Distinct contributions of low- and high-frequency neural oscillations to speech comprehension

Anne Kösem & Virginie van Wassenhove

Published online: 18 Oct 2016.

Submit your article to this journal ⍰

Article views: 971

View related articles ⍰

View Crossmark data ⍰

Citing articles: 8 View citing articles ⍰

Routledge
Taylor & Francis Group

REGULAR ARTICLE

Distinct contributions of low- and high-frequency neural oscillations to speech comprehension

Anne Kösem[a,b] and Virginie van Wassenhove[c]

[a]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands; [b]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [c]Cognitive Neuroimaging Unit, CEA DSV/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin Center, Gif/Yvette, France

ABSTRACT

In the last decade, the involvement of neural oscillatory mechanisms in speech comprehension has been increasingly investigated. Current evidence suggests that low-frequency and high-frequency neural entrainment to the acoustic dynamics of speech are linked to its analysis. One crucial question is whether acoustical processing primarily modulates neural entrainment, or whether entrainment instead reflects linguistic processing. Here, we review studies investigating the effect of linguistic manipulations on neural oscillatory activity. In light of the current findings, we argue that theta (3–8 Hz) entrainment may primarily reflect the analysis of the acoustic features of speech. In contrast, recent evidence suggests that delta (1–3 Hz) and high-frequency activity (>40 Hz) are reliable indicators of perceived linguistic representations. The interdependence between low-frequency and high-frequency neural oscillations, as well as their causal role on speech comprehension, is further discussed with regard to neurophysiological models of speech processing.

Introduction

Auditory speech presents slow temporal fluctuations in the 1–10 Hz range that are closely linked to the syllabic and prosodic properties of the acoustic signals. Both syllabic and prosodic temporal cues are crucial for speech comprehension (Greenberg, Carvey, Hitchcock, & Chang, 2003; Poeppel, 2003; Poeppel, Idsardi, & van Wassenhove, 2008; Rosen, 1992); it has thus been proposed that the brain finds encoding strategies to track the slow temporal fluctuations of speech. In line with this, distinct neural responses have been observed to follow the slow speech rhythms. Neural oscillations in the delta (1–3 Hz) and theta (3–8 Hz) ranges have been shown to temporally correlate with the slow temporal dynamics of speech (Ahissar et al., 2001; Boemio, Fromm, Braun, & Poeppel, 2005; Ding & Simon, 2012, 2013a, 2013b; Ding, Melloni, Zhang, Tian, & Poeppel, 2016; Doelling, Arnal, Ghitza, & Poeppel, 2014; Gross et al., 2013; Kayser, Ince, Gross, & Kayser, 2015; Luo & Poeppel, 2012, 2007; Luo, Liu, & Poeppel, 2010; Millman, Johnson, & Prendergast, 2015; Millman, Prendergast, Hymers, & Green, 2013; Nourski et al., 2009; Zion Golumbic et al., 2013; Zoefel & VanRullen, 2015a). In addition, the power of high-frequency activity (>40 Hz) has also been observed to follow phrasal and syllabic rhythms (Ding et al., 2016; Kubanek, Brunner, Gunduz, Poeppel, & Schalk, 2013; Mesgarani & Chang, 2012; Nourski et al., 2009; Pasley et al., 2012; Zion Golumbic et al., 2013).

While the neural entrainment patterns in the delta, theta, and high-frequency ranges have consistently been observed in human data during speech listening, their underlying mechanism is still unclear. Considering that the neural tracking of slow auditory rhythms has also been reported for non-speech stimuli (Henry & Obleser, 2012; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Stefanics et al., 2010), one first source of debate is whether these responses are distinctive of linguistic processing in speech, or whether they reflect a generic processing of acoustic features. As a follow-up to recent critical reviews (Ding & Simon, 2014; Zoefel & VanRullen, 2015b), we argue that the different neural responses are not only reflecting sound processing, but linguistic processing as well. More specifically, a body of recent experimental reports hint at the conclusion that different neural markers do not relate to the same processing stages of speech, namely: (1) theta oscillations would primarily reflect early sound analysis and phonological processing; (2) delta oscillations would reflect the encoding of sound properties as well as

abstract syntactic structures; and (3) high gamma activity would represent phonological, lexical and syntactic information, and may reflect the output of speech analysis.

In the second part of this opinion paper, we discuss and re-interpret existing work with regard to two main neurophysiological models of speech perception, which we will refer to as oscillatory models and gain models. So far, there is no clear empirical evidence that observed neural oscillations could either originate from intrinsic oscillatory mechanisms or from non-oscillatory based mechanisms. We discuss how the two hypotheses could be disentangled and make some experimental predictions. We further focus on one property of oscillatory models, that is, the coupling between neural responses at different frequencies. We highlight that the dissociation between delta, theta, and gamma speech-tracking are inconsistent with the hypothesized directionality in cross-frequency coupling between the three neural markers.

