Spindle-slow oscillation coupling during sleep predicts sequence-based language learning

Zachariah R. Cross[1*], Randolph F. Helfrich[2], Andrew W. Corcoran[1,3], Mark J. Kohler[1, 4], Scott Coussens[1], Lena Zou-Williams[1], Matthias Schlesewsky[1], M. Gareth Gaskell[5], Robert T. Knight[6,7], Ina Bornkessel-Schlesewsky[1]

[1]Cognitive and Systems Neuroscience Research Hub, University of South Australia, Adelaide, Australia.
[2]Center for Neurology, Hertie-Institute for Clinical Brain Research, University of Tübingen, Germany.
[3]Cognition and Philosophy Laboratory, Monash University, Melbourne, Australia.
[4]Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, Australia.
[5]Department of Psychology, University of York, York, United Kingdom.
[6]Department of Psychology, UC Berkeley, Berkeley, California, USA.
[7]Helen Wills Neuroscience Institute, UC Berkeley, Berkeley, California, USA.

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

Manuscript details

Number of pages:	36
Number of figures:	8
Abstract word count:	173
Introduction word count:	2,599
Discussion word count:	2,513
Data available at:	TBC
Code available at:	TBC

Author Contribution: Conceptualization by Z.C., I.B-S., M.S. and M.J.K. Data curation by Z.C. and L.Z.W. Data pre-processing and analysis by Z.C., S.C., R.F.H., and A.W.C. Writing – original draft by Z.C.; writing – review & editing by all authors; Supervision of project by I.B-S., M.S., M.J.K and R.F.H. Visualization by Z.C.; Funding acquisition by I.B-S. and Z.C.

Funding: Preparation of this work was supported by Australian Commonwealth Government funding under the Research Training Program (RTP; number 212190) and Maurice de Rohan International Scholarship awarded to ZC. IB-S is supported by an Australian Research Council Future Fellowship (FT160100437). AWC and LZ-W are supported by Australian Government RTP scholarships. RTK is supported by an NIH RO1NS21135, while RFH is supported by a Feodor-Lynen Fellowship by the Alexander-von-Humboldt Foundation. This work was also supported by a UK ESRC grant (ES/N009924/1) awarded to Lisa Henderson and Gareth Gaskell.

Conflict of interest

The authors declare no competing financial interests.

ABSTRACT

Language is one of the most defining human capabilities, involving the coordination of brain networks that generalise the meaning of linguistic units of varying complexity. On a neural level, neocortical slow oscillations and thalamic spindles during sleep facilitate the reactivation of newly encoded memory traces, manifesting in distinct oscillatory activity during retrieval. However, it is currently unknown if the effect of sleep on memory extends to the generalisation of the mechanisms that subserve sentence comprehension. We address this question by analysing electroencephalogram data recorded from 35 participants during an artificial language learning task and an 8hr nocturnal sleep period. We found that a period of sleep was associated with increased alpha/beta power and improved behavioural performance. Phase amplitude coupling analyses also revealed that spindle-slow oscillation coupling predicted the consolidation of sequence-based word orders, which was associated with distinct patterns of oscillatory activity during sentence processing. Taken together, this study presents converging behavioural and neurophysiological evidence for a role of sleep in the consolidation of higher order language learning and associated oscillatory neural activity.

Keywords: Sleep and memory; language learning; sentence processing; neural oscillations; cross-frequency coupling

1. INTRODUCTION

The human brain is highly adept at extracting statistical regularities from its environment. This process of statistical learning is pivotal in generating knowledge of our surroundings and in deriving predictions that are critical for perceptual, cognitive and motor computations (Durrant, Cairney, & Lewis, 2016; Toth et al., 2017). Probabilistic contingencies are encountered at multiple spatiotemporal scales, as exemplified by differences in the complexity of generalisations (e.g., concrete vs. abstract categories), and in the way discrete units of information are combined (e.g., local associations vs. networks built on local associations; Fiser & Aslin, 2002; Perruchet & Pacton, 2006). From this perspective, predictions about sensory input that follow conditional regularities not only rely on mechanisms of associative memory, but also complex, hierarchical dependencies built at least in part on local statistical information. The successful generalisation of such complex environmental regularities has been shown to rely on sleep-based memory consolidation (e.g., Lewis, Knoblich, & Poe, 2018; Lutz, Wolf, Hubner, Born, & Rauss, 2018). However, it remains unclear if sleep-related memory processing extends to rules containing both complex structure and meaning, as in language at the sentence level (Cross et al., 2018). Here, we investigated whether sleep predicts the generalisation of an artificial modified miniature language (MML; Cross et al., 2020a). Using a nocturnal sleep paradigm with full-scalp electroencephalography, non-rapid eye movement (NREM) sleep spindle-slow oscillation coupling events were used to predict sensitivity to various word order rules and task-related oscillatory activity.

1.1. Sleep-based memory consolidation

A plethora of behavioural and neuroimaging evidence (for review, see Rasch & Born, 2013) demonstrates that sleep plays an active role in memory by consolidating and generalising mnemonic information. This dynamic account of the sleeping brain is supported by the Active System Consolidation hypothesis (ASC; Born & Wilhelm, 2012; Klinzing, Niethard, & Born, 2019). Core to ASC is that sleep facilitates repeated reactivation of encoded memory representations (Rasch & Born, 2013). This reactivation is dependent on cortical glutamatergic synapses, which become weak during prolonged wakefulness (Kavanau, 1997; Rasch & Born, 2013). ASC posits that, due to a disengagement from external input, sleep provides a suitable environment for internal processes to facilitate localised synaptic downscaling and the transfer of information into long-term memory (Born & Wilhelm, 2012). This hypothesis is supported by electrophysiological evidence that learned sequences are replayed during sleep, potentially via the coupling between fast and slow oscillatory events (Helfrich et al., 2019; Staresina et al., 2015). Such oscillatory coupling mechanisms are central to ASC, whereby thalamic spindles are coupled to the upstate of slow oscillations, providing a temporal receptive window for the replay of hippocampal memory traces and their transfer to the neocortex for long-term storage (Helfrich et al., 2018; Maingret et al., 2016). Critically, the transfer of newly encoded information from the hippocampal complex to the neocortex enables semanticization and generalisation of mnemonic information, enabling the neocortex to gradually learn the (statistical) regularities of sensory input (McClelland, McNaughton, & O'Reilly, 1995; Moscovitch et al., 2016), a process known to support language learning and comprehension (e.g., Davis & Gaskell, 2009).

1.2. Sleep and language learning

Mechanisms of sleep-based memory consolidation have been associated with the generalisation of various types of information, with recent work showing sleep-based effects on language, including novel-word learning (Bakker, Takashima, van Hell, Janzen, & McQueen, 2015; James, Gaskell, Weighall, & Henderson, 2017; Mirković & Gaskell, 2016) as well as the generalisation of grammatical rules (Batterink, Oudiette, Reber, & Paller, 2014; Nieuwenhuis et al., 2013). In addition to improving

behavioural performance, effects of sleep on language learning manifest in distinct oscillatory activity (for review: Schreiner & Rasch, 2016) and event-related potentials (ERPs). For example, increased scalp-recorded theta power (~ 4–7 Hz) occurs during the recognition of newly learned words compared to unlearned words after sleep (Bakker et al., 2015; Schreiner, Goldi, & Rasch, 2015), while the interaction between NREM and rapid eye movement (REM) sleep modulates the amplitude of language-related ERPs (N400, late positivity) during the processing of novel grammatical rules (Batterink et al., 2014). These findings are consistent with the idea that sleep changes the qualitative representation of learned information (Wagner et al., 2004), abstracting rules from the corpus of learned information and strengthening associations between complex regularities (Lewis, Knoblich & Poe, 2018). As such, sleep may facilitate language learning by abstracting the meaning (e.g., for words) and structural regularities (e.g., of the underlying grammar) of newly learned linguistic information, enabling the brain to generalize newly learned information to novel linguistic input. Taken together, by demonstrating sleep-related consolidation effects for linguistic stimuli of varying complexity, these findings have begun to elucidate the role of sleep-related memory consolidation in various aspects of language learning (Rasch, 2017).

Importantly, however, language comprehension involves much more than the recollection of single words. For example, during sentence processing, individual items (words) are retrieved from long-term memory and incrementally combined; the interpretation of the combined elements is constrained by semantic and grammatical regularities (Bornkessel & Schlesewsky, 2006; Hagoort, 2016; Meyer et al., 2015). Further, the learning and consolidation of the rules that govern sentence processing may be supported by sleep-based memory processing, as has been evidenced with less complex aspects of language (e.g., single word learning; James, Gaskell, Weighall, & Henderson, 2017; Mirković & Gaskell, 2016) and sequence learning more generally (Lutz et al., 2018). However, there is limited empirical evidence to date supporting a link between sleep-related oscillatory activity and sentence-level learning. Further, it is also unclear how an effect of sleep may manifest in task-related neural dynamics during sentence comprehension. Interestingly, aspects of NREM sleep predicts the extraction of temporal-based rules (Lerner & Gluck, 2019), akin to sequence-based processing present in many languages of the world (e.g., English, Dutch; Bates et al., 2001; Bornkessel-Schlesewsky et al., 2011; MacWhinney et al., 1984). From this perspective, sleep may benefit the consolidation of sequence-based regularities in language, compared to rules that do not draw on sequential information (Bornkessel-Schlesewsky et al., 2015), and this may be achieved through neurophysiological mechanisms present during NREM sleep (Born & Wilhelm, 2012), such as spindle-slow oscillation coupling (Hahn et al., 2020; Kim et al., 2017; Mikutta et al., 2019).

1.3. The neurobiology of sentence processing: a cross linguistic perspective

Understanding who is undertaking the action (i.e., which sentence participant is the "actor") is a crucial aspect of sentence comprehension (Bornkessel-Schlesewsky & Schlesewsky, 2016). There are a number of information sources ("cues") within a language or sentence that are used to identify the actor. For example, in English, the most prominent cue is the order of the words, where the first participant in the sentence is generally the actor (e.g., *the boy pushed the girl* indicates that the boy is the one doing the pushing; Bates et al., 2001; MacWhinney et al., 1984). However, in other languages, additional cues may play a more important role during actor identification, such as animacy (i.e., the more animate participant is typically the one undertaking the action) or word forms or endings known as case marking (Bornkessel-Schlesewsky et al., 2011). For example, the German sentence *Das Mädchen schubste der Junge,* verbatim *the-accusative girl pushed the-nominative boy,* means the same as the English example above (i.e. the boy is the one doing the pushing). This is made clear by the form of the articles. Accordingly, languages such as English are argued to be sequence-dependent (i.e., use fixed word order rules), whereas languages that rely on other cues such as animacy for interpretation

and rely less on the order of words are thought to be less sequence-dependent (i.e., rely on word-order-independent cues and thus permit a more flexible word order; MacWhinney et al., 1984; Bornkessel & Schlesewsky 2006; Bornkessel-Schlesewsky & Schlesewsky 2013; Bornkessel-Schlesewsky, Schlesewsky et al., 2015).

While research has focussed on the distinction between the processing of fixed and flexible word orders and their neuroanatomical substrates (e.g., Bornkessel-Schlesewsky et al., 2011; Kandylaki et al., 2016; Tune et al., 2016), considerably less is known regarding the electrophysiological mechanisms subserving their initial encoding and generalisation (cf. Cross et al., 2018). Given that sleep plays a critical role in memory consolidation and neural plasticity, sleep may play a key role in forming the knowledge necessary for sentence comprehension, and this may be reflected in neural activity that indexes both successful (sleep-based) memory consolidation, as well as the processing of fixed and flexible word order rules. Fixed word orders require a linear order-based processing strategy (i.e., the linear position of each element in the sentence), and thus likely rely on sequence-based processing for accurate interpretation (Bornkessel-Schlesewsky, Schlesewsky et al., 2015). By contrast, flexible word orders rely more heavily on sequence-independent cues (e.g., animacy) in order to combine non-adjacent elements in a sentence. Critically, research on sleep-based sequence learning has shown that, relative to wake, sleep increases prediction strength for visual sequences (Lutz et al., 2018), as well as facilitates the abstraction and integration of statistical regularities in continuous speech streams (Earle & Myers, 2015; Fenn et al., 2013) and simple grammatical rules that follow temporal regularities (Batterink & Paller, 2015; Batterink et al., 2014; Rasch, 2017). Such sleep-based effects on language learning at the sentence level may be observed in endogenous neural oscillatory activity, given the prominence of neural oscillations during both wake and sleep states (e.g., Buzsáki, 1996; Canolty & Knight, 2010; Helfrich et al., 2019). Neural oscillations not only support basic physiological mechanisms (e.g., regulation of neuronal population activity), but also domain-general cognitive computations, including attention (Helfrich & Knight, 2016), predictive coding (Bastos et al., 2020), and memory encoding/retrieval (Hasselmo, & Stern, 2014; Osipova et al., 2006; Tort, Komorowski, Manns, Koell, & Eichenbaum, 2009), and domain-specific processes, including language (for reviews, see Hauk, Giraud, & Clarke, 2017, and Meyer, 2018).

1.4. Neural oscillations as windows onto the sleep-based consolidation of language

While the study of the electrophysiology of sentence comprehension has largely focussed on ERPs (for overviews, see Bornkessel-Schlesewsky & Schlesewsky, 2016, 2019; Osterhout & Holcomb, 1995), recent evidence demonstrates that sentence processing is associated with dynamic activity in different frequency bands (e.g., Bastiaansen et al., 2002; Kaufeld et al., 2020; Kepinska, Pereda, Caspers, & Schiller, 2017; Lam, Schoffelen, Udden, Hulten, & Hagoort, 2016; Roehm et al., 2001; 2004). For example, linear beta power (~ 13 – 30 Hz) increases have been observed during the presentation of grammatical sentences (Bastiaansen & Hagoort, 2015; Lam et al., 2016; Wang et al., 2012), argued to reflect the propagation of top-down predictions during sentence comprehension (Lam, et al., 2016; Lewis & Bastiaansen, 2015; Wang et al., 2012). From a language learning perspective, beta power may increase linearly as a function of learning, reflecting the accumulation of knowledge based on repeated exposure to the regularities of the linguistic input (Cross et al., 2018). As such, sentences containing violations (e.g., ungrammatical word orders) may elicit a decrease in beta power, reflecting mismatches between sensory input and internal representations (i.e., knowledge) of the rules. Further, given the role of sleep in supporting memory (Rasch & Born, 2013), prediction strength (i.e., reflecting greater knowledge of the language) may be enhanced after sleep, manifesting in increased beta power relative to an equivalent period of wake.

Similar research (Kepinska, Pereda, Caspers, & Schiller, 2017) has demonstrated increased alpha power during the learning of an artificial language in individuals with high language learning

aptitudes, possibly reflecting more efficient and less effortful processing demands due to accelerated learning (Cross et al., 2018; Kepinska et al., 2017). These findings accord with proposals that alpha oscillations modulate the activation of task-relevant cortical regions (Ichihara-Takeda et al., 2015; Jiang, van Gerven, & Jensen, 2015; Mathewson et al., 2011), facilitating the flow of information through thalamo-cortical networks (Klimesch, Sauseng, & Hanslmayr, 2007), promoting successful long-term memory encoding and retrieval (Klimesch, 1997, 1999; Klimesch, Doppelmayr et al., 1999; Klimesch, Schimke, et al., 1990). However, how alpha activity changes as a function of sleep-associated consolidation remains unclear. If sleep strengthens memory for temporal sequences and statistical regularities, and increased alpha power reflects more efficient information processing during sentence processing, then alpha power should increase after a period of sleep and coincide with enhanced behavioural outcomes for successfully learned grammatical rules.

