bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Oscillatory and aperiodic neural activity jointly predict grammar learning

Zachariah R. Cross[1,2*], Andrew W. Corcoran[1,3], Matthias Schlesewsky[1,2], Mark. J. Kohler[1,4], Ina Bornkessel-Schlesewsky[1,4]

[1]Cognitive and Systems Neuroscience Research Hub, University of South Australia, Adelaide, Australia
[2]School of Psychology, Social Work and Social Policy, University of South Australia, Adelaide, Australia.
[3]Cognition and Philosophy Laboratory, Monash University, Melbourne, Australia.
[4]Adelaide Brain & Cognitive Development Laboratory, Adelaide University, Adelaide, Australia.

*Corresponding author: Tel. +61 8 8302 4375, e-mail: Zachariah.Cross@mymail.unisa.edu.au

Manuscript details

Number of pages:	31
Number of figures:	7
Abstract word count:	250
Introduction word count:	1,309
Discussion word count:	2,372
Data available at:	TBC
Code available at:	TBC

Funding: Preparation of this work was supported by Australian Commonwealth Government funding awarded to ZRC under the Research Training Program (RTP; number 212190). IB-S is supported by an Australian Research Council Future Fellowship (FT160100437). AWC is supported by an Australian Government RTP scholarship.

Acknowledgements: We thank Isabella Sharrad, Lena Zou-Williams, Erica Wilkinson, Nicole June and Angela Osborn for help with data collection. Thank you also to the participants.

Conflict of interest

The authors declare no competing financial interests.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ABSTRACT

Memory formation involves the synchronous firing of neurons in task-relevant networks, with recent models postulating that a decrease in low frequency oscillatory activity underlies successful memory encoding and retrieval. To date, this relationship has predominantly been investigated using objects (e.g., faces, natural scenes); however, considerably less is known about the oscillatory correlates of complex rule learning (e.g., language acquisition). Further, recent work has shown that aperiodic (non-oscillatory) $1/f$ activity is functionally and behaviourally relevant, yet its interaction with oscillatory activity during complex rule learning remains virtually unknown. Using spectral decomposition and power-law exponent estimation of human EEG data, we show for the first time that $1/f$ and oscillatory activity jointly influence the learning of different word order rules of a miniature language system. Fixed word order rules were associated with an increased power-law exponent (i.e. steeper $1/f$ slope) compared to flexible word order rules. We also show that stronger anterior beta synchronisation predicts fixed word order rule learning and subsequent behavioural performance, while stronger theta/alpha synchronisation is associated with the learning of flexible word order rules. These results also revealed nonlinear differences between word order rules as a function of time and sensor space. Moreover, we demonstrated that inter-individual variations in spectral power across the learning task interacted with aperiodic activity to influence subsequent behavioural performance. Together, these results suggest that $1/f$ activity plays an important role in higher-order cognition, including language processing, and that grammar learning is modulated by different word order permutations, which manifest in distinct oscillatory profiles.

Keywords: language learning • neural oscillations • aperiodic activity • sentence processing • $1/f$ activity • EEG

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

1. INTRODUCTION

Understanding the neural mechanisms underlying the encoding and retrieval of new information is a fundamental endeavour in cognitive neuroscience. Memory supports many essential cognitive functions, from learning the distinction between semantic categories (e.g., animal vs. human) to complex (motor) sequences, such as learning how to drive a car or speak a new language. However, while research has revealed a neocortical-hippocampal network subserving the encoding and retrieval of memory (Greenough, 1984; McClelland et al., 1995; Norman & O'Reilly, 2003), considerably less is known about the temporal mechanisms underlying memory formation, such as electrophysiological patterns of neural synchronisation (i.e., excitation) and desynchronization (i.e., inhibition; Fellner et al., 2019; Hanslmayr & Staudigl, 2014; Hanslmayr et al., 2012).

While a broad literature has related neural (de)synchronisation with memory encoding and retrieval of image and word stimuli (Parish, Hanslmayr, & Bowman, 2018), considerably less is known about oscillatory activity during the encoding and retrieval of complex sequences, such as in language (cf. de Diego-Balaguer, Fuentemilla, & Rodriguez-Fornells, 2011; Kepinska, Pereda, Caspers, & Schiller, 2017). Further, the few studies examining sequence and artificial language learning report mixed findings relative to (episodic) word and image paradigms: while alpha/beta desynchronisation predicts encoding of words and images (Griffiths, et al., 2019), alpha/beta and theta synchronisation is associated with sequence (Crivelli-Decker et al., 2018) and language learning (e.g., Kepinska et al., 2017). This apparent inconsistency might be accounted for by stimulus heterogeneity; however, another possible source of divergence may lie in the mixture of oscillatory power with aperiodic activity (Ouyang et al., 2020; Wen & Liu, 2016), which to the best of our knowledge has not been addressed in previous studies involving complex rule learning paradigms. Here, we aim to better characterise the (de)synchronisation patterns underlying the encoding and retrieval of language-related rules. We also investigate whether task-related aperiodic activity interacts with oscillatory power in individualised frequency bands to influence language learning.

1.1. Oscillatory correlates of artificial language learning

Previous research examining spectral dynamics during (artificial) language learning has revealed that increased alpha/beta power (i.e., synchronisation) predicts sensitivity to grammatical violations (Kepinska et al., 2017), while increased gamma phase coherence between frontal, temporal and parietal cortices is associated with successful learning (De Diego-Balaguer et al., 2011). An increase in theta synchronisation is also associated with poorer learning of auditory rules (De Diego-Balaguer et al., 2011).During visual grammar learning, theta synchronisation increases early in the task and declines after prolonged exposure (Kepinska et al., 2017).While these findings are inconsistent with images and word stimuli, studies using complex sequence processing have reported linear increases in beta power for predicable sequences, which coincide with increased theta activity in task-related cortical regions (Crivelli-Decker et al., 2018; Wang et al., 2019), and occipital alpha power (Wang et al., 2019). These data are also in line with the supposed role of active, top-down mechanisms in the processing of incoming sensory input, akin to a predictive-coding account of brain function (Friston, 2010; Kikuchi et al., 2018; Rao & Ballard, 1999). Under this framework,

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

higher cortical areas send predictions to downstream regions via alpha/beta oscillations (Friston, 2019; Friston et al., 2020; Kikuchi, et al., 2018), which are compared to error signals passed from lower levels of the hierarchy. From this perspective, greater alpha/beta synchronization during grammar learning may reflect the encoding of complex structured sequences, while theta synchronization may facilitate the binding of individual elements based on sequence rules; however, this hypothesis has not been tested in the context of artificial language learning.

1.2. Aperiodic activity as a functionally relevant source of neural activity

In addition to oscillatory activity, the brain exhibits a $1/f$-like power distribution, which is often removed from the signal to isolate transient task-related oscillations (Donoghue et al., 2020; Haller et al., 2018; He, 2014; Lendner et al., 2019). However, aperiodic activity has recently been implicated in a variety of higher-order cognitive computations (Fellner et al., 2019), partially explaining individual differences in theta activity during memory encoding and recall performance (Sheehan et al., 2018) and processing speed over and above that of alpha activity (Ouyang et al., 2020). Work on prediction during language has also shown that higher aperiodic, but not oscillatory activity influences the N400 event-related potential and performance accuracy (Dave et al., 2018), attesting to the utility of examining global power spectrum changes alongside narrow-band responses to predict behavioural outcomes.

The slope (or power-law exponent) of the 1/f power distribution has also been shown to vary parametrically by task difficulty: it is at its steepest for easy tasks and becomes progressively shallower with increasing task difficulty (He et al., 2010). In this way, aperiodic activity may serve as a proxy for inter-individual differences in neural "efficiency" during online information processing (He, 2014). For example, in the context of artificial language learning, a shallower $1/f$ slope during grammar acquisition may be associated with higher behavioural performance, explaining behavioural gains over and above that of oscillatory activity – an empirically testable prediction.

Taken together, understanding aperiodic brain activity seems critical if a full understanding of the neurobiology of cognition is to be achieved (He et al., 2010). Further, oscillatory and aperiodic activity appear to play separable functional roles in higher-order cognition, such as memory formation (Sheehan et al., 2018) and linguistic prediction (Dave et al., 2018). However, it is currently unknown whether oscillatory and aperiodic activity interact during the initial stages of complex sequence learning, and whether any such interaction influences behavioural outcomes. Clarifying the (separable) roles of oscillatory and aperiodic components of the electroencephalogram (EEG) power spectrum may also bridge the diverging results reported in studies utilising image and word stimuli and artificial grammar paradigms, lending support to the idea that neuronal oscillations differentially contribute to the formation of memory of varying complexity.

1.3. The present study

To better characterise the neural mechanisms underlying complex rule learning, we examined delta, theta, alpha and beta power during an artificial language learning task. We also studied the interaction between oscillatory and aperiodic activity in order to characterise how

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

patterns of (de)synchronisation and aperiodic fluctuation influence the generalisation of different word order rules characteristic of many natural languages. We used the artificial miniature language Mini Pinyin (Cross, Zou-Williams, Wilkinson, Schlesewsky, & Bornkessel-Schlesewsky, 2020a), which contains two main sentence constructions, namely fixed and flexible word orders. The key distinction between these word orders relates to the cues used for interpretation. Fixed word orders rely heavily on linear order-based processing (e.g., the linear position of each element in the sentence), and thus likely relies on predictive sequence processing for accurate interpretation. Flexible word orders rely more heavily on sequence-independent cues (e.g., animacy) in order to combine non-adjacent elements in a sentence. Importantly, the distinction between fixed and flexible word orders not only mirrors differences between natural languages (Bates et al., 2001; Bornkessel-Schlesewsky et al., 2011; MacWhinney et al., 1984; Muralikrishnan et al., 2015), but also basic and dissociable components of the neurobiology of human language (e.g., Bornkessel-Schlesewsky et al., 2015), including the capability of the brain to generalise the meaning of ordinal sequences (the *when* in fixed word order rules) and establishing semantic schemas of unordered items (the *what* in flexible word order rules; Cross et al., 2018a; Friston & Buzsaki, 2016).

We implemented generalised additive and linear mixed-effects analyses to model dynamic changes in oscillatory and aperiodic activity during the learning of fixed and flexible word order rules, and subsequent performance on a judgement task. We hypothesised that: (1) alpha/beta synchronisation would increase across the learning task, with stronger synchronisation predicting improved behavioural performance; (2) higher theta synchronisation over the duration of the learning task would be associated with improved behavioural performance on the judgement task, and; (3) variations in oscillatory activity would differentially predict sensitivity to fixed and flexible word orders. We also explored aperiodic effects on learning-related oscillatory dynamics, given that recent studies suggest aperiodic activity plays a functional role in memory formation and language processing.

2. METHOD

2.1. Participants

Data from 34 right-handed healthy, monolingual, native English-speakers (mean age = 25.3, SD = 7.13; 17 female) were analysed from a study examining the role of sleep in language learning (for a detailed description of sample demographics, see Cross et al., 2020b). Ethics approval was granted by the University of South Australia's Human Research Ethics committee (I.D: 0000032556).

2.2. Stimuli and design

Stimuli were based on the modified miniature language Mini Pinyin (for a detailed description of the language, see Cross et al., 2020a; see also Cross et al., 2020b), which contains grammatical rules present in a number of natural languages (see Figure 1A and 1B for example sentence constructions and vocabulary items). Briefly, each sentence in Mini Pinyin contains two noun phrases and a verb phrase, while each noun is associated with a different classifier:

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

human nouns are preceded by *ge*, while animals, and small and large objects are preceded by *zhi*, *xi* and *da*, respectively.

Mini Pinyin includes two main sentence types based on whether the sentence contains the coverb *ba*. When *ba* is present, the sentence contains a fixed word order, in that the first noun phrase is invariably the Actor (the active, controlling participant) and the sentence must be verb-final. From this perspective, accurate sentence processing is dependent on the linear position of the words. When *ba* is not present, the sentence contains a flexible word order, in that the first noun phrase can either be the Actor or the Undergoer (the affected participant); however, the sentence must be verb-medial. As such, accurate sentence interpretation is based more heavily on the animacy status of the noun phrases rather than word order.