Theta, delta, and gamma oscillations track speech at distinct processing stages

Neural entrainment and intelligibility: acoustic analysis of speech

While speech perception naturally relies on acoustic signals, it is also informed by linguistic knowledge. It comes as a direct question whether neural entrainment mainly reflects the acoustical analysis of rhythmic sounds, or whether they could also be the hallmark of linguistic processing. To test this, several reports have compared the properties of the entrained neural responses when listening to speech in contrast to neural entrainment to speech-like unintelligible acoustic streams. Among all neural responses tracking speech dynamics, theta oscillations – whose dynamics directly echo the syllabic rhythms – are the neural responses that have received most scrutinized interest (Ding & Simon, 2014; Ghitza, 2011; Giraud & Poeppel, 2012; Peelle & Davis, 2012; Poeppel, 2003; Poeppel et al., 2008). For now, existing reports present contradicting results, either associating stronger low-frequency entrainment with improved speech intelligibility (Ahissar et al., 2001; Ding & Simon, 2013a; Doelling et al., 2014; Gross et al., 2013; Peelle, Gross, & Davis, 2013; Pérez, Carreiras, Gillon Dowens, & Duñabeitia, 2015; Rimmele, Zion Golumbic, Schröger, & Poeppel, 2015) or showing no change in neural entrainment responses between intelligible speech and non-intelligible speech-like sounds (Howard & Poeppel, 2010; Millman et al., 2015; Peña & Melloni, 2012; Zoefel & VanRullen, 2015a). We argue below that the main differences between studies could be accounted for by the acoustical properties of the stimuli. In other words,

low-frequency neural entrainment would primarily reflect acoustic processing.

In findings supporting that low-frequency neural entrainment is related to some extent to linguistic processing, the manipulations of speech intelligibility was done by applying important modifications of the speech acoustics. For instance, the manipulations to render speech unintelligible included changes in speed (Ahissar et al., 2001), the loss of spectral complexity using noise-vocoded speech (Peelle et al., 2013; Rimmele et al., 2015), changes in the temporal dynamics of the speech envelope using time-reversed speech (Gross et al., 2013; Mai, Minett, & Wang, 2016), or the removal of strong acoustic edges (Doelling et al., 2014), and the addition of background noise (Ding & Simon, 2013a). Yet, temporal as well as spectro-temporal characteristics of sounds are known to drive neural entrainment (Ding, Chatterjee, & Simon, 2013); reducing the spectro-temporal complexity of speech could thus be predicted to have consequences on neural entrainment that are independent of linguistic processing.

In contrast, when variations in acoustic properties between intelligible and unintelligible speech stimuli were controlled for, no clear link was observed between theta entrainment and non-acoustic speech analysis. For instance, specific noise could be added to speech signals to remove the energetic fluctuations and spectral content that yield the entrainment to sounds at the level of the cochlea (Zoefel & VanRullen, 2015c). Despite the absence of amplitude fluctuations in the speech envelope, strong entrainment of the electroencephalography (EEG) response was observed when participants listened to the manipulated sounds. The observed neural oscillatory response was interpreted as entrainment to speech features processed beyond the cochlea (Zoefel & VanRullen, 2015a). However, and importantly, the strength of neural entrainment was equally strong when the manipulated speech signals were time-reversed thus on-intelligible (Zoefel & VanRullen, 2015a). These findings suggest that differences in the neural entrainment to speech and time-reversed speech could originate from the processing of acoustical temporal dynamics, and not intelligibility. One other way to reduce the acoustic differences between speech and non-intelligible speech is to contrast brain responses while listening to distinct languages. Neural oscillations have been shown to track to a lesser extent non-native or foreign compared to native speech for certain types of languages contrasts (e.g. English vs. Spanish) (Pérez et al., 2015), but not for others (Spanish vs. Italian vs. Japanese) (Peña & Melloni, 2012). As distinct languages differ with regard to the temporal fluctuations of the speech envelope (e.g. phonetic and stress patterns), the differences in reports could be

explained by non-linguistic acoustic processing. Consistent with this, word and pseudo-word sequences of similar phonological structure are shown to yield similar neural entrainment patterns (Mai et al., 2016). Finally, to rule out any acoustical confounds, Millman et al. (2015) have investigated stimuli with the exact same acoustical information but distinct linguistic content. To do so, the authors used noise-vocoded speech that was initially unintelligible but could become intelligible after the presentation of a clear speech utterance (Millman et al., 2015). In this study, no differences were observed in the strength of neural entrainment before and after the training, suggesting that neural entrainment was uniquely driven by the acoustical processing of the signals. It should be noted that no study so far has reported that intelligible speech processing could occur without neural entrainment. Hence, the observed low-frequency entrainment may be the hallmark of a necessary (but not sufficient, and not language-related) mechanism used for speech comprehension.

To conclude, it remains unclear using manipulations of speech intelligibility whether neural entrainment to speech has a functional role in encoding linguistic information essentially due to potential acoustical confounds in the experimental designs.

