The Role of Modality-specific Brain Regions in Statistical Learning: Insights from Intracranial Neural Entrainment

Daniela Herrera-Chaves[1], Greydon Gilmore[1], Mohamad Abbass[1], Lyle Muller[1], Ana Suller-Marti[1], Seyed M. Mirsattari[2], Stefan Köhler[1*], and Laura Batterink[1*]

Abstract

■ Statistical learning (SL) is a powerful mechanism that supports the ability to extract regularities from environmental input. Yet, its neural underpinnings are not well understood. Previous EEG studies of SL have found that the brain tracks regularities by synchronizing its activity with the presented stimuli—a phenomenon known as neural entrainment. However, EEG lacks the spatial resolution to unveil the specific brain regions where this process takes place. In our study, 18 patients with drug-resistant epilepsy who were implanted with intracranial electrodes for presurgical investigation listened to a continuous speech stream containing embedded trisyllabic words. Neural entrainment was measured at the syllable and word frequencies, with the latter providing an online index of learning. SL was further assessed through both explicit and implicit behavioral measures. Behaviorally, we found evidence of learning at the group level in both tasks. At the neural level, our analyses revealed three temporal tuning profiles: 25% of contacts showed entrainment at the syllable frequency, 11% of contacts showed entrainment at both the word and syllable frequencies, and 4% showed entrainment only to the word frequency. Word entrainment, indicating sensitivity to word structures, was most commonly found in auditory and language-related regions, including insula, middle temporal gyrus, superior temporal gyrus, and supramarginal gyrus. In contrast, evidence for neural entrainment in the hippocampus was weak. Overall, these results support the idea that speech-based SL is largely supported by modality-specific brain regions. ■

INTRODUCTION

The external world bombards us with a continuous flow of sensory input. As complex as this input may be, it contains structure in the form of regularities that repeat over time. Statistical learning (SL) is a fundamental learning mechanism that supports the extraction of such regularities from the environment based on their distributional patterns across time and space (Frost, Armstrong, & Christiansen, 2019). In one of the first demonstrations of SL, Saffran, Aslin, and Newport (1996) exposed infants to a continuous speech stream containing repeating nonsense words that offered no acoustic cues to indicate word boundaries (i.e., no pauses or tone changes between words). As such, the only cues to discover word boundaries were the co-occurrence statistics between neighboring syllables—namely, that syllables within words occurred together more often than syllables across words. Following only 2 min of exposure, a looking-time test revealed that infants were able to distinguish the words in the stream from recombined foil items, suggesting that they had become sensitive to the statistical regularities in the input. Numerous studies have since replicated and expanded upon this initial finding (Isbilen & Christiansen, 2022; Frost et al., 2019), demonstrating that SL is not limited to auditory speech segmentation but also supports the extraction of visual (Fiser & Aslin, 2001, 2002a, 2002b), tactile (Conway & Christiansen, 2005), and nonlinguistic auditory patterns (Moser et al., 2021; Saffran, Johnson, Aslin, & Newport, 1999). Furthermore, in addition to infants (e.g., Choi, Batterink, Black, Paller, & Werker, 2020; Fló et al., 2019), evidence of SL has also been found in children (e.g., Moreau, Joanisse, Mulgrew, & Batterink, 2022; Arciuli & Simpson, 2011; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), adults of different ages (e.g., Wang, Köhler, & Batterink, 2023; Saffran et al., 1997; Saffran, Newport, et al., 1996), and even nonhuman species (e.g., Arnon et al., 2025; Boros et al., 2021). Thus, SL is currently conceptualized as a universal process that likely supports many different aspects of perception and cognition (Bogaerts, Frost, & Christiansen, 2020; Sherman, Graves, & Turk-Browne, 2020).

While SL studies have traditionally focused on assessing the behavioral outcomes of SL (Isbilen & Christiansen, 2022), in recent years there has been a growing interest in elucidating the neural mechanisms active during the

[1]University of Western Ontario, [2]Mayo Clinic, Jacksonville, FL
*Stefan Köhler and Laura Batterink contributed equally to the study.

Journal of Cognitive Neuroscience 38:5, pp. 906–926
https://doi.org/10.1162/JOCN.a.2411

learning process itself. Studies using fMRI have commonly found the engagement of modality-specific sensory regions in SL. For instance, studies using visual stimuli have found activation in occipital areas, including lateral occipital cortex and ventral occipitotemporal cortex (Karuza et al., 2017; Turk-Browne, Scholl, Chun, & Johnson, 2009). Other studies using linguistic auditory stimuli found activation in the superior temporal gyrus (STG) when participants were exposed to artificial language streams, as contrasted with a control condition (Karuza et al., 2013; Cunillera et al., 2009; McNealy, Mazziotta, & Dapretto, 2006). Activation was also observed in other regions along the temporal cortex, including middle temporal gyrus (MTG; Karuza et al., 2013; McNealy et al., 2006), posterior temporal gyrus (Schneider, Scott, Legault, & Qi, 2024), and transverse temporal gyrus (McNealy et al., 2006). Furthermore, a recent study found engagement of the left posterior temporal gyrus during a linguistic auditory SL task and a passive story listening task, but not during a nonlinguistic auditory SL task (Schneider et al., 2024). This suggests that linguistic auditory SL may recruit hubs within the language network that are not engaged by nonlinguistic auditory stimuli, possibly reflecting domain specificity along with modality specificity in SL.