Theta activity (~ 4 – 7 Hz) has also been observed during sentence-level processing. Hippocampal theta power has been shown to increase for constrained (e.g., *she locked the door with the [key]*) relative to unconstrained contexts (Piai et al., 2016; e.g., *she walked in here with the …*), while increased frontal theta power has been reported during embedded-antecedent retrieval in German (e.g., English translation: *during the rehearsal, the conductor[feminine] that the pianist[masculine] irritated interrupted the prelude before the concert, and he complained*; Meyer, Grigutsch, Schmuck, Gaston, & Friederici, 2015). Increases in theta power have also been observed during the learning of trisyllabic sequences in an artificial language (de Diego-Balaguer, Fuentemilla, & Rodriguez-Fornells, 2011). From this perspective, theta oscillations may reflect contextual-based computations during language processing, which may be strengthened during sleep via reactivation of stimulus-relevant memory traces (Cross et al., 2018). As such, an increase in theta power may be observed during the processing of flexible word order rules, and this effect would likely be enhanced after sleep compared to wake.

Taken together, these lines of evidence suggest that during sentence processing, alpha oscillations coordinate information processing in neural networks that govern attention and memory (Klimesch et al., 2007; Klimesch, 2012), while theta oscillations may index the binding of information to form coherent sentential memory representations (Cross et al., 2018; de Diego-Balaguer et al., 2011; Piai et al., 2016). Finally, beta oscillations have been argued to support prediction-based processing during language comprehension (Lewis & Bastiaansen, 2015; Lewis et al., 2015). However, while the oscillatory correlates of fluent sentence processing are being increasingly well-characterised, the mechanisms underlying long-term consolidation of sentence-level regularities, including during sleep-associated consolidation, remain unclear.

1.5. The current study

To better characterise the sleep-based and task-related neural mechanisms underlying complex rule learning in language, we analysed EEG during an 8-hour nocturnal sleep period and sentence judgement tasks. Participants learned the artificial miniature language Mini Pinyin (Cross et al., 2020a), which contains two main sentence constructions, namely fixed and flexible word orders.

We used non-linear directional cross-frequency coupling analyses in combination with phase-dependent correlation measures to quantify the temporal relationship between slow oscillations and spindles during NREM sleep (for a similar approach, see Helfrich et al., 2018; Staresina et al., 2015). We also quantified spectral activity in the theta, alpha and beta bands using complex Morlet wavelets during sentence judgement tasks. These estimates were used to independently predict behavioural performance and were analysed in combination with spindle-slow oscillation coupling metrics to determine whether task-related EEG is modulated by mechanisms of sleep-associated memory consolidation. We hypothesised that task-related theta, alpha and beta activity and NREM spindle-slow

oscillation coupling would interact in their effect on language learning and sentence processing. Specifically, we predicted that:

H^1: Behavioural sensitivity (d' prime) to fixed word orders (i.e., sequence-based rules) would be improved after sleep compared to an equivalent period of wake;

H^2: Immediately after learning, higher behavioural sensitivity to grammatical rules would be associated with reduced beta and alpha power, while an increase in alpha and beta power would be predictive of behavioural performance after sleep compared to wake;

H^3: Higher behavioural sensitivity to flexible word orders relative to fixed word orders would be associated with an increase in task-related theta power, and this effect would be accentuated after sleep compared to wake;

H^4: An increase in behavioural sensitivity to fixed word order rules would be predicted by stronger spindle-slow oscillation coupling, and;

H^5: Spindle-slow oscillation coupling will interact with task-related oscillatory activity to predict higher behavioural sensitivity to grammatical rules.

2. MATERIALS AND METHODS

2.1. Participants

We recruited 36 right-handed participants who were healthy, monolingual, native English-speakers (16 male) ranging from 18 – 40 years old (mean = 25.4, SD = 7.0). Participants were randomly assigned to either a Sleep (n = 18) or Wake condition. All participants reported normal or corrected-to-normal vision, no history of psychiatric disorder, substance dependence, or intellectual impairment, and were not taking medication that influenced sleep or neuropsychological measures. All participants provided informed consent and received a $120 honorarium. One participant from the Sleep condition was removed from the analysis due to technical issues during the experimental tasks and sleep period, resulting in a total sample size of 35 (mean age = 25.4, SD = 7.10; 16 males; Sleep n = 17). Ethics approval was granted by the University of South Australia's Human Research Ethics committee (I.D: 0000032556).

2.2. Screening and control measures

The Flinders Handedness Survey (FLANDERS; Nicholls, Thomas, Loetscher, & Grimshaw, 2013) was used to screen handedness, while the Pittsburgh Sleep Quality Index (PSQI; Buysse, Reynolds, Monk, Berman, & Kupfer, 1988) screened for sleep quality. PSQI scores ranged from 1-5 (M = 2.9, SD = 1.33) out of a possible range of 0 – 21, with higher scores indicating worse sleep quality. Prospective participants with scores > 5 were unable to participate. As an additional control, the Stanford Sleepiness Scale (SSS) was administered at the beginning and end of the experiment to measure self-perceived sleepiness.

2.3. Electroencephalography

The electroencephalogram (EEG) was recorded during the learning and sentence judgement tasks and sleep durations 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 reference was located at FCz, with EEG signals re-referenced to linked mastoids offline. The ground electrode was located at AFz. The electrooculogram (EOG) was recorded via electrodes located at the outer canthus of each eye (horizontal EOG) and above and below participants' left eye (vertical EOG). Sub-mental electromyography (EMG) was added to facilitate accurate scoring of sleep periods. 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.

2.4. Vocabulary and structure of Mini Pinyin

Stimuli consisted of sentences from a modified miniature language modelled on Mandarin Chinese, named Mini Pinyin (Cross et al., 2020a). Mini Pinyin contains 32 transitive verbs, 25 nouns, 2 coverbs, and 4 classifiers. The nouns included 10 human entities, 10 animals and 5 objects (e.g., *apple*). Each category of noun was associated with a specific classifier, which always preceded each of the two noun phrases in a sentence. As illustrated in Figure 1B, *ge* specifies a human noun, *zhi* for animals, and *xi* and *da* for small and large objects, respectively. Overall, Mini Pinyin contains 576 unique sentences (288 grammatical, 288 ungrammatical) which are divided into two equivalent sets (see Cross et al., 2020a for a complete description of the stimuli; for the complete set of stimuli, visit: https://osf.io/f5ksg/?view_only=738200b63e78433eb50e667ea772c08f*)*.

Here we focussed on a subset of sentence conditions to investigate the mechanisms underlying the learning of different word order rules, which fundamentally differs between the languages of the world (for review, see Bates, Devescovi, & Wulfeck, 2001). As discussed earlier, languages like English and Dutch rely primarily on word order, while languages like German and Turkish rely more on cues such as case marking and animacy (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky et al., 2015; MacWhinney, Bates, & Kliegl, 1984). From this perspective, Mini Pinyin enables a comparison between sentences with differing word orders (see Figure 1A), and the influence sleep may have on the respective consolidation of fixed and flexible word order rules. The subset of stimuli in the current analysis contained 96 sentences in the sentence learning task and 144 sentences in the grammaticality judgement tasks. The remaining sentences were considered fillers. These filler sentences included sentences that violated classifier-noun pairs, and thus were not suitable for testing predictions regarding fixed and flexible word order processing (for a full description of all sentence conditions present in Mini Pinyin, please see Cross et al., 2020a).

As is apparent in Figure 1A, sentences that do not contain the coverb *ba* (i.e., actor-verb-undergoer, AVU; undergoer-verb-actor, UVA) yield a flexible word order, such that understanding *who is doing what to whom* is not dependent on the ordering of the noun phrases. Instead, determining *who is doing what to whom* is facilitated by animacy cues. For instance, in the UVA condition, *the bear* is interpreted as the actor despite the first noun phrase being *the apple*, since it is implausible for an apple to eat a bear. Therefore, both AVU and UVA are grammatical constructions. By contrast, sentences such as A*ba*UV yield a fixed word order, such that the inclusion of *ba* strictly renders the first noun phrase as the actor. Note that the positioning of the verb is critical in sentences with and without a coverb. With the inclusion of a coverb, the verb must be placed at the end of the sentence, while the verb must be positioned between the noun phrases in constructions without a coverb.

(A) Grammatical and Ungrammatical Sentence Constructions

	Grammatical		Ungrammatical
Fixed		**Fixed**	
AbaUV	zhi xiong ba xi pingguo chile '(animal) bear ba (small object) eats.'	UbaAV	xi pingguo ba zhi xiong chile '(small object) apple ba (animal) bear eats.'
		AbaVU	zhi xiong ba chile xi pingguo '(animal) bear ba eats (small object) apple.'
Flexible		**Flexible**	
AVU	zhi xiong chile xi pingguo '(animal) bear eats (small object) apple.'	AUV	zhi xiong xi pingguo chile '(animal) bear (small object) apple eats.'
UVA	xi pingguo chile zhi xiong '(small object) apple eats (animal) bear.	UAV	xi pingguo zhi xiong chile '(small object) apple (animal) bear eats.'

English Translation: the bear eats the apple

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

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

Figure 1. (A) Example of grammatical and ungrammatical fixed and flexible word order sentences. Classifiers and nouns are coded in blue, while verbs are red. The coverb *ba* is coded in green. For the ungrammatical sentences (right), the point of violation in the sentence is underlined. The direct English translation for each sentence construction is provided below (i.e., *the bear eats the apple*). **(B)** A sample of the linguistic elements present in Mini Pinyin and their English translation. Note that *ba* does not have a specific meaning, but when present in a sentence, instantiates a strict actor-undergoer-verb word order.

2.5. Experimental protocol

Participants received a paired picture-word vocabulary booklet containing the 25 nouns and were asked to maintain a minimum of 7hrs sleep per night (see Figure 2A for a portion of nouns from the vocabulary booklet). Participants were required to learn the 25 nouns to ensure that they had a basic vocabulary of the nouns in order to successfully learn the 32 transitive verbs. They were asked to record periods of vocabulary learning in an activity log. Participants were instructed to study the booklet for at least fifteen minutes per day and were informed that they would need to pass a vocabulary test before commencing the main experimental protocol. After approximately one week, participants returned to complete the main experimental session, where EEG was recorded during a sentence learning task, and baseline and delayed sentence judgement tasks.

2.6. Vocabulary test

Participants completed a vocabulary test by translating the nouns from Mini Pinyin into English using a keyboard, as illustrated in Figure 3C. Each trial began with a 600ms fixation cross, followed by the visual presentation of the noun for up to 20s. Prospective participants who scored < 90% were unable to complete the main experimental EEG session. As such, all 36 participants included in the current paper obtained over 90% correct on the vocabulary test. The proportion of individuals who did not pass the vocabulary test was small (e.g., approximately less than 5 cases); however, the exact number was not recorded.

2.7. Sentence learning

Sentence and picture stimuli were presented using OpenSesame (Mathot, Schreij, & Theeuwes, 2012). During sentence learning, pictures were used to depict events occurring between two entities. The pictures and entities shown during the learning task were combinations of the static pictures shown in the vocabulary booklet (for an example of booklet versus sentence learning picture stimuli, see Figure 2A and 2B, respectively).

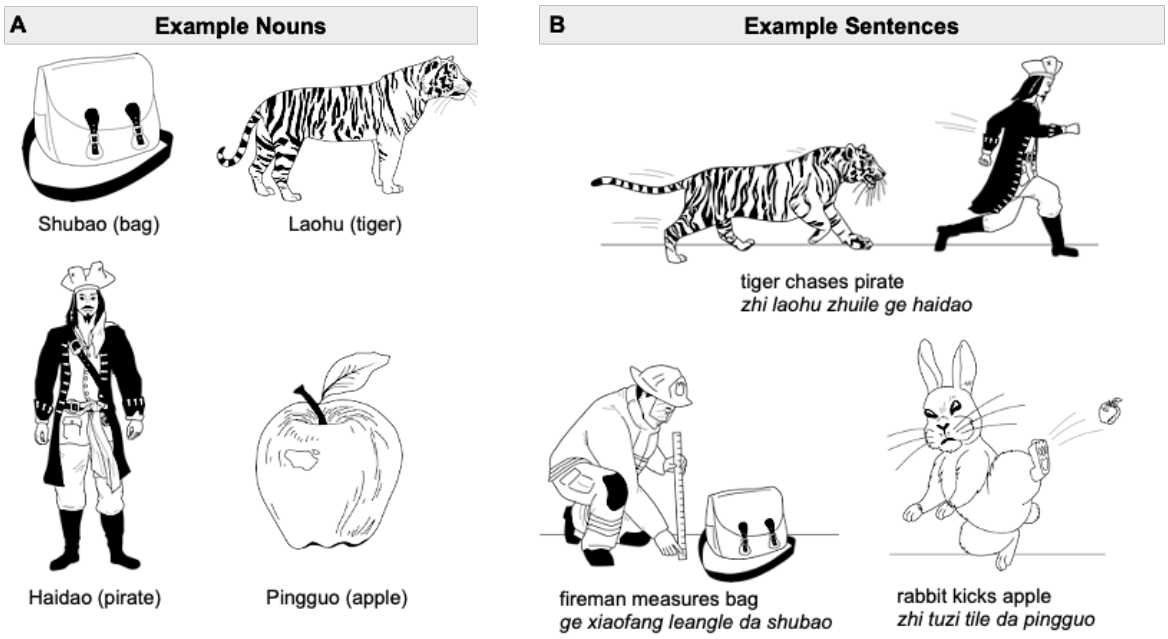

Figure 2. (A) Portion of the 25 illustrations used in the vocabulary booklet, which included human, animal and inanimate objects (i.e., bag, apple). **(B)** Portion of the illustrations used in the sentence learning task, illustrating the interaction between two entities. Note that the entities used in sentence learning task are the same as the illustrations used in the vocabulary booklet.

While participants were aware that they would complete sentence judgement tasks at a later point, no explicit description of or feedback regarding grammatical rules was provided during the learning task, since explicit instructions may differentially impact memory processes (Batterink, Reber, & Paller, 2015; Batterink & Neville, 2013; Weber et al., 2016). Further, participants trained on an artificial language without explicit instruction show more native-like neural signatures of language processing, suggesting closer approximation to natural language learning (Friederici, Steinhauer, & Pfeifer, 2002; Morgan-Short, Steinhauer, Sanz, & Ullman, 2012).