Theta, delta, and gamma oscillations represent distinct linguistic features

Cortical entrainment to speech reflects to a great extent the bottom-up analysis of acoustic signals. Yet, neural entrainment can also reflect top-down modulations of sound processing by attention (Lakatos et al., 2008) or prediction (Kayser et al., 2015; Stefanics et al., 2010). The attentional modulations of low-frequency oscillations could be particularly relevant for speech processing in adverse conditions to track the relevant signal and to discard irrelevant distracting information (Ding & Simon, 2012; Zion Golumbic et al., 2013). In addition to non-speech-specific top-down modulations, recent reports have suggested that linguistic processing could modulate neural entrainment. As will now be discussed, theta entrainment may not only reflect acoustic analysis but also possibly phonological processing. It has recently been shown that EEG theta entrainment to natural speech could be best predicted when speech was parameterized as the combination of low-level acoustic and phonetic features (Di Liberto, O'sullivan, & Lalor, 2015). The phonetic labelling of sounds did not improve the modelisation of the neural oscillatory response to time-reversed speech, suggesting that these labels efficiently captured speech-specific content. The results thus indicated that theta oscillations are sensitive to the phonetic features of

speech. Consistent with these reports, theta (3–8 Hz) oscillations are also recruited during phonemic restoration (Riecke et al., 2012; Riecke, Esposito, Bonte, & Formisano, 2009; Strauß, Kotz, Scharinger, & Obleser, 2014; Sunami et al., 2013). Theta oscillations could also be recruited for the identification of consonants in syllables (Ten Oever & Sack, 2015), as well as the processing of co-articulation cues that could guide the parsing of words within speech streams (Kösem, Basirat, Azizi, & van Wassenhove, 2016).

Delta oscillations have been associated with the processing of non-speech-specific attentional and predictive modulations of auditory processing (Lakatos et al., 2008; Nozaradan, Peretz, Missal, & Mouraux, 2011; Stefanics et al., 2010; Zion Golumbic et al., 2013) but they do not serve the same purpose as theta oscillations. Findings have suggested that oscillatory responses can be elicited in the absence of acoustic landmarks delimiting speech units in the acoustic signal (Buiatti, Peña, & Dehaene-Lambertz, 2009; Ding et al., 2016). More importantly, recent report suggest that delta oscillations may reflect abstract language processing (Ding et al., 2016): in this study, brain signals presented rhythmic fluctuations that corresponded to sub-harmonics of the syllabic rhythm (4 Hz), which were not present in the speech acoustics. Crucially, the observed delta frequency corresponded to the syntactic complexity of the heard speech signal. Series of phrases constituted of two monosyllabic words were associated with a neural entrainment at 2 Hz, and sentences composed of four monosyllabic words induced a delta response at 1 Hz. Hence, delta oscillations could be linked to the combination or unification of word units into a longer and more abstract linguistic structure. The process of syntactic unification is attributed to the synergic interaction between left inferior frontal cortices and left posterior temporal areas (Hagoort & Indefrey, 2014). Consistent with this idea, recent findings have demonstrated top-down communication using transfer entropy measures of delta band responses between the left Inferior Frontal Gyrus and the left auditory cortices (Park, Ince, Schyns, Thut, & Gross, 2015).

In addition to low-frequency entrainment, high-frequency activity (>40 Hz) has also been shown to follow speech dynamics (Ding et al., 2016; Kubanek et al., 2013; Mesgarani & Chang, 2012; Nourski et al., 2009; Pasley et al., 2012; Peña & Melloni, 2012; Zion Golumbic et al., 2013). Several pieces of evidence suggest that high-frequency power is modulated by variations in linguistic processing at different stages of speech processing. First, the content of heard and attended speech signals can be decoded from high-frequency power (Mesgarani & Chang, 2012; Pasley et al., 2012; Zion Golumbic et al., 2013). More than pure acoustic processing, high-

frequency activity may reflect the representations of speech sounds considering that phonological features can be decoded from high-frequency power of intracranial recordings in auditory cortices (Mesgarani, Cheung, Johnson, & Chang, 2014). The power of high-frequency neural activity is also modulated by lexical knowledge so that high-frequency power is stronger for words compared to pseudo-words (Cibelli, Leonard, Johnson, & Chang, 2015; Hannemann, Obleser, & Eulitz, 2007; Mai et al., 2016) and for native compared to unknown languages (controlled for phonologic and phono tactic repertoires) (Peña & Melloni, 2012). It should be noted here that lexical modulations were observed as a direct modulation of high-frequency oscillatory power, and not on the tracking of speech dynamics by high-frequency power. However, additional evidence suggests that high-frequency activity tags the timing of abstract linguistic speech units and structures. For instance, high-frequency activity has been shown to reflect the result of top-down word parsing: high-frequency oscillatory power dynamics systematically delineated the boundaries of perceived monosyllabic words in ambiguous speech streams (Kösem et al., 2016). High-frequency activity also represented, concurrently with delta oscillations, the segmentation of longer abstract linguistic structures (phrases and sentences) in continuous speech (Ding et al., 2016).

All in all, the reviewed data suggest that neural tracking may conjointly decode the acoustic and the distinctive linguistic features of the speech signals. However, these findings also suggest that theta, delta, and high-frequency oscillatory activity may relate to different types of speech processing: theta entrainment may process acoustic and phonological information of the signal, delta oscillations may reflect the combinatorial processes underlying sentence unification, and high-frequency activity may be associated with the output of phonological as well as lexical and syntactic processing.

