Oscillatory and aperiodic neural activity jointly predict language learning

Abbreviated Title: Neural correlates of language learning

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

[1]Cognitive and Systems Neuroscience Research Hub, University of South Australia, Adelaide, Australia
[2]Cognition and Philosophy Laboratory, Monash University, Melbourne, Australia.
[3]Faculty of Health and Medical Sciences, Adelaide University, Adelaide, Australia.

Corresponding author: Zachariah R. Cross, Room H1-04, University of South Australia Magill Campus, St Bernards Road, Magill, SA 5072, Australia. E-mail: Zachariah.Cross@unisa.edu.au

Manuscript details

Number of pages:	30
Number of figures:	6
Abstract word count:	199
Introduction word count:	518
Discussion word count:	1,587
Data available at:	TBC
Code available at:	TBC

Acknowledgements: Preparation of this work was supported by Australian Commonwealth Government funding awarded to ZRC and AWC under the Research Training Program. IB-S is supported by an Australian Research Council Future Fellowship (FT160100437). We thank Isabella Sharrad, Lena Zou-Williams, Erica Wilkinson, Nicole Vass and Angela Osborn for help with data collection. Thank you also to the participants.

Conflict of interest statement: The authors declare no competing financial interests.

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. However, to date, this relationship has been investigated primarily with face and image stimuli; considerably less is known about the oscillatory correlates of complex rule learning, as in language. Further, recent work has shown that non-oscillatory ($1/f$) activity is functionally relevant to cognition, yet its interaction with oscillatory activity during complex rule learning remains unknown. Using spectral decomposition and power-law exponent estimation of human EEG data (17 females, 18 males), we show for the first time that $1/f$ and oscillatory activity jointly influence the learning of word order rules of a miniature artificial language system. Fixed word order rules were associated with a steeper $1/f$ slope, while flexible word order rules were associated with a shallower slope. We also show that increased alpha/beta power predicts fixed relative to flexible word order rule learning and 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.

SIGNIFICANCE STATEMENT

Characterising the neurobiological basis of cognition requires a deep understanding of how neural populations communicate. Here we show that $1/f$ and oscillatory activity interact during a complex language learning task to influence sentence processing. In addition, we observe that this interaction manifests on multiple spatiotemporal scales, with $1/f$ activity becoming less topographically distributed over the time course of learning. These results significantly add to our understanding of how aperiodic and oscillatory activity interact during higher-order cognitive processing, providing novel insights into how neural communication supports the learning and retrieval of complex language-related rules.

Keywords: language learning • neural oscillations • aperiodic activity • sentence processing • EEG • artificial grammar

1. INTRODUCTION

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 a broad literature has related neural oscillatory dynamics (i.e., [de]synchronisation of neural populations) to the encoding and retrieval of images and words (e.g., 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 in the human EEG predicts encoding of words and images (e.g., 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 has not been addressed in studies on the neural basis of complex rule learning to date.

Electrophysiological brain activity exhibits a $1/f$-like power distribution, which is often removed from the signal to isolate transient task-related oscillations (Donoghue et al., 2020; He, 2014; Lendner et al., 2019). However, this aperiodic component 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 the aperiodic slope – but not oscillatory activity – influences the N400 event-related potential and performance accuracy (Dave et al., 2018).

These findings suggest that aperiodic brain activity plays a critical functional role in the neurobiology of cognition (He et al., 2010); however, it is currently unknown whether oscillatory and aperiodic activity interact during memory encoding of information beyond single words and images, such as rule-based sequence learning, and whether any such interaction influences behavioural

outcomes. Clarifying the (separable) roles of oscillatory and aperiodic components of the EEG power spectrum may also bridge diverging results reported in studies using image and word stimuli and artificial grammar paradigms, lending support to the idea that neural oscillations differentially contribute to memory formation.

To better characterise the neural mechanisms underlying complex rule learning, we examined fluctuations in delta, theta, alpha and beta power during an artificial language learning task. We also modelled the interaction between oscillatory and aperiodic activity to characterise how 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), implementing generalised additive and linear mixed-effects analyses to model dynamic changes in oscillatory and aperiodic activity during the learning of various word order rules, and subsequent performance on a sentence judgement task.

2. METHOD

2.1. Participants

Data from 36 right-handed healthy, monolingual, native English-speakers were used from a study examining the effect of sleep on language learning (Cross et al., 2021). One participant was excluded from analysis for not having electroencephalography recorded during the sentence learning task due to experimenter error. The final sample size was 35 (M = 25.3, SD = 7.13; 17 female). Ethics approval was granted by the University of South Australia's Human Research Ethics committee (I.D.: 0000032556).

2.2. Stimuli and experimental 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., 2021), 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, and

each noun is associated with a different classifier: 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 coverbs *ba* and *bei*. Here we focus on 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 in the action described by the sentence) and the sentence must be verb final. In this context, 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 (i.e., verb must be positioned between the noun phrases). As such, accurate sentence interpretation is based more heavily on the animacy status of the noun phrases rather than word order. These manipulations are illustrated below, with a flexible word order shown in 1a and 1b and fixed (i.e., with the coverb *ba*) shown in 2:

(1)

(a) ge yisheng dale da piqiu.
(human) doctor hit (object) ball.
“the doctor hits the ball.”
(b) da xianjiao chile zhi laoshu.
(object) banana eat (animal) rat.
“the rat eats the banana.”

(2)

ge xiaofang ba da shubao liangle.
(human) firefighter ba (object) bag measure.
“the firefighter measures the bag.”

The experiment contained three phases: a vocabulary test, a learning phase, and a grammaticality judgement task. Approximately seven 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 participants who attained a score > 90% were eligible to undertake the learning and judgement task phases of the experiment. All 35 participants achieved 90% accuracy on the vocabulary task and thus completed the main experimental session. Overall, 576 unique sentences (288 grammatical, 288 ungrammatical) were created and divided into two equivalent sets.

