Induced alpha and beta electroencephalographic rhythms covary with single-trial speech intelligibility in competition

Vibha Viswanathan[1]✉, Hari M. Bharadwaj[2], Michael G. Heinz[3,4] & Barbara G. Shinn-Cunningham[1]

Neurophysiological studies suggest that intrinsic brain oscillations influence sensory processing, especially of rhythmic stimuli like speech. Prior work suggests that brain rhythms may mediate perceptual grouping and selective attention to speech amid competing sound, as well as more linguistic aspects of speech processing like predictive coding. However, we know of no prior studies that have directly tested, at the single-trial level, whether brain oscillations relate to speech-in-noise outcomes. Here, we combined electroencephalography while simultaneously measuring intelligibility of spoken sentences amid two different interfering sounds: multi-talker babble or speech-shaped noise. We find that induced parieto-occipital alpha (7–15 Hz; thought to modulate attentional focus) and frontal beta (13–30 Hz; associated with maintenance of the current sensorimotor state and predictive coding) oscillations covary with trial-wise percent-correct scores; importantly, alpha and beta power provide significant independent contributions to predicting single-trial behavioral outcomes. These results can inform models of speech processing and guide noninvasive measures to index different neural processes that together support complex listening. These results can inform models of speech processing and guide noninvasive measu

In contrast to induced alpha, which has been implicated in auditory attention, prior work suggests that the beta (13–30 Hz) rhythm may relate to maintenance of the current sensorimotor state[22] and sensorimotor predictive coding[23,24]. More generally, motor-theory accounts of speech recognition posit that sensorimotor integration between fronto-motor areas controlling articulation (e.g., inferior frontal gyrus and premotor cortex)

of each epoch was 5.25 s. This epoch period included all of the five keywords on which participants were scored for every sentence.

For each subject and experimental condition, the EEG response spectrogram in each epoch was calculated

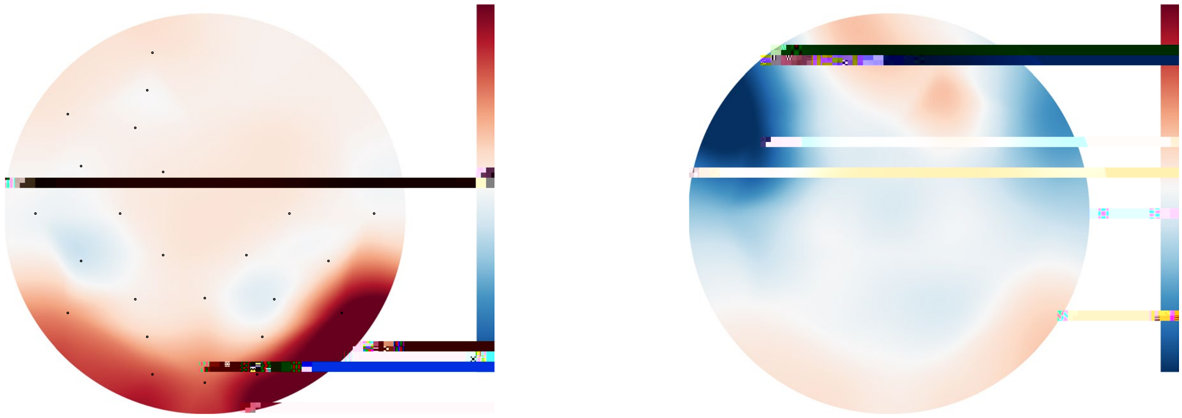
for this effect separately for each experimental condition by comparing the histogram of number of keywords correct per sentence against the distribution under the null hypothesis of independent outcomes across different keywords. Under the null hypothesis, the performance on any particular keyword in a sentence (i.e., whether or not the word was reported correctly) has a Bernouilli distribution with parameter p = average proportion correct score for the particular condition. Moreover, the probability of reporting correctly x keywords out of a total of 5 keywords per sentence is binomial with parameters n = 5, and p = average proportion correct score for the condition. Assuming independent outcomes across different sentences, the probability that M sentences out of a total of 1050 sentences per condition (pooled over all six subjects; each subject performed 175 sentences per condition) had x keywords correct is also binomial, with parameters n = 1050 and p = probability that x keywords per sentence are correct. We compared this final probability distribution (which models the null distribution of independent keyword outcomes in each condition) with the histogram of number of keywords correct per sentence. Specifically, we generated p-values describing, for each experimental condition, the likelihood of observing the actual correlation across keywords within each trial, assuming that performance on the words was truly independent. Note that our Bernouilli-trial null model is *not* modeling whether each word reported by the participant is in the keyword list; instead, it is modeling whether each keyword is present in the list of words reported by the participant.

We used a multinomial linear regression model to test whether beta power contributes additionally to predicting percent-correct score over the contribution of alpha power alone, and vice-versa (i.e., whether alpha power contributes additional predictive power over that contributed by beta power alone). The percent-correct score in different trials across subjects was the response (treated as an ordered factor variable with six levels: 0, 20, 40, 60, 80, and 100); the predictors were log alpha power, log beta power, and condition (factor variable with two levels). Likelihood-ratio Type II tests were used for statistical testing by calculating the deviance (i.e., −2 times log likelihood-ratio) and comparing it to a chi-squared distribution[51].