The experiment contained three phases, involving a vocabulary test, a learning phase, and a grammaticality judgement task. Approximately 7 days before the sentence learning phase, participants received a paired picture-word vocabulary booklet containing the 25 nouns. Participants were required to learn the 25 nouns to ensure that they had a basic vocabulary. Prior to the learning and judgement tasks, participants completed the vocabulary test on a desktop computer by typing in translations of the nouns from Mini Pinyin to English. Only prospective participants who attained a score > 90% were eligible to move onto the learning and judgement task phases of the experiment.

Three grammatical sentence structures were included in the learning phase. While no explicit instruction was given to participants, a picture was shown prior to each sentence illustrating events occurring between two entities, which was then described in the subsequently presented sentence. The learning task contained four blocks with 96 grammatical picture-sentence pairs that were presented via rapid visual serial presentation (RSVP). A further 156 novel sentences (50% grammatical, 50% ungrammatical) were presented during the judgement task which occurred immediately after the learning phase. The ungrammatical sentences violated either the position of the Actor or verb in fixed word order sentences (e.g., Actor-ba-Verb-Undergoer [AbaVU] instead of AbaUV) or the position of the verb in flexible word order sentences (e.g., AUV instead of AVU; see Figure 1A for an illustration and full list of ungrammatical constructions).

During the learning phase, each picture was presented for 4000ms, while each corresponding sentence was presented on a word-by-word basis, with each word presented for 700ms with an inter-stimulus interval (ISI) of 200ms. Across four blocks, each grammatical construction was presented 32 times, with stimuli pseudo-randomised, such that no sentence of the same construction followed each other. During the judgement task, novel grammatical and ungrammatical sentences were presented word-by-word with a presentation time of 600ms and an ISI of 200ms. Participants responded via a button press based on whether the sentence followed the rules of Mini Pinyin. The assignment of grammatical/ungrammatical response buttons was counterbalanced across participants. Response time windows were presented for a maximum of 4000ms. Each participant received feedback on whether their response was correct or incorrect (see Figure 1C and 1D for a schematic of the learning and judgement tasks,

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

respectively). Both the learning and judgement tasks were created in OpenSesame (Mathot et al., 2012) and performed on a desktop computer.

(A) Grammatical and Ungrammatical Sentence Constructions

Grammatical:

AVU: ge shuishou zhuole zhi maomi
UVA: zhi maomi zhuole ge shuishou
AbaUV: ge shuishou ba zhi maomi zhuole

English Translation: the sailor captures the cat

Ungrammatical:

UbaAV: zhi maomi ba ge shuishou zhuole
AbaVU: ge shuishou ba zhuole zhi maomi
AUV: ge shuishou zhi maomi zhuole
UAV: zhi maomi ge shuishou zhuole

(B) Sample of Linguistic Elements from Mini Pinyin and English Translations

Classifier	:	ge (human), zhi (animal), da (large object), xi (small object)
Noun	:	shuishou (sailor), maomi (cat), junma (pirate), pingguo (apple)
Coverb	:	ba (actor-undergoer-verb)
Verb	:	zhoule (capture), xile (wash), zhaole (photograph), chile (eat)

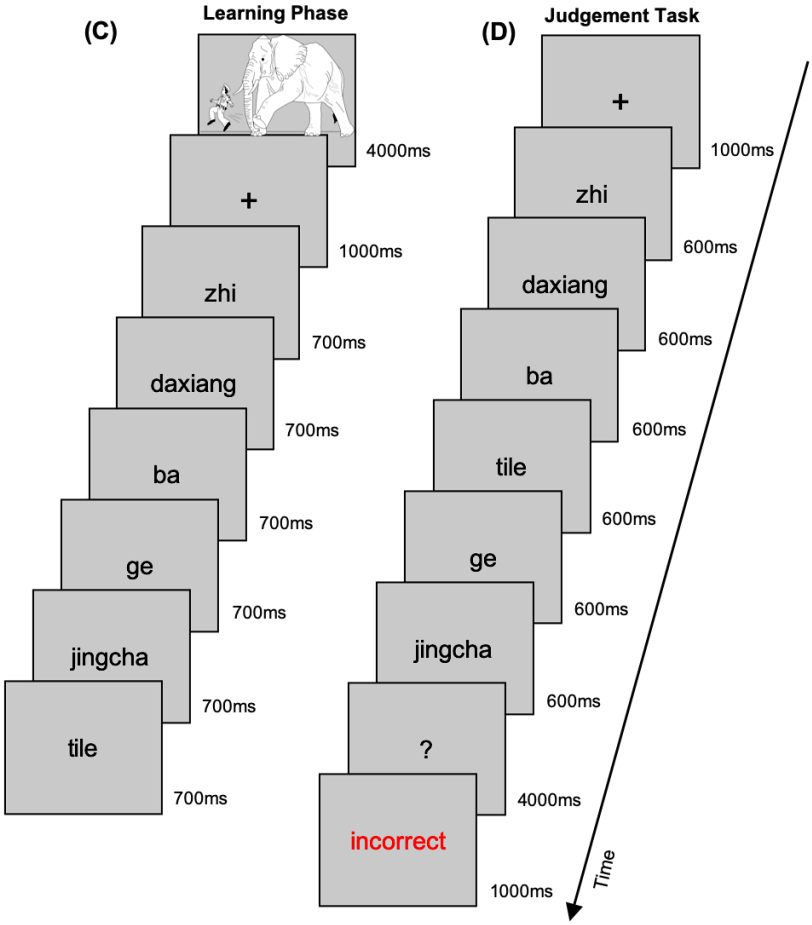

Figure 1: Summary of Mini Pinyin. (A) Summary of the grammatical (left) and ungrammatical (right) sentence constructions. (B) A portion of linguistic elements used in the sentence examples provided in (A).

2.3. EEG recording and pre-processing

The EEG was recorded during the learning task using a 32-channel BrainCap with sintered Ag/AgCI electrodes (Brain Products, GmbH, Gilching, Germany) mounted according to the extended International 10-20 system. The online reference was located at FCz. The ground electrode was located at AFz. The electrooculogram (EOG) was recorded via electrodes

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

located at the outer canthus of each eye and above and below participants' left eye. The EEG was amplified using a BrainAmp DC amplifier (Brain Products GmbH, Gilching, Germany) using an initial band-pass filter of DC – 250 Hz with a sampling rate of 1000 Hz. Electrode impedances were kept below 10kΩ. EEG was also recorded during 2 minutes of eyes-open and 2 minutes of eyes-closed resting state periods immediately before the learning task and after the judgement task.

EEG analysis was performed in MATLAB 2017b (The MathWorks, Natick, USA) using custom scripts in conjunction with the Fieldtrip toolbox (Oostenveld et al., 2011). EEG data were re-referenced offline to the average of both mastoids and band-pass filtered from 1 – 40 Hz. Data were then epoched from -200ms to 13s relative to the onset of each picture-sentence pair for both fixed and flexible sentences and corrected for ocular artefacts using Infomax Independent Component Analysis (Bell & Sejnowski, 1995; implemented in runica.m). Components demonstrating clear EOG artefacts were removed and electrodes showing strong impedance were interpolated with surrounding electrodes.

2.4. EEG data analysis

The aim of the analysis was to characterise the oscillatory and aperiodic mechanisms underlying the initial encoding of complex grammatical rules. To this end, we computed five spectral features from the EEG during the learning phase: mean power density within the delta, theta, alpha, and beta bands, and the slope of the $1/f$ spectral distribution (i.e. power-law exponent). These metrics were used to predict whether trial-level variation during the learning task predicted behavioural performance on the judgement task. We also tested whether interactions between oscillatory and aperiodic activity afford unique information predicting behavioural performance.

2.4.1. Spectral decomposition and power-law exponent χ estimation

The power-law scaling exponent χ, which summarises the rate of decay of the power spectrum in double-logarithmic co-ordinates, was estimated using the Irregular-Resampling Auto-Spectral Analysis toolbox (IRASA v1.0; Wen & Liu, 2015). Briefly, this technique seeks to separate oscillatory from aperiodic (random fractal) components by iteratively resampling the power spectrum across a range of non-integer factors h and their reciprocals $1/h$. This procedure shifts any narrowband components away from their original location along the frequency spectrum while leaving the distribution of the fractal component intact. The median of the resampled spectral estimates is then calculated in order to strip the spectrum of any redistributed narrowband peaks. For a more detailed treatment of the IRASA method, see Wen and Liu (2016).

Channel data from the 7-13 s time window of each trial epoch (i.e., the trial period corresponding to word presentation) were divided into ten 2 s segments using a sliding window (400 ms step width) and passed to amri_sig_fractal.m. They were Hanning-windowed, detrended, and anti-alias filtered per default settings, with the output frequency range set to 1 – 40 Hz. Following Muthukumaraswamy and Liley (2018), an extended set of resampling

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

factors (h = 1.1 to 2.9 in steps of 0.05, $h \neq 2$) was used to reduce interference from fractal-oscillatory interactions.

Once the fractal component had been recovered from the power spectrum, the power-law exponent χ was estimated using amri_sig_plawfit.m. This function fits a linear regression to a region of the double-log transformed fractal spectrum (here, 1 – 35 Hz) after rescaling the frequency spectrum to achieve equally-spaced intervals in log-space. The negative of the regression slope coefficient was taken as the χ exponent.

2.4.2. Spectral band power estimation

In order to quantify narrowband changes in spectral power independent of underlying changes in aperiodic activity, power estimates were derived from spectra following fractal component subtraction. This residual, 'oscillatory spectrum' was averaged across epoch segments, half-wave rectified (negative values set to zero), and divided into the four frequency bands of interest. Notably, the limits of each frequency band were adapted for each participant on the basis of their resting-state EEG. Specifically, band limits for each of the four frequency bands of interest were calculated according to the harmonic frequency architecture proposed by Klimesch (2012; 2013), in which the centre frequency of each successive band constitutes a harmonic series scaled in relation to individual alpha frequency (IAF). To avoid the potential overlap of neighbouring frequency bands, we determined lower and upper frequency bounds using the following formulae:

$$f_1 = f_c - f_c/4,$$

$$f_2 = f_c + f_c/2,$$

where f_c is the centre frequency, f_1 the lower bound, and f_2 the higher bound of a given frequency band.

IAF estimates used to derive f_c were obtained from a set of parieto-occipital electrodes (P3/P4/O1/O2/P7/P8/Pz/Iz) using the restingIAF package (v1.0.3; Corcoran et al., 2019; see also Cross et al. 2018b). This method applies a Savitzky-Golay filter (frame width = 11 bins, polynomial order = 5) to smooth and differentiate the power spectrum prior to estimating a weighted average of the spectral peak frequencies identified across channels within a specified frequency range (here, 7—14 Hz). A minimum of 3 channel estimates was required to return an IAF for a given recording. Estimates derived from pre- and post-session eyes-closed resting states were then averaged for each participant using meanIAF.m. For further details on this algorithm, see Corcoran and colleagues (2018).

Having determined the individualised bounds of the delta, theta, alpha, and beta bands according to the IAF, power within each of these bands was quantified using the mean power density metric proposed by Westfall (1990):

$$P_k = \frac{1}{1 + k_{f_2} - k_{f_1}} \sum_{i=k_{f_1}}^{k_{f_2}} p(f_i)\,,$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

where $p(f_i)$ is the power estimate of the ith frequency bin, and f_1 and f_2 index the lower and upper bounds of the individualised frequency band k, respectively. An advantage of this approach is that power estimates are scaled by spectral range, thus controlling for differing frequency bandwidths both within and between individuals. Note that this formula was applied to power estimates that had been normalised to the mean of the oscillatory spectrum.