Potential neural origins of the observed oscillatory responses

The functional role of neural oscillations in speech processing

From a mechanistic perspective, what can be concluded from the observed modulations of low- and high-frequency oscillatory activity during speech listening? One crucial question concerns the origins of the neural entrainment patterns observed in magnetoencephalography (MEG), EEG, and intracranial recordings. Is oscillatory activity present in electrophysiological data because neural oscillations are mechanistically involved in speech processing, or does it reflect the by-product

of other mechanisms (e.g. gain control mechanisms) at stake during speech processing? For now, both types of models have received empirical support, and could in principle explain to a similar extent the reported findings. Here, we discuss briefly the two kinds of models and suggest how the two models could be tested against each other.

Oscillatory models (Ding & Simon, 2014; Ghitza, 2011; Giraud & Poeppel, 2012; Peelle & Davis, 2012; Poeppel, 2003; Poeppel et al., 2008) propose that neural entrainment to speech originates from the recruitment of intrinsic neural oscillatory mechanisms, which have a causal influence on speech perception. Neural oscillations would specifically serve as parsers of the acoustic signals into a neural code for speech – namely, discretized informational units relevant for additional linguistic processing. In these models, intrinsic low-frequency neural oscillations in the delta (1–4 Hz) and theta (4–8 Hz) ranges track the phrasal and syllabic rates of speech, respectively. Low-frequency oscillations impose periods of inhibition and excitation in high-frequency neural dynamics (Fries, 2009; Schroeder & Lakatos, 2009a) yielding cross-frequency coupling, that is, increase in high-frequency activity at particular phases of low-frequency activity (Akam & Kullmann, 2014; Canolty et al., 2006; Canolty & Knight, 2010; Lakatos et al., 2005). High-frequency activity would thus reflect the read-out of speech processing (Ghitza, 2011; Giraud & Poeppel, 2012; Mesgarani et al., 2014; Poeppel, 2003; Poeppel et al., 2008). Because high-frequency activity is constrained by low-frequency entrainment, this cross-frequency-coupling results in the discretization of the neural representation of the auditory input. Crucially, these models assume a hierarchical coupling between the distinct neural oscillations being entrained to the speech dynamics, namely: low-frequency entrainment should be the master mechanism driving high-frequency dynamics (Schroeder & Lakatos, 2009b; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008).

Alternatively, neural entrainment to speech could reflect evoked responses controlled by gain mechanisms (Kerlin, Shahin, & Miller, 2010) or the outcome of temporal coherence analyses (Ding & Simon, 2012; Shamma, Elhilali, & Micheyl, 2011). Here, the observed neural oscillations would reflect the summated individual evoked response to each relevant acoustic feature in speech. Because speech is rhythmic, any magnification of speech-evoked responses would lead to enhanced oscillatory activity in auditory cortices. For instance, an increase in speech intelligibility would be the result of an increased excitability of the neural network encoding for the relevant speech features (Ding & Simon, 2012; Hillyard, Hink, Schwent, & Picton, 1973; Kerlin et al.,

2010; Woldorff et al., 1993). Hence, the observed enhancement of neural oscillatory activity could be the by-product of the overall enhancement of the evoked response to the stimulus. In this scheme, brain oscillations do not confer any mechanistic benefits to the encoding of speech. In other words, gain models would predict the emergence of oscillatory brain patterns during speech listening (Ding & Simon, 2014), although the nature of the speech decoding mechanisms would not be oscillatory *per se*.

Behavioural and neuroscientific evidence so far, though mostly interpreted within the oscillatory framework, could be explained by the two mechanistic theories. As discussed earlier, enhanced neural oscillatory activity observed for intelligible and/or attended speech signals could either originate from the stronger recruitment of parsing oscillatory mechanisms, or from the applied gain on the evoked response to the selected acoustic features. Other pieces of evidence interpreted in favour of oscillatory models include the fading of neural entrainment for fast rhythmic stimulation (Ahissar et al., 2001; Nourski et al., 2009), and the impairment of speech recognition when the acoustic information is presented at too fast or too slow rates (Brungart & van Wassenhove, 2007; Brungart, Iyer, Simpson, & van Wassenhove, 2008; Ghitza, 2014; Ghitza & Greenberg, 2009). These data suggest that the rate of auditory information must be constrained between 2 and 10 Hz so that speech gets efficiently parsed and becomes understandable (Boemio et al., 2005; Poeppel, 2003; Poeppel et al., 2008). The evidence for restricted temporal sensitivities to acoustic features in auditory speech and the existence of temporal encoding windows in audiovisual speech (van Wassenhove, 2013; van Wassenhove, Grant, & Poeppel, 2007) are compatible with oscillatory models hypothesizing that speech parsing is constrained by theta oscillations (Ghitza, 2011; Giraud & Poeppel, 2012). However, these results could also be explained by non-oscillatory mechanisms. Stimulus-evoked responses have been suggested to reflect 50–300 ms long packets of neural activity, that provide increased information on the stimulus as the packet progresses in time (Luczak, McNaughton, & Harris, 2015).This view suggests that the evoked response to ongoing stimulation is discrete, and that packets of neural activity would constitute fundamental units of neural code: sensory information that will be presented within the time period of the packet will be integrated as one unit in information. Because the packets have limited durations, the encoding accuracy of the stimulus' content may be potentially impaired if auditory information is presented at a too fast rate or a too slow rate. This would eventually yield reduced observed oscillatory

activity and speech intelligibility. In other words, alternative mechanisms could make the assumption that the neural code to speech is discrete and has limited temporal parsing windows, this without the recruitment of intrinsic neural oscillators.