Each picture corresponded to multiple sentence variations (i.e., AVU, UVA, AbaUV), similar to grammatical conditions in Figure 3A. Picture-sentence pairs were presented to participants as correct language input. Participants were presented with a fixation cross for 1000ms, followed by the picture illustrating the event between two entities for 4000ms. A sentence describing the event in the picture was then presented on a word-by-word basis. Each word was presented for 700ms followed by a 200ms ISI. This pattern continued for 128 sentence-picture combinations, 96 of which are reported in the current analysis, until the end of the task, which took approximately 40 minutes. The 96 sentences included in this analysis included the flexible (i.e., AVU, UVA) and fixed (i.e., AbaUV) sentence constructions. Sentences considered as fillers contained a coverb that was not *ba*, and thus were not relevant to testing the predictions posited in the current analysis. During this task, participants were required to learn the structure of the sentences and the meaning of the verbs, classifiers and the coverb *ba*. Stimuli were pseudo-randomised, such that no stimuli of the same construction followed each other,

and each sentence contained a different combination of nouns and verbs. This was done to encourage learning of the underlying grammatical rules rather than episodic events of individual sentences. Further, the two lists of sentences were counterbalanced across participants and testing session (baseline, delayed). Following the sentence learning task, participants completed the baseline sentence judgement task.

2.8. Baseline and delayed judgement tasks

The baseline sentence judgement task taken ~0 hr after learning provided a baseline to control for level of encoding, while the delayed judgement task took place ~12hrs after the learning and baseline judgement tasks. During both judgement tasks, 288 sentences without pictures (144 grammatical, 144 ungrammatical), 156 of which are reported here, were presented word-by-word with a presentation time of 600ms and an ISI of 200ms. The 156 included sentences included a combination of grammatical and ungrammatical flexible (i.e., AVU, UVA) and fixed (i.e., AbaUV) sentence constructions, while the 132 sentences that were considered fillers contained coverbs that were not *ba*, and classifier-noun pair violations, and thus were not relevant to testing the predictions of the current analysis. Thus, while these additional sentences can be used to test other hypotheses related to language learning, they are considered fillers for the purposes of the current analysis. Participants received feedback on whether their response was correct or incorrect during the baseline but not the delayed judgement task. This was to ensure that participants were able to continue learning the language without explicit instruction. Figures 3A and 3B illustrate the sequence of events in the sentence learning and baseline judgement tasks, respectively.

Participants were instructed to read all sentences attentively and to judge their grammaticality (yes/no) via a button-press. As a cue for judgment, a question mark appeared in the centre of the monitor for 4000ms after the offset of the last word. Two lists of sentence stimuli were created, which were counterbalanced across participants and the baseline and delayed sentence judgement tasks. Half of the sentences were grammatical, with each of the grammatical constructions shown an equal number of times. The other half of the sentences were ungrammatical constructions. Stimuli were pseudo-randomised, such that no stimuli of the same construction followed each other.

2.9. Main experimental procedure

For the wake condition, participants completed the vocabulary test and EEG setup at ~08:00hr. The learning task was administered at ~09:00hr, followed by the baseline judgement task, with EEG recorded during both the learning and judgement task. Participants then completed the behavioural control tasks and were free to leave the laboratory to go about their usual daily activities, before returning for EEG setup and the delayed judgement task at ~21:00hr the same day. EEG was also recorded during the delayed judgement task.

Participants in the sleep condition arrived at ~20:00hr to complete the vocabulary test and EEG setup before completing the learning task at ~21:00hr, followed by the baseline judgement task, with EEG recorded during both the learning and judgement tasks. Participants were then given an 8hr sleep opportunity from 23:00hr – 07:00hr. Polysomnography was continuously recorded and later scored. After waking, participants were disconnected from the head box and given a ~1hr break to alleviate sleep inertia before completing the delayed judgement task and behavioural control tasks. During this time, participants sat in a quiet room and consumed a small meal. Resting-state EEG recordings were obtained during quiet sitting with eyes open and eyes closed for two minutes, respectively. See Figure 3D for a schematic of the experimental protocol.

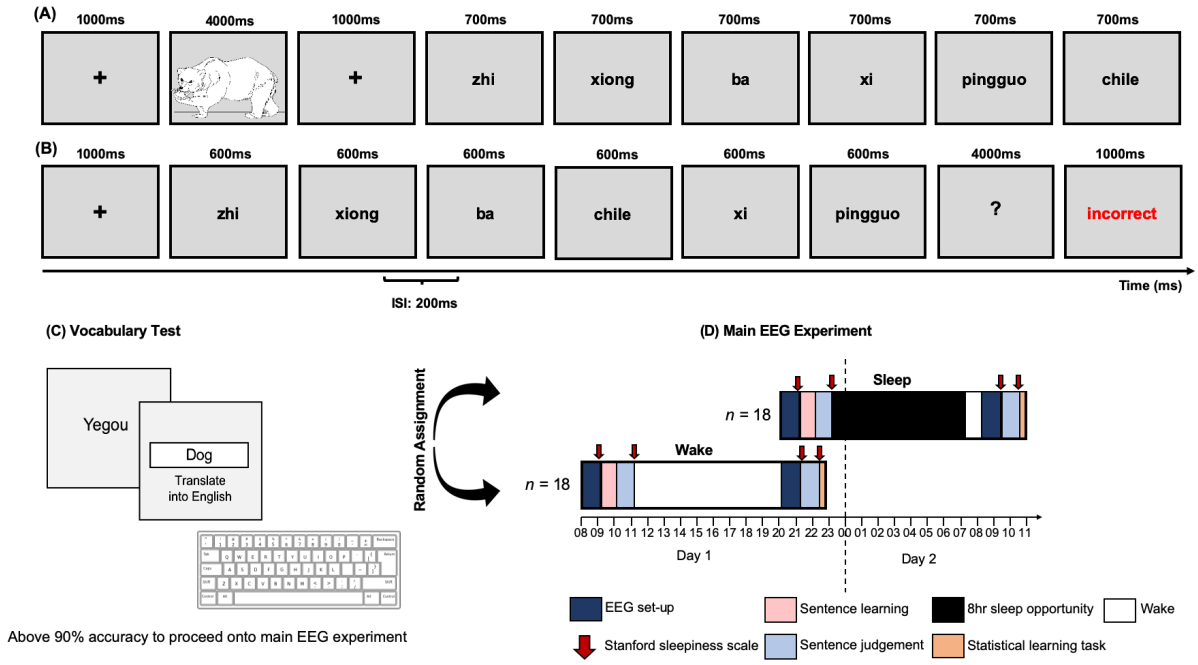

Figure 3. (A) Schematic representation of a single trial of a grammatical sentence during the sentence learning task. **(B)** Schematic representation of a single trial during the baseline sentence judgement task. This sentence is a violation of the grammatical rule presented in (A), whereby the verb *chile* is positioned in the middle of the sentence when it should be positioned at the end of the sentence. Here, the participant incorrectly categorised this sentence as grammatical, and thus received feedback indicating that their response was incorrect. **(C)** Schematic diagram of the vocabulary test, which required participants to translate the nouns (e.g., *yegou*) into English (e.g., *dog*) using a keyboard. **(D)** Experimental protocol representing the time course of the conditions (sleep, wake) and testing sessions (sentence learning task, baseline and delayed sentence judgement tasks). After completing the vocabulary test, participants were randomly assigned to either the sleep or wake conditions, with each participant only completing one of the two conditions. Time is represented along the x-axis, while each coloured block corresponds to a different task during the experimental protocol.

3. DATA ANALYSIS

3.1. Behavioural analysis

Three measures of behavioural performance were calculated. For behavioural models, grammaticality ratings were calculated on a trial-by-trial basis, determined by whether participants correctly identified grammatical and ungrammatical sentences. For EEG models, memory performance was quantified using the sensitivity index (d') from signal detection theory (Stanislaw & Todorov, 1999). Hit Rate (HR) and False Alarm rate (FA) were computed to derive d', defined as the difference between the z transformed probabilities of HR and FA (i.e., d' = z[HR] – z[FA]), with extreme values (i.e., HR and FA values of 0 and 1) adjusted using the recommendations of Hautus (1995). To examine the influence of sleep EEG on memory for complex linguistic rules, differences scores were calculated by subtracting d' values from the baseline sentence judgement task from d' values from the delayed sentence judgement task, with positive values indicating an improvement in performance.

3.2. EEG recording and pre-processing

Task-related EEG analyses during the baseline and delayed sentence judgement tasks were performed using MNE-Python (Gramfort et al., 2013; for an analysis of the sentence learning task, please see Cross et al. 2020c). EEG data were re-referenced offline to the average of both mastoids and filtered with a digital phase-true finite impulse response (FIR) band-pass filter from 0.1 – 40 Hz to remove slow signal drifts and high frequency activity. Data segments from -0.5 – 6.5s relative to the onset of each sentence were extracted and corrected for ocular artefacts using Independent Component

Analysis (fastica). Bad electrodes were visually inspected based on the number of epochs dropped; epochs were dropped when they exceeded a 150 μV peak-to-peak amplitude criterion or were identified as containing recordings from flat channels (i.e., < 5 μV). Electrodes containing more than 20% of excluded epochs were interpolated with surrounding electrodes (evoked.interpolate_bads).

3.2.1. Task-related time frequency analysis.

To determine the individualised ranges used to define each frequency band, individual alpha frequency (IAF) was estimated from participants' pre- and post-experiment resting-state EEG recording. IAFs were estimated from an occipital-parietal cluster (P3/P4/O1/O2/P7/P8/Pz/Iz) using philistine.mne.savgol_iaf (see Corcoran, Alday, Schlesewsky, & Bornkessel-Schlesewsky, 2018; see also Cross et al., 2020b) implemented in MNE (philistine.mne). IAF-adjusted frequency bandwidths were calculated using the golden mean algorithm (Klimesch, 2012).

We conducted task-related time-frequency analyses by convolving the pre-processed EEG with a family of complex Morlet wavelets using the MNE function tfr_morlet. Frequencies ranging between 3.5 and 30 Hz were analysed using wavelet cycles that were adjusted for each individually defined frequency band of interest. The mother wavelet for each frequency band was defined as the centre frequency value divided by four (e.g., 10 Hz / 4 = 2.5 Hz: 25 Hz / 4 = 6.25 Hz). Relative power change values in the post-stimulus interval were computed as a relative change from a baseline interval spanning -1s to the onset of each sentence. Relative power changes were then extracted for each sentence within the following frequency ranges: theta (~ 4 – 7 Hz), alpha (~ 8 – 12 Hz) and beta (~ 13 – 30 Hz). 500ms was added to the beginning and end of each sentence epoch to avoid edge artefacts.

3.2.2. Quantification of the broadband spectral profile.

As differences in task-evoked spectral power between sleep and wake conditions at delayed testing may reflect sleep-related changes in global neural synchrony (Freeman & Zhai, 2009; Voytek et al., 2015), it is important to distinguish such broadband spectral dynamics from narrowband changes in (putative) oscillatory components. To this end, we used the irregular-resampling auto-spectral analysis toolbox (IRASA v1.0; Wen & Liu, 2015) to estimate the $1/f$ power-law exponent characteristic of background spectral activity.

IRASA aims to isolate the aperiodic (random fractal) component of neural time series data via a process that involves resampling the signal at multiple non-integer factors h and their reciprocals $1/h$. Because this resampling procedure systematically shifts narrowband peaks away from their original location along the frequency spectrum, averaging the spectral densities of the resampled series attenuates peak components, while preserving the $1/f$ distribution of the fractal component. The (negative) exponent summarising the slope of aperiodic spectral activity is then calculated by fitting a linear regression to the estimated fractal component in log-log space. For a full mathematical description of IRASA, see Wen and Liu (2016).

Trial data from each analysed electrode were epoched using a 2 s sliding window (50% overlap) and submitted to the toolbox function amri_sig_fractal. Data were detrended and anti-alias filtered as per default settings; the output frequency range was set to 1–40 Hz. Following Muthukumaraswamy and Liley (2018), an extended set of resampling factors was selected (h = 1.1 to 2.9 in steps of 0.05, $h \neq 2$) in order to mitigate incomplete component separation due to fractal-oscillatory interaction. A power-law function was subsequently fit to each fractal estimate within the frequency range 2.5 – 30 Hz (amri_sig_plawfit). These functions were averaged across epochs for each trial, sentence condition, electrode and sentence judgement tasks (i.e., baseline, delayed sentence judgement tasks), and the inverse slope of this function taken as the trial-level estimate of the $1/f$ exponent.

Briefly, the $1/f$ exponent is thought to reflect the excitation–inhibition balance of neural populations (Gao, Peterson, & Voytek, 2017), with a flatter exponent indexing "noisier" population spiking (Donoghue et al., 2020; Voytek et al., 2015). While the functional significance of the $1/f$ exponent is still being investigated, it becomes flatter with age, explaining age-related cognitive decline (Voytek et al., 2015). It also varies across states of consciousness (e.g., Lendner et al., 2020) and predicts individual differences in processing speed (Ouyang et al., 2019) and language learning (Cross et al., 2020c). From this perspective, in order to control for any effect of aperiodic activity on language learning across sleep and wake, the $1/f$ exponent was included as a covariate in all statistical models. This approach reduces the amount of variance in the residual error term, increases statistical power, and has been adopted as an alternative baselining approach in ERP research (Alday, 2019).

3.2.3. Sleep parameters and sleep EEG analyses.

Sleep data were scored by two sleep technicians (Z.R.C and S.C.) according to standardised criteria (Berry et al., 2012) within Compumedics Profusion 3 software (Melbourne, Australia). The EEG was viewed with a high-pass filter of 0.3 Hz and a low-pass filter of 35 Hz. The following sleep parameters were calculated: total sleep time, sleep onset latency, wake after sleep onset, and time (minutes) and percent of time spent in each sleep stage (N1, N2, N3 and R). Spindle-slow oscillation coupling strength and spindle-slow oscillation phase-dependent correlation measures were extracted via the YASA toolbox (Vallat, & Jajcay, 2020) implemented in MNE-Python based on published algorithms (Helfrich et al., 2018; Staresina et al., 2015).

Briefly, the EEG data were re-referenced to linked mastoids and filtered from 0.1 – 30 Hz using a digital phase-true FIR band-pass filter. Data were then epoched into 30 s bins and subjected to a multivariate covariance-based artifact rejection procedure. This approach estimates a reference covariance matrix for each sleep stage and rejects epochs that deviate too far from this reference, where deviation is established using Riemannian geometry (Barachant, Andreev, & Congedo, 2013; Barthélemy, Mayuad, Ojeda, & Congedo, 2019; for a detailed description, visit https://github.com/raphaelvallat/yasa/blob/master/notebooks/13_artifact_rejection.ipynb).

For slow oscillations, continuous NREM EEG data were filtered using a digital phase-true FIR band-pass filter from 0.3 – 2 Hz with a 0.2 Hz transition band to detect zero crossing events that were between 0.3 – 1.5 s in length, and that met a 75 to 500 microvolt criterion. These artifact-free epochs were then extracted from the raw EEG signal. For sleep spindles, the signal was filtered between 12 – 16 Hz with a wide transition bandwidth of 1.5 Hz, while the amplitude was calculated by applying a Hilbert transform which was then smoothed with a 200ms moving average. Events were only included if they were within 0.5 – 2 s. Artifact-free events were then defined as 4s (± 2 s) peak-locked sleep-spindle epochs.