2.5. Statistical analysis

We used *R* v.3.6.2 (R Core Team, 2019) and the packages lme4 v.1.1.21 (Bates et al., 2015), lmerTest v.3.0.1 (Kuznetsova et al., 2017), effects v.4.0.3 (Fox et al., 2019), car v.3.0.2 (Fox et al., 2011), tidyverse v.1.3.0 (Wickham et al., 2019), lattice v.0.20-38 (Sarkar, 2008), itsadug v.2.3 (van Rij et al., 2016), mgcv v.1.8-31 (Wood, 2006), mgcViz v.0.1.4 (Fasiolo et al., 2019) and rgl v.0.1.3 (Nenadic & Greenacre, 2007). Raincloud plots were produced to visualise behavioural data using the code provided by (Allen et al., 2019). For linear models, contrasts for categorical variables were sum-to-zero contrast coded, with coefficients reflecting differences to the grand mean (Schad et al., 2020).

2.5.1. Generalised additive mixed models (GAMMs)

Generalized additive models (GAMs) are a nonparametric extension of the standard linear regression model that substitute a linear predictor variable x with a smooth function $f(x)$ (Hastie & Tibshirani, 1987, 1990; Wood, 2017). Generalized additive mixed models (GAMMs; Lin & Zhang, 1999) constitute a further extension that incorporates random effects components within the GAM framework (Pedersen et al., 2019; Wood, 2013). Together, these innovations offer an elegant solution to the problem of autocorrelation amongst residuals induced by (1) attempting to fit linear models to non-linear relationships, and (2) non-independence (or nesting) of observations (e.g. repeated measures within subjects or items; (Baayen et al., 2008).

Here, GAMMs were constructed to investigate how the power-law exponent χ, and the mean power density P for each kth frequency band delta, theta, alpha, and beta, fluctuate over (experimental) time and (sensor) space during artificial grammar learning. Trial-level χ and P_k estimates were modelled as a function of experimental time (trial number), sensor space (2D Cartesian co-ordinates), and sentence type (Fixed vs. Flexible). Random factor smooth interactions were included to account for individual differences in the functional relationship between spectral features and time (see Baayen et al., 2017, for a similar approach). Each GAMM took the following general form:

$$g\big(\mathbb{E}(Y_i)\big) = \beta_0 + \beta_1 type_i + f(trial_i, topo.x_i, topo.y_i, \text{by} = type_i) +$$

$$f_{subject_i}(trial_i) + \epsilon, \quad \epsilon_i \sim (0, \Lambda\sigma^2)\,,$$

where $\mathbb{E}(Y_t)$ is the expected value of the ith observation of spectral feature Y, g is the link function, β_0 is the model intercept, $\beta_1 type$ is the main effect of sentence type, $f(\,.\,,\text{by} = type)$ is the tensor product interaction between predictors for each level of the Type factor, $f_{subject}$ is the by-subject factor smooth on Trial, and $\mathbf{\Lambda}$ is a correlation matrix modelling the autocorrelation of residuals ε. Note that marginal smooths for topographic x and y co-ordinates were treated as isotropic (i.e. assumed to share a common scale).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

GAMMs were estimated using the bam() function of the *R* package mgcv (v.1.8-31; Wood, 2011). Models were fit using the Fast REML method. Since power estimates were bounded at 0 and showed heavy positive-skewness, all power models were initially fit with a log-linked Tweedie distribution (Wood, Pya, & Säfken, 2016). Model residuals were then evaluated for evidence of additional nonlinearity/autocorrelation structure, and re-fit with an increased basis dimension/first-order autoregressive process (AR(1)) as required.[1]

Each finalised model was fit with tensor product interaction smooths in order to enable ANOVA-decomposition of main effect and interaction components (Wood, Scheipl, & Faraway, 2013). All tensor product smooths were fit using low rank thin plate regression splines as their basis function (Wood, 2003, 2017). Factor smooths were fit with a first-derivative penalty in order to shrink subject-level smooths towards the population-level. An additional shrinkage penalty was imposed on the smoothing penalty null space to enable automated model reduction (see Marra & Wood, 2011). As Type was entered as an ordered factor with Fixed assigned as the reference level, model terms involving a Sentence Type interaction assess the difference between Fixed and Flexible condition splines (see van Rij et al., 2016).

2.5.2 Linear mixed-effects models

Linear mixed effects models (LMM) are extensions of the traditional linear model, but include a combination of fixed and random effects as predictor variables (Harrison et al., 2018). The inclusion of random effects structures (e.g., by-participant and by-item) enable the modelling of inter- and intra-individual variance, particularly for subjects, items and time in repeated-measures designs (Meteyard & Davies, 2020). Here, fixed effects included Delta, Theta, Alpha and Beta power, Sentence Type (Fixed, Flexible), χ-exponent, topographical factors Laterality and Sagittality and their interaction. Subject-level factor smooths from corresponding GAMMs were also included as a fixed effect in the LMM to investigate whether individual differences in spectral feature dynamics over the duration of the learning task predict performance for Fixed and Flexible sentences during the judgement task. Proportion correct on the judgement task was specified as the dependent variable, while intercepts were grouped by participant. More complex random effect structures involving random slopes by participant did not converge. Type II Wald χ^2-tests were used to provide *p*-value estimates, while an 83% confidence interval (CI) threshold was adopted, which corresponds to the 5% significance level with non-overlapping estimates (Austin & Hux, 2002; MacGregor-Fors & Payton, 2013).

3. RESULTS

3.1. Task performance

The results on the judgement task are visualised in Figure 2 using raincloud plots (Allen et al., 2019). Participants performed moderately on the judgement task, with a mean accuracy of 62.28% (range: 8.33 – 100%) and mean reaction time of 878.08 ms (range: 254.75 – 2076.83 ms). Specifically, there is clearly a high degree of inter-individual variability across both fixed and flexible sentences; however, fixed grammatical sentences had a higher degree of accuracy

[1] Note that inclusion of the AR(1) model required the assumption of Gaussian-distributed random errors.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and faster responses overall. For a detailed report and interpretation of these behavioural data, see Cross et al. (2020b).

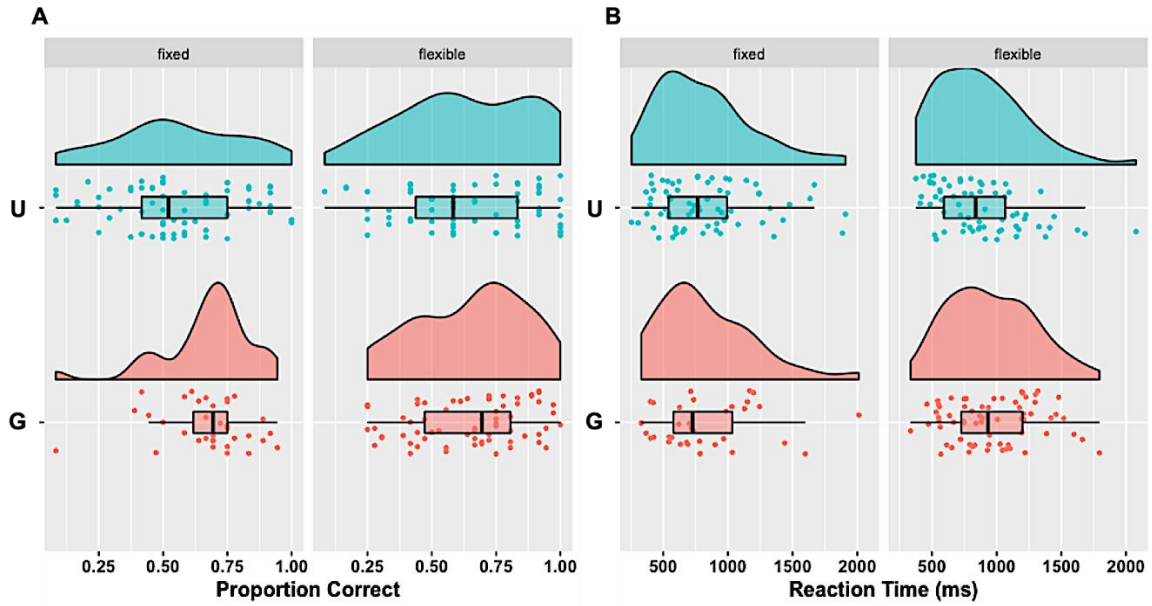

Figure 2: Behavioural responses during the sentence judgement task. (A) Mean proportion of correct responses (x-axis) for Grammatical (G) and Ungrammatical (U) Fixed (left) and Flexible (right) sentence types. (B) Mean reaction time (ms; x-axis) for Grammatical (G) and Ungrammatical (U) Fixed (left) and Flexible (right) sentence types. Individual data points represent the mean for each participant.

3.2. Neurophysiological results

3.2.1. Aperiodic and oscillatory changes across time and space during language learning

Here, we examine how aperiodic and oscillatory activity changes across the learning task (for full model summaries, see the supplementary material). Average change in activity across learning task trials is illustrated in Figure 3.

The χ-exponent summarising the slope of the aperiodic spectral component was on average 0.01 n.u. lower (i.e., shallower) in the Flexible compared to Fixed word order condition (se = 0.002, t = -4.82, p < .001). The χ-exponent also showed significant nonlinear differences between Fixed and Flexible word order rules over time and space. The mean χ-exponent decreased almost linearly for the first third of Fixed rule trials, increased slightly over the middle third, and declined more steeply over the final third of trials (edf = 2.803, F = 3.18, p < .001). A similar qualitative pattern was observed during Flexible rule learning; however, changes in slope were more pronounced, differing significantly from the Fixed condition (edf = 3.512, F = 6.62, p < .001). These effects were further modulated by topography, with the χ-exponent higher over centro-posterior regions for early Fixed trials (edf = 3.041, F = 1.49, p < .001), becoming more laterally pronounced during later trials (edf = 8.532, F = 2.11, p < .001). Time by laterality effects were enhanced during Flexible trials (edf = 2.466, F = 1.05, p < .001), while posterior sensor activity showed a U-shaped relative difference relative to Fixed trials (edf = 2.029, F =1.64, p <.001).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Mean delta power density was on average 0.03 n.u. higher during Flexible word order processing compared to Fixed word order processing during the learning task (se = 0.006, t = 4.88, $p < .001$). Mean power density also showed significant nonlinear differences between sentence types as a function of trial time and sensor space. While mean delta power tended to decline almost linearly during Fixed trials (edf = 1.995, F = 2.08, $p < .001$), a more complex pattern was evident during Flexible trials (edf = 3.731, F = 11.44, $p < .001$). These dynamics were further modulated by topography, with mean delta power density lower over lateral regions at the beginning of the session, but higher towards the end of the session, for Flexible relative to Fixed word order rules (edf = 2.243, F = 0.74, $p < .001$). Mean delta power over posterior sensors showed a similar U-shaped pattern of difference during the Flexible word order trials as that observed for the χ-exponent (edf = 2.867, F = 2.681, $p < .001$).

Mean theta power density during the learning task was on average 0.06 n.u. lower during Flexible word order processing compared to Fixed word order processing (se = 0.007, t = -7.50, $p < .001$). Further, mean theta power decreased nonlinearly across the learning task for Fixed word order rules (edf = 3.100, F = 4.03, $p < .001$), showing a steep decrease over the first quarter of trials, increasing gradually over the second quarter, and then steadily decreasing until the end of the task. Flexible word order rules evoked a similar pattern of theta band activation, with a significantly more pronounced nonlinear effect (edf = 1.866, F = 1.51, p = .03). Mean theta density was further modulated by topography, showing a decline in power density over centro-anterior sensors, and an increase in density over posterior sensors, during the middle period of the learning phase (edf = 5.543, F = 0.38, $p < .001$; no difference between word-order conditions; p = .239).