We focus here on a subset of sentence conditions to investigate the mechanisms underlying the learning of different word order permutations (for EEG analyses during the sentence judgement task, see Cross et al., 2021). While no explicit instructions were given to participants in regard to the structure of the miniature language, a picture was shown prior to each sentence illustrating an event occurring between two entities, which was then described in the subsequently presented sentence. The learning task contained four blocks with 128 grammatical picture-sentence pairs overall (96 of which were included in the subset analysed here) that were presented via rapid visual serial presentation. The subset contained a further 156 novel sentences (50% grammatical, 50% ungrammatical) that were presented during the judgement task, which occurred immediately after the learning phase. The remaining sentences were considered fillers. The ungrammatical sentences induced a violation at 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 the four blocks, each grammatical construction was presented 32 times, with stimuli pseudo-randomised such that no sentences 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 to indicate whether the sentence conformed to 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. Participants received feedback on whether their

response was correct or incorrect (see Figure 1C and 1D for a schematic of the learning and judgement tasks, 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)

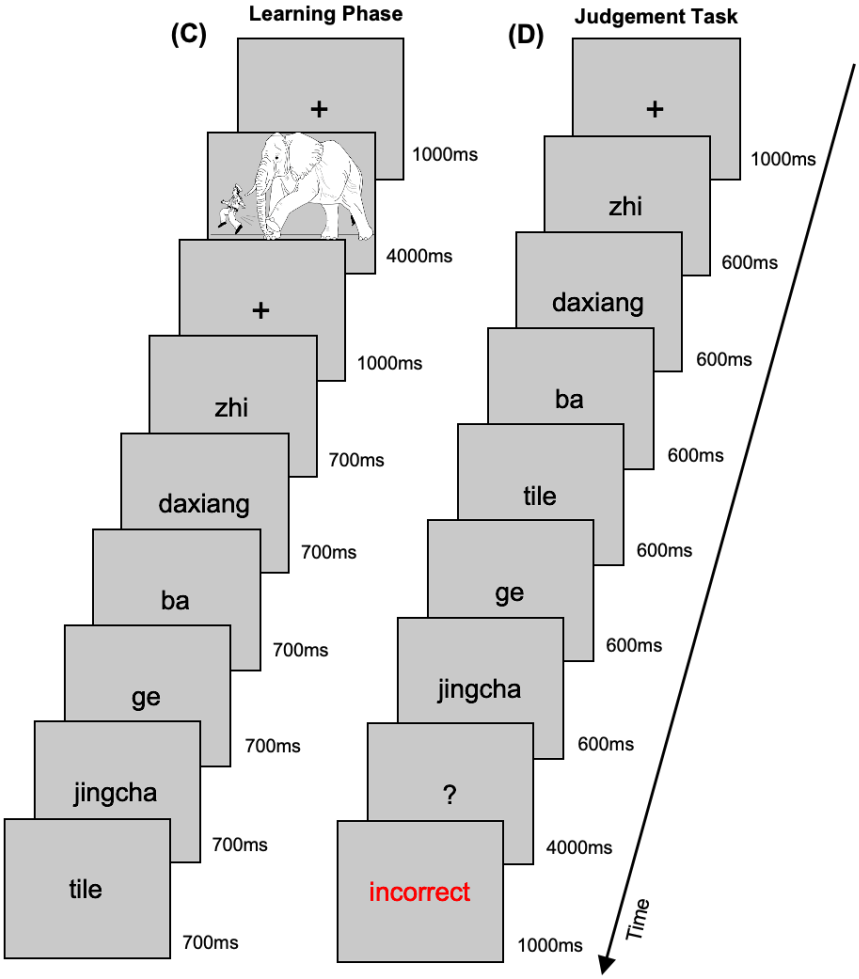

Fig 1 | (A) Summary of the grammatical (left) and ungrammatical (right) sentence constructions. (B) A portion of linguistic elements used in the sentence examples provided in (A). (C) Schematic of the sequence of events occurring in the sentence learning phase. (D) Schematic of the sequence of events occurring in the sentence judgement task.

2.3. EEG recording and pre-processing

Participants' EEG was recorded using a 32-channel BrainCap with sintered Ag/AgCl 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 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) with an initial band-pass filter of DC – 250 Hz and a sampling rate of 1000 Hz. Electrode impedances were kept below 10kΩ. EEG was also recorded during two minutes of eyes-open and two 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 using a two-pass Butterworth IIR filter (implemented in ft_preprocessing). 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 artefacts were visually inspected and subsequently interpolated with surrounding electrodes based on spherical spline interpolation (Perrin et al., 1989).

2.4. EEG data analysis

The aim of the analysis was to characterise the oscillatory and aperiodic dynamics underlying the initial encoding of complex grammatical rules. To this end, we computed differences between fixed

and flexible word order sentences in the following five spectral features of the EEG recorded during the learning phase: mean power density within individualised delta, theta, alpha, and beta bands, and the (inverse) slope of the $1/f$ spectral distribution (i.e., power-law exponent). We then used these metrics to investigate whether trial-level variation in oscillatory power and $1/f$ slope 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, 2016). 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 narrowband peaks. For a more detailed treatment of the IRASA method, see Wen and Liu (2016).

Trial data were divided into two non-overlapping 6s segments corresponding to picture and sentence presentation phases, respectively. These segments were further subdivided into 2s epochs (50% overlap) and separately passed to amri_sig_fractal.m for spectral parameter estimation. Following Muthukumaraswamy and Liley (2018), an extended set of resampling factors (h = 1.1 to 2.9 in steps of 0.05, $h \neq 2$) was used to mitigate the influence of fractal-oscillatory interactions on component separation.

Once the fractal component had been recovered from each power spectrum, it was parameterised using amri_sig_plawfit.m. This function rescales the frequency spectrum to achieve equally-spaced intervals in log-space before fitting a linear regression to a subregion of the double-log

transformed fractal spectrum (here, 1 – 30 Hz). 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, mean power densities were estimated following the subtraction of the fractal component of the PSD. This residual, 'oscillatory spectrum' was averaged across epochs within each segment, 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, the boundaries of each frequency band 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 the 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 (based on the IAF-scaled harmonic series), f_1 the lower bound, and f_2 the higher bound of a given frequency band.