In addition to modality- and domain-specific cortical areas, there are a number of regions that appear to support SL across modalities and domains. For example, both visual and auditory SL studies have found engagement of the BG, particularly the caudate nucleus (Karuza et al., 2013; Turk-Browne et al., 2009) and putamen (Karuza et al., 2013, 2017; McNealy et al., 2006), which is perhaps not surprising given the noted similarities between SL and implicit learning (Batterink, Paller, & Reber, 2019; Christiansen, 2019; Perruchet & Pacton, 2006). The inferior frontal gyrus (IFG), a region that has been implicated in sequential structure learning (e.g., Schapiro, Rogers, Cordova, Turk-Browne, & Botvinick, 2013; Petersson, Folia, & Hagoort, 2012), has also been implicated in visual (Turk-Browne et al., 2009), linguistic auditory (Karuza et al., 2013), and nonlinguistic auditory SL (Abla & Okanoya, 2008).

However, another important brain region that has attracted considerable interest as a potential domain-general hub of SL is the hippocampus. Evidence from a neural network model that simulates known properties of the hippocampus supports the idea that this structure may contribute to SL (Schapiro, Turk-Browne, Botvinick, & Norman, 2017). Within this computational model, the simulated trisynaptic pathway, which projects from entorhinal cortex to dentate gyrus and then through CA3 and CA1, was able to support the generation of nonoverlapping representations of highly similar episodes (a process known as pattern separation). Conversely, the monosynaptic pathway, which projects from entorhinal cortex directly to CA1, was demonstrated to support learning of regularities across episodes (SL). While the evidence for the specific involvement of the monosynaptic pathway in SL is still limited (but see Wang, Rosenbaum, et al., 2023, for discussion), broader evidence for hippocampal involvement in SL comes from numerous fMRI studies. Two studies found greater BOLD activation in the hippocampus to structured versus random visual sequences (Ellis et al., 2021; Turk-Browne et al., 2009), while a study using visual–spatial stimuli showed that activation of the hippocampus and other medial temporal lobe regions was associated with behavioral performance in a subsequent recognition task (Karuza et al., 2017). Furthermore, an fMRI study using multivariate pattern analysis found that exposure to pairs of images that appear sequentially (image A is always followed by image B) led to an increase in representational similarity of corresponding activation for these two items in the hippocampus as learning took place (Schapiro, Kustner, & Turk-Browne, 2012). However, so far, activation of the hippocampus has generally not been found in the auditory-linguistic domain (Schneider et al., 2024; Orpella et al., 2022; Karuza et al., 2013; McNealy, Mazziotta, & Dapretto, 2010; McNealy et al., 2006), suggesting that hippocampal involvement could be limited to visual SL paradigms.

Alongside studies that have focused on understanding the brain regions involved in SL, another line of work has used neural entrainment as an online index of SL. Neural entrainment can be defined broadly as the temporal alignment of neural activity with regularities in a stimulus stream (Obleser & Kayser, 2019). By presenting individual stimuli within a structured input stream at a fixed rate, neural entrainment can be induced at frequencies corresponding both to the individual stimuli (e.g., syllables) and to the underlying statistical structure (e.g., words). Neural entrainment at the frequency of statistical regularities, in the absence of sensory cues to their boundaries, is interpreted to reflect the perceptual binding of items into structured word units (e.g., Batterink & Paller, 2017). Numerous EEG studies of linguistic SL have indeed found neural entrainment at the frequency of the embedded words in continuous speech, which generally tends to increase over the course of learning (see Sjuls, Harvei, & Vulchanova, 2024, for a review). In addition, a subset of these studies have found that entrainment to words correlates with performance on implicit and/or explicit behavioral measures of learning (e.g., Batterink, 2020; Choi et al., 2020; Batterink & Paller, 2017, 2019; Buiatti, Pena, & Dehaenelambertz, 2009). Neural entrainment to the statistical structure has also been shown to correlate with individual language-related abilities, including phonological awareness (Zhang, Riecke, & Bonte, 2021) and spelling skills (Ringer, Sammler, & Daikoku, 2024), suggesting that sensitivity to structure at the neural level may relate to other aspects of language more broadly.

A handful of studies have incorporated neural entrainment approaches with intracranial EEG (iEEG), providing insights into where in the brain entrainment occurs. Henin et al. (2021) recorded iEEG data in patients with epilepsy who were presented with both auditory and visual

statistical streams and found entrainment to words in modality-specific sensory cortices (i.e., temporal cortices) as well as parietal and frontal regions including IFG. No entrainment effects were found in the hippocampus. An additional representational similarity analysis revealed, however, that the hippocampus uniquely represents the identity of the bound statistical units, showing more similar patterns of neural activity for stimuli that belonged to the same statistical unit than those that did not. Another iEEG study examined visual SL of both exemplar-level and category-level regularities (Sherman et al., 2023). In the exemplar-level condition, six exemplar images of natural landscapes were organized in three arbitrary pairs (e.g., an image of a canyon followed by a mountain), whereas in the category-level condition, regularities occurred at category level (e.g., unique images of canyons were presented for each trial). Significant neural entrainment at the pair frequency was found in both types of regularities throughout visual cortex, in addition to frontal and temporal cortex, but it was not possible to examine the involvement of the hippocampus due to insufficient coverage. An additional iEEG study examined the contributions of the hippocampus and auditory cortex in a linguistic auditory SL task (Ramos-Escobar et al., 2022). Results indicated increased spectral power at the syllable frequency in auditory cortex while the hippocampus showed increased power at the word frequency, suggesting that these two regions process different levels of regularities.