For alpha activity, mean power density was on average 0.14 n.u. lower during Flexible word order processing compared to Fixed word order processing (se = 0.022, t = -6.10, $p <$ .001). Mean alpha power also showed a significant nonlinear effects of trial time over Fixed and Flexible word order rule learning: alpha power increased steeply over the first third of Fixed word order trials, levelled off over the middle third, and increased again towards the end of the learning task (edf = 3.442, F = 9.96, $p < .001$); a more complex pattern of activity was observed for Flexible word order trials (edf = 2.738, F = 4.31, $p < .001$). This effect was also modulated by topography: alpha activity was significantly higher over central sensors in the initial stages of Fixed word order learning, becoming more pronounced over lateral sensors during the middle of the session (edf = 5.405, F = 0.89, p = .003); while additional laterality differences towards the beginning and end of the learning phase were noted for Flexible word order trials (edf = 1.282, F = 0.40, p = .005). Alpha power density further varied as a function of sagittality over trial time (edf = 6.212, F = 1.62, $p < .001$), although this effect did not significantly differ between Fixed and Flexible word order conditions (p = .174).

Finally, mean beta power density was on average 0.09 n.u. higher during Flexible word order compared to Fixed word order trials (se = 0.012, t = 7.61, $p < .001$). Mean beta power density showed significant nonlinear changes over time for both Fixed and Flexible word order rule learning: beta power decreased over the first third of Fixed word order trials, steadily increasing thereafter (edf = 1.182, F = 0.74, $p < .001$). Beta power showed the opposite pattern during the first two-thirds of Flexible word order trials, before increasing at a faster rate

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

towards the end of the session (edf = 3.503, F = 8.97, p < .001). These dynamics were further modulated by topography, with mean beta power density higher over lateral (edf = 2.177, F = 0.68, p = .001) and posterior (edf = 5.854, F = 2.43, p < .001) sensors during early Fixed word order trials, and decreasing towards the end of the session. Significant deviations from this pattern were also noted during Flexible word order trials for laterality (edf = 7.556, F = 1.88, p < .001) and sagittality (edf = 4.253, F = 1.16, p < .001).

Together, these data illustrate a complex relationship between spectral activity in the range of 1 – 35 Hz during grammar learning (see Figure 4 for a visualisation of mean spectral power changes over stimulus presentation). Low frequency activity (i.e., < ~8 Hz) shares a similar spatiotemporal pattern to aperiodic activity, where mean power density tends to decrease over the course of grammatical learning. Higher frequency activity (i.e. ~8-35 Hz), by contrast, tends to increase over time, with greater variability in the temporal dynamics associated with Fixed and Flexible word order learning observed in the beta band.

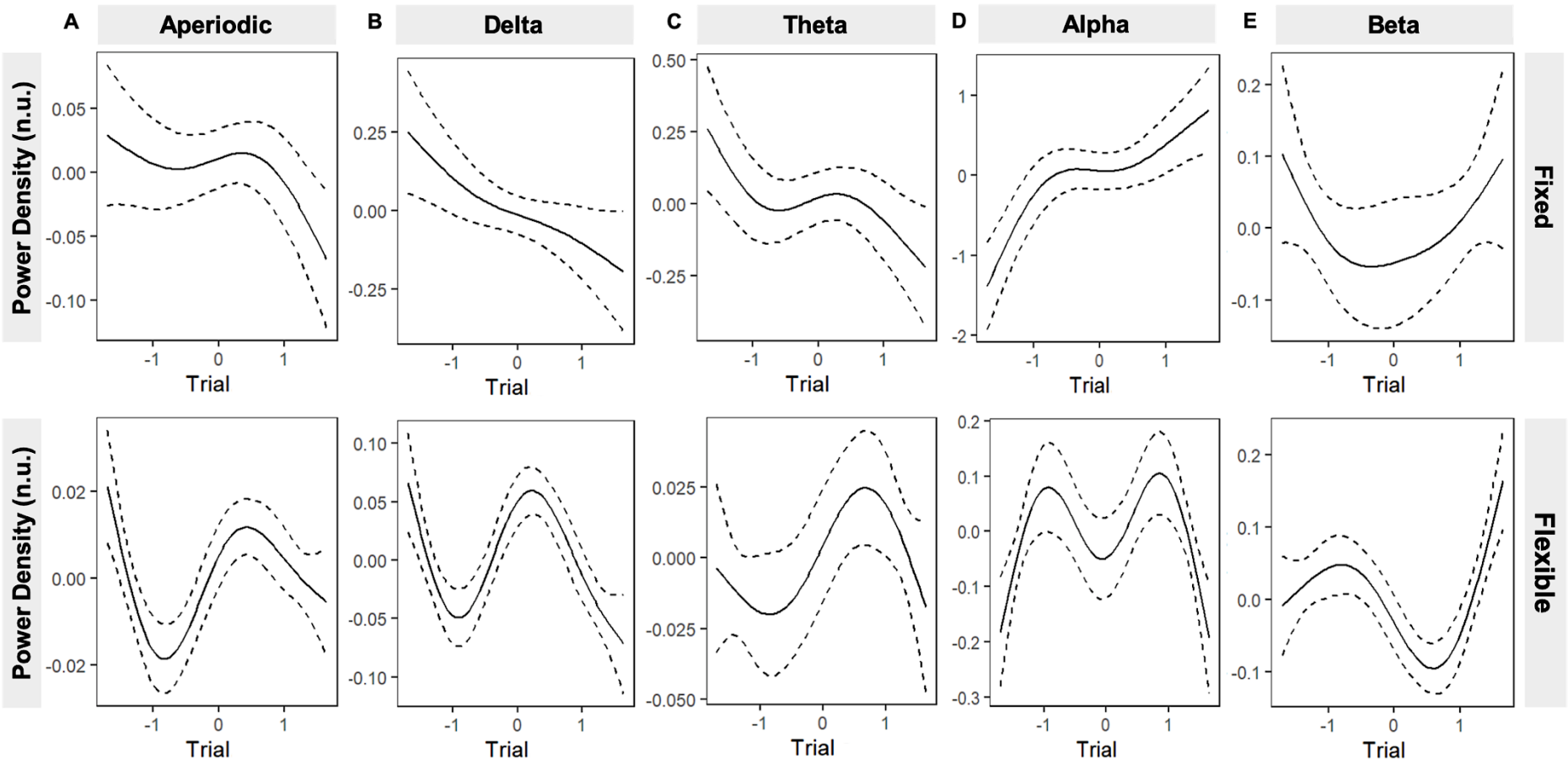

Figure 3. Modelled effects for changes in aperiodic (A), delta (B), theta (C), alpha (D) and beta (E) activity across the learning task for Fixed (top row) and Flexible (bottom row) word order rules. Note that the bottom row displays *difference* waves, where a horizontal line at 0 would entail no deviation from the Fixed waveform.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

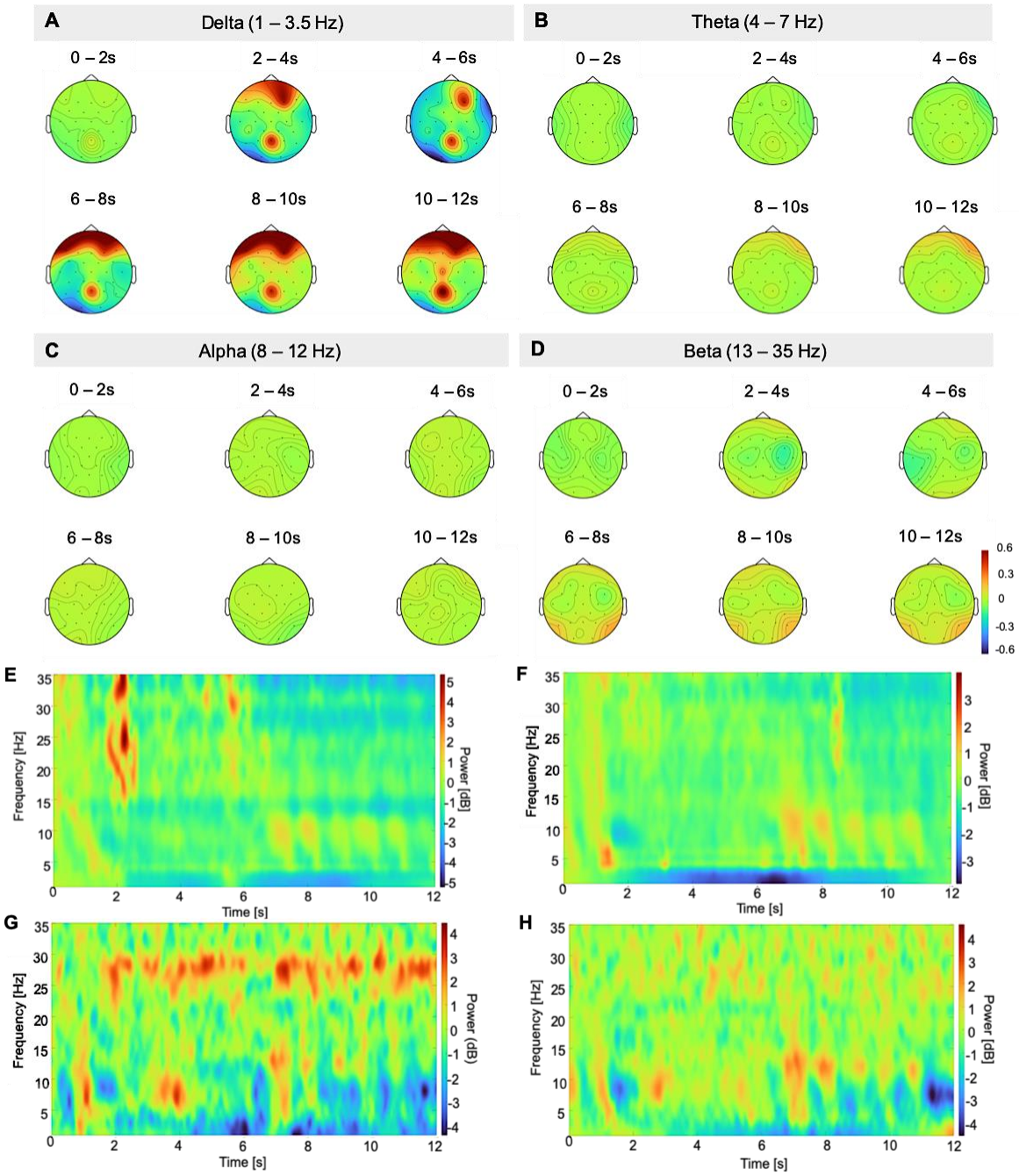

Figure 4: Power spectral modulations as a function of stimulus presentation averaged across the learning task. Each topographical plot represents differences in grand mean delta (A), theta (B), alpha (C) and beta (D) spectral power density between fixed and flexible word order sentences. Note that 0 – 6 s is during the presentation of the picture, while 6 – 12 s is during the presentation of the grammatical sentence presented via RSVP. (E) represents the grand event-related spectral perturbation (ERSP) based on the mean spectral power density values across participants and electrodes for fixed word order sentences, while (F) represents the same for flexible word order sentences. (G) and (H) represent single participant ERSP plots across all channels for fixed and flexible sentences, respectively. Time is represented on the x-axis, spanning the presentation of the picture (2 – 6 s) and the grammatical sentence presented via RSVP (6 – 12 s).

3.2.2. Learning-related oscillatory correlates of behavioural performance

Here, we examine whether power spectral density estimates in the individualised bands during the learning task predict behavioural performance on the sentence judgement task (for full model summaries, see the supplementary material). The first model focussed on activity in

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

the delta band. There was a significant Delta x Sagittality x Sentence Type interaction ($\chi2(2) = 39.36$, $p < .001$). As is illustrated in Figure 5A, increased delta power predicted an increase in the proportion of correct responses for Flexible word orders, while the reverse was observed for Fixed word orders. All other three- and four-way interactions were nonsignificant. By contrast, the theta model demonstrated a significant Theta x Sagittality x Laterality x Sentence Type interaction ($\chi2(4) = 13.10$, $p = .01$): an increase in theta power in posterior regions predicted an increase in the proportion of correct responses for Flexible word orders, while predicting a decrease in performance for Fixed word orders. This effect was most pronounced at midline posterior regions, weakening in central regions and disappearing in central- and right-anterior regions.