IAF estimates used to determine 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. 2020b). This method applies a Savitzky-Golay filter (frame width = 11 bins, polynomial order = 5) to smooth and differentiate the power spectrum prior to estimating the peak frequency within a specified frequency range (here, 7—14 Hz). Peak estimates were averaged across channels, with a minimum of 3 channel estimates 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 IAF-anchored bounds for the delta, theta, alpha, and beta bands, power within each band was quantified using the mean power density metric proposed by Westfall (1990):

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

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.

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 deviation from the grand mean (Schad et al., 2020).

2.5.1. Generalised additive mixed models

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). 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 exponent χ, and the mean power density P for each kth frequency band (delta, theta, alpha, and beta) fluctuate during artificial grammar learning. Trial-level χ and P_k estimates from the sentence processing phase of each trial were modelled as a

function of experimental time (trial number), sensor space (2D Cartesian co-ordinates), and sentence type (fixed, flexible). Estimates from the preceding image presentation phase were treated as a baseline measure of spectral activity. 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, Corcoran, Macefield, & Hohwy, 2021, for similar approaches). Each GAMM took the following general form:

$$Y_i = \beta_0 + \beta_1 type_i + \beta_2 baseline_i + f(trial_i, topo.x_i, topo.y_i, \text{by} = type_i) +$$

$$f_{subject_i}(trial_i) + \epsilon,$$

where Y_i is the ith observation of spectral feature Y, β_0 is the model intercept, $\beta_1 type$ is a factor encoding the main effect of Sentence Type, $\beta_2 baseline$ is a covariate encoding the corresponding observation for Y during the baseline period, $f(\,.\,,$ by $= type)$ is the tensor product interaction between the experimental time (*trial*) and sensor space (*topo.x, topo.y*) covariates for each level of Sentence Type, $f_{subject}$ is the by-subject factor smooth on Trial, and ε is a Gaussian error term. Note that marginal smooths for sensor space co-ordinates were treated as isotropic (i.e., assumed to share a common scale).

GAMMs were estimated using the bam() function of the *R* package mgcv (v.1.8-33; Wood, 2011). Models were fit using the Fast REML method. P_k estimates for both the baseline and sentence processing period were $\log_{10}$ transformed prior to model inclusion. Models were 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). Type was entered as an ordered factor with Fixed assigned as the reference level, hence 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

The relationship between aperiodic and oscillatory power during grammar learning with behavioural performance on the judgement task was assessed using linear mixed-effects models. Behavioural performance was operationalised using the discrimination index (d'). d' is defined as the difference between the z transformed probabilities of hit rate (HR) and false alarm rate (FA; i.e., d'=z[HR] – z[FA]). These models took the following general form:

$$d'_{\ i} = \beta_0 + \beta_1 P_{k_i} * \beta_2 bexp_i * \beta_3 type_i * \beta_4 lat_i * \beta_5 sag_i + \beta_6 blnP_{k_i} + \beta_7 bln_{bexp} + subject_{0i} + \epsilon,$$

where P_k is mean power density in the frequency band of interest (i.e., delta, theta, alpha, beta), *bexp* is the exponent of the aperiodic $1/f$ slope, *type* refers to Sentence Type (fixed, flexible), *sag* is Sagittality (anterior, central, posterior) and *lat* refers to Laterality (left, midline, right). Subject ID (*subject*) was modelled as a random effect on the intercept. More complex random effect structures involving random slopes by participant did not converge. Baseline (pre-stimulus) oscillatory and aperiodic activity were added into the models as covariates, represented by $blnP_k$ and bln_{bexp}, respectively (Alday, 2019).

Type II Wald χ^2-tests from the ANOVA tables from the *car* package (Fox et al., 2011) were used to provide *p*-value estimates. An 83% confidence interval (CI) threshold was adopted for visualisations, 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. Participants performed moderately well on the judgement task, with a mean d' score of 1.02 (range: -1.20 – 3.78) and mean reaction time of 878.08 ms (range: 254.75 – 2076.83). There is clearly a high degree of inter-individual variability across both fixed and flexible sentences; however, flexible sentences had greater variability in d' scores, while fixed grammatical sentences had faster responses overall. For a detailed report and interpretation of these behavioural data, see Cross et al. (2020a).

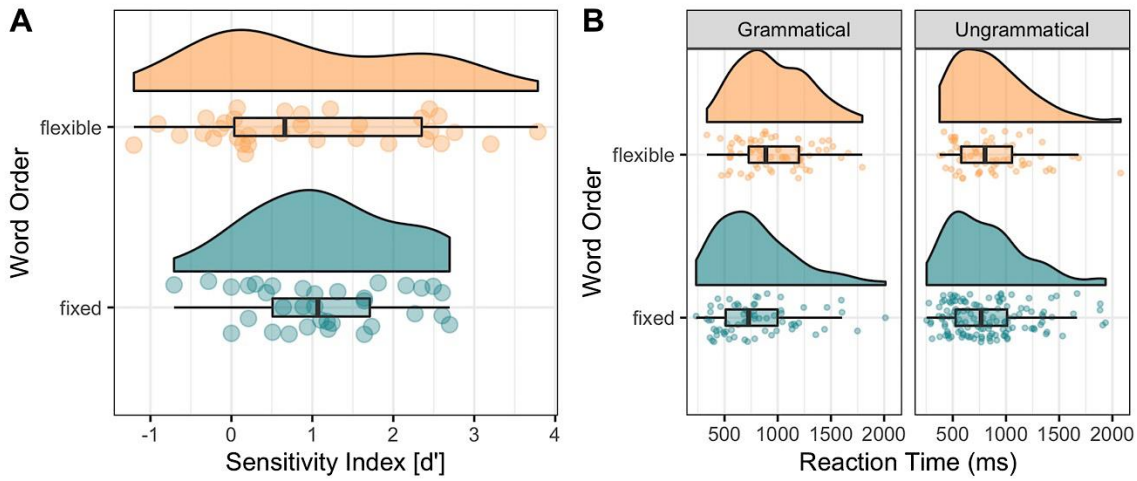

Fig 2 | Raincloud plots (Allen et al., 2019) illustrating the behavioural responses during the sentence judgement task. (A) Mean d' scores (x-axis) for Fixed and Flexible sentence types. (B) Mean reaction time (ms; x-axis) for Grammatical (left) and Ungrammatical (right) Fixed and Flexible sentence types. 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