Taken together, results from previous studies have provided crucial insights into which brain structures support SL and how the brain rapidly tracks statistical regularities. However, due to limitations of any given neuroimaging method and conflicting results across studies, the involvement of the hippocampus in SL remains unclear, particularly for auditory linguistic stimuli. Furthermore, prior iEEG studies have generally not demonstrated behavioral evidence of learning in their patients. Specifically, Sherman et al. (2023) did not include any behavioral measures, while the patients in the studies by Ramos-Escobar et al. (2022) and Henin et al. (2021) performed at chance on a two-alternative forced-choice recognition measure. Henin and colleagues also assessed patients' RTs on a 1-back cover task during exposure and found generally faster RTs to structured compared with random exposure streams, though these overall faster RTs cannot necessarily be taken as evidence of regularity learning per se. In the absence of behavioral evidence, it is not clear whether the neural entrainment effects found in these studies can be safely attributed to the learning process. Moreover, given theoretical discussions about a potential role of the hippocampus that may be limited to the explicit retrieval of statistical regularities, inclusion of both implicit and explicit behavioral measures would help constrain interpretation of entrainment effects.

In the current iEEG study, we aimed to contribute to our understanding of the SL process by characterizing which brain regions show sensitivity to online statistical structure in input. To address limitations from previous studies, we recruited a sample of patients with epilepsy who had ample electrode coverage of the hippocampus (in addition to other regions) and included sensitive behavioral measures to assess both implicit and explicit learning outcomes of SL. We presented patients with a continuous speech stream that contained four embedded trisyllabic words. With this setup, entrainment to syllables provides an index of sensory-level processing, and entrainment to words offers an index of SL (Henin et al., 2021; Batterink & Paller, 2017). After the exposure phase, SL was assessed explicitly through a familiarity rating task and implicitly through a RT-based target detection task. Based on prior evidence of the engagement of modality-specific sensory regions in SL, we predicted that primary and associative auditory regions would show significant neural entrainment to both the syllable and word frequencies. Furthermore, based on the results from some fMRI (Karuza et al., 2017; Schapiro et al., 2012; Turk-Browne et al., 2009), lesion (Covington, Brown-Schmidt, & Duff, 2018; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014), and iEEG studies (Ramos-Escobar et al., 2022), in addition to the computational model developed by Schapiro et al. (2017), we predicted that entrainment to the word frequency would be observed in the hippocampus. Lastly, based on the findings from previous fMRI studies using linguistic auditory stimuli on their SL paradigms (Schneider et al., 2024; Karuza et al., 2013; Cunillera et al., 2009; McNealy et al., 2006) and on literature on language lateralization more broadly (Fedorenko, Ivanova, & Regev, 2024), we predicted that neural entrainment to the word frequency would be lateralized to the left hemisphere in regions within the language network, particularly temporal cortices. We also expected to observe significant evidence of SL on our implicit and explicit behavioral tasks, allowing us to confirm that learning occurred in the majority of patients.

METHODS

Participants

We tested 18 patients with drug-resistant epilepsy (eight women, age range: 19–51 years, mean age = 31.2 years) who had undergone implantation of intracranial depth electrodes (i.e., stereoelectroencephalography [sEEG]) for presurgical investigation (see Table 1 for patient information). Patients were recruited from the Epilepsy Monitoring Unit at University Hospital in London, Ontario, Canada, and provided signed informed consent according to the Western University research ethics board. Sample size was determined by previous iEEG studies (Sherman et al., 2022, 2023; Henin et al., 2021) and patient availability during our data collection period (July 2023 to November 2024). All patients completed the four experimental tasks without interruptions.

Table 1. General Patient Information

Participant ID	*Age (years)*	*Sex*	*Handedness*	*No. of Contacts*
1	25	F	R	104
2	19	F	R	123
3	51	F	R	154
4	39	M	L	97
5	34	F	R	126
6	20	M	R	118
7	35	F	R	104
8	23	M	R	121
9	24	F	R	129
10	36	M	R	132
11	27	F	R	117
12	25	M	R	146
13	26	M	R	128
14	26	F	L	126
15	25	M	L	131
16	39	M	R	116
17	51	M	R	135
18	36	M	R	154

The number of contacts indicates the electrode contacts that went into further stages of analyses after preprocessing (see below). M = male; F = female; R = right; L = left.

Stimuli

The structured language consisted of 12 unique syllables originally used by Batterink and Paller (2019). The syllables were recorded by a male native English speaker who spoke with neutral intonation. Each syllable was spoken in isolation to avoid co-articulation between syllables and had an approximate duration of 220–250 msec from onset to offset. Each syllable was extracted into its own sound file, with the beginning of each sound file coinciding with the onset of the syllable. These 12 syllables were arranged to form four trisyllabic nonsense words (*bafuko, regeme, fetisu, rupuni*), used for the structured exposure stream and subsequent tasks. A second set of stimuli for the random exposure stream consisted of 12 additional syllables (*bi, bu, da, do, go, ku, la, pa, pi, ro, ti, tu*), created using Google Cloud's text-to-speech synthesizer. We note that these syllables, while synthesized, sounded very similar to natural human-produced speech. We chose to create the syllables for the random condition with a speech synthesizer because it offered an efficient, straightforward, and well-controlled way to match the structured syllables on features such as syllable consistency, articulation, and pitch. Syllables were produced by a female voice to minimize any interference with the structured stimuli set and were of similar duration to structured syllables (220–250 msec). Syllables were presented in a pseudorandom order to create the random exposure stream, which served as a control condition for the structured exposure stream.

Experimental Procedure

All patients performed the four tasks in the same order (see Figure 1). The tasks were carried out on a desktop computer placed on a portable cart designed for research testing.