Software accessibility. Stimulus presentation was controlled using custom MATLAB (The MathWorks, significantly with percent-correct score within condition [$F(5,2093) = 6.4915$, $p = 5.346\mathrm{e}{-06}$]. Posthoc analyses revealed that the largest contribution to the main effect of score on alpha and beta power came from the linear term, indicating that alpha and beta power increased with score ($T = 4.216$, $p = 2.59\mathrm{e}{-05}$ for alpha; $T = 4.173$, $p = 3.13\mathrm{e}{-05}$ for beta). The quadratic term was also significant in predicting alpha power, but carried a negative coefficient ($T = -2.521$, $p = 0.0118$) in line with the plateauing of alpha power with increasing score seen in Fig. 5A. In the model for beta power, the cubic term was also significant ($T = 3.003$, $p = 0.00271$), in line with the U-shaped trend seen over the 20–100% range of scores in Fig. 5B. We conducted pairwise t-tests to compare the changes in alpha and beta power for each step increase in percent-correct score. These sequential-difference-contrast analyses showed that alpha and beta power increased when percent-correct score increased from 0 to 20 ($T = 1.994$, $p = 0.0463$ for alpha; $T = 4.249$, $p = 2.24\mathrm{e}{-05}$ for beta); however, for the successive steps (20 to 40, 40 to 60, etc.), the increase in power was not significant for either alpha or beta. Figure 5A also suggests thee

and beta power each make significant independent contributions to predicting trial outcome (i.e., trial-wise speech intelligibility).

Because of our observation that pre- and during-stimulus power were correlated for both alpha and beta, and that the average power across both periods covaries with single-trial behavioral outcomes, we wished to further understand the temporal evolution of the two rhythms over the time course of the trial. While it is well-established that alpha enhancement begins in the preparatory period (for example, before stimulus onset but after cueing subjects to "stay still listen now" as in our study), the temporal dynamics of the beta rhythm during speech perception in noise is not as well studied. Thus, we contrasted the scalp topographic maps between the during- and pre-stimulus periods for both beta and alpha (Fig. 9) to obtain further insight. Figure 9 shows that parieto-occipital alpha power is stronger in the during-stimulus period (Panel A), consistent with maintaining an increasing attentional focus on the target speech. However, the scalp topomap difference between during- and pre-stimulus periods in the beta band (Panel B) shows regions of reduced power fronto-centro-laterally in both hemispheres, and regions of increased power fronto-medially. This suggests that the beta power observed in the present study consists of two functionally distinct components. The fronto-centro-lateral component of beta, which is stronger in the pre-stimulus period (i.e., before stimulus onset but after the "stay still and listen



power in parieto-occipital EEG channels and beta power in frontal channels significantly covary with, and importantly independently contribute to, single-trial speech intelligibility in our speech-in-noise tasks (Figs. 5, 8). These results are consistent with the posited role of the parieto-occipital alpha rhythm in auditory selective attention[10–14,16–18] and the frontal beta rhythm in maintenance of the current sensorimotor state[22] and sensorimotor predictive coding[23,24] that is thought to stabilize speech representation in adverse listening conditions[25–29]. The interpretation that some combination of these top-down effects influences single-trial behavioral outcome is also supported by the observed correlation in performance across words within a trial (Fig. 7).

Our results are in line with prior reports of a positive correlation between alpha power in parietal EEG channels and speech intelligibility in noise (e.g., across SNRs as quantified in the during-stimulus period by Hall et al.[20], and across individuals as quantified in the pre-stimulus period by Alhanbali et al.[21]). However, at least at first glance, our results appear to be at odds with other reports (e.g., by Obleser and Weisz[63] and Becker et al.[64], who used noise-vocoded speech in quiet, and Dimitrijevic et al.[65], who used digits in noise) that better comprehension is associated with alpha suppression (rather than a power increase) in the late during-stimulus period in temporal brain regions and central EEG channels. This discrepancy may be explained in part by the existence of multiple neural generators of task-related alpha (i.e., alpha power in the parieto-occipital and central EEG channels may reflect two different mechanisms of alpha[18]). Moreover, some of these studies presented speech in quiet rather than with simultaneous competing sounds, which could evoke different mechanisms[63,64].

Foxe and Snyder[10] distinguish between parieto-occipital alpha seen in an unaroused state (e.g., when visual stimuli are ignored[15]) and that seen in selective attention across different stimuli (especially spatial selective attention, where alpha power is lateralized according to the hemifield of focus;[16–18]). In the present study, the target speech and masker sources were both presented diotically rather than spatially separated; thus, even though it required selective attention, our task did not involve any spatial focus of attention. It may be that the alpha in the current study, which covaries with trial-wise speech intelligibility, reflects an overall suppression of the visual scene and focus of auditory attention, rather than a mechanism specific to stimulus selection. Another possibility is that there may be a common mechanism in play across the parieto-occipital alpha seen in the two cases. Indeed, the frontoparietal attention network becomes active during spatial attention and working memory for auditory stimuli as well as for visual inputs, even though many earlier studies assume it is strictly a visuospatial processing network[66–70]. Thus, future studies should disambiguate between the different mechanisms by which the alpha rhythm may mediate suppression of sensory distractors[10–14], especially for co-localized sources like those used in the current study.