The alpha model yielded a significant Alpha x Sagittality x Sentence Type interaction ($\chi2(2) = 167.93$, $p < .001$). As is clear from Figure 5C, an increase in anterior alpha power predicted an increase in the proportion of correct responses for Flexible, but not Fixed, word orders. All other three- and four-way interactions were nonsignificant. Finally, the beta model demonstrated a significant Beta x Sagittality x Laterality x Sentence Type interaction ($\chi2(2) = 77.95$, $p < .001$). An increase in beta power was associated with higher proportion of correct response for Fixed word order sentences but was associated with a decrease in performance for Flexible word orders. While this effect was similar across all regions, the effect was most pronounced over anterior midline channels.

Together these results suggest that an increase in posterior low-frequency activity (i.e., < 8 Hz) during grammar learning is associated with an increase in accurate responses for flexible word order sentences, while an increase in anterior alpha power showed a similar behavioural effect. Higher frequency activity showed a more complex pattern, with an increase in anterior beta power during grammar learning predicting an increase in accurate responses for fixed word order sentences, but a decrease in performance for flexible word order sentences.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

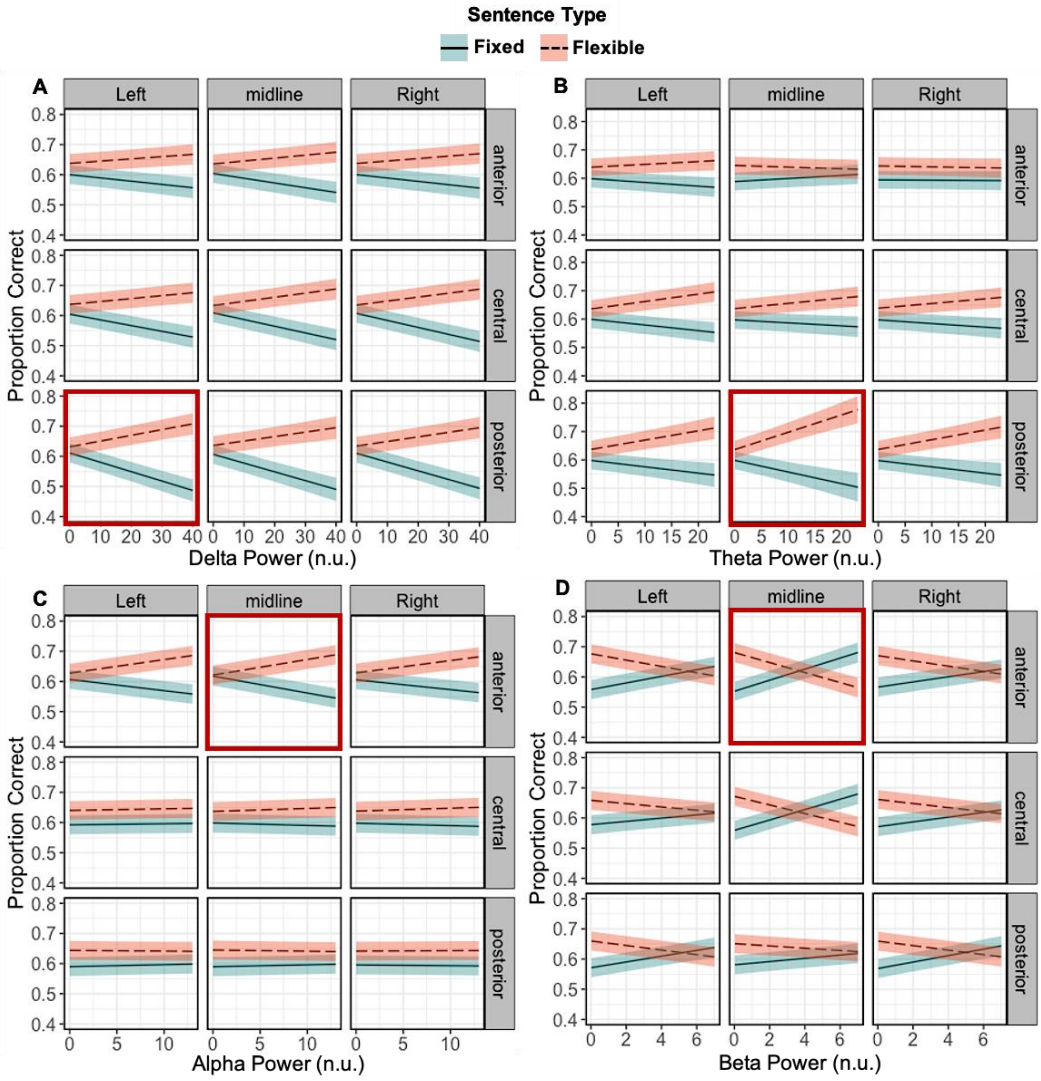

Figure 5: Estimated marginal means for proportion of correct responses (y-axes) by power in the delta, theta, alpha and beta bands (x-axes), sentence type (fixed = blue solid line, flexible = dashed orange line), sagittality and laterality. Shaded regions indicate the 83% CI. The red highlight indicates ROIs showing the largest effect between fixed and flexible word orders.

3.2.3. Interactions between oscillatory and aperiodic activity influences behaviour

Next, we tested whether $1/f$ slope and oscillatory activity interact during learning to influence performance on the judgement task. While the Delta x χ-exponent x Sentence Type interaction was nonsignificant ($\chi2(1) = 1.83$, $p = .17$), activity in the Theta ($\chi2(1) = 53.21$, $p < .001$), Alpha ($\chi2(1) = 291.00$, $p < .001$) and Beta ($\chi2(1) = 77.58$, $p < .001$) bands interacted significantly with χ-exponent and Sentence Type. As is clear from Figure 6, increased theta power and decreased χ-exponent predicted performance for fixed word order sentences, and this effect reversed with high χ-exponent. By contrast, decreased alpha power and low χ-exponent predicted an increase in performance for fixed word order sentences, while the reverse was observed for flexible word orders. Finally, an increase in beta power and low χ-exponent positively predicted fixed, but not flexible word order performance. This effect diminished as the χ-exponent increased. Together, these results suggest that learning-related oscillatory and aperiodic activity selectively influence the learning and processing of different word orders rules.

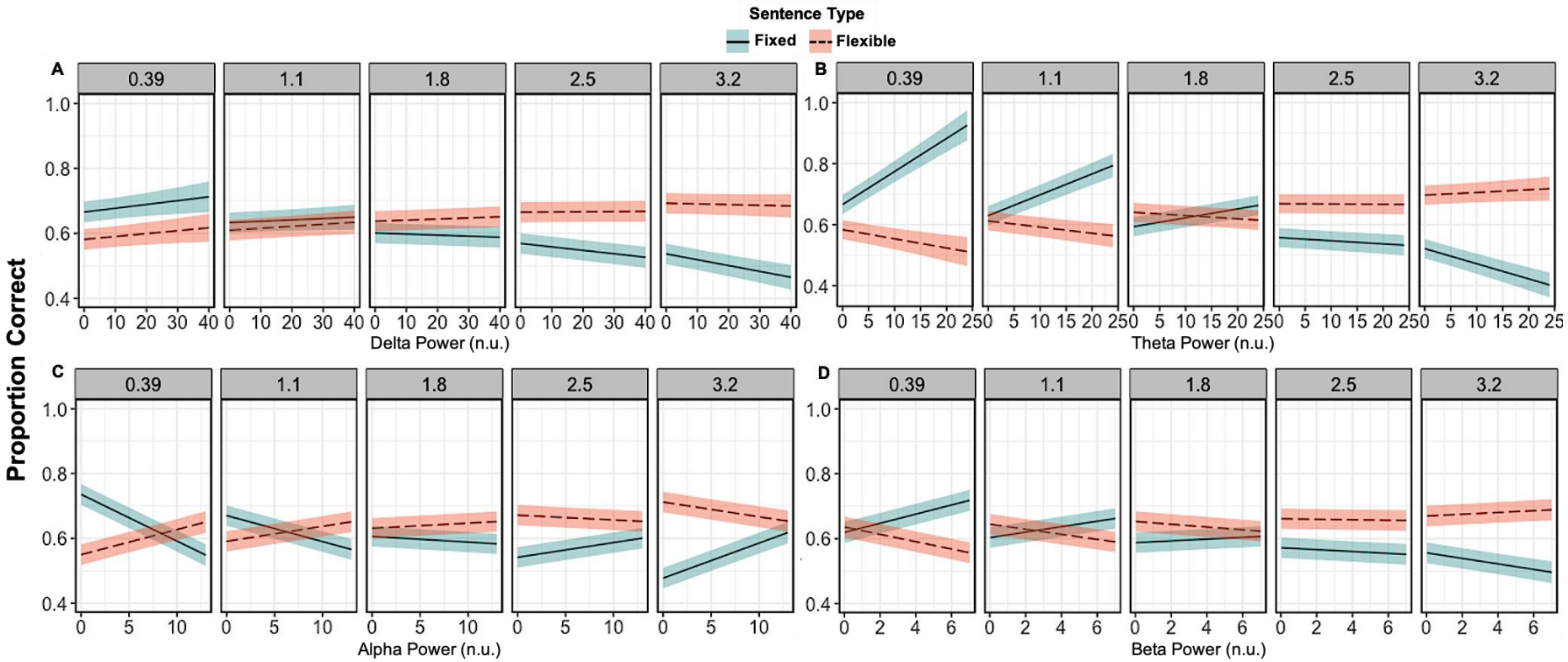

Figure 6: Estimated marginal means for proportion of correct responses (y-axes) by global spectral power in the delta, theta, alpha and beta bands (x-axes, $1/f$ slope (faceted from low to high) and sentence type (fixed = blue solid line, flexible = dashed orange line). Shaded regions indicate 83% confidence intervals.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

3.2.4. Individual variability in neural activity across learning influences sentence processing

Here, we examined whether inter-individual differences in spectral power variations across the learning task predict performance on the judgement task. In line with our aim, we focus on interactions between variations in (spectral) Power, Trial and Sentence Type. Variations in spectral power are random effect estimates of change in spectral activity for each subject from the GAMM models, which reflect intra-individual differences in power estimates across time. For each frequency band, Wald tests (see Table 1) revealed no significant interactions with Sentence Type and Trial. However, as is apparent from Table 1, each frequency band yielded a significant interaction with Sentence Type. Figure 7 shows that this effect is clearest for the delta and beta bands. An increase in delta power was associated with higher performance for flexible word order rules, while lower delta power predicted an increase in performance for fixed word order sentences. This (de)synchronisation pattern was inverted for the beta band: while an increase in beta power predicted higher performance for fixed word order rules, a reduction in power was associated with improved performance for flexible word order rules. The theta and alpha bands, by contrast, show the weakest differentiation between fixed and flexible word order processing, which changed over the course of the experiment, as shown in Figure 7C and 7D, respectively. Finally, aperiodic activity demonstrated a similar pattern to delta, albeit stronger, with a lower χ-exponent (i.e., shallower $1/f$ slope) predicting improved performance for fixed word order sentences, but worse performance for flexible word orders. This effect is also opposite to the influence of beta power on fixed and flexible word orders.

Table 1: Type II Wald χ^2 tests for interaction terms examining the relationship between sentence type, trial and variation in spectral power on proportion of correct responses.