A question related to the debate on the mechanistic nature behind the observed neural oscillations in electrophysiological recordings, is whether neural entrainment carries temporal predictions (Morillon & Schroeder, 2015). Gain control and oscillatory models differ in their views regarding this question. Neural oscillations are somewhat epiphenomenal in gain models and do not relate to temporal predictive mechanisms. In oscillatory models of speech processing, neural oscillations mechanistically track the dynamics the speech envelope via phase resetting at specific speech onsets. As such, neural entrainment could constitute a non-predictive evoked mechanism if phase resetting occurs at each sound onset. Alternatively, phase reset may not be constantly instantiated, and recent work showed that the endogenous control of oscillatory phase during perception may provide a neural code for event timing (Kösem, Gramfort, & van Wassenhove, 2014; Samaha, Bauer, Cimaroli, & Postle, 2015; van Wassenhove, 2016). Hence one possible role of intrinsic oscillators would be, via the tuning to the timing of current sensory input, to predict the arrival of future events (Morillon & Schroeder, 2015; van Wassenhove, 2016). It is difficult to test oscillatory and gain models against each other, as well as predictive and non-predictive accounts of neural entrainment during speech processing, as they all hypothesise that oscillatory activity should be observed during the presentation of the speech signals. However, the models make different predictions on the presence of oscillatory activity after speech stimulation, or when speech temporal characteristics are suddenly changing. Predictive oscillatory models suggest that neural entrainment that followed sensory information presented at instant t should last at instant $t +$ 1 even in the absence or change of stimulation. By analogy, neural entrainment should work as a bicycle wheel: even if the pedalling has stopped, the wheel continues turning. In contrast, gain control models predict that neural oscillation responses would vanish in the absence of any external input, or would tune immediately to the changes of dynamics of new sensory input. Evoked oscillatory models (systematic phase-reset) would also predict that oscillations phase-reset directly and adapt to the dynamics of the new sensory input; but oscillatory activity should persist in the absence of new stimulation. There are glimpses of evidence that neural entrainment could last for a few cycles after visual/tACS stimulation (Helfrich et al., 2014; Spaak, de

Lange, & Jensen, 2014; Zaehle, Rach, & Herrmann, 2010) but additional research is needed to test whether this remising oscillatory response has consequences on speech perception. Another way to test oscillatory models is to recruit the imputed endogenous oscillators without external sound stimulation. Manipulations of sound processing via theta stimulation with tACS have successfully been reported (Riecke, Formisano, Herrmann, & Sack, 2015; Riecke, Sack, & Schroeder, 2015) and tACS studies involving speech stimuli could thus provide additional support for oscillator models.

Coupling between low- and high-frequency neural tracking of speech

Currents oscillatory models of speech processing propose that different neural oscillators decode distinct linguistic units but also, and crucially, make the assumption that the three oscillators (delta, theta, and gamma) are hierarchically coupled to each other, and that oscillators at lower frequencies constrain activity at higher frequency. Recent computational models have provided evidence that phase-amplitude coupling between theta and high-frequency oscillations could provide a relevant encoding strategy to delineate syllabic boundaries within speech in a bottom-up fashion (Ghitza, 2011; Hyafil, Giraud, Fontolan, & Gutkin, 2015; Shamir, Ghitza, Epstein, & Kopell, 2009). However, additional mechanisms are required to explain the observed top-down modulations of neural oscillations based on linguistic knowledge. If the coupling holds, then theta and high-frequency dynamics should undergo similar modulations by linguistic processing. Specifically, high-frequency dynamics should always be constrained so that the maximum of high-frequency power occurs at a specific phase of the low-frequency oscillation (Canolty et al., 2006; Schroeder & Lakatos, 2009a). Yet, as discussed in the first section, high-frequency power is indicative of linguistic manipulations while theta activity is unchanged (Mai et al., 2016; Peña & Melloni, 2012), and the dynamics of high-frequency power may follow the perceived parsing of speech sequences while theta oscillations remain stable in comparison (Kösem et al., 2016). Although more evidence is needed to confirm the observed coupling behaviour between high- and low-frequency neural dynamics during speech processing, the data is in line with reports in the visual domain, showing that distinct phases of theta-modulated high-frequency lead to distinct perceptual outcomes (Landau, Schreyer, van Pelt, & Fries, 2015; Maris, Womelsdorf, Desimone, & Fries, 2013).