Structured Exposure Task

Patients were instructed to listen to an "alien" message that researchers needed help to decipher, and importantly, they were not told that the audio contained four words nor that the words were trisyllabic. Patients then listened passively to a continuous stream made up of the four trisyllabic words (*bafuko, regeme, fetisu, rupuni*), concatenated in a predefined pseudorandom order with the constraint that the same word could not appear consecutively. To enable neural entrainment analyses, the individual syllables in the speech stream appeared at a fixed rate of 300 msec (3.33 Hz). Consequently, the trisyllabic words appeared at a rate of 900 msec (1.11 Hz). Each word appeared 90 times throughout the speech stream, resulting in a total duration of 5.4 min. Critically, there were no pauses or auditory cues between words. Thus, the only indication of word boundaries was the statistical co-occurrence regularities among the syllables.

Target Detection Task

This task was designed as an implicit measure of patients' knowledge of the statistical regularities present in the structured exposure, as it does not require the intentional retrieval of the learned regularities. In this task, participants were asked to detect specific target syllables as quickly as possible, which were embedded within shorter snippets of the artificial language. Because syllables that occur in later (i.e., second and third positions) within a word are more predictable, they should elicit faster RTs than syllables in the first position (Wang, Köhler, & Batterink, 2023; Batterink et al., 2019; Batterink & Paller, 2017; Batterink, Reber, Neville, & Paller, 2015).

Prior to the beginning of each stream, the target syllable for that given stream (e.g., "re") appeared written on-screen and was played twice. The stream was then initiated and contained the four words from the structured language repeated four times each in a pseudorandom order. The written form of the syllable remained on screen throughout the duration of the stream. Each stream contained four targets, and across the task, a total of 36 streams were presented, organized into three blocks. Each of the 12 syllables of the stimulus inventory served as the target for a stream three times. This yielded a total of 144

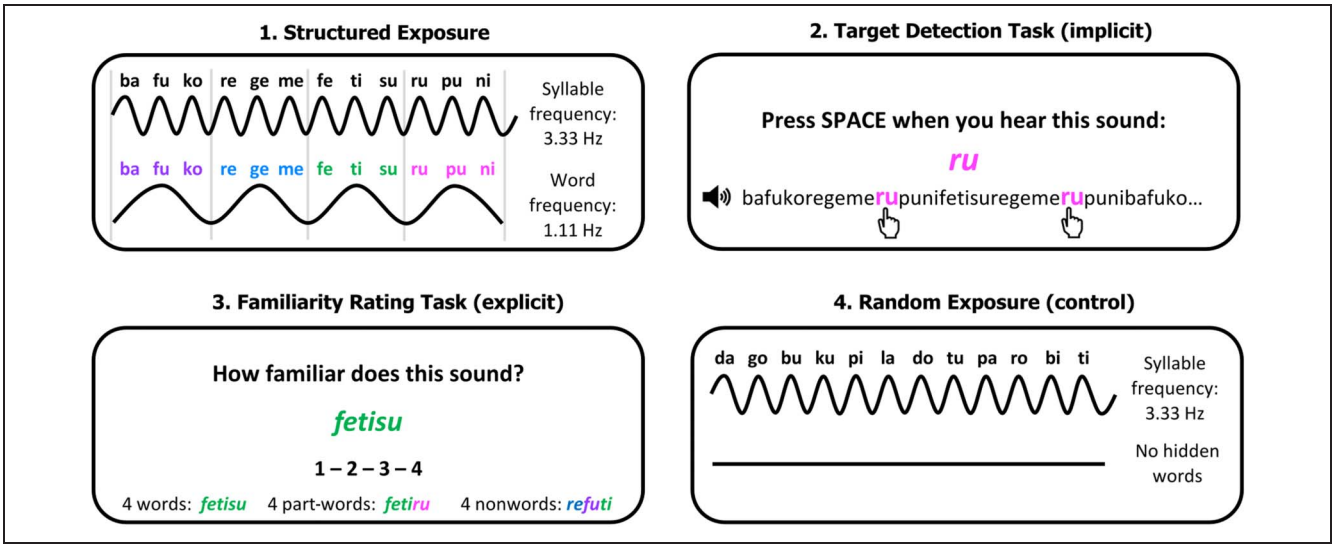

Figure 1. Task design and experimental procedure. Patients first listened to a 5.4-min structured speech stream that contained four embedded trisyllabic words. Next, they performed the target detection task, where they listened to shorter snippets of the structured exposure stream and reacted to a specific syllable in each stream. Then, patients completed the familiarity rating task, where they rated the familiarity of the three word types (word, part-word, and nonword) on a scale from 1 to 4. Finally, patients listened to a control speech stream in which new syllables were presented in a pseudorandom order with no underlying word structures.

targets, 48 in each triplet position (first, second and third). Stimulus timing parameters were identical to those in the exposure task, resulting in an individual stream duration of 14.4 sec. Block order was randomized across participants.

Before starting the task, patients completed two practice trials that used a different voice and different syllables from the main task. After each practice trial, the average RT and number of hits was shown on screen. No feedback was given on the main task.

Familiarity Rating Task

This task assessed patients' explicit recognition of the word structures. Here, patients were presented with three different types of word structures: words, part-words, and nonwords. Words were the same ones that appeared in the two previous tasks: *bafuko, regeme, fetisu, rupuni*. Part-words were reconfigured foils that had two syllables from the same word along with a syllable from a different word[1] (e.g., *bafu* + *ni*), and nonwords were made up of syllables taken from three different words (e.g., *fe* + *pu* + *ko*). There were 12 unique trials overall (four words, four part-words, and four nonwords). On each trial, patients listened to either a word, part-word, or nonword and were then instructed to indicate how familiar it sounded to them on a scale from 1 to 4, with 4 being the most familiar. SL is indexed by higher ratings to words compared with part-words and nonwords.