Neurophysiological signals are non-stationary, showing dynamic changes over time as a function of endogenous and exogenous factors (e.g., Donoghue, Schaworonkow & Voytek, 2021), such as attentional fluctuations and the complexity of incoming sensory information. However, neurophysiological signals are typically analysed using linear models, which often do not capture non-linear modulations in neural activity, particularly over time. Here, we examine how aperiodic and oscillatory dynamics evolve over time during language learning, focusing specifically on the way in which spectral activity varies across sentence types (fixed vs flexible word orders; for full model summaries, see the supplementary material). Estimated changes in aperiodic and oscillatory spectral activity across learning task conditions is illustrated in Figure 3 (for topographical effects, see Figure 4).

The χ-exponent GAMM revealed that the $1/f$ slope was steeper overall for Flexible compared to Fixed word order sentences (estimated degrees of freedom [edf] = 3.24, $F = 31.80$, $p < .001$). Smooth terms on the Fixed condition revealed a bimodal pattern of variation that peaked around the first and third quarter of trials. Exponent values for Flexible word order sentences were more variable: highest

at the beginning of the session, exponents declined over the first third of trials, increased over the second third, and declined thereafter. This model further revealed significant topographic differences between conditions, with Flexible word orders evoking higher exponents over frontal regions compared to Fixed word order sentences (edf = 4.86, $F = 3.07$, $p < .001$; see Figure 4).

Mean delta power was higher on average during Flexible compared to Fixed word order trials (edf = 3.64, $F = 6.67$, $p < .001$). Smooth terms revealed this difference was predominantly driven by an increase in mean power density during the middle third of Flexible word order trials. This interaction was significantly modulated by sagittality, with between-condition differences (Fixed versus Flexible) in mean power density increasing over frontal electrodes (edf = 1.15, $F = 1.20$, $p = .01$). Similarly, theta power was lower on average during Flexible compared to Fixed rule learning (edf = 1.41, $F = 1.50$, $p = .01$). Here, however, the nonlinear pattern of spectral fluctuation was similar across both conditions, with power increasing over the first half of the session, decreasing over the next quarter, and then increasing again until the end of the task. This pattern of activity varied as a function of topography, with the difference between conditions being accentuated across more left-lateralised and posterior sites, as illustrated in Figure 4 (edf = 5.02, $F = 0.11$, $p = .02$).

Alpha power tended to increase over the course of the learning task, although this pattern was interrupted by a marked decline during the second quarter of the session (edf = 3.62, $F = 14.58$, $p < .001$). Flexible word orders evoked less alpha power than fixed word orders, especially during early and mid-session trials (edf = 3.60, $F = 4.74$, $p < .001$). This difference was most pronounced over left-lateralised and frontal sites (edf = 2.71, $F = 0.89$, $p < .001$), and least pronounced over right-lateralised and posterior sites, where both Fixed and Flexible sentences had a similar profile of increased alpha power (edf = 1.13, $F = 0.00$, $p = .43$). Finally, beta power was higher on average during Flexible than Fixed word order rules. Similar to the alpha model, beta power tended to increase over the course of the session, albeit with more variance around this trajectory (edf = 2.97, $F = 4.73$, $p < .001$). Flexible trials evoked markedly greater power density over the second and final quarters of the learning task (edf = 3.42, $F = 8.73$, $p < .001$). Power was lowest over central posterior electrodes, particularly toward the end of the learning phase for Fixed word order sentences (edf = 1.50, $F = 0.22$, $p = .04$). Bilateral

frontal regions also demonstrated an increase in the power difference between Fixed and Flexible trials (edf = 4.53, $F = 0.80$, $p < .001$).

Taken together, these data illustrate dynamic changes in both aperiodic and oscillatory activity as a function of different word order rules during learning. The aperiodic slope demonstrated dynamic non-linear changes over time, while delta power tended to decrease over time. By contrast, theta, alpha and beta power tended to increase across learning, with differences between different word order rules.

