three sequences of each CV phoneme type.

To investigate the mechanism underlying the correlation, we use the experimental auditory stimuli to drive both the NGNMM and a phase-resetting oscillator model. We compute the entrainment of the model activity to the periodic stimuli and identify any correlation with the sharpness of the phonemes. Both the NGNMM and the oscillator model reproduce the experimentally observed correlation. When the phase modulation that creates the phase-resetting behaviour in the oscillator model is removed, the correlation disappears. This suggests that a phase-resetting mechanism may underlie the correlation, and may therefore be facilitated by the dynamics of the NGNMM. Understanding which features of the NGNMM allow this may enable the identification of biological correlates of this phenomenon.

Methods

The experimental audio stimuli, near-isochronous $\sim$4Hz sequences of 20 CV phonemes, were transformed to form sound envelopes for each sequence. Each sequence comprised a single consonant (1 of 15) paired with 20 randomly selected vowels, e.g. *"be, ba, bo, bo, bu,..."*. Additive noise, constructed out of a random selection of distorted phoneme envelopes, was added to the stimuli at particular signal-to-noise ratios. This resulted in 60 trials for each stimulus, of which there were three for each of the 15 CV phonemes.

We used the envelopes to drive three different models (Figure 1). First, the NGNMM, with the drive applied in the same way as in (A. Byrne et al., 2017) and parameters set such that the NGNMM was quiescent under no stimulation, but oscil-

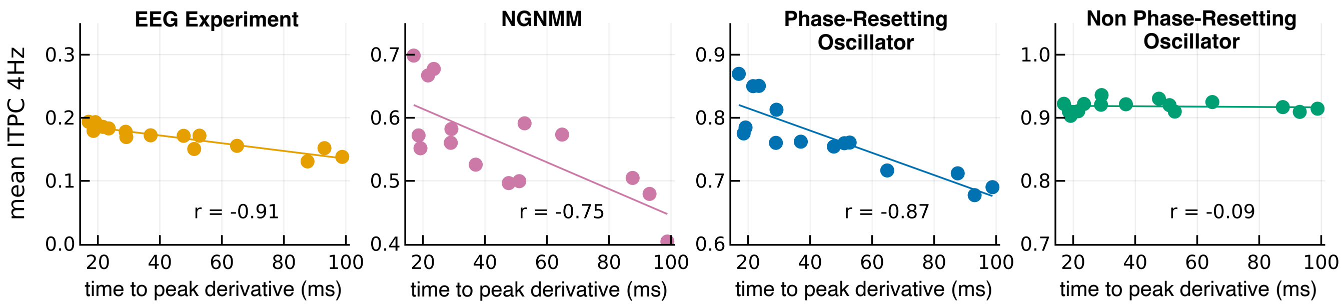

Figure 2: Mean 4Hz ITPC vs envelope sharpness. Each dot represents a different consonant-vowel condition. The NGNMM and phase-resetting oscillator reproduce the correlation seen in the EEG experiment. The non-phase-resetting oscillator does not.

lated at approximately 4Hz under small stimulus. Second, we used the oscillator model from (Oganian et al., 2023):

$$\begin{aligned}\dot{\theta} &= 2\pi F - \frac{cs(t)}{r}\sin(\theta)\,,\\ \dot{r} &= r(1-r^2) + cs(t)\cos(\theta)\,,\end{aligned} \quad (1)$$

with a stimulus s(t) that forces a reset towards $\theta = 0$. The stimuli were normalised to have a consistent integral over the samples. We added 1/f noise to the activity with a signal-to-noise ratio of 0.1. Finally, we created an oscillator model without phase-resetting by removing the sine and cosine terms.

The Inter Trial Phase Coherence (ITPC) of the model activity across the 60 trials of a given stimulus was used to evaluate the strength of the entrainment to the periodic stimuli, as in (Cucu et al., 2022). For each CV phoneme, three sets of 60 trials were conducted, and the mean ITPC calculated. We used the measure of phoneme sharpness specified in (Cucu et al., 2022). We tested a range of signal-to-noise ratios and drive strengths, and present a typical example in Figure 2.

Results

The NGNMM demonstrates robust entrainment to the 4Hz stimuli, and further reproduces the negative correlation observed in experiment (EEG: $r = -0.91$, NGNMM: $r = -0.75$, Figure 2 **a**, **b**).The phase-resetting oscillator model also produces a strong correlation between the entrainment strength and phoneme sharpness ($r = -0.9$, Figure 2 **c**), but when the phase-resetting dynamics are removed, the correlation is no longer present ($r = -0.094$, Figure 2 **d**).

Discussion

The correlation is absent when there is no phase-resetting, showing that phase-resetting is required. The NGNMM also produces the correlation, indicating it also resets phase. We are interested in the neurological origin of this experimentally observed speech processing phenomenon. By taking advantage of the biophysical basis of the NGNMM's construction, and through comparison to the phase-resetting model as a minimal working example, it may be possible to identify the key neuro-dynamical components that enable phoneme-sharpness specific tuning of neural entrainment to speech.

Acknowledgments

AS is funded by a UK Research and Innovation grant (EP/S022937/1). This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/.

References

Byrne, A., Brookes, M. J., & Coombes, S. (2017). A mean field model for movement induced changes in the beta rhythm. *Journal of Computational Neuroscience*, *43*, 143–158.

Byrne, Á., O'Dea, R. D., Forrester, M., Ross, J., & Coombes, S. (2020). Next-generation neural mass and field modeling. *Journal of Neurophysiology*, *123*(2), 726–742.

Byrne, Á., Ross, J., Nicks, R., & Coombes, S. (2022). Mean-field models for eeg/meg: from oscillations to waves. *Brain Topography*, *35*(1), 36–53.

Cucu, M. O., Kazanina, N., & Houghton, C. (2022). Syllable-initial phonemes affect neural entrainment to consonant-vowel syllables. *Frontiers in Neuroscience*, *16*, 826105.

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, *23*(11), 913–926.

Oganian, Y., Kojima, K., Breska, A., Cai, C., Findlay, A., Chang, E. F., & Nagarajan, S. S. (2023). Phase alignment of low-frequency neural activity to the amplitude envelope of speech reflects evoked responses to acoustic edges, not oscillatory entrainment. *Journal of Neuroscience*, *43*(21), 3909–3921.

Pietras, B., Devalle, F., Roxin, A., Daffertshofer, A., & Montbrió, E. (2019). Exact firing rate model reveals the differential effects of chemical versus electrical synapses in spiking networks. *Physical Review E*, *100*(4), 042412.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.