Random Exposure Task

As in the structured exposure task, the 12 syllables created for the random condition were presented at a fixed rate of 300 msec (3.33 Hz). However, the syllables were not arranged into words and were instead presented in a pseudorandom order with no underlying statistical structure. Here, the only constraint was that the same syllable could not appear consecutively. Each syllable was repeated 90 times, resulting in a total duration of 5.4 min. As in the structured exposure task, participants passively listened to the speech stream without interruptions or cover tasks.

All experimental tasks were presented using PsychoPy (Peirce et al., 2019).

iEEG Recordings

iEEG data were collected using NATUS NeuroWorks EEG Software, with a sampling rate of 2048 Hz. The signals were recorded from Ad-Tech Medical Instrument Corporation depth electrodes for sEEG. Electrodes were 0.86 mm in diameter, and each electrode had 10 contacts/channels with 3-, 4-, 5-, or 6-mm spacing in between contacts. During data acquisition, signals were referenced to a subdermal electrode implanted at the top of the head (frontoparietal galea). A separate channel was used to insert markers (or "triggers") that aligned the stimuli from our experimental tasks with the iEEG recordings. More specifically, a pulse was sent to the trigger channel at every syllable onset in both exposure conditions (structured and random). The location of electrodes was based solely on each patient's clinical needs without consideration of research goals of this or any other research study.

iEEG Data Preprocessing

Current analysis of the iEEG recordings was limited to the structured and random exposure conditions. All preprocessing was performed using the EEGLAB toolbox (Version 2023.0) in MATLAB (Version 2023b; The MathWorks). First, the data were downsampled to 512 Hz. The pulse peaks from the trigger channel corresponding to syllable onsets were identified with a peak-finding function in MATLAB and marked to create events in the data. The data were then notch-filtered at 60 Hz and its first three harmonics (120, 180, and 240 Hz) and band-pass filtered at 0.2–250 Hz with the *pop_eegfiltnew* function in EEGLAB. A baseline correction was applied using the *pop_rmbase* function. Electrode contacts localized in cerebrospinal fluid and skull, as well as empty channels, were removed. Then, the data were visually inspected, and any contact showing excessive noise was also removed. From a total of 2410 initial electrode contacts across participants (M = 133.9 per participant, range = 110–160), 149 were removed (M = 8.28, range = 1–17). The final data set consisted of 2261 electrode contacts, including 813 within gray matter, 1334 within white matter, 107 crossing gray and white matter, and seven crossing gray matter and CSF (see Supplemental Information A for overall coverage across the brain).

We opted to include data from all the above contacts, given that even contacts classified as fully within white matter are often near the boundary of gray matter and thus record activity that has spread from these neighboring gray matter regions (Mercier et al., 2022). We also note that electrode contacts distant from gray matter were categorized as "unclassified" and therefore did not contribute to region-specific analyses. We nonetheless conducted a supplemental analysis that focuses only on gray matter contacts (see Supplemental Information B) and found very similar results to our main analysis.

Given that noise artifacts spanning multiple channels that would warrant epoch removal are rare in iEEG recordings, all epochs were maintained in the data sets without epoch-wise trial rejection. The data were re-referenced to the common average of the remaining contacts across all electrodes using the *pop_reref* function. Finally, using the events created from the trigger channel, the data were segmented into nonoverlapping epochs of 10.8 sec, time-locked to the onset of every 36 syllables (12 words in the structured condition), yielding a total of 30 epochs in each condition.

Electrode Localization

All steps pertaining to electrode localization were performed by the clinical team at the Epilepsy Monitoring Unit at University Hospital in London, Ontario. The processing pipeline involved electrode contact localization, brain tissue segmentation, and atlas fitting. First, the location of electrode contacts was performed semi-automatically in 3D Slicer using the SEEG Assistant module (Narizzano et al., 2017). The entry and target points of each electrode were manually defined on the postoperative CT image. These entry/target labels were then provided to the SEEGA algorithm, which automatically segmented the individual electrode contacts. Then, brain tissue segmentation and atlas fitting were carried out to obtain information on the particular areas of the brain that individual electrode contacts were localized in. For this step, the pre-operative T1 MRI scans were nonlinearly registered to the Montreal Neurological Institute template (Fonov et al., 2011; Fonov, Evans, McKinstry, Almli, & Collins, 2009) using NiftyReg (Modat et al., 2010). Then, an anatomical mask was generated by applying the inverse transform to the T1 scan using the *antsApplyTransforms* algorithm from Advanced Normalization Tools 2.2.0 (Tustison et al., 2021).

Data Analysis

Behavioral Data

Target detection task. For each patient, average RTs were calculated for the syllables in the first, second, and third positions within words. Keyboard responses were considered "hits" if they occurred within a 0–1200 msec time window after the onset of the target syllable (a criterion employed in prior studies: Wang, Köhler, et al., 2023; Wang, Rosenbaum, et al., 2023; Batterink & Paller, 2017, 2019). All other responses were considered false alarms and were not included in RT analyses. In addition to RT, a "hit rate" was calculated as the number of correctly detected targets divided by the total number of targets. Finally, the total number of false alarms was also computed.