	Delta		**Theta**		**Alpha**		**Beta**		**1/*f***	
Parameter	χ^2	*p*	χ^2	*p*	χ^2	*p*	χ^2	*p*	χ^2	*p*
Type	240.88	**<.001**	239.77	**<.001**	239.98	**<.001**	242.45	<.001	243.04	**<.001**
Power	.01	.89	.00	.98	0.02	.86	.06	.79	.06	.80
Trial	.00	.95	.00	.99	.00	.96	.02	.87	.00	.92
Type x Power	75.50	**<.001**	6.52	**.01**	22.09	**<.001**	176.31	<.001	209.31	**<.001**
Type x Trial	.29	.58	.06	.80	.18	.67	.81	.36	.60	.43
Power x Trial	.00	.92	.00	.98	.00	.98	.00	.98	.00	.95
Type x Power x Trial	3.02	.08	2.37	.12	.57	.45	.31	.57	3.97	**.04**

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

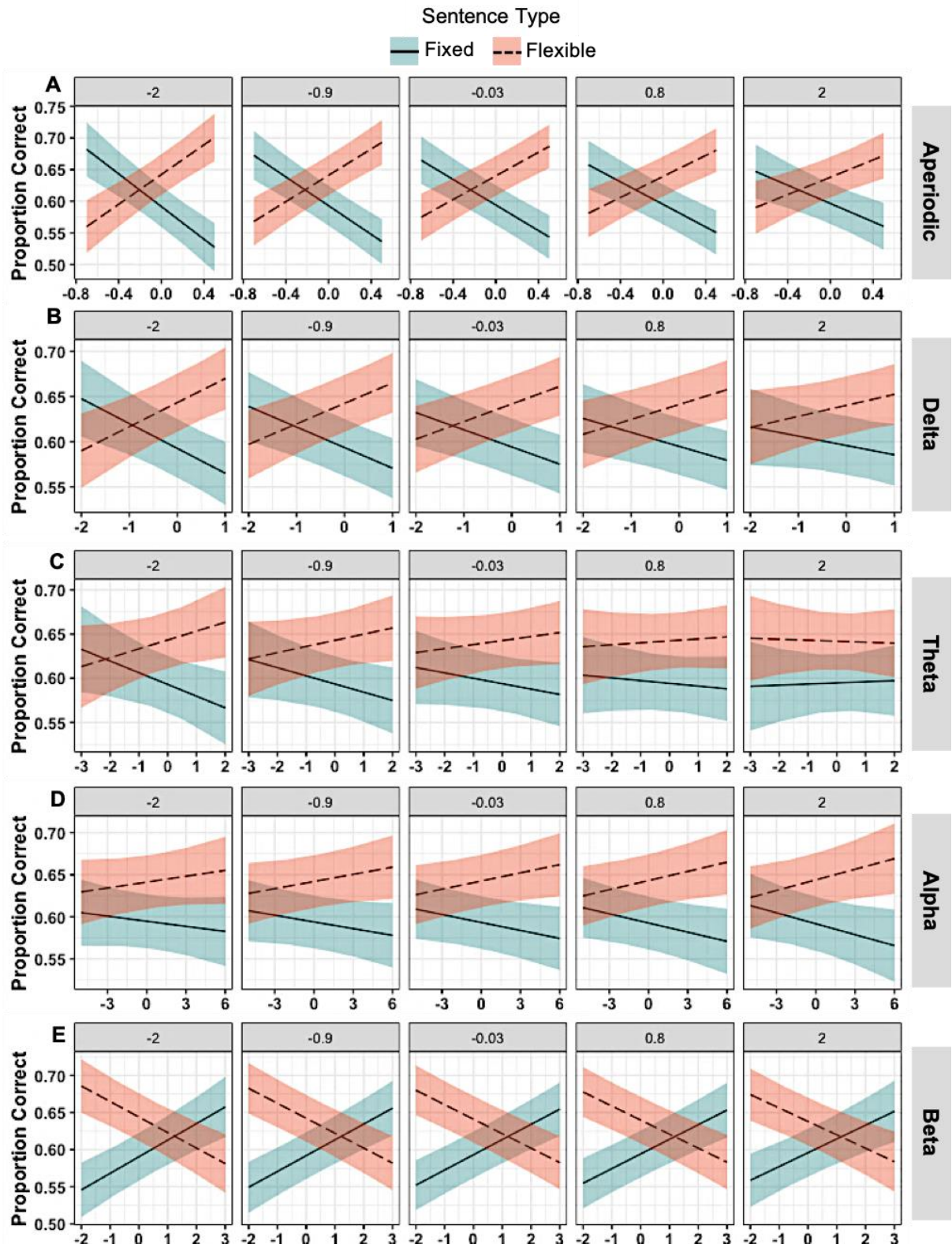

Figure 7: Estimated marginal means for proportion of correct responses (y-axes) by global variations in aperiodic activity and spectral power in the delta, theta, alpha and beta bands (x-axes) and Sentence Type (fixed = blue solid line, flexible = dashed orange line) averaged across each trial of the learning task. Shaded regions indicate 83% confidence intervals.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

4. DISCUSSION

The current experiment tested how neural oscillations and aperiodic activity differentially contribute to higher-order cognitive processing. Using a miniature artificial language, we analysed how spectral power in the 1 – 35 Hz range is modulated during the learning and processing of different word order permutations. We demonstrated that aperiodic and oscillatory activity interacted to influence fixed word order processing to a greater degree than flexible word orders. Our results further reveal that: (1) fixed word order processing was associated with an increase in the $1/f$ slope, while flexible word order sentences were associated with a shallower slope; (2) fixed and flexible word order sentences were associated with a nonlinear modulations in in delta, theta, alpha and beta activity, and; (3) aperiodic and low frequency (i.e., < 8 Hz) activity yielded a similar pattern of activity during language learning, including their spatiotemporal profiles. Moreover, we demonstrated that inter-individual variations in spectral power across the learning task interacted with aperiodic activity to influence subsequent behavioural performance. To our knowledge, this is the first study to systematically characterise the contribution of oscillatory and aperiodic neural activity during language learning and word order processing. Below, we discuss how our findings relate to existing literature, and how the combined study of aperiodic and oscillatory activity is necessary if a full understanding of the neurobiology of cognition is to be achieved.

4.1. Neural oscillations, language learning and sentence processing

To date, only a few studies have examined the neural oscillations involved in (artificial) grammar learning (e.g., de Diego-Balaguer et al., 2011; Kepinska et al., 2017), demonstrating that alpha/beta synchronisation and gamma phase synchrony between different cortical areas predicts successful learning. While we did not examine gamma activity due to potential EEG artifacts (e.g., volume conduction currents, saccades etc; Buzsaki & Schomburg, 2015; Kovach et al., 2011; Whitham et al., 2007), we found that alpha/beta activity differentially contribute to the learning and processing of fixed and flexible word order rules. Increased anterior alpha synchronisation predicted an increase in performance for flexible word orders, while desynchronization predicted fixed word order processing. By contrast, the strongest beta effects were across central anterior sensors, with beta synchronisation predicting higher performance for fixed sentences, while beta desynchronization was associated with higher performance for flexible word orders.

These findings are consistent with previous work (e.g., Kepinska et al., 2017), but reveal more fine-grained patterns of (de)synchronisation between word order variations, which may be explained by cue-integration-based models of native language processing (Bates et al., 2001; Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky et al., 2015; Kaufeld et al., 2020; Martin, 2016). Under this framework, cues that are differentially weighted according to the conditional probabilities of the language are integrated to comprehend incoming linguistic input (e.g., sentences). Here, fixed word order sentences contained linear order-based cues, which are analogous to English, while flexible word orders required animacy-based cues for interpretation. From this perspective, and in line with previous work on sequence processing (Crivelli-Decker et al., 2018; Kikuchi, et al., 2018; Wang et al., 2019), beta synchronisation likely reflected the propagation of top-down predictions during the learning of fixed word order

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

rules. Specifically, in fixed sentences, the first noun is invariably the Actor, and as such, predictions are constrained to anticipating that the second noun will be the Undergoer, while also containing a verb-final construction. As such, due to the strong sequence dependence in fixed word orders, precision-weighted predictions would likely increase linearly across the sentence, manifesting in beta synchronisation (Arnal, 2012; Cross et al., 2018a; Lewis & Bastiaansen, 2015).

The inverse relationship with flexible word order processing – which was predicted by greater beta desynchronization – can also be explained under this framework. Given that flexible word orders can contain either Actor-first or Undergoer-first constructions, predictions cannot be based on the linear position of the words, and instead must be driven by the integration of (non-adjacent) animacy-based cues to arrive at an accurate sentential percept. Given that our sample consisted of native monolingual English speakers (a language that relies heavily on word order cues; Bates et al., 2001; Bornkessel-Schlesewsky, et al. 2011; MacWhinney et al., 1984), beta desynchronization during flexible word order processing may have reflected prediction errors and internal model updating. That is, beta desynchronization during the learning of flexible word order rules may have reflected internal generative (predictive) model updating based on mismatches with predicted and actual sensory input, while beta synchronisation during fixed word order processing likely reflected the accumulation of top-down predictions based on our sample of native English speakers' preference for word-order based cues. Importantly, this interpretation is consistent with temporal sequence learning paradigms, where beta power has been shown to increase for fixed sequences relative to "random" sequences (Crivelli-Decker et al., 2018).

Alpha activity showed a similar pattern; however, increased anterior alpha synchronisation predicted flexible word order processing, while desynchronization predicted fixed word order processing. Frontal alpha synchronisation during language processing may reflect goal-directed processing and the allocation of enhanced attentional resources, which may have been required for the successful learning of flexible word order rules (Kepinska et al., 2017), given that they deviate from the canonical English word order (Bates et al., 2001). This interpretation is in line with evidence demonstrating that alpha oscillations reflect rhythmic cortical gating by alternating the activation of task-relevant cortical regions while actively inhibiting the processing of task-irrelevant information (Chapeton et al., 2019; de Vries et al., 2020; Gallotto et al., 2020; Klimesch, 2012; Jensen & Mazaheri, 2010). From this perspective, greater alpha synchronisation may have facilitated the extraction of flexible word order rules by suppressing task-irrelevant input and optimising cortical communication in a selectively precise manner, promoting the encoding and consolidation of non-canonical grammatical rules.

Finally, we recently proposed that the combination of (animacy-based) non-adjacent elements is reflected in increased theta activity (Cross et al., 2018a; see also Cross et al., 2020b). Here, we found that theta synchronisation predicted improved performance for flexible but not fixed word order rules, an effect that parametrically decreased from posterior to anterior regions. Increased theta synchronisation has been shown to predict memory for temporally close stimuli (e.g., in temporal associative memory tasks; for review, see Herweg et al., 2020),

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and in encoding novel grammatical rules (e.g., Kepinska et al., 2017). The increase in theta power reported here for flexible word orders may have reflected the learning and integration of rules governing conceptual relations between words, rather than word order-based rules. Indeed, such theta effects have been reported during native sentence processing, with theta synchronisation localised to left temporal regions (Lam et al., 2016). These effects are also consistent with the general memory literature: retrieval of language (e.g., single words), shape and face stimuli elicit higher theta synchronisation (Bastiaansen et al., 2002; Klimesch et al., 2008; Klimesch et al., 2010; Mormann et al., 2005; Osipova et al., 2006), with these effects manifesting the strongest over medial temporal and prefrontal cortices (Guderian & Düzel, 2005), likely reflecting the activation of relevant memory traces and executive control processes, respectively. Theta activity has also been shown to decrease for fixed relative to flexible sequences, which were associated with increased theta synchronisation over frontal areas (Crivelli-Decker et al., 2018), highlighting the role of theta in binding non-sequential elements to form a coherent percept.

Together, these results build upon the handful of studies examining the oscillatory correlates of grammar learning, but also offer more fine-grained insights into the mechanisms underlying grammar learning and sentence processing, including: (1) beta activity likely reflects the accumulation and propagation of top-down precision-weighted predictions, which are modulated by different word order permutations; (2) alpha activity reflects more domain-general mechanisms of attention during learning and sentence processing, allowing the brain to prioritise the encoding of novel grammatical rules, and; (3) theta oscillations support the encoding of novel grammatical rules by binding conceptual relations between individual elements in a sentence, particularly for flexible word orders. The pattern of findings in the current study also speak strongly to the idea that neural oscillations differentially contribute to the formation of memory of varying complexity, such that the encoding and retrieval of image and word stimuli elicit distinct (de)synchronisation patterns compared to artificial grammar learning paradigms; however, future research should systematically study such task- and stimulus-related differences.