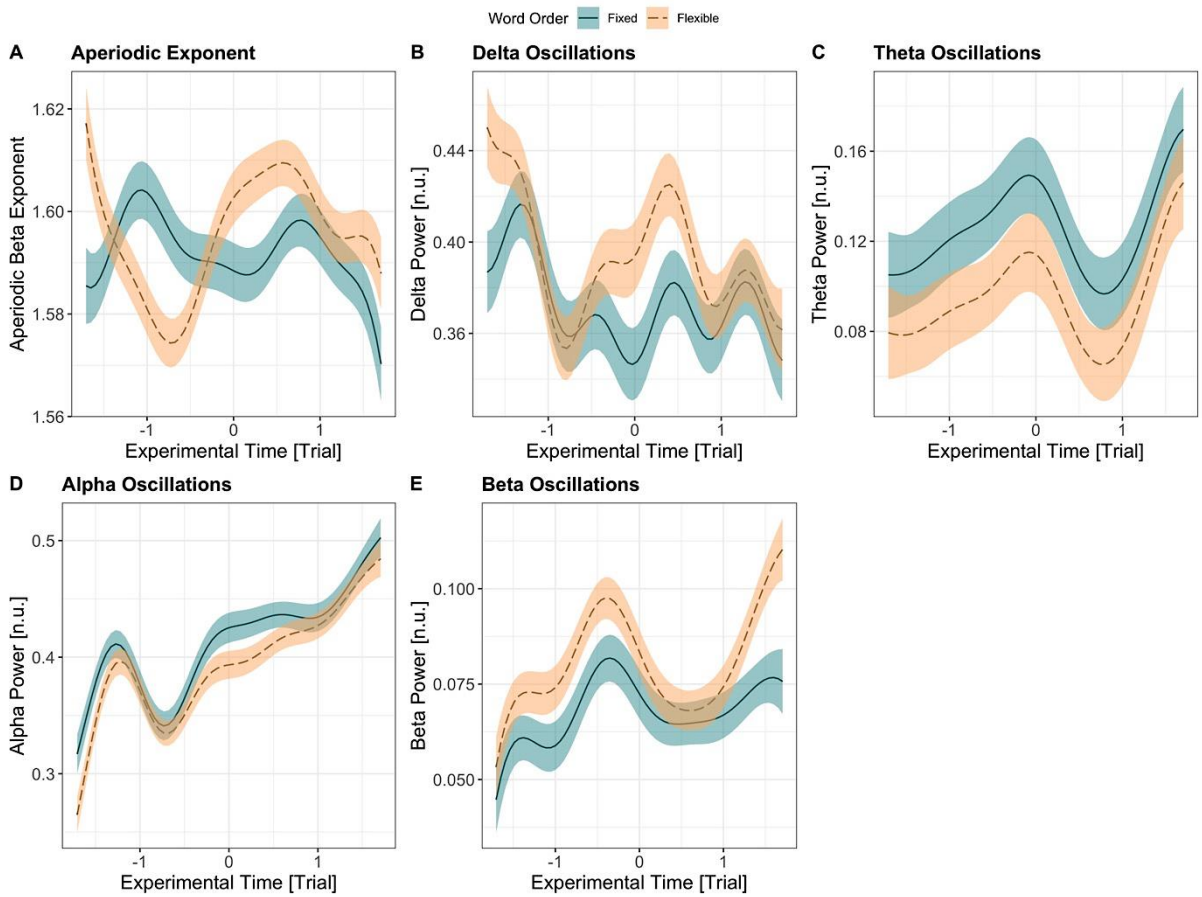

Fig 3 | Modelled effects at channel Cz for changes in aperiodic slope (A), delta (B), theta (C), alpha (D) and beta (E) activity across the learning task for fixed (solid line) and flexible (dashed line) word order rules. Trial number (scaled; -1 = start of the task, 1 = end of the task) is represented on the x-axis.

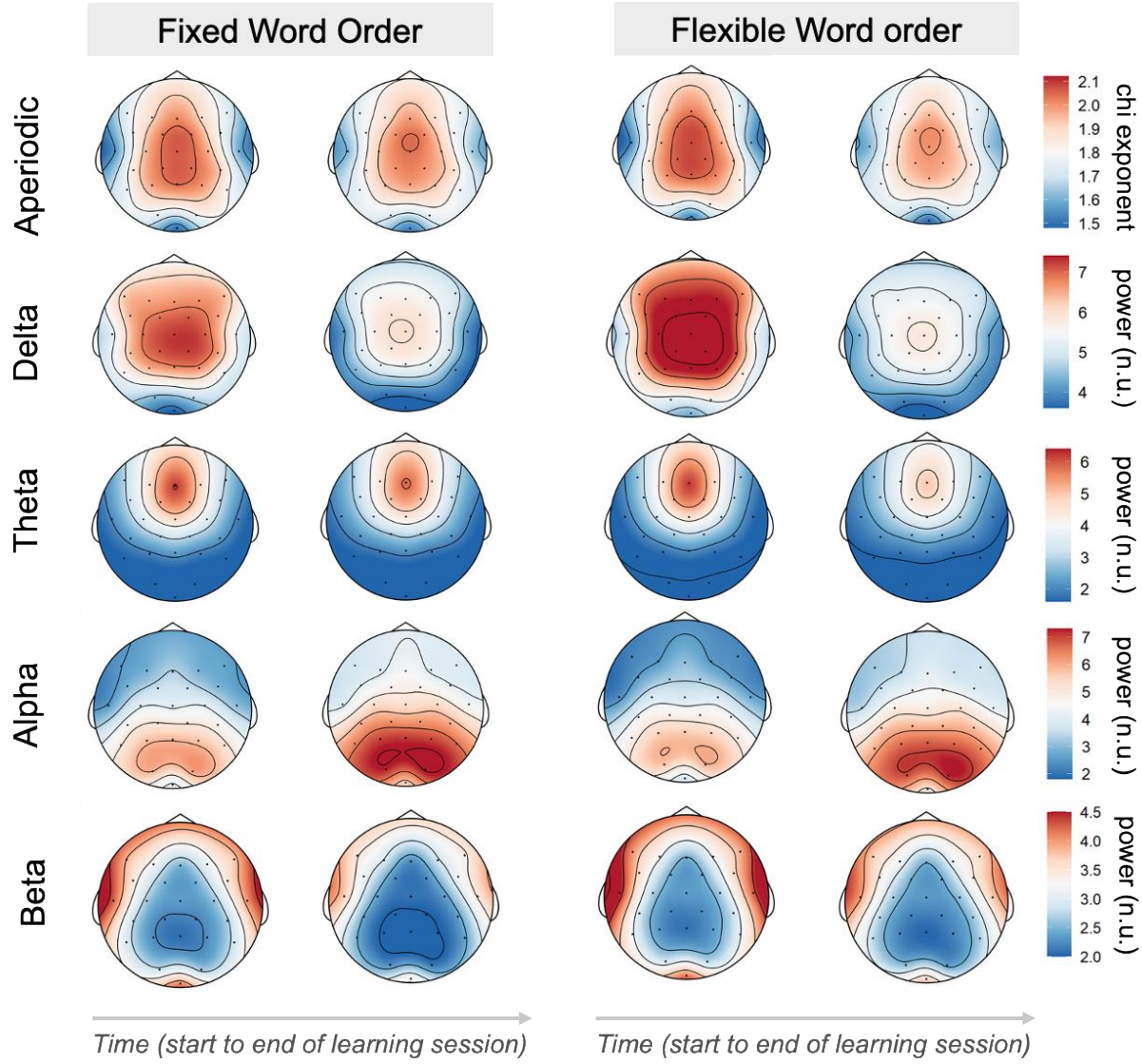

Fig 4 | Difference in topographical distribution of aperiodic and oscillatory activity between fixed and flexible word order sentences at the beginning and end of the sentence learning task.