To determine whether the precise temporal relationship between spindles and slow oscillations influences the consolidation of complex grammatical rules, we calculated an event-locked cross-frequency coupling metric (for a detailed description of this method, see Helfrich et al., 2018). In brief, we first filtered the normalized slow oscillation trough-locked data into the slow oscillation component (0.1 – 1.25 Hz) and extracted the instantaneous phase angle after applying a Hilbert transform. Then we filtered the same trials between 12 – 16 Hz and extracted the instantaneous amplitude from the Hilbert transform. For every participant at channel Cz and epoch, we detected the maximal sleep spindle amplitude and corresponding slow oscillation phase angle. The mean circular direction (phase) and resultant vector length (mean vector length [MVL]; coupling strength) across all NREM events were then determined using circular statistics implemented in the pingouin package (Vallat, 2018). A scale

of 0 – 1 was used for MVL, with 1 indicating that each coupled spindle occurred at the same phase of the slow oscillation, and 0 indicating that each coupled spindle occurred at a different phase of the slow oscillation (Denis et al., 2020).

3.3. Statistical analysis

Data were imported into *R* version 4.0.2 (R Core Team, 2020) and analysed using linear mixed-effects models fit by restricted maximum likelihood (REML) estimates using *lme4* (Bates, 2010). For the behavioural model, we used a logit mixed-effects model, modelling response choice (1 = correct; 0 = incorrect) as a binary outcome variable, with 156 observations per subject for each testing session (baseline, delayed). This model also factored in by-item and by-participant differences by modelling them as random effects on the intercept, which a d' measure would not be able to capture, due to d' being an aggregate metric. Further, logit mixed-effects models are standard in the (neuro)psycholinguistic literature and are argued to be the preferred mode of analysing data from repeated measure designs, as they account for within- and between-subject variance, as well as variance introduced by items (i.e., individual sentences; Baayen, Davidson, & Bates, 2008; Judd, Westfall & Kenny, 2012; Van Dongen, Olofsen, Dinges, & Maislin, 2004). As such, a logit mixed model is particularly suited to the current data due to: (a) inter-individual differences in language learning and performance on the sentence judgement tasks; (2) the use of categorical response variables (i.e., binomial correct/incorrect responses) as a measure of language learning, and; (3) the ability to account for item variability, given that items may vary in familiarity across participants and thus influence learning outcomes (Baayen et al., 2008; Quené, Huub, & Bergh, 2008). Further, the use of a trial-based outcome variable in our main statistical models allows for more fine-grained analyses of by-item and by-participant variability (e.g., Cross et al., 2020a) which are lost in aggregated variables, such as proportion correct or d'. The behavioural model took the following form:

$$\text{Logit}(response_i) = \beta_0 + \beta_1 grammaticality_i * \beta_2 type_i * \beta_3 condition_i * \beta_4 session_i + \beta_5 sss_i + subject_{0i} + item_{0i} + \epsilon,$$

where *grammaticality* encodes sentence grammaticality (grammatical, ungrammatical), *type* refers to word order (fixed, flexible), *condition* is sleep versus wake, *session* is baseline and delayed testing, and *sss* refers to self-perceived sleepiness estimated from the Stanford Sleepiness Scale. Random effects on the intercept for Subject and Item were included to account for systematic variance amongst participants and sentence stimuli, respectively (Judd, Westfall, & Kenny, 2012; Van Dongen, Olofsen, Dinges, & Maislin, 2004). Asterisks denote interaction terms; pluses denote additive terms. Further, the following structure was used for the EEG models, where we were interested in predicting behaviour from task-related oscillatory activity, and which did not include trial-based response accuracy:

$$dprime_i = \beta_0 + \beta_1 power_i * \beta_2 aperiodic_i * \beta_3 condition_i * \beta_4 session_i * \beta_5 type_i + subject_{0i} + \epsilon,$$

where *power* is spectral power from a specific frequency band (i.e., theta, alpha, beta), *aperiodic* refers to the $1/f$ exponent estimated from the task-related EEG (see S1 in the supplementary material for full model summaries and visualisations of the influence of $1/f$ on putative oscillatory activity), *condition* is sleep versus wake, *session* is baseline versus delayed testing, and *type* is sentence word order (fixed, flexible). Subject was modelled as a random effect on the intercept, while d' was specified as the outcome. Type II Wald $\chi 2$-tests from the *car* package (Fox, 2011) were used to provide *p*-value estimates, while effects were plotted using the package *effects* (Fox et al., 2019) and *ggplot2* (Wickham, 2016). Categorical factors were sum-to-zero contrast coded, meaning that factor level estimates were compared to the grand-mean (Schad et al., 2020). Further, an 83% confidence interval (CI) threshold was used given that this approach corresponds to the 5% significance level with non-overlapping estimates (Austin & Hux, 2002; MacGregor-Fors & Payton, 2013).

Finally, rather than adding sagittality and laterality as fixed effects of topographical distribution, we selected *a priori* regions of interest (ROI) based on the published literature. This approach was adopted to (1) reduce computational complexity from additional parameters; (2) reduce the probability of encountering a Type I or Type II error; and (3) avoid model nonconvergence due to over-parameterization. Previous studies adopted cluster-based permutation methods, as well as more traditional statistical analysis approaches to isolate ROIs during artificial language learning (e.g., de Diego-Balaguer et al., 2011; Kepinska, Pereda, et al., 2017) and native sentence processing (e.g., Lewis, Lemhfer, Schoffelen, & Schriefers, 2016; Meyer et al., 2015). For the alpha band, we selected a posterior cluster of electrodes (P3/P4/O1/O2/Pz/Iz), given the prominence of alpha power over occipital-parietal electrodes, particularly during visual language experiments (e.g., Kepinska, Pereda, et al., 2017; Kielar, Meltzer, Moreno, Alain, & Bialystok, 2014; Lam et al., 2016). For beta, we selected an anterior-central and anterior-left electrode cluster (FC1/FC2/Fz/F4/FC6/F8), which has been reported as a significant ROI for beta activity during higher-order language tasks (e.g., Bastiaansen & Hagoort, 2015; de Diego-Balaguer et al., 2011; Kepinska, Pereda, et al., 2017). Finally, a central-left electrode cluster (T7/CP5/C3/CP1/Cz/CP2) was chosen for the theta band based on studies reporting sleep-related memory effects (Koster, Finger, Kater, Schenk, & Gruber, 2016), as well as memory-related computations during language processing (Kielar et al., 2014; Meyer et al., 2015)

4. RESULTS

4.1. Behavioural results

Participants showed a moderate degree of accuracy on the sentence judgement tasks (mean accuracy = 63.35%, see Table 1).

Table 1. Percent correct and d' by condition (sleep, wake), sentence judgement tasks (baseline, delayed), grammaticality (grammatical, ungrammatical) and sentence type (fixed, flexible). Standard deviations (SD) are given in parentheses. Differences scores (delayed – baseline testing) for d' are given at the bottom of the table. d' values of 0 indicate no change in performance from baseline to delayed testing.

Condition	Session	Grammaticality	Sentence Type	Correct (SD)	d' (SD)
Sleep	Baseline	Grammatical	Flexible	65.20 (18.41)	0.79 (1.24)
			Fixed	67.81 (12.88)	0.91 (0.79)
		Ungrammatical	Flexible	59.07 (22.78)	
			Fixed	48.36 (13.61)	
	Delayed	Grammatical	Flexible	58.33 (32.95)	0.99 (1.89)
			Fixed	71.57 (22.19)	1.49 (1.34)
		Ungrammatical	Flexible	66.42 (23.82)	
			Fixed	50.65 (21.32)	
Wake	Baseline	Grammatical	Flexible	63.19 (23.11)	1.11 (1.34)
			Fixed	67.59 (20.63)	1.40 (0.91)
		Ungrammatical	Flexible	69.68 (24.37)	
			Fixed	52.00 (15.32)	
	Delayed	Grammatical	Flexible	67.70 (25.95)	1.41 (1.48)
			Fixed	62.61 (26.80)	1.20 (1.42)
		Ungrammatical	Flexible	69.74 (27.42)	
			Fixed	52.35 (18.53)	
d' difference scores [delayed – baseline]					
Sleep			Flexible		0.20 (1.28)
			Fixed		0.59 (1.03)
Wake			Flexible		0.30 (0.62)
			Fixed		-0.19 (1.01)

Statistical analysis revealed that the Condition (sleep versus wake; $\chi2(1) = .07$, $p = .79$) and Session (baseline versus delayed; $\chi2(1) = .03$, $p = .85$) effects on the probability of a correct response were nonsignificant. There was a main effect of Grammaticality, such that the probability of a correct response was higher for grammatical sentences ($\chi2(1) = 58.86$, $p < .001$). The main effect of Type was also significant, with the probability of a correct response being greater for flexible word orders ($\chi2(1) = 6.49$, $p = .01$). See Table 2 and Figure 4 for a summary of interaction effects.

Table 2. Type II Wald χ^2 tests for interaction terms examining the relationship between grammaticality (grammatical, ungrammatical), sentence type (fixed, flexible), condition (sleep, wake) and testing session (baseline and delayed judgement tasks) on the probability of a correct response.

Parameters	***χ2***	***df***	***p***
Grammaticality x Type	141.39	1	**<.001**
Grammaticality x Condition	6.63	1	**.01**
Type x Condition	10.08	1	**.001**
Grammaticality x Session	4.66	1	**.03**
Type x Session	0.23	1	.62
Condition x Session	0.007	1	.93
Grammaticality x Type x Session	1.32	1	.24
Grammaticality x Type x Condition	2.09	1	.14
Grammaticality x Condition x Session	1.18	1	.27
Type x Condition x Session	9.68	1	**.001**
Grammaticality x Type x Condition x Session	10.19	1	**.001**

Most importantly, the Type x Condition x Session and Grammaticality x Type x Condition x Session interactions were significant. As illustrated in Figure 4A, the likelihood of a correct response for ungrammatical sentences with a flexible word order was higher at baseline testing for the wake condition relative to the sleep condition, while the likelihood of a correct response was higher for grammatical fixed word orders at delayed testing for the sleep condition.

To compensate for differences in baseline performance, Figure 4B presents changes in performance between baseline and delayed testing. Here, both the high-level interactions can be understood. For the fixed order sentences (left facet) there is a general improvement in performance after sleep and a general worsening of performance after wake. This is consistent with a consolidation benefit of sleep for fixed order aspects of the novel grammar, supporting H^1. For the flexible order sentences (right-hand side), the general pattern is for a small benefit of wake on performance, but no overall change across sleep. However, the four-way interaction suggests that the pattern of performance for the sleep condition is more complex: for ungrammatical sentences there is an improvement in performance, but this is mirrored by an equal decrement in performance for the grammatical sentences. This aspect of the data was not predicted, but perhaps makes sense in terms of a criterion shift. The three-way interaction shows that sleep benefits generalisation for fixed (but not flexible) word orders. Given this selective improvement in grammatical knowledge across sleep, participants may have become more confident in their knowledge of the fixed word orders, and hence less inclined to endorse flexible word orders, leading to more correct rejections of ungrammatical sentences, but also more incorrect endorsements of grammatical sentences.

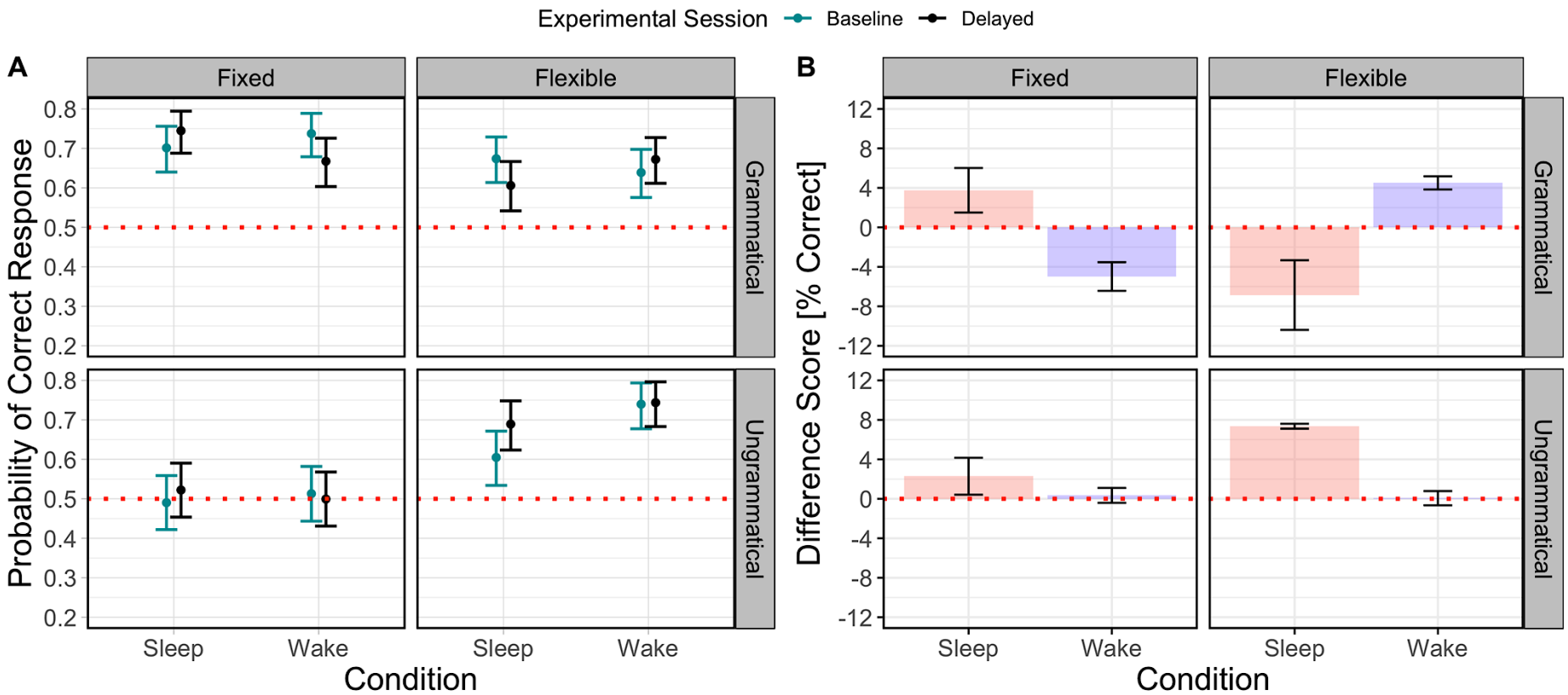

Figure 4. Visualisation of the behavioural results. (A) Modelled relationship between the probability of correct response (y-axis; higher values indicate a higher probability of a correct response), condition (sleep, wake; x-axis), grammaticality (grammatical = top panel; ungrammatical = bottom panel), sentence type (left facet = fixed, right facet = flexible), and sentence judgement task (blue left = baseline, right black = delayed). Bars represent the 83% confidence interval **(B)** Difference scores (percentage correct at delayed – baseline; positive values indicate improvements in performance) between the sleep and wake conditions on grammatical (top panel) and ungrammatical (bottom row) fixed word order sentences (facetted on the left), and grammatical and ungrammatical flexible word order sentences (facetted on the right). Bars represent the standard error of the mean. The red dashed line in **(A)** indicates chance-level performance, while the red dashed line in **(B)** indicates no change in performance from baseline to delayed testing.