4.2. Aperiodic and oscillatory interactions support (language) learning and memory

Using Irregular-Resampling Auto-Spectral Analysis (IRASA; Wen & Liu, 2016), we estimated the $1/f$ slope during artificial grammar learning with the aim of characterising the influence of dynamic alterations in aperiodic activity on higher-order cognitive processing. This is the first study to examine aperiodic activity and its interaction with oscillatory power in the context of language learning, with two critical findings emerging: (1) learning-related aperiodic activity interacted with oscillatory power to modulate behavioural performance for fixed word orders to a stronger degree than flexible word order rules, and; (2) 1/f slope and beta activity showed an inverse relationship with flexible and fixed word order processing. These findings speak strongly to the fact that $1/f$-like activity should not be filtered from the signal, but instead be analysed in combination with transient oscillatory activity to predict behaviour.

This claim is supported by a number of studies demonstrating the influence of aperiodic activity on a range of cognitive computations, including processing speed (Ouyang et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2020), memory (Sheehan et al., 2018) and prediction in language (Dave et al., 2018). From a neurophysiological perspective, $1/f$-like neural activity has been proposed to encode information relating to intrinsic brain function (Muthukumaraswamy and Liley, 2018), including the balance between excitation/inhibition (Gao et al., 2017), likely reflecting glutamate and GABA synaptic inputs into inter- and intra-cortical networks (Dave et al., 2018; Gao et al., 2017). Based on this perspective, Dave et al. (2018) argued that aperiodic activity influences prediction in language by modulating the strength of predictions of upcoming linguistic information via population spiking synchrony (Engel et al., 2001). This interpretation applies to our finding that aperiodic and beta activity showed an inverse association with performance for fixed and flexible word orders: beta synchronisation predicted accurate responses for fixed sentences, while beta desynchronization predicted performance for flexible word orders. By contrast, a steeper $1/f$ slope during learning was associated with improved accuracy for flexible sentences, while a shallower slope predicted improved accuracy for fixed word orders. This relationship becomes more complex when examining the interactive influence of beta and $1/f$ activity on behaviour: when the $1/f$ slope is shallow, the effect of beta activity on fixed and flexible word order processing is strongest; however, when the $1/f$ is steep, the effect of beta activity on accurate sentence processing for both fixed and flexible word orders diminishes. We believe that this pattern of results can be explained by integrating two perspectives; namely the "spectral fingerprints" hypothesis (Hanslmayr & Staudigl, 2014; Keitel & Gross, 2016; Siegel et al., 2012; Watrous et al., 2015; Womelsdorf et al., 2014) and models of hierarchical predictive coding (Friston, 2010, 2019, 2020).

The "spectral fingerprints" hypothesis argues that power changes in different frequency bands reflects distinct stages of memory and information processing (Fellner et al., 2019; Keitel & Gross, 2016), rather than reflecting a "spectral tilt" between lower and higher frequencies. For example, decreases in alpha/beta and increases in gamma power during memory retrieval have been shown to occur on different temporal scales and in different brain areas, providing evidence against proposals that a change in the tilt of the power spectrum solely drives memory computations (Fellner et al., 2019). Further, increases in high frequency gamma activity has been proposed to reflect the propagation of bottom-up sensory signals (Lewis et al., 2015; Richter, Thompson, Bosman, & Fries, 2017), while a decrease in alpha/beta power is thought to index top-down prediction errors (Bressler & Richter, 205; Friston, 2019; Samaha, Bauer, Cimaroli, & Postle, 2015). From this perspective, a steeper $1/f$ slope may reflect the maintenance of top-down predictions that allow comprehenders to generate expectations of incoming stimuli, minimizing prediction errors. This interpretation also holds for interactions observed with aperiodic and oscillatory activity in the theta and alpha bands, and as such, provides evidence that $1/f$-like activity dynamically influences cortical excitability across the frequency spectrum to minimize prediction error and to maximize precision-weighted predictions during language learning and sentence processing.

Taken together, we have demonstrated for the first time that oscillatory and aperiodic activity jointly contribute to the learning of higher-order language. There are, of course, several open questions that arise from the present results. First, how does the interaction between oscillatory and aperiodic activity relate to individual differences in atypical populations, such

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

as those with schizophrenia and age-related pathologies, including Alzheimer's disease? Previous research has shown that cognitive deficits characteristic of schizophrenia may be better explained by changes in the $1/f$ slope than irregularities in the canonical frequency bands (Peterson et al., 2018), and that $1/f$ activity mediates age-related deficits in working memory (Voytek et al., 2015); however, the interaction between aperiodic and oscillatory activity during more complex cognitive computations, such as sequence learning and language processing, remain unknown. Second, we did not examine oscillatory and aperiodic activity during the judgement task (cf. Cross et al., 2020b). Doing so would have provided a more thorough analysis of the aperiodic and oscillatory components related to grammar learning, how these dynamics change over time, and whether they manifest differently during retrieval. Lastly, given that we used scalp-recorded EEG, we are unable to make any neuroanatomical inferences. Future research using techniques with greater spatiotemporal resolution, such as magnetoencephalography and intracranial EEG, would be able to better isolate potential neuroanatomical generators of aperiodic and oscillatory activity during grammar learning, and by extension, complex cognitive processing more generally.

4.3. Conclusion

In this study we connected neural oscillations, a mainstream electrophysiological measure, with the $1/f$ slope, and in doing so, have provided further evidence that aperiodic activity plays an important role in higher-order cognition, including language learning. We have also demonstrated that grammar learning is modulated by different word order permutations, which manifest in distinct oscillatory profiles during incremental sentence processing. Future work would benefit from examining how and if these interactions emerge in (age-related) pathologies, and whether patterns of aperiodic and oscillatory activity during language learning and sentence processing are generated by specific neuroanatomical networks. Such work will provide a better understanding of the neurobiology of cognition in both health and disease.

REFERENCES

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, *4*.

Arnal, L. H. (2012). Predicting "When" Using the Motor System's Beta-Band Oscillations. *Front Hum Neurosci*, *6*, 225. https://doi.org/10.3389/fnhum.2012.00225

Austin, P. C., & Hux, J. E. (2002). A brief note on overlapping confidence intervals. *Journal of Vascular Surgery*, *36*(1), 194–195.

Baayen, H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, *94*, 206–234.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Changes in theta and beta oscillations as signatures of novel word consolidation. *J Cogn Neurosci*, *27*(7), 1286–1297. https://doi.org/10.1162/jocn_a_00801

Bastiaansen, M. C., Van Berkum, J. J., & Hagoort, P. (2002). Syntactic processing modulates the θ rhythm of the human EEG. *Neuroimage*, *17*(3), 1479–1492.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Green, P., & Bolker, M. B. (2015). Package 'lme4.' *Convergence*, *12*(1), 2.

Bates, E., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: A Cross-Language Perspective. *Annual Review of Psychology*, *52*(1), 27.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159.

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genc, S., Philipp, M., Roehm, D., & Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain Lang*, *117*(3), 133–152. https://doi.org/10.1016/j.bandl.2010.09.010

Bornkessel-Schlesewsky, I., Schlesewsky, M., Small, S. L., & Rauschecker, J. P. (2015). Neurobiological roots of language in primate audition: Common computational properties. *Trends Cogn Sci*, *19*(3), 142–150. https://doi.org/10.1016/j.tics.2014.12.008

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychol Rev*, *113*(4), 787–821. https://doi.org/10.1037/0033-295X.113.4.787

Bressler, S. L., & Richter, C. G. (2015). Interareal oscillatory synchronization in top-down neocortical processing. *Current Opinion in Neurobiology*, *31*, 62-66. doi: https://doi.org/10.1016/j.conb.2014.08.010

Buzsaki, G., & Schomburg, E. W. (2015). What does gamma coherence tell us about inter-regional neural communication? *Nat Neurosci*, *18*(4), 484–489. https://doi.org/10.1038/nn.3952

Chapeton, J. I., Haque, R., Wittig, J. H., Inati, S. K., & Zaghloul, K. A. (2019). Large-Scale Communication in the Human Brain Is Rhythmically Modulated through Alpha Coherence. *Curr Biol*, *29*(17), 2801-2811 e5. https://doi.org/10.1016/j.cub.2019.07.014

Corcoran, A. W., Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2019). restingIAF [Computer software]. doi.org/10.5281/zenodo.2575868

Corcoran, A. W., Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2018). Toward a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology*. https://doi.org/10.1111/psyp.13064

Crivelli-Decker, J., Hsieh, L. T., Clarke, A., & Ranganath, C. (2018). Theta oscillations promote temporal sequence learning. *Neurobiol Learn Mem*. https://doi.org/10.1016/j.nlm.2018.05.001

Cross, Z. R, Helfrich, R. F., Kohler, M. J., Corcoran, A. W., Coussens, S., Zou-Williams, L., Schlesewsky, M., Gaskell, M. G., Knight, R. T., & Bornkessel-Schlesewsky, I.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

(2020b). Slow wave-spindle coupling during sleep predicts language learning and associated oscillatory activity. *BioRxiv*. doi: https://doi.org/10.1101/2020.02.13.948539

Cross, Z. R., Kohler, M. J., Schlesewsky, M., Gaskell, M. G., & Bornkessel-Schlesewsky, I. (2018a). Sleep-Dependent Memory Consolidation and Incremental Sentence Comprehension: Computational Dependencies during Language Learning as Revealed by Neuronal Oscillations. *Frontiers in Human Neuroscience*, *12*. *https://doi.org/10.3389/fnhum.2018.00018*

Cross, Z. R., Santamaria, A., Corcoran, A. W., Alday, P. M., Coussens, S., & Kohler, M. J. (2018b). Alpha oscillations prior to encoding preferentially modulate memory consolidation during wake relative to sleep. *bioRxiv*, 202176.

Cross, Z. R., Zou-Williams, L., Wilkinson, E., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2020a). Mini Pinyin: A modified miniature language for studying language learning and incremental sentence processing. *PsyArXiv*. doi: 10.31234/osf.io/b9erz

Dave, S., Brothers, T. A., & Swaab, T. Y. (2018). 1/f neural noise and electrophysiological indices of contextual prediction in aging. *Brain Res*, *1691*, 34–43. https://doi.org/10.1016/j.brainres.2018.04.007

de Diego-Balaguer, R., Fuentemilla, L., & Rodriguez-Fornells, A. (2011). Brain Dynamics Sustaining Rapid Rule Extraction from Speech. *J Cogn Neurosci*, *23*(10), 3105–3120.

de Vries, I. E. J., Slagter, H. A., & Olivers, C. N. L. (2020). Oscillatory Control over Representational States in Working Memory. *Trends Cogn Sci*, *24*(2), 150–162. https://doi.org/10.1016/j.tics.2019.11.006

Donoghue, T., Dominguez, J., & Voytek, B. (2020). https://doi.org/10.1101/2020.01.11.900977

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, *2*, 12.

Fasiolo, M., Nedellec, R., Goude, Y., & Wood, S. N. (2019). Scalable visualization methods for modern generalized additive models. *Journal of Computational and Graphical Statistics*, 1–9.

Fellner, M. C., Gollwitzer, S., Rampp, S., Kreiselmeyr, G., Bush, D., Diehl, B., Axmacher, N., Hamer, H., & Hanslmayr, S. (2019). Spectral fingerprints or spectral tilt? Evidence for distinct oscillatory signatures of memory formation. *PLoS Biol*, *17*(7), e3000403. https://doi.org/10.1371/journal.pbio.3000403

Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., & Heilberger, R. (2011). *Package "car": Companion to applied regression*.

Fox, J., Weisberg, S., Friendly, M., Hong, J., Andersen, R., Firth, D., Taylor, S., & Fox, M. J. (2019). *Package 'effects.'*

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nat Rev Neurosci*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. J. (2019). Waves of prediction. *PLoS Biol*, *17*(10), e3000426. https://doi.org/10.1371/journal.pbio.3000426

Friston, K., & Buzsaki, G. (2016). The Functional Anatomy of Time: What and When in the Brain. *Trends Cogn Sci*, *20*(7), 500–511. https://doi.org/10.1016/j.tics.2016.05.001

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Friston, K., Parr, T., Yufik, Y., Sajid, N., Price, C. J., & Holmes, E. (2020). Generative models, language and active inference. *PsyArXiv Preprints*. doi: 10.31234/osf.io/4j2k6

Friston, Karl. (2018). Does predictive coding have a future? *Nature Neuroscience*, *21*(8), 1019–1021.