3.2.2. Task-related aperiodic and oscillatory activity are dynamically related during language learning

Neurophysiological signals are dominated by transient oscillatory and broadband aperiodic activity; however, in the study of the oscillatory correlates of higher-order language processing, aperiodic activity is rarely considered, with little known regarding its influence on task-related oscillatory activity (cf. Cross et al., 2021). Here, we examined the associations between task-related oscillations in individualised (i.e., anchored on participants' IAF) frequency bands and aperiodic activity during the learning of different word order rules. There was a significant positive association

between delta power and the aperiodic slope ($\beta = 0.61$, $p < .001$, $R^2 = .23$); however, this did not vary by sentence type. A similar pattern was observed for theta power ($\beta = 0.62$, $p = .003$, $R^2 = .15$); however, there was no association between alpha power and the aperiodic slope ($\beta = -0.52$, $p = .26$, $R^2 = .01$). Finally, there was a significant negative association between task-evoked beta power and the aperiodic slope ($\beta = -0.84$, $p < .001$, $R^2 = .57$), but as with the lower frequency bands, this effect did not vary by sentence type (for a visualisation of these associations, see Figure 5). These results indicate that differing levels of aperiodic activity may lead to false inferences about oscillatory components. Based on this, we now examine whether such aperiodic and (putative) oscillatory activity interact to predict performance on the sentence judgement task.

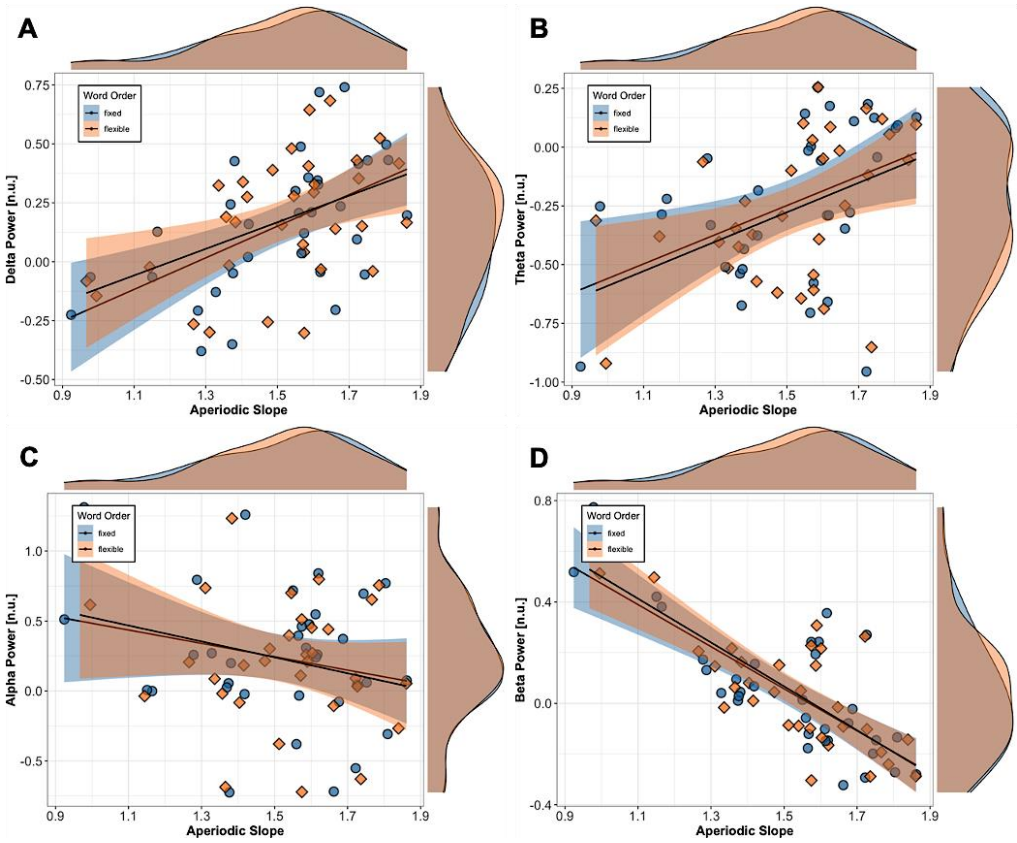

Fig 5 | Association between task-related aperiodic slope and oscillatory power in the delta (A), theta (B), alpha (C) and beta (D) bands during the sentence learning task. The aperiodic slope is represented on the x-axis (higher values indicate a steeper slope), while oscillatory power is represented on the y-axis (higher values indicate higher power). The density of observations for frequency band power and the aperiodic slope are indicated on the margins of each plot, while the fixed and flexible word order sentences are coded in blue and orange, respectively.