Unlike parieto-occipital alpha, the functional role of frontal beta in auditory perception is less understood. That parieto-occipital alpha is associated with attentional focus and is present even before stimulus onset (Figs. 2, 3A) is well documented[10–19]. However, in the current study we find that frontal beta power in both the pre-stimulus and during-stimulus periods covaries with single-trial speech-in-noise outcomes (Supplementary Figs. S1 and S2; statistics given in Results). Although our results about during-stimulus beta may potentially be explained by invoking the predictive coding theory[22–24,30], the role of pre-stimulus frontal beta is less clear. The scalp topomap result shown in Fig. 9B suggests that the beta power observed in the present study consists of two functionally distinct components. The fronto-centro-lateral component is stronger in the pre-stimulus period and may reflect a mechanism that suppresses neuronal processing of new movements, favoring maintenance of the current sensorimotor state[22,61,62,71–74]; in the present study this motor suppression may begin as subjects prepare for the upcoming stimulus after being cued to "stay still and listen now". In contrast, the fronto-medial component of beta may be a network mechanism spanning fronto-motor and auditory areas for top-down prediction/anticipation that may be active during both pre- and during-stimulus periods[22–24,30,75]. Our behavioral manipulations and 32-channel EEG recordings cannot further disambiguate between these two components of beta. Nevertheless, we find that although pre-stimulus beta covaried with during-stimulus beta and during-stimulus power levels were similar to pre-stimulus levels (Fig. 4B), during-stimulus frontal beta power contributed significant additional predictive power to predict within-condition percent-correct score over the contribution of pre-stimulus power alone (Supplementary Table S2). Future experiments should be designed to dissociate beta rhythms associated with motor suppression during attentional engagement from beta activity associated with dynamic predictive coding mechanisms. In particular, high-density EEG or magnetoencephalography (MEG) recordings along with source-space analysis can be used to probe which specific beta-band mechanisms relate to speech understanding in competition (e.g., beta-band synchrony between auditory and fronto-motor areas would imply a different mechanism from beta activity that is confined to motor cortex).

Our current results show that trial-by-trial variations in alpha and beta power are correlated, even within subject (Fig. 6A; statistics given in "Results"). Prior studies have also reported that oscillatory activity within the alpha and beta bands are correlated[76], even though they may represent distinct functions. Despite being correlated, alpha and beta power each provide significant independent contributions to predicting single-trial percent-correct score (Fig. 8). Moreover, there are individual differences in the overall magnitude of alpha and beta power across trials (Fig. 6A). Comparing these neural individual differences (Fig. 6A) to the individual differences in behavioral performance (Fig. 1) leads us to hypothesize that a greater alpha or beta power for an individual subject might relate to greater average performance for that subject; however, we are unable to test this specific hypothesis due to the low statistical power (just 6 subjects) in our study to conduct such an analysis of individual differences. Rather, we relate trial-by-trial fluctuations in alpha and beta power to trial-wise variations in behavioral outcome. Our data (Fig. 6B) suggests that there may also be individual differences in the

21. Alhanbali, S., Munro, K. J., Dawes, P., Perugia, E. & Millman, R. E. Associations between pre-stimulus alpha power, hearing level and performance in a digits-in-noise task. *Int. J. Audiol.* **61**(3), 197–204 (2022).
22. Engel, A. K. & Fries, P. Beta-band oscillations—signalling the status quo?. *Curr. Opin. Neurobiol.* **20**(2), 156–165 (2010).
23. Arnal, L. H. & Giraud, A. L. Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* **16**(7), 390–398 (2012).
24. Lewis, A. G. & Bastiaansen, M. A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex* **68**, 155–168 (2015).
25. Hickok, G., Houde, J. & Rong, F. Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron* **69**(3), 407–422 (2011).
26. Skipper, J. I., Devlin, J. T. & Lametti, D. R. The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain Lang.* **164**, 77–105 (2017).
27. Pulvermüller, F. Neural reuse of action perception circuits for language, concepts and communication. *Prog. Neurobiol.* **160**, 1–44 (2018).
28. Adank, P., Davis, M. H. & Hagoort, P. Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia* **50**(1), 77–84 (2012).
29. Du, Y., Buchsbaum, B. R., Grady, C. L. & Alain, C. Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc. Natl. Acad. Sci. USA* **111**(19), 7126–7131 (2014).
30. Cope, T. E. *et al.* Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nat. Commun.* **8**(1), 1–16 (2017).
31. Park, H., Ince, R. A., Schyns, P. G., Thut, G. & Gross, J. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr. Biol.* **25**(12), 1649–1653 (2015).
32. Keitel, A., Ince, R. A., Gross, J. & Kayser, C. Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *Neuroimage* **147**, 32–42 (2017).
33.