To explain the observed differences between low- and high-frequency neural entrainment to speech, one extreme hypothesis would be to consider that there is no actual coupling between delta, theta, and high-frequency oscillations. Low- and high-frequency entrainment mechanisms would co-occur independently and decode distinct properties of the speech sounds. This theory would speak against theta oscillations as the "master oscillator" mechanism during the encoding of linguistic content (Scott & McGettigan, 2013). Another possibility is that cross-frequency-coupling mechanisms mainly operate between distinct brain areas (Fontolan, Morillon, Liegeois-Chauvel, & Giraud, 2014; Park et al., 2015), and different brain sources preferentially track speech dynamics through either low- or of high-frequency channels (Ding et al., 2016). Finally, an alternative mechanism could consist in maintaining the coupling between low-frequency and high-frequency oscillations, but varying the phase of coupling to provide an additional code for perception (Jensen, Bonnefond, & VanRullen, 2012; Jensen, Gips, Bergmann, & Bonnefond, 2014; Kösem et al., 2014; Ng, Logothetis, & Kayser, 2013; Panzeri, Brunel, Logothetis, & Kayser, 2010). Hence, complex coupling schemes provide potential explanation of the current experimental data (Maris, Fries, & van Ede, 2016; van Wassenhove, 2016). Future research will give us promising insight on the origins of the distinct brain oscillatory responses to speech, how they are interacting with each other, and what their predictive impact is on speech comprehension.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCiD

Anne Kösem http://orcid.org/0000-0002-2692-9999
Virginie van Wassenhove http://orcid.org/0000-0002-2569-5502

References

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13367–13372. doi:10.1073/pnas.201400998
Akam, T., & Kullmann, D. M. (2014). Oscillatory multiplexing of population codes for selective communication in the

mammalian brain. *Nature Reviews Neuroscience*, 15(2), 111–122. doi:10.1038/nrn3668

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8(3), 389–395. doi:10.1038/nn1409

Brungart, D., Iyer, N., Simpson, B., & van Wassenhove, V. (2008). *The effects of temporal asynchrony on the intelligibility of accelerated speech*. International Conference on Auditory-Visual Speech Processing (AVSP), Moreton Island, Australia.

Brungart, D., & van Wassenhove, V. (2007). *The effects of temporal acceleration and deceleration on AV speech perception*. International Conference on Auditory-Visual Speech Processing (AVSP), Hilvarenbeek, The Netherlands.

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage*, 44(2), 509–519. doi:10.1016/j.neuroimage.2008.09.015

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Berger, M. S., … Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, 313(5793), 1626–1628. doi:10.1126/science.1128115.High

Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11), 506–515. doi:10.1016/j.tics.2010.09.001

Cibelli, E. S., Leonard, M. K., Johnson, K., & Chang, E. F. (2015). The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening. *Brain and Language*, 147, 66–75. doi:10.1016/j.bandl.2015.05.005

Di Liberto, G. M., O'sullivan, J. A., & Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Current Biology*, 25(19), 2457–2465. doi:10.1016/j.cub.2015.08.030

Ding, N., Chatterjee, M., & Simon, J. Z. (2013). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage*, 88C, 41–46. doi:10.1016/j.neuroimage.2013.10.054

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*. Advance online publication. doi:10.1038/nn.4186

Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences*, 109(29), 11854–11859. doi:10.1073/pnas.1205381109

Ding, N., & Simon, J. Z. (2013a). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *The Journal of Neuroscience*, 33(13), 5728–5735. doi:10.1523/JNEUROSCI.5297-12.2013

Ding, N., & Simon, J. Z. (2013b). Power and phase properties of oscillatory neural responses in the presence of background activity. *Journal of Computational Neuroscience*, 34(2), 337–343. doi:10.1007/s10827-012-0424-6

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, 8, 13367. doi:10.3389/fnhum.2014.00311

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, 85(Pt 2), 761–768. doi:10.1016/j.neuroimage.2013.06.035

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature Communications*, 5, 4694. doi:10.1038/ncomms5694

Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Review of Neuroscience*, 32, 209–224. doi:10.1146/annurev.neuro.051508.135603

Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, 2, 130. doi:10.3389/fpsyg.2011.00130

Ghitza, O. (2014). Behavioral evidence for the role of cortical θ oscillations in determining auditory channel capacity for speech. *Frontiers in Psychology*, 5, 751. doi:10.3389/fpsyg.2014.00652

Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, 66(1–2), 113–126. doi:10.1159/000208934

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517. doi:10.1038/nn.3063

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech – a syllable-centric perspective. *Journal of Phonetics*, 31(3–4), 465–485. doi:10.1016/j.wocn.2003.09.005

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, 11(12), e1001752. doi:10.1371/journal.pbio.1001752

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience*, 37, 347–362. doi:10.1146/annurev-neuro-071013-013847

Hannemann, R., Obleser, J., & Eulitz, C. (2007). Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Research*, 1153, 134–143. doi:10.1016/j.brainres.2007.03.069

Helfrich, R. F., Schneider, T. R., Rach, S., Trautmann-Lengsfeld, S. A., Engel, A. K., & Herrmann, C. S. (2014). Entrainment of brain oscillations by transcranial alternating current stimulation. *Current Biology*, 24(3), 333–339. doi:10.1016/j.cub.2013.12.041

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 109(49), 20095–20100. doi:10.1073/pnas.1213390109

Hillyard, S., Hink, R., Schwent, V., & Picton, T. (1973). Electrical signs of selective attention in the human brain. *Science*, 182(4108), 177–180.

Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, 104(5), 2500–2511. doi:10.1152/jn.00251.2010

Hyafil, A., Giraud, A., Fontolan, L., & Gutkin, B. (2015). Neural cross-frequency coupling: Connecting architectures, mechanisms, and functions. *Trends in Neurosciences*, 38(11), 725–740. doi:10.1016/j.tins.2015.09.001

Jensen, O., Bonnefond, M., & VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends in Cognitive Sciences*, *16*(4), 200–206. doi:10.1016/j.tics.2012.03.002

Jensen, O., Gips, B., Bergmann, T. O., & Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends in Neurosciences*, *37*(7), 357–369. doi:10.1016/j.tins.2014.04.001

Kayser, S. J., Ince, R. A. A., Gross, J., & Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *The Journal of Neuroscience*, *35*(44), 14691–14701. doi:10.1523/JNEUROSCI.2243-15.2015

Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party". *The Journal of Neuroscience*, *30*(2), 620–628. doi:10.1523/JNEUROSCI.3631-09.2010

Kösem, A., Basirat, A., Azizi, L., & van Wassenhove, V. (2016). High frequency neural activity predicts word parsing in ambiguous speech streams. *Journal of Neurophysiology*. Advance online publication. doi:10.1152/jn.00074.2016

Kösem, A., Gramfort, A., & van Wassenhove, V. (2014). Encoding of event timing in the phase of neural oscillations. *NeuroImage*, *92*, 274–284. doi:10.1016/j.neuroimage.2014.02.010

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., & Schalk, G. (2013). The tracking of speech envelope in the human cortex. *PloS One*, *8*(1), e53398. doi:10.1371/journal.pone.0053398

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113. doi:10.1126/science.1154735

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. doi:10.1152/jn.00263.2005

Landau, A. N., Schreyer, H. M., van Pelt, S., & Fries, P. (2015). Distributed attention Is implemented through theta-rhythmic gamma modulation. *Current Biology*, *25*(17), 2332–2337. doi:10.1016/j.cub.2015.07.048

Luczak, A., McNaughton, B. L., & Harris, K. D. (2015). Packet-based communication in the cortex. *Nature Reviews Neuroscience*, *16*(12), 745–755. doi:10.1038/nrn4026

Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology*, *8*(8), e1000445. doi:10.1371/journal.pbio.1000445

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010. doi:10.1016/j.neuron.2007.06.004

Luo, H., & Poeppel, D. (2012). Cortical oscillations in auditory perception and speech: Evidence for two temporal windows in human auditory cortex. *Frontiers in Psychology*, *3*, 170. doi:10.3389/fpsyg.2012.00170

Mai, G., Minett, J. W., & Wang, W. S.-Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, doi:10.1016/j.neuroimage.2016.02.064

Maris, E., Fries, P., & van Ede, F. (2016). Diverse phase relations among neuronal rhythms and their potential function. *Trends in Neurosciences*, *39*(2), 86–99. doi:10.1016/j.tins.2015.12.004

Maris, E., Womelsdorf, T., Desimone, R., & Fries, P. (2013). Rhythmic neuronal synchronization in visual cortex entails spatial phase relation diversity that is modulated by stimulation and attention. *NeuroImage*, *74*, 99–116. doi:10.1016/j.neuroimage.2013.02.007

Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, *485*(7397), 233–236. doi:10.1038/nature11020

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, *343*(6174), 1006–1010. doi:10.1126/science.1245994

Millman, R. E., Johnson, S. R., & Prendergast, G. (2015). The role of phase-locking to the temporal envelope of speech in auditory perception and speech intelligibility. *Journal of Cognitive Neuroscience*, *27*(3), 533–545. doi:10.1162/jocn_a_00719

Millman, R. E., Prendergast, G., Hymers, M., & Green, G. G. R. (2013). Representations of the temporal envelope of sounds in human auditory cortex: Can the results from invasive intracortical "depth" electrode recordings be replicated using non-invasive MEG "virtual electrodes"? *NeuroImage*, *64*, 185–196. doi:10.1016/j.neuroimage.2012.09.017

Morillon, B., & Schroeder, C. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, *1337*(1), 26–31. doi:10.1111/nyas.12629

Ng, B., Logothetis, N., & Kayser, C. (2013). EEG phase patterns reflect the selectivity of neural firing. *Cerebral Cortex*, *23*(2), 389–398. doi:10.1093/cercor/bhs031

Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., … Brugge, J. F. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *The Journal of Neuroscience*, *29*(49), 15564–15574. doi:10.1523/JNEUROSCI.3065-09.2009

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234–10240. doi:10.1523/JNEUROSCI.0411-11.2011

Panzeri, S., Brunel, N., Logothetis, N. K., & Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends in Neurosciences*, *33*(3), 111–120. doi:10.1016/j.tins.2009.12.001

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology*, *25*(12), 1649–1653. doi:10.1016/j.cub.2015.04.049

Pasley, B. N., David, S. V., Mesgarani, N., Flinker, A., Shamma, S. A., Crone, N. E., … Chang, E. F. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology*, *10*(1), e1001251. doi:10.1371/journal.pbio.1001251

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*, 320. doi:10.3389/fpsyg.2012.00320

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. doi:10.1093/cercor/bhs118

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of Cognitive Neuroscience*, *24*(5), 1149–1164. doi:10.1162/jocn_a_00144

Pérez, A., Carreiras, M., Gillon Dowens, M., & Duñabeitia, J. A. (2015). Differential oscillatory encoding of foreign speech. *Brain and Language*, *147*, 51–57. doi:10.1016/j.bandl.2015.05.008