3.2.3. Interactions between oscillatory and aperiodic activity modulate language learning

Given the association between task-related oscillatory and aperiodic activity during the learning of complex linguistic rules, we now examine whether the $1/f$ slope and oscillatory activity interact during learning to influence behavioural performance on the sentence judgement task. For the delta model, there was a significant Power x Sentence Type ($\chi2(1) = 4.54$, $p = .03$) and $1/f$ Slope x Sentence Type interaction ($\chi2(1) = 40.94$, $p < .001$). As shown in Figure 6A, when the $1/f$ slope was steepest and delta power was low, d' was higher for flexible relative to fixed word orders. By contrast, when the $1/f$ slope was shallow, performance for flexible word orders was at chance, indicated by a d' estimate of 0, while d' estimates for fixed word orders remained at 1, indicating moderate performance. For the theta model, there was also a significant Power x Sentence Type ($\chi2(1) = 14.09$, $p < .001$) and $1/f$ Slope x Sentence Type interaction ($\chi2(1) = 57.18$, $p < .001$). Here, when the $1/f$ slope was shallow, d' scores were higher for fixed sentences, while the inverse was observed for flexible sentences (Figure 6B). When the $1/f$ slope was steepest, performance for both fixed and flexible sentences increased, independent of theta activity. All higher-order interactions were non-significant (for full model summaries see the supplementary material).

For the alpha model, there was a significant three-way Power x $1/f$ Slope x Sentence Type interaction ($\chi2(1) = 39.05$, $p < .001$). As depicted in Figure 6C, when the $1/f$ slope was shallow, as alpha power increased, d' scores were higher for flexible word orders, while there was no association between alpha activity and d' for fixed word order sentences. When the $1/f$ slope was steepest, this pattern of results reversed, such that as alpha power increased, so did d' scores for fixed word order sentences. Similarly, the beta model yielded a significant Power x $1/f$ Slope x Sentence Type interaction ($\chi2(1) = 22.20$, $p < .001$). When the $1/f$ slope was low, there was no association between beta power and flexible word orders; however, decreased beta power was associated with higher d' scores for fixed word order sentences (Figure 6D). By contrast, when the $1/f$ slope was steep, increased beta power predicted higher d' scores for fixed but lower d' scores for flexible word order sentences, respectively.

Together, these results suggest that when there is a steeper $1/f$ slope, increased alpha and beta power were associated with better behavioural performance, and thus better learning outcomes for fixed

relative to flexible word order sentences. Further, when the $1/f$ slope was steep, there was a general increase in performance for both fixed and flexible word order sentences, relative to when the $1/f$ slope was shallow.

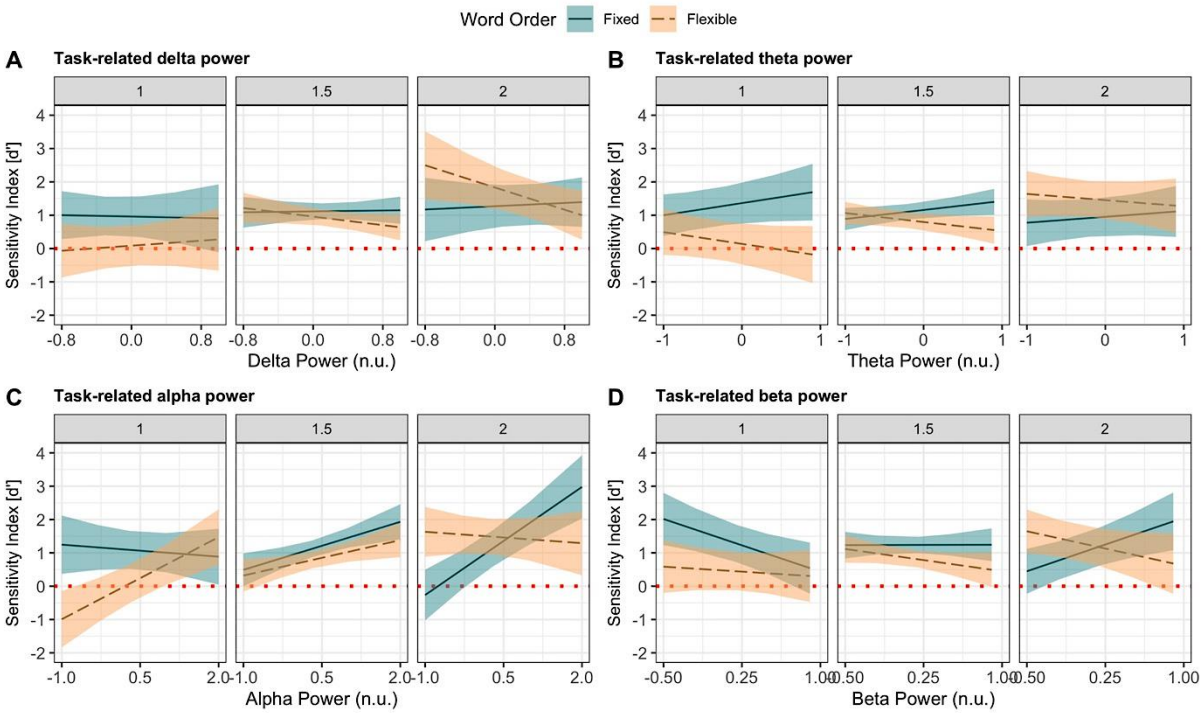

Fig 6 | Visualisation of the relation between behavioural performance, aperiodic slope, and oscillatory activity. Modelled effects of task-related oscillatory activity (x-axis; higher values indicate greater power) on d' scores (y-axis; higher values indicate better performance) for fixed and flexible word order sentences (fixed = solid line; flexible = dashed line). Task-related aperiodic $1/f$ slope estimates are faceted from shallow (left) to steep (right). The red dashed line indicates chance-level performance, while the shaded regions indicate the 83% confidence interval. The x-axis reflects scaled single-trial oscillatory power estimates, with negative values (i.e., -1) reflecting a decrease in power and positive values reflecting an increase in power.

4. DISCUSSION

Here, we estimated the $1/f$ slope during artificial grammar learning to characterise the influence of dynamic alterations in aperiodic and oscillatory activity on higher-order cognition. 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 both fixed and flexible word orders; (2) the

$1/f$ slope becomes steeper during the learning of complex rules, but this effect differed on the type of rules being learned. These findings speak strongly to the view that $1/f$ activity should not be baseline subtracted from neurophysiological signals but analysed in conjunction with oscillatory activity to predict behaviour.