4.2. Neurophysiological results

4.2.1. Neuronal oscillations and incremental sentence processing.

Based on the clear differences in behavioural performance between a period of sleep versus wake on fixed and flexible word order sentences, we now examine whether task-evoked oscillatory power during the baseline and delayed sentence judgement tasks between the sleep and wake conditions predicts behaviour. Further, the $1/f$ exponent was modelled as a covariate, as changes in aperiodic activity partially predicted differences in task-evoked oscillatory activity during the judgement tasks in both the sleep and wake conditions (see S1 in the supplementary material for full model summaries and visualisations of these effects). The effects of each model are illustrated in Figure 5.

The first model focussed on oscillatory activity in the theta band. Our analysis revealed a significant Power x Condition x Session x Type interaction ($\chi2(1) = 5.18$, $p = .03$): while higher task-evoked theta power at baseline testing predicted a decrease in d' scores in the sleep and wake conditions for fixed word order sentences, higher theta power predicted increased d' scores for flexible word order sentences, with this effect being stronger for the wake condition. By contrast, higher theta power at delayed testing predicted an increase in d' scores for sleep and wake conditions for both fixed and flexible word order sentences. This suggests that greater theta power after a delay period regardless of whether participants slept correlates with enhanced sensitivity to grammatical violations, partially supporting H^3.

For the alpha band, we observed a significant Power x Condition x Session x Type interaction ($\chi2(1) = 11.90$, $p < .001$): while there was no association between alpha power and performance at baseline testing for either the sleep or wake conditions and fixed and flexible word order sentences, higher alpha power at delayed testing predicted an increase in d' scores for fixed word order sentences for both the sleep and wake conditions. By contrast, higher alpha power predicted an increase in d' scores for the sleep condition for fixed word order sentences, while the inverse was observed for the wake condition.

Finally, for the beta band, there was also a significant Power x Condition x Session x Type interaction ($\chi2(1) = 5.17$, $p = .02$): while lower beta power predicted an increase in d' scores for the sleep and wake conditions for flexible word order sentences at baseline testing, this effect was not present for the sleep group for fixed word order sentences. By contrast, at delayed testing, higher beta power predicted an increase in d' scores for both the sleep and wake conditions for flexible word order sentences, while there was a negative association between beta power and d' for the wake condition for fixed word order sentences (for full model summaries, see the supplementary materials).

Taken together, our analyses demonstrate that sleep relative to an equivalent period of wake results in a shift in the relationship between alpha/beta power and performance for flexible word order rules: directly after learning, lower beta power predicted an increase in d' scores, while an increase in alpha/beta power predicted behaviour after a delay period containing sleep, supporting H^2. This effect was similar but less pronounced for theta activity, indicating a delay period after learning rather than sleep *per se* may influence the effect of task-evoked theta activity on sentence processing. For a visualisation of the change in spectro-temporal activity across each of the sentence conditions and scalp, see Figure 6.

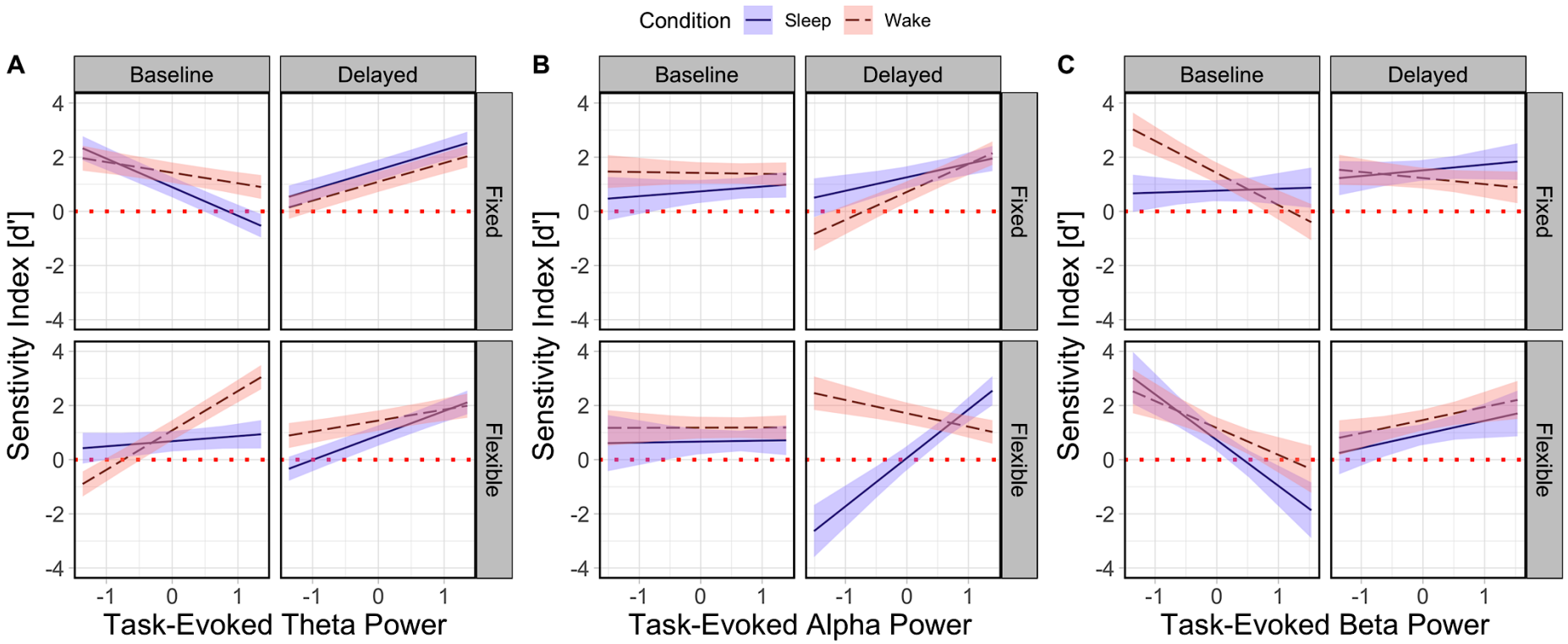

Figure 5. Visualisation of the EEG and behavioural results. Modelled effects of task-related oscillatory activity (x-axis; higher values indicate greater power) on d' scores (y-axis; higher values indicate better performance) for the sleep and wake conditions (sleep = purple solid line; wake = dashed pink line) for fixed (top panel) and flexible (bottom panel) sentences. The baseline sentence judgement task is facetted on the left, while the delayed sentence judgement task is facetted on the right. The red dashed line indicates chance-level performance, while the shaded regions indicate the 83% confidence interval.

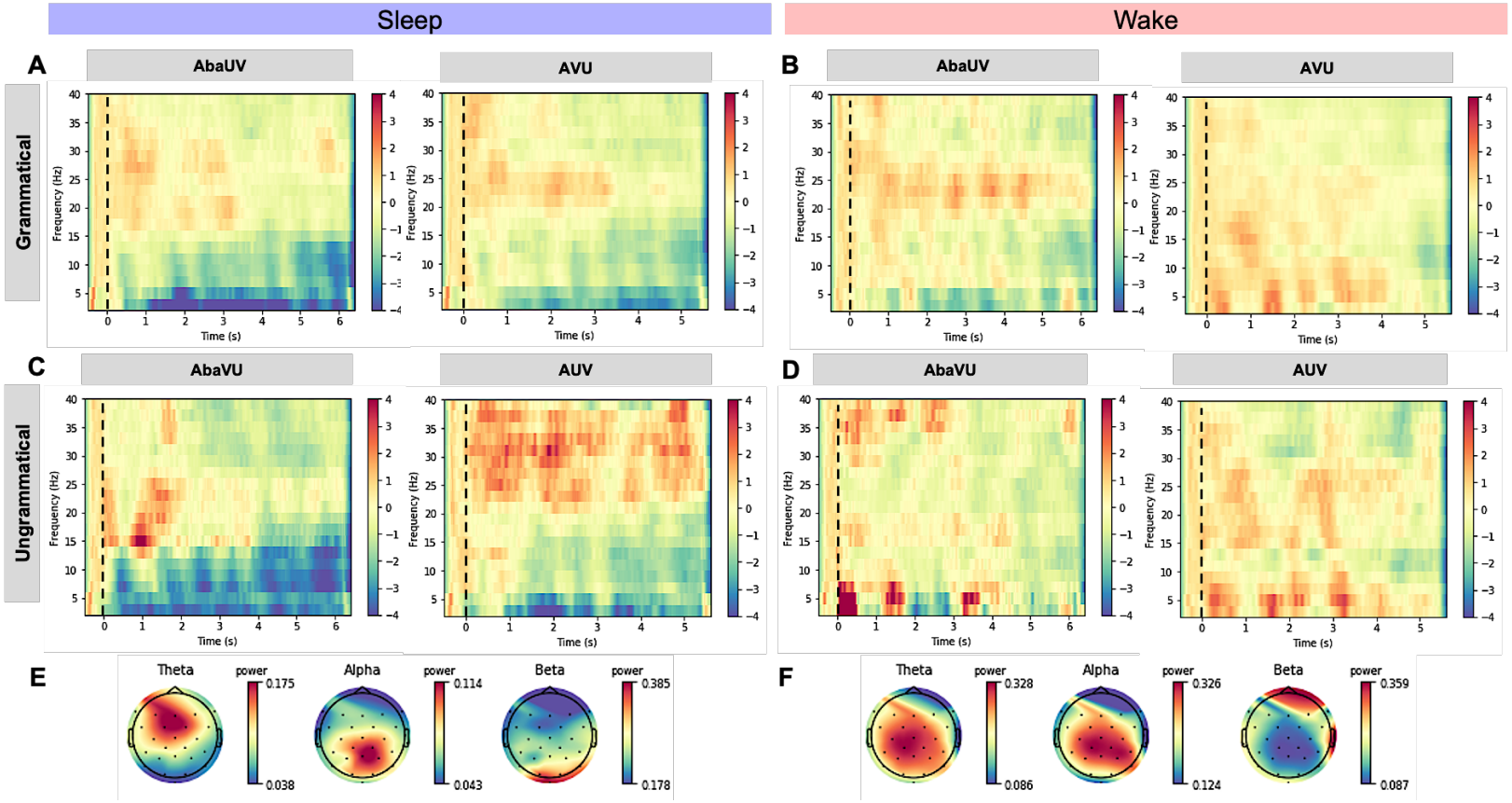

Figure 6. Task-related spectral dynamics between the sleep and wake conditions. **(A)** Spectral power dynamics for grammatical fixed (left) and flexible (right) sentences for the sleep condition. **(B)** Spectral power dynamics for grammatical fixed (left) and flexible (right) sentences for the wake condition. **(C)** Spectral power dynamics for ungrammatical fixed (left) and flexible (right) sentences for the sleep condition. **(D)** Spectral power dynamics for ungrammatical fixed (left) and flexible (right) sentences for the wake condition. The onset of each sentence is illustrated by the dashed black line, while relative power (baselined 200ms prior to sentence onset) across the sentence is represented on the x-axis (1 – 6 seconds). Warmer colours indicate an increase in power (µV), while cooler colours indicate a decrease in power. Values at approximately 0 indicate no change in power relative to the baseline interval. Frequency (Hz) is represented on the y-axis. Topographical changes in power in theta, alpha and beta power during grammatical sentence processing for the sleep **(E)** and wake **(F)** conditions are illustrated on the bottom panel. Warmer colours indicate power increases in that topographical region for that frequency band, while cooler colours indicate a decrease in power.

4.2.2. Associations between NREM neurophysiology and memory retention.

Thus far we have observed clear differences between the sleep and wake conditions on the relationship between frequency-band power and behavioural performance, which supported the hypothesis that sensitivity to fixed word order rules would be improved after sleep relative to wake (H^1), and that improved performance would be associated with higher task-related beta and alpha power, and that this effect would be accentuated after sleep compared to wake (H^2). A logical next step is to test H^4, namely that differences in performance between the sleep and wake conditions would be modulated by NREM spindle-slow oscillation coupling. Based on recent research (e.g., Helfrich et al., 2018; Holz et al., 2012; Mikutta et al., 2019), we focussed on two cardinal markers of sleep-related memory consolidation: (1) the coupling strength between maximal spindle amplitude and the phase of the slow oscillation, and (2) spindle onset relative to the phase of the slow oscillation (for a summary of typical sleep parameters and their correlation with d' difference scores, see Table 3). While Phase did not significantly predict changes in behavioural performance across sleep (Phase x Type: $\chi2(1) = .03$, $p = .85$), there was a significant Coupling Strength x Type interaction on d' difference scores ($\chi2(1) = 8.21$, $p = .004$; see Figure 7C for a visualisation of this effect).

As is illustrated in Figure 7C, stronger spindle-slow oscillation coupling predicted higher d' difference scores for fixed word order sentences. These results suggest that the rules governing fixed word order sentences benefit from stronger sleep-based consolidation and may be prioritized over the rules governing flexible word order sentences – a finding consistent with the behavioural effects observed in section 4.1 "Behavioural Results". Figure 7 further illustrates an exemplary full-night multi-taper spectrogram, comodulagram and preferred phase of spindle-slow oscillation coupling for Stage 2 and SWS.

Table 3. Descriptive statistics for sleep parameters and correlations with the difference between d' at delayed and baseline testing for fixed and flexible word orders.

Sleep Parameter	**Mean Minutes (SD)**	**% in Stage (SD)**	**Correlations with d' (Delayed – Baseline)**			
			Fixed		**Flexible**	
			r	***p***	***r***	***p***
TST	400.00 (67.02)		**-.44**	**.006**	.30	1.00
SOL	15.23 (12.23)		**.45**	**.006**	-.47	.36
WASO	52.64 (55.60)		**.41**	**.008**	-.19	1.00
N1	38.05 (29.47)	10.05 (8.21)	.12	.21	.10	1.00
N2	196.30 (46.29)	49.52 (10.36)	.26	.21	.33	.45
SWS	104.23 (42.27)	25.84 (9.60)	.02	.43	-.48	.07
REM	61.30 (39.39)	14.57 (8.56)	**-.46**	**.006**	.04	1.00

Note. SD = standard deviation. TST = total sleep time; SOL = sleep onset latency; WASO = wake after sleep onset; N1 = stage 1; N2 = stage 2; SWS = slow wave sleep; REM = rapid eye movement sleep. Correlations represent the relationship between percent of time spent in each sleep stage and d'. Significance values are Holm-Bonferroni corrected (Holm, 1979). For a visualisation of the correlations between sleep parameters and d' difference scores, please see Figure S5 in the supplementary material.