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as "asymmetric sampling in time". *Speech Communication*, *41*(1), 245–255. doi:10.1016/S0167-6393(02)00107-3

Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 1071–1086. doi:10.1098/rstb.2007.2160

Riecke, L., Esposito, F., Bonte, M., & Formisano, E. (2009). Hearing illusory sounds in noise: The timing of sensory-perceptual transformations in auditory cortex. *Neuron*, *64*(4), 550–561. doi:10.1016/j.neuron.2009.10.016

Riecke, L., Formisano, E., Herrmann, C. S., & Sack, A. T. (2015). 4-Hz transcranial alternating current stimulation phase modulates hearing. *Brain Stimulation*, *8*(4), 777–783. doi:10.1016/j.brs.2015.04.004

Riecke, L., Sack, A. T., & Schroeder, C. E. (2015). Endogenous delta/theta sound-brain phase entrainment accelerates the buildup of auditory streaming. *Current Biology*, *25*(24), 3196–3201. doi:10.1016/j.cub.2015.10.045

Riecke, L., Vanbussel, M., Hausfeld, L., Başkent, D., Formisano, E., & Esposito, F. (2012). Hearing an illusory vowel in noise: Suppression of auditory cortical activity. *The Journal of Neuroscience*, *32*(23), 8024–8034. doi:10.1523/JNEUROSCI.0440-12.2012

Rimmele, J. M., Zion Golumbic, E., Schröger, E., & Poeppel, D. (2015). The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. *Cortex*, *68*, 144–154. doi:10.1016/j.cortex.2014.12.014

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *336*(1278), 367–373. doi:10.1098/rstb.1992.0070

Samaha, J., Bauer, P., Cimaroli, S., & Postle, B. R. (2015). Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(27), 8439–8444. doi:10.1073/pnas.1503686112

Schroeder, C. E., & Lakatos, P. (2009a). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. doi:10.1016/j.tins.2008.09.012

Schroeder, C. E., & Lakatos, P. (2009b). The gamma oscillation: Master or slave? *Brain Topography*, *22*(1), 24–26. doi:10.1007/s10548-009-0080-y

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3), 106–113.

Scott, S. K., & McGettigan, C. (2013). Do temporal processes underlie left hemisphere dominance in speech perception? *Brain and Language*, *127*(1), 36–45. doi:10.1016/j.bandl.2013.07.006

Shamir, M., Ghitza, O., Epstein, S., & Kopell, N. (2009). Representation of time-varying stimuli by a network exhibiting oscillations on a faster time scale. *PLoS Computational Biology*, *5*(5), e1000370. doi:10.1371/journal.pcbi.1000370

Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neurosciences*, *34*(3), 114–123. doi:10.1016/j.tins.2010.11.002

Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local entrainment of α oscillations by visual stimuli causes cyclic modulation of perception. *The Journal of Neuroscience*, *34*(10), 3536–3544. doi:10.1523/JNEUROSCI.4385-13.2014

Stefanics, G., Hangya, B., Hernadi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *The Journal of Neuroscience*, *30*(41), 13578–13585. doi:10.1523/JNEUROSCI.0703-10.2010

Strauß, A., Kotz, S. A., Scharinger, M., & Obleser, J. (2014). Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *NeuroImage*, *97*, 387–395. doi:10.1016/j.neuroimage.2014.04.005

Sunami, K., Ishii, A., Takano, S., Yamamoto, H., Sakashita, T., Tanaka, M., … Yamane, H. (2013). Neural mechanisms of phonemic restoration for speech comprehension revealed by magnetoencephalography. *Brain Research*, *1537*, 164–173. doi:10.1016/j.brainres.2013.09.010

Ten Oever, S., & Sack, A. T. (2015). Oscillatory phase shapes syllable perception. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(52), 15833–15837. doi:10.1073/pnas.1517519112

van Wassenhove, V. (2013). Speech through ears and eyes: Interfacing the senses with the supramodal brain. *Frontiers in Psychology*, *4*, 388. doi:10.3389/fpsyg.2013.00388

van Wassenhove, V. (2016). Temporal cognition and neural oscillations. *Current Opinion in Behavioral Sciences*, *8*, 124–130. doi:10.1016/j.cobeha.2016.02.012

van Wassenhove, V., Grant, K., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598–607. doi:10.1016/j.neuropsychologia.2006.01.001

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, *90*(18), 8722–8726. doi:10.1073/pnas.90.18.8722

Zaehle, T., Rach, S., & Herrmann, C. S. (2010). Transcranial alternating current stimulation enhances individual alpha activity in human EEG. *PLoS One*, *5*(11), e13766. doi:10.1371/journal.pone.0013766

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., … Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, *77*(5), 980–991. doi:10.1016/j.neuron.2012.12.037

Zoefel, B., & VanRullen, R. (2015a). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage*, *124*, 16–23. doi:10.1016/j.neuroimage.2015.08.054

Zoefel, B., & VanRullen, R. (2015b). The role of high-level processes for oscillatory phase entrainment to speech sound. *Frontiers in Human Neuroscience*, *9*, 360. doi:10.3389/fnhum.2015.00651

Zoefel, B., & VanRullen, R. (2015c). Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *The Journal of Neuroscience*, *35*(5), 1954–1964. doi:10.1523/JNEUROSCI.3484-14.2015