Only few studies have examined the neural oscillations involved in grammar learning (e.g., de Diego-Balaguer et al., 2011; Kepinska et al., 2017), demonstrating that alpha/beta synchronisation and gamma phase synchrony between cortical areas predicts learning. Here, we found that alpha/beta differentially contribute to the learning and processing of fixed and flexible word orders. Increased alpha power predicted an increase in performance for fixed word orders when the aperiodic $1/f$ was steep, while a decrease in alpha power predicted higher performance for flexible word orders when aperiodic activity was high. Similarly, increased beta power was associated with higher performance for fixed sentences, while reduced beta power was associated with higher performance for flexible word orders.

These findings are consistent with previous work (e.g., Kepinska et al., 2017), but reveal fine-grained patterns of spectral activity between word order variations, which may be explained by cue-integration-based models of 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 probabilities of the language are integrated to comprehend incoming linguistic input (e.g., sentences). Here, fixed word orders 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), increased beta power likely reflected the propagation of top-down predictions during the learning of fixed word orders. 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. Therefore, due to the strong sequence dependence in fixed word orders, precision-weighted predictions would likely increase linearly across the sentence, manifesting in increased beta power (Arnal, 2012; Cross et al., 2018; Lewis & Bastiaansen, 2015).

The inverse relationship with flexible word order processing – which was predicted by a reduction in beta power– can also be explained under this framework. Given that flexible word orders 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), a reduction in beta power during flexible word order processing likely reflected prediction errors and internal model updating. That is, beta desynchronization during the learning of flexible word orders may have reflected internal model updating based on mismatches with predicted and actual sensory input, while an increase in beta power 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 increases for fixed relative to "random" sequences (Crivelli-Decker et al., 2018).

Alpha activity showed a similar pattern: when the $1/f$ slope was steep increased alpha power predicted fixed word order processing, while reduced alpha power predicted flexible word order processing. Alpha power reductions during language comprehension may reflect goal-directed processing and enhanced allocation of attentional resources, which is required for the successful learning of flexible word orders (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 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, a decrease in alpha power likely 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.

We also found that an increase in theta power predicted performance for fixed but not flexible word orders. Theta power predicts memory for temporally close stimuli (e.g., in temporal associative

memory tasks; for review, see Herweg et al., 2020), and in encoding novel grammatical rules (e.g., Kepinska et al., 2017). Increases in theta power reported here for fixed word orders may have reflected the learning and integration of word-order-based rules, which involve the decoding and combination of words that are adjacent in a sentence. Indeed, such theta effects have been reported during native sentence processing (Lam et al., 2016). These effects are also consistent with the general memory literature: retrieval of language (e.g., single words), and shape/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 over medial temporal and prefrontal cortices (Guderian & Düzel, 2005), indexing the activation of relevant memory traces and executive control processes, respectively.

We also demonstrated that aperiodic activity is associated with language learning and performance on a grammaticality judgement task. This finding is supported by several studies demonstrating the influence of aperiodic activity on a range of cognitive computations, including processing speed (Ouyang et al., 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 & 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: an increase in beta power predicted more sensitive behavioural responses for fixed sentences, while reduced beta predicted performance for flexible word orders. These findings can be explained by integrating two perspectives: the “spectral fingerprints” hypothesis (Hanslmayr & Staudigl, 2014; Keitel & Gross, 2016; Siegel et al., 2012; Watrous et al., 2015; Womelsdorf et al., 2014) and generalised predictive coding (Friston, 2010, 2018, 2019).

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 have 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 prediction errors (Bressler & Richter, 2015; 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 for incoming stimuli, thus minimizing prediction error at lower levels of the cortical hierarchy. 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 may partially reflect cortical excitability across the frequency spectrum to minimize prediction error during language learning and sentence processing.

Taken together, we have demonstrated that oscillatory and aperiodic activity jointly predict the learning of higher-order language. There are, of course, several open questions that arise from these results. For example, how do interactions between oscillatory and aperiodic activity relate to individual differences in atypical populations, such 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. 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.

Code Accessibility: All data and code will be made publicly available upon publication of this manuscript. This experiment was not preregistered.

REFERENCES

Alday, P. M. (2019). How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology*, *56*(12), e13451.

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.

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*

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. (2018). Toward a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology*, *55*(7), e13064. *https://doi.org/10.1111/psyp.1306*

Corcoran, A. W., Macefield, V. G., & Hohwy, J. (2021). Be still my heart: Cardiac regulation as a mode of uncertainty reduction. *Psychonomic Bulletin & Review*, 1-13. doi: https://doi.org/10.3758/s13423-021-01888-y

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. (2021). Spindle-slow oscillation coupling during sleep predicts sequence-based language learning. *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. (2018). 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., Chatburn, A., Alday, P. M., Coussens, S., & Kohler, M. J. (2020b). Individual alpha frequency modulates sleep-related emotional memory consolidation. *Neuropsychologica*, *148*. doi: https://doi.org/10.1101/202176

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. *Behavior Research Methods*. doi: https://doi.org/10.3758/s13428-020-01473-6

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.2019.11.00

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

Donoghue, T., Schaworonkow, N., & Voytek, B. (2021). Methodological considerations for studying neural oscillations. *European Journal of Neuroscience*. doi: https://doi.org/10.1111/ejn.15361

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. (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

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

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

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

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

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

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. (2013). An algorithm for the EEG frequency architecture of consciousness and brain body coupling. *Frontiers in human neuroscience*, *7*, 766. doi: https://doi.org/10.3389/fnhum.2013.00766

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

Klimesch, W.. (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

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

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, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and language, 148*, 51-63.

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*

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.

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).

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*

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.2561-17.201*

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. (2017). 1/f neural noise is a better predictor of schizophrenia than neural oscillations. *BioRxiv*, 113449.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184-187.

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 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/nrn313*

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

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.376*

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

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