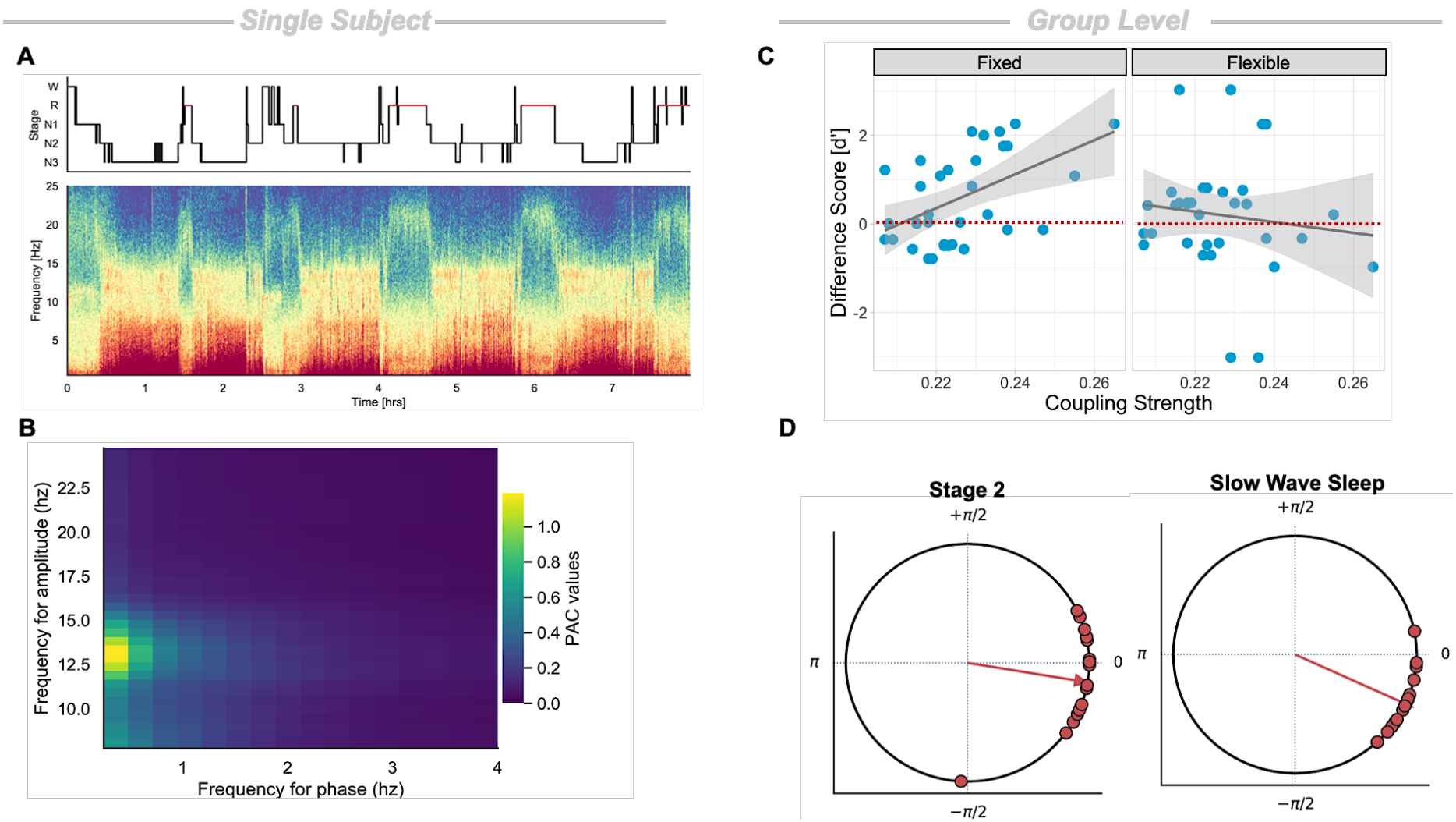

Figure 7. (A) Hypnogram and full-night multi-taper spectrogram for a single participant. **(B)** Comodulagram illustrating the frequency for phase (x-axis) and frequency for power (x-axis) during NREM sleep spindle-slow oscillation coupling epochs for a single participant. **(C)** Relationship between NREM (i.e., across N2 and SWS) spindle-slow oscillation strength and d' difference scores for fixed (left) and flexible (right) word order sentences (red dotted line indicates no change in performance from baseline to delayed testing). **(D)** Preferred phase of spindle-slow oscillation coupling for Stage 2 and **(D)** SWS (red circles indicate individual participants).

4.2.3. Spindle-slow oscillation coupling modulates task-evoked oscillatory activity and behavioural sensitivity to complex linguistic rules.

Next, we addressed whether NREM neurophysiology modulates task-evoked oscillatory activity during sentence processing, and whether this interaction predicts sensitivity to grammatical violations (H^5). Given that we observed a significant positive effect of spindle-slow oscillation coupling on d' difference scores for fixed word order sentences, we implemented a linear mixed-effects analysis for the theta, alpha and beta bands, with coupling strength, task-evoked power at the delayed testing session, sentence type (fixed, flexible) and the $1/f$ exponent as fixed effects. The dependent variable was difference scores in d' between baseline and delayed sentence judgement tasks, with positive values indicating an improvement in performance from baseline to delayed testing.

We observed a significant Coupling Strength x Power x Sentence Type interaction for the theta ($\chi2(1) = 6.24$, $p = .01$) and alpha ($\chi2(1) = 49.01$, $p < .001$) bands, but not for the beta band ($\chi2(1) = 3.23$, $p = .07$). These interactions are resolved in Figure 8, which illustrates a clear interaction between spindle-slow oscillation coupling strength and task-evoked oscillatory activity on sensitivity to fixed and flexible sentences. As is shown in Figure 8A (left facet), high spindle-slow oscillation coupling and an increase in task-evoked theta power was associated with improved performance from baseline to delayed testing for flexible sentences (dashed orange line), while the inverse was observed for fixed word order sentences (solid orange line). By contrast, when spindle-slow oscillation coupling was high, a decrease in alpha power predicted improved performance from baseline to delayed testing for fixed, while the inverse effect is observed for flexible word order sentences. This effect for both sentence types diminished considerably as alpha power increased (see Figure 8B). For a visualisation of the distribution of coupling strength values across Stage 2 and Stage 3, see Figure 8D.

Together, results demonstrate that theta and alpha task-related activity interact with spindle-slow oscillation coupling to differentially influence the processing of fixed and flexible word order rules, while task-related beta activity is not as strongly modulated by such sleep-related neural activity.

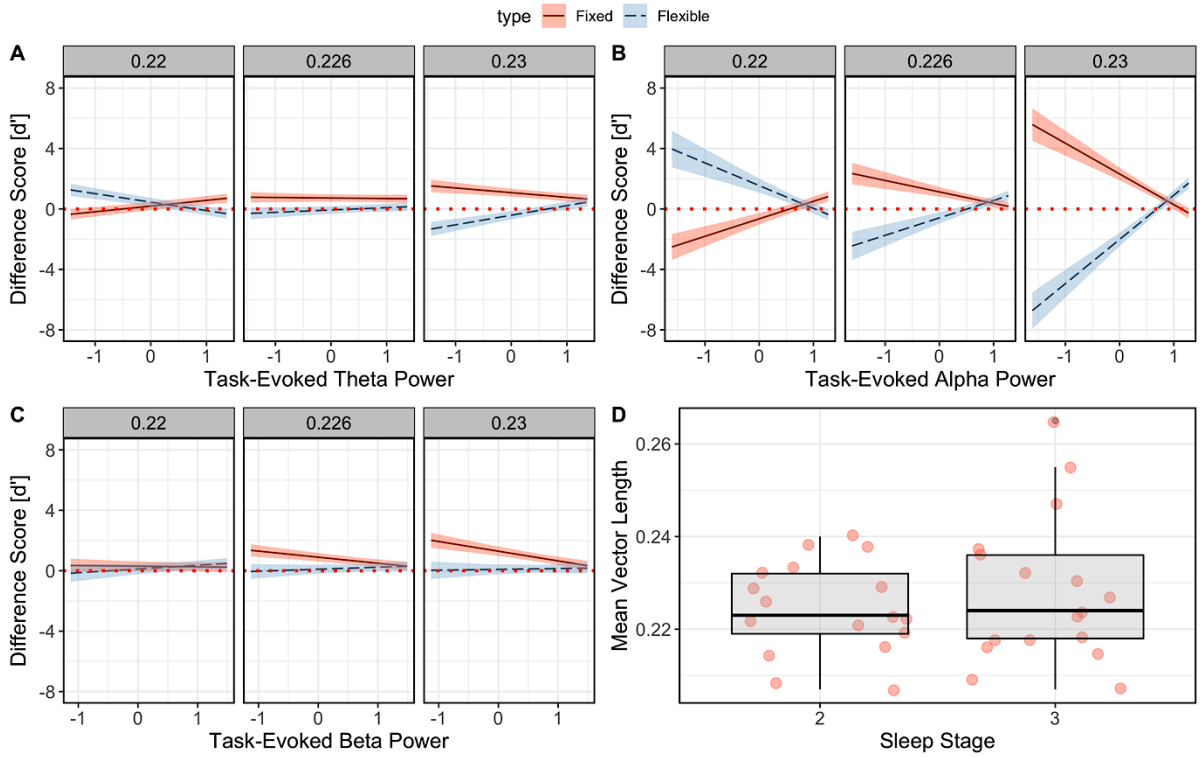

Figure 8. Modelled effects of task-evoked power in each frequency band (**A** = theta, **B** = alpha, **C** = beta) on d' difference scores (delayed – baseline sentence judgement task). Task-evoked theta, alpha and beta power are represented on the x-axis (left = lower power, right = higher power), while spindle-slow oscillation coupling

strength is faceted from low (left) to high (right). The lines represent sentence type: the solid orange lines represent fixed word order sentences, while the blue dashed line represents flexible word order sentences. The red dashed line indicates no change in performance from baseline to delayed testing. **(D)** Mean spindle-slow oscillation coupling strength during stage 2 and stage 3 (i.e., SWS). The individual data points represent single participant coupling metrics.

5. DISCUSSION

The coordination between slow oscillations and sleep spindles has been hypothesised to provide an optimal temporal receptive window for hippocampal-cortical communication during sleep-related memory consolidation (Helfrich et al., 2019; Staresina et al., 2015). Here, we show that the interaction between these cardinal markers of NREM sleep extends to the consolidation and generalisation of complex sentence-level regularities. Behaviourally, we demonstrated that a period of sleep compared to an equivalent period of wake benefits the consolidation of fixed relative to flexible word order rules, and that this effect is modulated by the strength of spindle-slow oscillation coupling. Our results further reveal that spindle-slow oscillation coupling is likely to be associated with changes in task-evoked theta and alpha oscillatory activity during sentence processing, possibly reflecting alterations in underlying memory-related neural populations.

5.1. Beyond single word learning: a role for sleep in consolidating word order rules

Using a novel complex modified miniature language paradigm (Cross et al., 2020a), we demonstrated that a period of sleep facilitates the extraction of fixed relative to flexible word order rules. Importantly, the key distinction between these word order permutations is that successful interpretation of fixed word order sentences relates to the sequential position of the noun phrases and verb (i.e., the first noun phrase is invariably the actor and the sentence is verb-final). By contrast, successful interpretation of flexible word order sentences depends more heavily on the animacy of the nouns. As such, fixed word order sentences, requiring a more sequential order-based interpretation, are more compatible with an English word-order-based processing strategy (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky et al., 2015; MacWhinney, Bates, & Kliegl, 1984). Critically, this sleep-based enhancement for fixed word order rules was predicted by stronger spindle-slow oscillation coupling (Figure 7C).

Sleep-related memory effects have been argued to be biased toward stimuli that follow temporal or sequence-based regularities compared to relational information (for review, see Lerner & Gluck, 2019). This is posited to occur via the hippocampal complex being biased toward encoding the temporal occurrences of sensory input (Durrant, Taylor, Cairney, & Lewis, 2011), which are replayed during SWS, potentially via spindle-slow oscillation coupling (e.g., Navarrete et al., 2020; Solano et al., 2020). Here, we provide evidence that supports this account. During learning, the hippocampus may have preferentially extracted the temporal regularities of fixed word order sentences. Sleep-associated memory processing, achieved via spindle-slow oscillation coupling, may have then selectively strengthened these memory traces, facilitating the formation of neural representations of linguistic rules. During subsequent wake, these newly established representations would have then been used by cortical networks to detect the temporal regularities in novel sentences, resulting in greater sensitivity to fixed word order rules.

Taken together, sleep-based consolidation of higher order language appears to favour sequence-based regularities, with mechanisms of sleep-related memory consolidation generalizing fixed over flexible word order rules. It is important to note, however, that our sample of participants were native monolingual speakers, and as such, may have preferentially consolidated the fixed word order rules at the expense of the flexible rules. While behavioural work is beginning to demonstrate sentence-level preferences of grammatical rules that are analogous to learners' native languages (e.g., Cross et al.,

2020a), less is known regarding the neural underpinnings of this phenomenon. This would be an interesting next step for future research. In the following, we discuss how the neurobiological processes underpinning the beneficial effect of spindle-slow oscillation coupling on memory consolidation extends to higher order language learning.

5.2. Spindle-slow oscillation coupling as a neurobiological marker of sleep-associated memory consolidation and higher-order language learning

Recent work has demonstrated that the coupling between slow oscillations and spindles predicts successful overnight memory consolidation (Helfrich et al., 2018; Mikutta et al., 2019; Muehlroth et al., 2019; Niknazar, Krishnan, Bazhenov, & Mednick, 2015). However, these studies often utilised old-new paradigms with single words (e.g., Helfrich et al., 2018; Mikutta et al., 2019) or word-image pairs (e.g., Muehlroth et al., 2019), leaving the functional relevance of these NREM oscillations for the generalisation of more complex linguistic information unknown. Moreover, such sleep-related memory effects are typically reported at the behavioural level, overlooking the critical influence that these effects may have on task-evoked neural activity as a proxy for changes in underlying task-related neural populations.

Adding to this body of literature, we have demonstrated that the generalisation of sequence-based rules is facilitated by the strength of NREM spindle-slow oscillation coupling. Mechanistically, during SWS, the cortex is highly synchronised during the up state of the slow oscillation, allowing effective interregional communication, particularly between the prefrontal cortex and hippocampal complex (Helfrich et al., 2019). It is during this slow oscillation up state that spindles induce an influx of Ca^{2+} into excitatory neurons, enabling synaptic plasticity and the generalisation and stabilisation of memory traces (Niknazar et al., 2015). From this perspective, the interaction between these cardinal markers of sleep-related memory processing appear to extend beyond the consolidation of single words and images and to the generalisation of sentence-level regularities.

In the following, we discuss how spindle-slow oscillation coupling, as a marker of sleep-associated memory consolidation, modulates task-related oscillatory activity, as well as the mechanisms such changes might reflect during sentence processing.

5.3. Theta oscillations bind linguistic elements in memory during sentence processing

Theta is the dominant frequency in the hippocampal complex and surrounding structures during wake (Covington & Duff, 2016; Duff & Brown-Schmidt, 2012). Oscillations in this frequency range are critical for memory function and coordinating hippocampal-cortical interactions, having been related to associative memory formation (Tort et al., 2009), tracking sequential rules (Crivelli-Decker et al., 2018) and predicting words based on contextual linguistic information (Piai et al., 2016). In the sleep and memory literature, increased theta oscillations have been reported for successfully remembered items, interpreted as reflecting a stronger memory trace induced by sleep-based consolidation (Koster et al., 2016; Schreiner et al., 2015). Here, it was demonstrated that an increase in theta oscillations predicted stronger sensitivity to both fixed and flexible word order rules after a 12hr delay period. This finding accords with the general memory literature, possibly reflecting the binding of temporally dispersed items in a (linguistic) sequence to generate a coherent sentential percept.