To examine whether each patient's RT facilitation was greater than what would be expected by the null hypothesis of no RT facilitation due to SL, we first calculated an "RT prediction score" (Batterink & Paller, 2019) by subtracting the average RT to syllables in the third position from the RT to syllables in the first position and dividing the difference by the average RT to the syllables in the first position: (avRT1 − avRT3) / avRT1. This computation adjusts for potential differences in baseline RTs between individuals, expressing facilitation to predictable targets as a proportion of RTs to the word-initial (least predictable) syllable targets. Then, we randomly shuffled the syllable position labels in our data and recalculated the RT prediction score. We repeated this shuffling process for 1000 iterations to create a null distribution of RT prediction scores for each patient. These shuffling analyses were performed in Python (Version 3.9.13) with the random seed set as 12345. Individual patients were considered to show significant RT facilitation if their true RT facilitation score exceeded the 95th percentile of the null distribution. Finally, to assess learning at the group level, a repeated-measures ANOVA was conducted with Syllable Position (1–3) as the within-subject factor. ANOVAs were conducted in JASP (Version 0.18).

Familiarity rating task. Average familiarity ratings were calculated for each word type (word, part-word, and nonword). To assess learning at the group level, a repeated-measures ANOVA was conducted with Word Type (word, part-word, nonword) as the within-subject factor. Because of the low number of trials in this task, we were unable to perform null hypothesis testing at the individual level.

iEEG Data

Our neural entrainment analyses were based on the two exposure tasks (structured and random conditions). We defined two frequencies of interest: the frequency of the individual syllables (3.33 Hz) and the frequency of the embedded words in the structured condition (1.11 Hz). Entrainment at the syllable frequency reflects sensory processing of the raw auditory input and was expected to appear in both exposure conditions. Entrainment at the word frequency, on the other hand, reflects processing of the statistical word structures and was therefore expected at significant levels only in the structured condition (Batterink & Paller, 2017, 2019). Neural entrainment was quantified by calculating intertrial phase coherence (ITC) across epochs within each condition. The fast Fourier transform was applied to each epoch to decompose the signal into its frequency components, and ITC was then computed across the 30 epochs across frequencies, including each frequency of interest (word and syllable). ITC was computed separately for each individual electrode contact.

Initial characterization. As a first step to characterizing which individual electrode contacts in each patient showed significant entrainment at our frequencies of interest, we generated surrogate data to contrast the ITC values against the null hypothesis of non-entrained neural activity. To do so, we generated 1000 surrogate data sets for each patient, for each condition, in which we shuffled each epoch onset by a random interval between −900 and 900 msec from the actual epoch onsets (following Moreau et al., 2022). By altering the precise temporal consistency across epochs, we eliminated the alignment between the neural activity and our stimuli, thus effectively simulating a scenario of non-phase-locked neural activity. Then, the ITC at each frequency bin was recomputed for each of the 1000 iterations, resulting in a null distribution of ITC values. The p values for the word frequency and the syllable frequency were computed by calculating the proportion of iterations in which the surrogate ITC exceeded the true (observed) ITC, such that lower p values reflect stronger observed entrainment relative to the null. The p values were subjected to the false discovery rate (FDR) correction for multiple comparisons at the electrode contact level, using a significance level of .05. All neural entrainment analyses were performed in MATLAB (Version 2023b) with the random seed set as 12345.

Linear mixed-effects modeling. Next, as a more comprehensive approach to characterizing entrainment effects at the group level, we ran linear mixed-effects (LME) models, which can be used to account for the hierarchical grouping of electrodes within participants and the unequal coverage of brain regions across participants, thus guarding against individual participants driving the effects (Mercier et al., 2022). Prior to fitting the models, we z scored each electrode contact's raw ITC (for each frequency in each condition) against the surrogate (shuffled) data to normalize ITC across electrode contacts and participants, producing zITC scores that were used as our dependent measure. Analyses were conducted using the LME4 package in R (Version 1.1-37; Bates, Mächler, Bolker, & Walker, 2015).

First, to assess the effect of condition (structured vs. random) on neural entrainment at our two frequencies of interest, we fit separate models for the word frequency and the syllable frequency with the structured condition coded as the reference. Both of these models included participant as a random effect and electrode as a nested random effect within participant to account for the different electrode configuration of each participant (zITC ~ Condition + (1 | Participant / Electrode)). We expected to observe a significant condition effect for zITC-Word, indicating overall stronger entrainment at the word frequency in the structured as compared with the random condition across electrodes and participants. Next, to assess which regions showed significant entrainment to the word frequency within the structured condition, we fit a separate model for zITC values at the word frequency within the structured condition (zITC ~ Region + (1 | Participant)).[2] Region was sum-coded (contr.sum) using the car package in R (Fox & Weisberg, 2019). This analysis highlights which brain regions are especially responsive to the statistical structure of the speech stream, characterizing which ones show stronger (or weaker) word entrainment relative to the average entrainment estimates across all brain regions. As a follow-up, to characterize whether entrainment within each region was significant (as assessed by zITC > 0), we used the emmeans package in R (Lenth, 2025) to extract each region's estimated marginal mean and then conducted one-sample t tests against zero (two-tailed).

Lateralization effects. Given the linguistic nature of the task, we hypothesized that neural entrainment to words would be left lateralized in some brain regions. To examine hemispheric lateralization in entrainment to the word frequency in the structured condition, we first conducted a chi-square test to assess whether the number of electrode contacts showing significant entrainment was greater than would be expected by chance in either the left or right hemisphere. We included only the brain regions that had sufficient coverage and word-entrained electrode contacts to fit the assumptions of the chi-square test (insula, MTG, STG, and supramarginal gyrus). These analyses were conducted in Python (Version 3.9.13). Next, following the same approach as in our main analysis, we

used LME models to further examine lateralization effects in these four regions. We fit a model for each region separately using zITC values within the structured condition (zITC ~ Laterality + (1 | Participant))[2], with right hemisphere coded as the reference.