Gallotto, S., Duecker, F., ten Oever, S., Schuhmann, T., de Graaf, T. A., & Sack, A. T. (2020). Relating alpha power modulations to competing visuospatial attention theories. *NeuroImage*, *207*, 116429.

Gao, R., Peterson, E. J., & Voytek, B. (2017). Inferring synaptic excitation/inhibition balance from field potentials. *Neuroimage*, *158*, 70-78. doi: 10.1016/j.neuroimage.2017.06.078

Greenough, W. T. (1984). Structural correlates of information storage in the mammalian brain: A review and hypothesis. *Trends in Neurosciences*, *7*(7), 229–233.

Griffiths, B. J., Mayhew, S. D., Mullinger, K. J., Jorge, J., Charest, I., Wimber, M., & Hanslmayr, S. (2019). Alpha/beta power decreases track the fidelity of stimulus-specific information. *Elife*, *8*. https://doi.org/10.7554/eLife.49562

Griffiths, B. J., Parish, G., Roux, F., Michelmann, S., van der Plas, M., Kolibius, L. D., Chelvarajah, R., Rollings, D. T., Sawlani, V., Hamer, H., Gollwitzer, S., Kreiselmeyer, G., Staresina, B., Wimber, M., & Hanslmayr, S. (2019). Directional coupling of slow and fast hippocampal gamma with neocortical alpha/beta oscillations in human episodic memory. *Proc Natl Acad Sci U S A*, *116*(43), 21834–21842. https://doi.org/10.1073/pnas.1914180116

Guderian, S., & Düzel, E. (2005). Induced theta oscillations mediate large- scale synchrony with mediotemporal areas during recollection in humans. *Hippocampus*, *15*(7), 901–912.

Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., Noto, T., Knight, R. T., Shestyuk, A., & Voytek, B. (2018). https://doi.org/10.1101/299859

Hanslmayr, S., & Staudigl, T. (2014). How brain oscillations form memories—A processing based perspective on oscillatory subsequent memory effects. *Neuroimage*, *85 Pt 2*, 648–655. https://doi.org/10.1016/j.neuroimage.2013.05.121

Hanslmayr, Simon, Staudigl, T., & Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: The information via desynchronization hypothesis. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00074

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E., ... & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. doi: 10.7717/peerj.4794

Hastie, T., & Tibshirani, R. (1987). Generalized additive models: Some applications. *Journal of the American Statistical Association*, *82*(398), 371–386.

Hastie, T., & Tibshirani, R. (1990). Exploring the nature of covariate effects in the proportional hazards model. *Biometrics*, 1005–1016.

He, B. J. (2014). Scale-free brain activity: Past, present, and future. *Trends Cogn Sci*, *18*(9), 480–487. https://doi.org/10.1016/j.tics.2014.04.003

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

He, B. J., Zempel, J. M., Snyder, A. Z., & Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron*, *66*(3), 353–369. https://doi.org/10.1016/j.neuron.2010.04.020

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human Memory. *Trends Cogn Sci*, *24*(3), 208–227. https://doi.org/10.1016/j.tics.2019.12.006

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186. doi: https://doi.org/10.3389/fnhum.2010.00186

Kaufeld, G., Bosker, H. R., Alday, P. M., Meyer, A. S., & Martin, A. E. (2020). Linguistic structure and meaning organize neural oscillations into a content-specific hierarchy. *BioRxiv*, 2020.02.05.935676. https://doi.org/10.1101/2020.02.05.935676

Keitel, A., & Gross, J. (2016). Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *PLoS Biol*, *14*(6), e1002498. https://doi.org/10.1371/journal.pbio.1002498

Kepinska, O., Pereda, E., Caspers, J., & Schiller, N. O. (2017). Neural oscillatory mechanisms during novel grammar learning underlying language analytical abilities. *Brain Lang*, *175*, 99–110. https://doi.org/10.1016/j.bandl.2017.10.003

Kikuchi, Y., Sedley, W., Griffiths, T. D., & Petkov, C. I. (2018). Evolutionarily conserved neural signatures involved in sequencing predictions and their relevance for language. *Current Opinion in Behavioral Sciences*, *21*, 145–153. https://doi.org/10.1016/j.cobeha.2018.05.002

Klimesch, W., Freunberger, R., Sauseng, P., & Gruber, W. (2008). A short review of slow phase synchronization and memory: Evidence for control processes in different memory systems? *Brain Res*, *1235*, 31–44. https://doi.org/10.1016/j.brainres.2008.06.049

Klimesch, Wolfgang, Freunberger, R., & Sauseng, P. (2010). Oscillatory mechanisms of process binding in memory. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1002–1014.

Klimesch, Wolfgang. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. http://dx.doi.org/10.1016/j.tics.2012.10.007

Kovach, C. K., Tsuchiya, N., Kawasaki, H., Oya, H., Howard, M. A., & Adolphs, R. (2011). Manifestation of ocular-muscle EMG contamination in human intracranial recordings. *Neuroimage*, *54*(1), 213–233. https://doi.org/10.1016/j.neuroimage.2010.08.002

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Lam, N. H., Schoffelen, J. M., Udden, J., Hulten, A., & Hagoort, P. (2016). Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *Neuroimage*. https://doi.org/10.1016/j.neuroimage.2016.03.007

Lendner, J. D., Helfrich, R. F., Mander, B. A., Romundstad, L., Lin, J. J., Walker, M. P., Larsson, P. G., & Knight, R. T. (2019). https://doi.org/10.1101/625210

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155–168. https://doi.org/10.1016/j.cortex.2015.02.014

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends Cogn Sci*, *15*(8), 343–351. https://doi.org/10.1016/j.tics.2011.06.004

Lin, X., & Zhang, D. (1999). Inference in generalized additive mixed modelsby using smoothing splines. *Journal of the Royal Statistical Society: Series b (Statistical Methodology)*, *61*(2), 381–400.

MacGregor-Fors, I., & Payton, M. E. (2013). Contrasting diversity values: Statistical inferences based on overlapping confidence intervals. *PLoS One*, *8*(2).

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 127–150. http://dx.doi.org/10.1016/S0022-5371(84)90093-8

Marra, G., & Wood, S. N. (2011). Practical Variable Selection for Generalized. Additive Models. *Computational Statistics & Data Analysis*, *55*, 2372–2387.

Martin, A. E. (2016). Language Processing as Cue Integration: Grounding the Psychology of Language in Perception and Neurophysiology. *Frontiers in Psychology*, *7*, 120. https://doi.org/10.3389/fpsyg.2016.00120

Mathot, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behav Res Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419.

Meteyard, L., & Davies, R. A. (2020) Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, 104092. doi: https://doi.org/10.1016/j.jml.2020.104092

Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C. E., & Fernández, G. (2005). Phase/amplitude reset and theta–gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus*, *15*(7), 890–900.

Muralikrishnan, R., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). Animacy-based predictions in language comprehension are robust: Contextual cues modulate but do not nullify them. *Brain Research*, *1608*, 108–137.

Muthukumaraswamy, S. D., & Liley, D. T. (2018). 1/f electrophysiological spectra in resting and drug-induced states can be explained by the dynamics of multiple oscillatory relaxation processes. *Neuroimage*, *179*, 582–595.

Nenadic, O., & Greenacre, M. (2007). Correspondence analysis in R, with two-and three-dimensional graphics: The ca package. *Journal of Statistical Software*, *20*(3).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*.

Osipova, D., Takashima, A., Oostenveld, R., Fernandez, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *J Neurosci*, *26*(28), 7523–7531. https://doi.org/10.1523/JNEUROSCI.1948-06.2006

Ouyang, G., Hildebrandt, A., Schmitz, F., & Herrmann, C. S. (2020). Decomposing alpha and 1/f brain activities reveals their differential associations with cognitive processing speed. *Neuroimage*, *205*, 116304. https://doi.org/10.1016/j.neuroimage.2019.116304

Packard, P. A., Rodriguez-Fornells, A., Bunzeck, N., Nicolas, B., de Diego-Balaguer, R., & Fuentemilla, L. (2017). Semantic Congruence Accelerates the Onset of the Neural Signals of Successful Memory Encoding. *J Neurosci*, *37*(2), 291–301. https://doi.org/10.1523/JNEUROSCI.1622-16.2016

Parish, G., Hanslmayr, S., & Bowman, H. (2018). The Sync/deSync Model: How a Synchronized Hippocampus and a Desynchronized Neocortex Code Memories. *J Neurosci*, *38*(14), 3428–3440. https://doi.org/10.1523/JNEUROSCI.2561-17.2018

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876.

Peterson, E. J., Rosen, B. Q., Campbell, A. M., Belger, A., & Voytek, B. (2018). https://doi.org/10.1101/113449

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Richter, C. G., Thompson, W. H., Bosman, C. A., & Fries, P. (2017). Top-down beta enhances bottom-up gamma. *The Journal of Neuroscience*, *37*(28), 6698-6711. doi: 10.1523/JNEUROSCI.3771-16.2017

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. *PNAS*, *112*(27), 8439-8444. doi: https://doi.org/10.1073/pnas.1503686112

Sarkar, D. (2008). *Lattice: Multivariate data visualization with R*. Springer Science & Business Media.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038.

Schreiner, T., Goldi, M., & Rasch, B. (2015). Cueing vocabulary during sleep increases theta activity during later recognition testing. *Psychophysiology*, *52*(11), 1538–1543. https://doi.org/10.1111/psyp.12505

Sheehan, T. C., Sreekumar, V., Inati, S. K., & Zaghloul, K. A. (2018). Signal Complexity of Human Intracranial EEG Tracks Successful Associative-Memory Formation across

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Individuals. *J Neurosci*, *38*(7), 1744–1755. https://doi.org/10.1523/JNEUROSCI.2389-17.2017

Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nat Rev Neurosci*, *13*(2), 121–134. https://doi.org/10.1038/nrn3137

van Rij, J., Wieling, M., Baayen, R. H., van Rijn, H., & van Rij, M. J. (2016). *Package 'itsadug.'*

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-Related Changes in 1/f Neural Electrophysiological Noise. *J Neurosci*, *35*(38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015

Wang, P., Maye, A., Daume, J., Xue, G., & Engel, A. K. (2019). Dynamic Predictions: Oscillatory Mechanisms Underlying Multisensory Sequence Processing. *BioRxiv*, 778969.

Watrous, A. J., Fell, J., Ekstrom, A. D., & Axmacher, N. (2015). More than spikes: Common oscillatory mechanisms for content specific neural representations during perception and memory. *Curr Opin Neurobiol*, *31*, 33–39. https://doi.org/10.1016/j.conb.2014.07.024

Wen, H. & Liu, Z. (2015). Separating fractal and oscillatory components in the power spectrum of neurophysiological signal [Computer software]. Purdue University Research Repository. doi: 10.4231/R7WQ01R7

Wen, H., & Liu, Z. (2016). Separating fractal and oscillatory components in the power spectrum of neurophysiological signal. *Brain Topography*, *29*(1), 13–26.

Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S., Broberg, M., Wallace, A., DeLosAngeles, D., Lillie, P., Hardy, A., Fronsko, R., Pulbrook, A., & Willoughby, J. O. (2007). Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin Neurophysiol*, *118*(8), 1877–1888. https://doi.org/10.1016/j.clinph.2007.04.027

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J., & Tiesinga, P. (2014). Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nat Neurosci*, *17*(8), 1031–1039. https://doi.org/10.1038/nn.3764

Wood, S. (2006). *R-Manual: The MGCV package*. Version 1.3–22. Technical Report.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society, Series B 65*, 95–114.

Wood, S. N. (2013). A simple test for random effects in regression models. *Biometrika 100*(4), 1005-1010.

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. 2nd ed. Boca Raton, FL: CRC Press

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.10.984971; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Wood, S. N., Pya, N. & Saefken, B. (2017). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association, 111*(516), 1548-1563.

Wood, S. N., Scheipl, F., & Faraway, J. J. (2013). Straightforward intermediate rank tensor product smoothing in mixed models. *Stat. Comput. 23*, 341-360