We also observed that spindle-slow oscillation coupling interacted with task-evoked theta power to predict sensitivity to grammatical rules. Specifically, when spindle-slow oscillation coupling strength was high, higher performance for flexible word order sentences was predicted by higher theta power, while the inverse was observed for fixed word order sentences. In line with systems

consolidation theory (Born & Wilhelm, 2012), NREM oscillatory activity contributes to the consolidation of newly encoded memory representations, which may manifest in stronger theta power during retrieval, indicating a stronger memory trace (Schreiner et al., 2015). Our results reveal that when spindle-slow oscillation coupling is high, the relationship between task-evoked theta activity and behavioural performance increasers, particularly for rules that are more flexible in nature.

5.4. Alpha oscillations reflect sleep-related effects on attention and perception

At delayed testing, an increase in alpha power predicted increased sensitivity for fixed and flexible word order rules, while the reverse was observed for the wake condition for flexible word order sentences. It is argued that one function of alpha power is to prevent competing memory traces from being concurrently activated during memory retrieval (Hanslmayr, Staudigl, & Fellner, 2012). Increased alpha power during memory processing has also been argued to reflect goal-directed inhibition and disengagement of task-irrelevant brain regions (Kepinska, Pereda, et al., 2017; Park et al., 2014). From this perspective, those who slept may have had greater access to more fine-tuned and integrated memory representations, while those who remained awake required more effortful processing (as reflected in decreased alpha power for flexible word order processing) in order to efficiently bind individual words to form a coherent sentential percept. This interpretation is in line with Kepinska et al. (2017), who reported that greater alpha power predicted enhanced behavioural performance on a language learning task for individuals with a greater language learning aptitude. As such, similar to those with a higher aptitude who have the ability to learn complex language-related regularities, sleep-based consolidation mechanisms may serve a functional role in fine-tuning networks involved in processing a newly learned language (Cross et al., 2018), requiring less attentional resources post-sleep to process incoming task-relevant stimuli. A decrease in alpha power may therefore reflect more effortful processing by suppressing the input of irrelevant cortical regions to efficiently retrieve encoded memory traces to interpret incoming linguistic information.

This interpretation is also in line with principles of predictive coding: recent work has demonstrated that sleep results in increased prediction accuracy for sequence knowledge (Lutz et al., 2018), and has been used to explain sleep-based consolidation effects of linguistic stimuli (Rauss & Born, 2017). Predictive coding accounts (e.g., Friston, 2010; Hobson & Friston, 2012; Rauss & Born, 2017) conceive of the brain as a prediction engine that infers the causes of its sensory states by modelling the statistical regularities of its environment. Further, given that sleep optimises cortical memory representations through a process of active consolidation and generalisation, the relationship between an increase in alpha power post-sleep and enhanced behavioural performance might reflect the propagation of top-down predictions during sentence processing, an interpretation supported by recent computational work, albeit not in the sentence processing domain (Alamia & VanRullen, 2019).

It is also important to note the differences between the current study and studies examining event-related changes in alpha activity in response to more general aspects of memory. Previous research has reported decreased alpha power during memory retrieval (Fellner et al., 2019; Hanslmayr, Staresina, & Bowman, 2016; Hanslmayr et al., 2012). These studies often employ episodic memory tasks of word-scene (Minarik, Berger, & Sauseng, 2018) and word-word pairs (Martín-Buro, Wimber, Henson, & Staresina, 2019), likely requiring different processing demands than incremental sentence processing of a newly learned language. Analogous to this research, we found that a decrease in alpha power predicted greater behavioural gains in the wake condition for flexible word order rules. As such, the correlation between increased alpha power on behavioural performance may manifest as a result of overnight consolidation processes, rather than a 10-minute delay period from learning to retrieval (e.g., Minarik et al., 2018). Functionally, this relationship between increased alpha power and enhanced

behavioural performance may index increased activity in thalamo-cortical and cortico-cortical feedback loops (e.g., Klimesch, 1997, 1999), and by extension, the retrieval of (sleep-based) language-related long-term memory (Cross et al., 2018).

5.5. Beta power indexes knowledge accumulation and prediction during sentence processing

Current models of beta oscillations and sentence processing (e.g., Lewis, Schoffelen, Schriefers, & Bastiaansen, 2016; Lewis et al., 2015) posit that during sentence comprehension, an increase in beta power reflects the maintenance of model predictions, while a decrease in beta power reflects prediction error signals. Based on emerging evidence for the role of sleep in forming internal predictive models of the world (Friston, 2010; Hobson & Friston, 2012; Lutz et al., 2018; Rauss & Born, 2017), we predicted that sleep-based memory consolidation, linked to reductions in synaptic weight (Tononi & Cirelli, 2014), would accentuate beta power during sentence processing relative to an equivalent period of wake, with this increase predicting greater sensitivity to grammatical violations (Cross et al., 2018). Results from the current study partially support this hypothesis: low beta power during the baseline sentence judgement task predicted increased sensitivity to grammatical rules for both the sleep and wake conditions, while higher beta power predicted improved behavioural performance at delayed testing for flexible word order rules, and marginally for fixed word order rules for the sleep condition.

These results can be interpreted based on the proposal that oscillatory beta activity reflects top-down predictions (Fontolan, Morillon, Liegeois-Chauvel, & Giraud, 2014; Lewis & Bastiaansen, 2015; Lewis et al., 2015), and that these predictive codes are strengthened with increasing rule exposure rather than a period of sleep *per se*. From this perspective, after adequate exposure to novel rules, an unpredicted change in the (linguistic) sensory input may result in a decrease in beta power (e.g., in response to a grammatical violation), or alternatively, beta power may increase in a direction-specific manner in response to incoming linguistic information with high levels of certainty (e.g., in response to grammatically correct sentences; Lewis & Bastiaansen, 2015; Lewis et al., 2015). This idea is supported by the results of the current study: during the baseline sentence judgement task, which occurred after only forty minutes of learning, a decrease in beta power predicted improved sensitivity to both fixed and flexible word order rules (Figure 5C), possibly reflecting repeated internal model updating, helping to form implicit temporal predictions of linguistic regularities. After repeated exposure and an adequate consolidation period (i.e., 12hr of sleep or wake), an increase in beta power then predicted improved sensitivity to grammatical rules. Beta oscillations have been proposed to synchronise task-relevant neural networks that encode predicted sensory inputs (Arnal & Giraud, 2012), and as such, an increase in beta power after a delay period may reflect the synchronisation of newly established neural populations engaged during the processing of grammatical rules with high levels of predictability.

5.6. Limitations and future directions

One major limitation of the current study is the lack of an adaptation night and further control conditions, making it difficult to disentangle circadian effects on sleep and memory (e.g., Tucker et al., 2017). While our experimental design is analogous to previous work, future studies should include groups in AM-PM (12h Wake), PM-AM (12h Sleep), PM-PM (24h Sleep early) and AM-AM (24h Sleep late), as recommended by Nemeth, Gerbier and Janacsek (2019). We did, however, model participants' sleepiness levels and the $1/f$ exponent in our statistical analyses, which partially controlled for potential time-of-day effects. Further, the evidence presented here is correlational and neuroanatomical inferences are unable to be drawn based on scalp-recorded EEG. However, this is the first study to relate sleep-based memory consolidation mechanisms to sentence-level oscillatory activity, and as such, has set the foundation for future work using techniques with greater spatial-

temporal resolution. For example, electrocorticography and stereoelectroencephalography would allow for a better characterization of task-evoked cortical dynamics and spindle-slow oscillation coupling between cortical regions and the hippocampal complex, respectively (e.g., Helfrich et al., 2018, 2019). This approach could be further complemented by demonstrating a selective reinstatement of memory traces during spindle-slow oscillation coupling using representational similarity analysis (Zhang, Fell, & Axmacher, 2018). Identifying stimulus-specific representations during the encoding of sentence-level regularities and tracking the replay of stimulus activity related to spindle-slow oscillation events would further demonstrate the critical role of sleep-based oscillatory mechanisms on higher-order language learning.

5.7. Conclusion

In summary, our results demonstrate that the temporal coupling between NREM slow oscillations and spindles subserves the consolidation of complex sentence-level rules. In particular, we demonstrated that spindle-slow oscillation coupling promotes the consolidation of sequence-based rules, and also modulates task-evoked oscillatory activity previously implicated in language learning (e.g., Diego-Balaguer et al., 2011; Kepinska et al., 2017) and sentence processing (Meyer, 2018). Taken together, these findings add to models of sleep-based memory consolidation (e.g., Born & Wilhelm, 2012; Lewis & Durrant, 2011) and help characterise how effects of sleep-related oscillatory dynamics on memory manifest in oscillatory activity during complex language-related combinatorial operations.

Acknowledgements: We thank Louise Kyriaki for helpful discussions and feedback on an earlier version of this manuscript. We also thank the research assistants at the Cognitive and Systems Neuroscience Research Hub. Particular thanks to Isabella Sharrad, Erica Wilkinson, Nicole Vass and Angela Osborn for help with data collection. Thank you also to the participants.

REFERENCES

Alamia, A., & VanRullen, R. (2019). Alpha oscillations and traveling waves: Signatures of predictive coding? *PLoS Biol*, *17*(10), e3000487. https://doi.org/10.1371/journal.pbio.3000487

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. *https://doi.org/10.1111/psyp.13451*

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in cognitive sciences*, *16*(7), 390-398.

Austin, P. C., & Hux, J. E. (2002). A brief note on overlapping confidence intervals. *Journal of Vascular Surgery*, *36*(1), 194-195. doi: 10.1067/mva.2002.125015

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.

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*

Barachant, A., Andreev, A., & Congedo, M. (2013, January). The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry. In TOBI Workshop lV (pp. 19-20).

Barthélemy, Q., Mayaud, L., Ojeda, D., & Congedo, M. (2019). The Riemannian potato field: a tool for online signal quality index of EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *27*(2), 244-255.

Bastiaansen, M. C. M., van Berkum, J. J. A., & Hagoort, P. (2002). Syntactic Processing Modulates the θ Rhythm of the Human EEG. *NeuroImage, 17*(3), 1479–1492. https://doi.org/10.1006/nimg.2002.1275

Bastiaansen, M., & Hagoort, P. (2015). Frequency-based Segregation of Syntactic and Semantic Unification during Online Sentence Level Language Comprehension. *J Cogn Neurosci*, *27*(11), 2095–2107. *https://doi.org/10.1162/jocn_a_00829*

Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N., & Miller, E. K. (2020). Layer and rhythm specificity for predictive routing. *NEURON-D-20-00146*.

Bates, D. M. (2010). lme4: Mixed-effects modeling with R.

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

Bates, E., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: A cross-language perspective. *Annual review of psychology*, *52*(1), 369-396.

Batterink, L. J., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, *65*, 169–179. https://doi.org/10.1016/j.neuropsychologia.2014.10.024

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory (Cold Spring Harbor, NY)*, *22*(11), 544–556.

Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience*, *25*(6), 936–951.

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C., & Vaughn, B. (2012). The AASM manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine*.

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychol Res*, *76*(2), 192–203. *https://doi.org/10.1007/s00426-011-0335-6*

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: a new dorsal–ventral stream model of sentence comprehension. *Brain and language*, *125*(1), 60-76.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in psychology*, *10*, 298.

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., ... & Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and language*, *117*(3), 133-152.

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-Schlesewsky, I., Staub, A., & Schlesewsky, M. (2016). The timecourse of sentence processing in the brain. In *Neurobiology of language* (pp. 607-620). Academic Press.

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

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1988). The Pittsburgh Sleep Quality Index: A New Instrument for Psychiatric Practice and Research. *Psychiatry Research*, *28*, 20.

Buzsáki, G. (1996). The Hippocampo-Neocortical Dialogue. *Cerebral Cortex*, *6*(2), 81–92. https://doi.org/10.1093/cercor/6.2.81

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

Congedo, M., Barachant, A., & Andreev, A. (2013). A new generation of brain-computer interface based on Riemannian geometry. *arXiv preprint arXiv:1310.8115*.

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*

Covington, N. V., & Duff, M. C. (2016). Expanding the language network: Direct contributions from the hippocampus. *Trends in Cognitive Sciences*, *20*(12), 869-870.

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., Corcoran, A. W., Schlesewsky, M., Kohler, M. J., & Bornkessel-Schlesewsky, I. (2020c). Oscillatory and aperiodic neural activity jointly predict grammar learning. *BioRxiv.* doi: https://doi.org/10.1101/2020.03.10.984971

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. *Neuropsychologia*. doi: *http://dx.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

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3773-3800.

de Diego-Balaguer, R., Fuentemilla, L., & Rodriguez-Fornells, A. (10). Brain Dynamics Sustaining Rapid Rule Extraction from Speech. *Journal of Cognitive Neuroscience*, *23*(10), 3105–3120. Retrieved from pbh. (95462200)

Denis, D., Kim, S. Y., Kark, S. M., Daley, R. T., Kensinger, E. A., & Payne, J. D. (2020). Slow oscillation-spindle coupling impairs emotional memory consolidation following stress. *BioRxiv*. doi: https://doi.org/10.1101/2020.06.08.140350

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., ... & Shestyuk, A. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature neuroscience*, *23*(12), 1655-1665.

Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in human neuroscience*, *6*, 69.

Durrant, S. J., Cairney, S. A., & Lewis, P. A. (2016). Cross-modal transfer of statistical information benefits from sleep. *Cortex*, *78*, 85–99. https://doi.org/10.1016/j.cortex.2016.02.011

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322–1331. *https://doi.org/10.1016/j.neuropsychologia.2011.02.015*

Earle, F. S., & Myers, E. B. (2015). Sleep and native language interference affect non-native speech sound learning. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1680.

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

Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2013). Sleep restores loss of generalized but not rote learning of synthetic speech. *Cognition*, *128*(3), 280-286.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458–467. https://doi.org/10.1037/0278-7393.28.3.458

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

Fox, J. (2011). *Tests for Multivariate Linear Models with the car Package*. 99.

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

Freeman, W. J., & Zhai, J. (2009). Simulated power spectral density (PSD) of background electrocorticogram (ECoG). *Cognitive Neurodynamics*, *3*(1), 97–103. https://doi.org/10.1007/s11571-008-9064-y

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proc Natl Acad Sci U S A*, *99*(1), 529–534. https://doi.org/10.1073/pnas.012611199

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

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

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., … Hamalainen, M. (2013). MEG and EEG data analysis with MNE-Python. *Front Neurosci*, *7*, 267. *https://doi.org/10.3389/fnins.2013.00267*

Hagoort, P. (2016). MUC (Memory, Unification, Control): A model on the neurobiology of language beyond single word processing. In *Neurobiology of language* (pp. 339-347). Academic Press.

Hahn, M. A., Heib, D., Schabus, M., Hoedlmoser, K., & Helfrich, R. F. (2020). Slow oscillation-spindle coupling predicts enhanced memory formation from childhood to adolescence. *Elife*, *9*, e53730.