Time course of neural entrainment. To explore how neural entrainment at our two frequencies of interest changed over the course of exposure, we used a sliding time window analysis to study these changes at a fine-grained scale (as in Batterink & Choi, 2021). We grouped each patient's data into "bundles" of five epochs each and recomputed the ITC for the epochs within each bundle, following the same procedure used in our main neural entrainment analysis. We then slid the bundle by one epoch and calculated the ITC for this new bundle. This process was repeated until the sliding bundle reached the last epoch, such that consecutive bundles contained epochs 1–5, then 2–6, then 3–7, and so on, finalizing in 26–30. Then, for each patient, we calculated a linear slope using MATLAB's *polyfit* function, representing the trajectory of ITC over time at the syllable frequency and word frequency, separately. To examine whether entrainment systematically increased or decreased over time at the group level, we conducted one-sample *t* tests to determine whether the mean slope at each frequency significantly differed from zero.

RESULTS

Behavioral Results

Target Detection Task

Overall, patients showed the expected decrease in RT as a function of Syllable Position (Figure 2A), reflecting facilitation due to SL (effect of Syllable Position: $F(2, 32) = 22.18$, $p < .001$, $\eta^2 = .581$; linear contrast: $p < .001$). Nearly all patients (16 out of 17 analyzed data sets) showed this trend at least numerically. Note that one patient had to be excluded from this analysis due to abnormally poor target detection performance and excessive false alarms (hit rate = 53.5%, number of false alarms = 99).

Patients achieved an overall average hit rate of 72.2%, which represents an acceptable level of performance, albeit somewhat lower than hit rates reported for neurotypical populations in which testing occurred outside of a clinical setting (89.1% in Batterink & Paller, 2017; 83.0% in Batterink & Paller, 2019; 87.6% in Wang, Köhler, et al., 2023), and those observed in older adults (88.6% in Wang, Köhler, et al., 2023). The mean number of false alarms across the task was 20.71 (range = 5–43).

Finally, null hypothesis testing at the individual level showed that 11 out of 17 patients had significant learning as evidenced by RT prediction scores that exceeded the 95th percentile of the null distribution. These results indicate that the majority of patients showed evidence for implicit SL, as reflected in their robust RT facilitation for the more predictable syllables within words.

Familiarity Rating Task

As a group, patients rated words as more familiar than part-words and nonwords (Figure 2B; effect of Word Type: $F(1.45, 24.62) = 9.64$, $p = .002$, $\eta^2 = .362$; linear contrast: $p < .001$). Out of the 18 patients, 11 showed descriptive evidence of explicit SL, as indicated by numerically higher familiarity ratings for words as compared with part-words and nonwords.

iEEG Neural Entrainment Results

Initial Characterization

We analyzed intracranial electrophysiological recordings from a total of 2261 electrode contacts across the 18 patients. The most densely covered areas included the insula, temporal lobe, and medial-temporal lobe, including the hippocampus. We examined neural entrainment at our two frequencies of interest (word and syllable) for each patient and each individual electrode contact

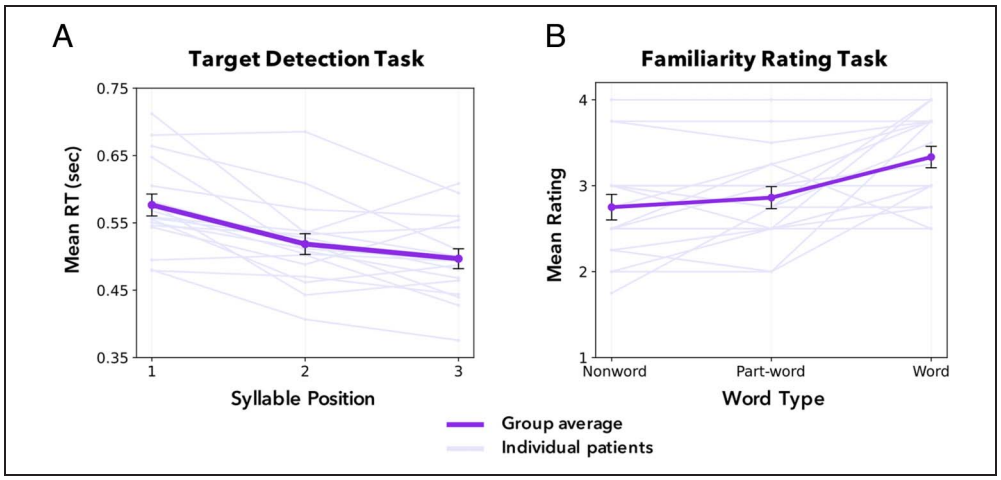

Figure 2. Results from target detection and familiarity rating tasks. (A) Target detection task. Data shown for 17 patients (one patient was excluded from this analysis). On average, patients had faster RTs to the syllables in the second and third positions within words compared with syllables in the first position. (B) Familiarity rating task. Data shown for all 18 patients. On average, patients rated words as more familiar than both part-word and nonwords foils. Bars indicate *SEM*.