Hanslmayr, S., Staresina, B. P., & Bowman, H. (2016). Oscillations and Episodic Memory: Addressing the Synchronization/Desynchronization Conundrum. *Trends Neurosci*, *39*(1), 16–25. https://doi.org/10.1016/j.tins.2015.11.004

Hanslmayr, S., 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

Hasselmo, M. E., & Stern, C. E. (2014). Theta rhythm and the encoding and retrieval of space and time. *Neuroimage*, *85*, 656-666

Hauk, O., Giraud, A. L., & Clarke, A. (2017). Brain oscillations in language comprehension.

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values ofd′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46-51.

Helfrich, R. F., & Knight, R. T. (2016). Oscillatory dynamics of prefrontal cognitive control. *Trends in cognitive sciences*, *20*(12), 916-930.

Helfrich, R. F., Lendner, J. D., Mander, B. A., Guillen, H., Paff, M., Mnatsakanyan, L., … Knight, R. T. (2019). Bidirectional prefrontal-hippocampal dynamics organize information transfer during sleep in humans. *Nat Commun*, *10*(1), 3572. https://doi.org/10.1038/s41467-019-11444-x

Helfrich, R. F., Mander, B. A., Jagust, W. J., Knight, R. T., & Walker, M. P. (2018). Old Brains Come Uncoupled in Sleep: Slow Wave-Spindle Synchrony, Brain Atrophy, and Forgetting. *Neuron*, *97*(1), 221-230 e4. https://doi.org/10.1016/j.neuron.2017.11.020

Hobson, J. A., & Friston, K. J. (2012). Waking and dreaming consciousness: Neurobiological and functional considerations. *Prog Neurobiol*, *98*(1), 82–98. *https://doi.org/10.1016/j.pneurobio.2012.05.003*

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Holz, J., Piosczyk, H., Feige, B., Spiegelhalder, K., Baglioni, C., Riemann, D., & Nissen, C. (2012). EEG Sigma and slow-wave activity during NREM sleep correlate with overnight declarative and procedural memory consolidation. *J Sleep Res*, *21*(6), 612–619. https://doi.org/10.1111/j.1365-2869.2012.01017.x

Ichihara-Takeda, S., Yazawa, S., Murahara, T., Toyoshima, T., Shinozaki, J., Ishiguro, M., … Nagamine, T. (2015). Modulation of alpha activity in the parieto-occipital area by distractors during a visuospatial working memory task: A magnetoencephalographic study. *J Cogn Neurosci*, *27*(3), 453–463. https://doi.org/10.1162/jocn_a_00718

James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neurosci Biobehav Rev*, *77*, 1–13. https://doi.org/10.1016/j.neubiorev.2017.01.054

Jiang, H., van Gerven, M. A., & Jensen, O. (2015). Modality-specific alpha modulations facilitate long-term memory encoding in the presence of distracters. *J Cogn Neurosci*, *27*(3), 583–592. *https://doi.org/10.1162/jocn_a_00726*

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of personality and social psychology*, *103*(1), 54.

Kandylaki, K. D., Nagels, A., Tune, S., Kircher, T., Wiese, R., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2016). Predicting "when" in discourse engages the human dorsal auditory stream: An fMRI study using naturalistic stories. *Journal of Neuroscience*, *36*(48), 12180-12191.

Kaufeld, G., Bosker, H. R., Ten Oever, S., Alday, P. M., Meyer, A. S., & Martin, A. E. (2020). Linguistic structure and meaning organize neural oscillations into a content-specific hierarchy. *Journal of Neuroscience*.

Kavanau, J. L. (1997). Memory, sleep and the evolution of mechanisms of synaptic efficacy maintenance. *Neuroscience*, *79*(1), 7-44.

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

Kielar, A., Meltzer, J. A., Moreno, S., Alain, C., & Bialystok, E. (2014). Oscillatory responses to semantic and syntactic violations. *J Cogn Neurosci*, *26*(12), 2840–2862. *https://doi.org/10.1162/jocn_a_00670*

Kim, S. Y., Pardilla-Delgado, E., & Alger, S. E. (2017). Enhancing memory consolidation through slow oscillation and spindle synchronization. *Journal of Neuroscience*, *37*(48), 11517-11519.

Klimesch, W. (1997). EEG-alpha rhythms and memory processes. *International Journal of psychophysiology*, *26*(1-3), 319-340.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169-195.

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. H. (1999). Paradoxical'alpha synchronization in a memory task. *Cognitive Brain Research, 7*(4), 493-501.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Res Rev*, *53*(1), 63–88. *https://doi.org/10.1016/j.brainresrev.2006.06.003*

Klimesch, W., Schimke, H., Ladurner, G., & Pfurtscheller, G. (1990). Alpha frequency and memory performance. *Journal of Psychophysiology*.

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

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience*, *22*, 1598-1610. doi: https://doi.org/10.1038/s41593-019-0467-3

Koster, M., Finger, H., Kater, M., Schenk, C., & Gruber, T. (2016). Neuronal oscillations indicate sleep-dependent changes in the cortical memory trace. *Journal of Cognitive Neuroscience*.

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., ... & Knight, R. T. (2020). An electrophysiological marker of arousal level in humans. *Elife*, *9*, e55092.

Lerner, I., & Gluck, M. A. (2019). Sleep and the extraction of hidden regularities: a systematic review and the importance of temporal rules. *Sleep medicine reviews*, *47*, 39-50.

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., Lemhfer, K., Schoffelen, J. M., & Schriefers, H. (2016). Gender agreement violations modulate beta oscillatory dynamics during sentence comprehension: A comparison of second

language learners and native speakers. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2016.06.031

Lewis, A. G., Schoffelen, J. M., Schriefers, H., & Bastiaansen, M. (2016). A Predictive Coding Perspective on Beta Oscillations during Sentence-Level Language Comprehension. *Front Hum Neurosci*, *10*, 85. https://doi.org/10.3389/fnhum.2016.00085

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain Lang*, *148*, 51–63. https://doi.org/10.1016/j.bandl.2015.01.003

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

Lewis, P. A., Knoblich, G., & Poe, G. (2018). How Memory Replay in Sleep Boosts Creative Problem-Solving. *Trends Cogn Sci*, *22*(6), 491–503. https://doi.org/10.1016/j.tics.2018.03.009

Lutz, N. D., Wolf, I., Hubner, S., Born, J., & Rauss, K. (2018). Sleep Strengthens Predictive Sequence Coding. *J Neurosci*, *38*(42), 8989–9000. *https://doi.org/10.1523/JNEUROSCI.1352-18.2018*

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. *https://doi.org/10.1016/S0022-5371(84)90093-8*

Maingret, N., Girardeau, G., Todorova, R., Goutierre, M., & Zugaro, M. (2016). Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nature neuroscience*, *19*(7), 959-964.

Martín-Buro, M. C., Wimber, M., Henson, R. N., & Staresina, B. P. (2019). https://doi.org/10.1101/708602

Mathewson, K. E., Lleras, A., Beck, D. M., Fabiani, M., Ro, T., & Gratton, G. (2011). Pulsed Out of Awareness: EEG Alpha oscillations represent a pulsed inhibition of ongoing cortical processing. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00099

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.

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609-2621.

Meyer, L., Grigutsch, M., Schmuck, N., Gaston, P., & Friederici, A. D. (2015). Frontal-posterior theta oscillations reflect memory retrieval during sentence comprehension. *Cortex*, *71*, 205–218. *https://doi.org/10.1016/j.cortex.2015.06.027*

Meyer, L., Grigutsch, M., Schmuck, N., Gaston, P., & Friederici, A. D. (2015). Frontal–posterior theta oscillations reflect memory retrieval during sentence comprehension. *Cortex*, *71*, 205-218.

Mikutta, C., Feige, B., Maier, J. G., Hertenstein, E., Holz, J., Riemann, D., & Nissen, C. (2019). Phase-amplitude coupling of sleep slow oscillatory and spindle activity correlates with overnight memory consolidation. *J Sleep Res*, *28*(6), e12835. https://doi.org/10.1111/jsr.12835

Minarik, T., Berger, B., & Sauseng, P. (2018). The involvement of alpha oscillations in voluntary attention directed towards encoding episodic memories. *Neuroimage*, *166*, 307–316. https://doi.org/10.1016/j.neuroimage.2017.10.064

Mirković, J., & Gaskell, M. G. (2016). Does Sleep Improve Your Grammar? Preferential Consolidation of Arbitrary Components of New Linguistic Knowledge. *PLoS One*, *11*(4), e0152489. https://doi.org/10.1371/journal.pone.0152489

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*(4), 14.

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annual review of psychology*, *67*, 105-134.

Muehlroth, B. E., Sander, M. C., Fandakova, Y., Grandy, T. H., Rasch, B., Shing, Y. L., & Werkle-Bergner, M. (2019). Precise Slow Oscillation-Spindle Coupling Promotes Memory Consolidation in Younger and Older Adults. *Sci Rep*, *9*(1), 1940. https://doi.org/10.1038/s41598-018-36557-z

Muthukumaraswamy, S. D., & Liley, D. TJ. (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. *https://doi.org/10.1016/j.neuroimage.2018.06.068*

Navarrete, M., Schneider, J., Ngo, H. V. V., Valderrama, M., Casson, A. J., & Lewis, P. A. (2020). Examining the optimal timing for closed-loop auditory stimulation of slow-wave sleep in young and older adults. *Sleep*, *43*(6), zsz315.

Nemeth, D., Gerbier, E., & Janacsek, K. (2019). *Four Pitfalls in Sleep and Memory Research and How to Avoid Them*. https://doi.org/10.20944/preprints201908.0208.v1

Nicholls, M. E., Thomas, N. A., Loetscher, T., & Grimshaw, G. M. (2013). The Flinders Handedness survey (FLANDERS): A brief measure of skilled hand preference. *Cortex*, *49*(10), 2914–2926. *https://doi.org/10.1016/j.cortex.2013.02.002*

Nieuwenhuis, I. L., Folia, V., Forkstam, C., Jensen, O., & Petersson, K. M. (2013). Sleep promotes the extraction of grammatical rules. *PloS one*, *8*(6), e65046.

Niknazar, M., Krishnan, G. P., Bazhenov, M., & Mednick, S. C. (2015). Coupling of Thalamocortical Sleep Oscillations Are Important for Memory Consolidation in Humans. *PLoS One*, *10*(12), e0144720. https://doi.org/10.1371/journal.pone.0144720

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*

Osterhout, L., & Holcomb, P. J. (1995). Event related potentials and language comprehension.

Ouyang, G., Hildebrandt, A., Schmitz, F., & Herrmann, C. S. (2019). 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

Park, H., Lee, D. S., Kang, E., Kang, H., Hahm, J., Kim, J. S., … Jensen, O. (2014). Blocking of irrelevant memories by posterior alpha activity boosts memory encoding. *Human Brain Mapping*, *35*(8), 3972–3987. https://doi.org/10.1002/hbm.22452

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Piai, V., Anderson, K. L., Lin, J. J., Dewar, C., Parvizi, J., Dronkers, N. F., & Knight, R. T. (2016). Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. *Proceedings of the National Academy of Sciences*, 201603312.

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*(4), 413-425.

Rasch, B. (2017). Sleep and language learning. *Brain Lang*, *167*, 1–2. https://doi.org/10.1016/j.bandl.2017.02.002

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiol Rev*, *93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Rauss, K., & Born, J. (2017). A Role of Sleep in Forming Predictive Codes. In *Cognitive Neuroscience of Memory Consolidation* (pp. 117–132). Springer.

Roehm, D., Schlesewsky, M., Bornkessel, I., Frisch, S., & Haider, H. (2004). Fractionating language comprehension via frequency characteristics of the human EEG. *NeuroReport, 15*(3), 409–412.

Roehm, Dietmar, Klimesch, W., Haider, H., & Doppelmayr, M. (2001). The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neuroscience Letters, 310*, 137–140.

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., & Rasch, B. (2016). The beneficial role of memory reactivation for language learning during sleep: A review. *Brain Lang*. https://doi.org/10.1016/j.bandl.2016.02.005

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

Solano, A., Riquelme, L., Perez-Chada, D., & Della-Maggiore, V. (2020). Local coupling between sleep spindles and slow oscillations supports the consolidation of motor memories. *bioRxiv*.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, *31*(1), 137-149.

Staresina, B. P., Bergmann, T. O., Bonnefond, M., van der Meij, R., Jensen, O., Deuker, L., … Fell, J. (2015). Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nat Neurosci*, *18*(11), 1679–1686. https://doi.org/10.1038/nn.4119

Tononi, G., & Cirelli, C. (2014). Sleep and the price of plasticity: From synaptic and cellular homeostasis to memory consolidation and integration. *Neuron*, *81*(1), 12–34. https://doi.org/10.1016/j.neuron.2013.12.025

Tort, A. B., Komorowski, R. W., Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009). Theta–gamma coupling increases during the learning of item–context associations. *Proceedings of the National Academy of Sciences*, *106*(49), 20942-20947.

Toth, B., Janacsek, K., Takacs, A., Kobor, A., Zavecz, Z., & Nemeth, D. (2017). Dynamics of EEG functional connectivity during statistical learning. *Neurobiol Learn Mem*, *144*, 216–229. https://doi.org/10.1016/j.nlm.2017.07.015

Tucker, M. A., Morris, C. J., Morgan, A., Yang, J., Myers, S., Pierce, J. G., … Scheer, F. A. J. L. (2017). The Relative Impact of Sleep and Circadian Drive on Motor Skill Acquisition and Memory Consolidation. *Sleep*, *40*(4). *https://doi.org/10.1093/sleep/zsx036*

Tune, S., Schlesewsky, M., Nagels, A., Small, S. L., & Bornkessel-Schlesewsky, I. (2016). Sentence understanding depends on contextual use of semantic and real world knowledge. *NeuroImage*, *136*, 10-25.

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software, 3*(31), 1026, *https://doi.org/10.21105/joss.01026*

Vallat, R., & Jajcay, N. (2020). YASA (yet another spindle algorithm). doi: *https://zenodo.org/record/4244889#.X8zv1RMzalM*

Van Dongen, H. P., Olofsen, E., Dinges, D. F., & Maislin, G. (2004). Mixed-model regression analysis and dealing with interindividual differences. *Methods in Enzymology, 384*, 139-171.

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*

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, *427*(6972), 352-355.

Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J. M., Magyari, L., … Bastiaansen, M. (2012). Beta oscillations relate to the N400m during language comprehension. *Hum Brain Mapp*, *33*(12), 2898–2912. https://doi.org/10.1002/hbm.21410

Weber, K., Christiansen, M. H., Petersson, K. M., Indefrey, P., & Hagoort, P. (2016). FMRI Syntactic and Lexical Repetition Effects Reveal the Initial Stages of Learning a New Language. *J Neurosci*, *36*(26), 6872–6880. https://doi.org/10.1523/JNEUROSCI.3180-15.2016

Wen, H., & Liu, Z. (2015). Separating Fractal and Oscillatory Components in the Power Spectrum of Neurophysiological Signal. 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 Topogr*, *29*(1), 13–26. *https://doi.org/10.1007/s10548-015-0448-0*

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Zhang, H., Fell, J., & Axmacher, N. (2018). Electrophysiological mechanisms of human memory consolidation. *Nature Communications*, *9*(1), 1–11.