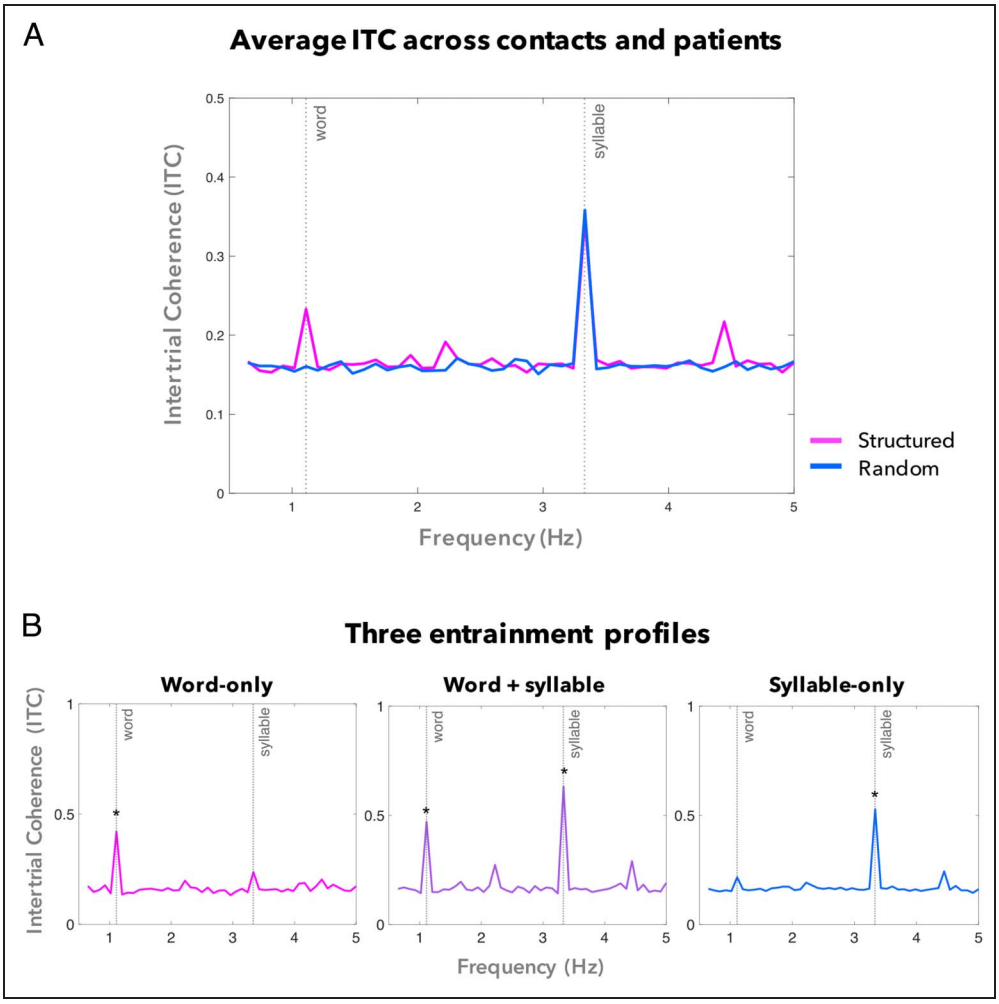

Figure 3. Main neural entrainment results. (A) Average ITC as a function of frequency across all patients and all electrode contacts. Entrainment to the word frequency (1.11 Hz) was present strongly in the structured condition but not in the random condition. Entrainment to the syllable frequency (3.33 Hz) was strong in both conditions, as expected. (B) Averaged plots of the electrode contacts across patients that showed entrainment to the word frequency only, word + syllable frequencies, and syllable frequency only in the structured condition. Significant ITC values were calculated at the individual level ($p < .05$, FDR-corrected). Only the electrode contacts that had significant entrainment to each profile were included in the plots.

separately. Across all electrode contacts pooled across all patients, we observed entrainment to the syllable frequency in both the structured and random conditions, while robust entrainment at the word frequency was observed only in the structured condition (Figure 3A). This overall pattern is expected based on prior EEG studies (e.g., Batterink & Choi, 2021; Batterink & Paller, 2017) and iEEG work (Henin et al., 2021).

At the individual electrode contact level, we found different temporal tuning responses among the contacts that entrained to our task. Namely, in the structured condition, 25.4% of contacts showed significant entrainment ($p < .05$, FDR-corrected) to only the syllable frequency (563 out of 2261), 11.1% of contacts showed significant entrainment to both the word and the syllable frequency (252 out of 2261), and 4.1% of contacts showed significant entrainment to the word frequency alone (92 out of 2261). In contrast, in the random condition, less than 1% of contacts (10 out of 2261) entrained to the word frequency only, less than 1% entrained to both frequencies (six out of 2261), and 36.8% entrained to the syllable frequency only (833 out of 2261). Interestingly, a significantly greater number of contacts showed entrainment at the syllable frequency in the random condition compared with the structured condition (McNemar's $\chi^2(1, n = 2261) = 86.39$, $p < .001$), echoing findings from scalp EEG which sometimes show stronger average syllable-level entrainment to random sequences (e.g., Batterink & Paller, 2017; Buiatti et al., 2009). Figure 3B displays the average ITC of the electrode contacts across patients that showed each of these three entrainment profiles.

Next, we capitalized on the spatial resolution afforded by iEEG to localize the specific brain regions involved in tracking the statistical regularities in our task. Figure 4A displays the whole-brain distribution of all responsive electrode contacts pooled across patients for both conditions. Since our main interest was entrainment to the word frequency as a neural marker of SL, we isolated the electrode contacts that entrained to word-only and word + syllable frequencies (see Figure 4B and Table 2). We found that these contacts were primarily localized within MTG (69 out of 390, 17.6%), STG (44 out of 131, 33.6%), insula (85 out of 387, 22.0%), and supramarginal gyrus (23 out of 66, 34.8%). Other responsive contacts in less densely covered regions were localized in frontal cortex, putamen, and transverse temporal gyrus (i.e., Heschl's gyrus; see

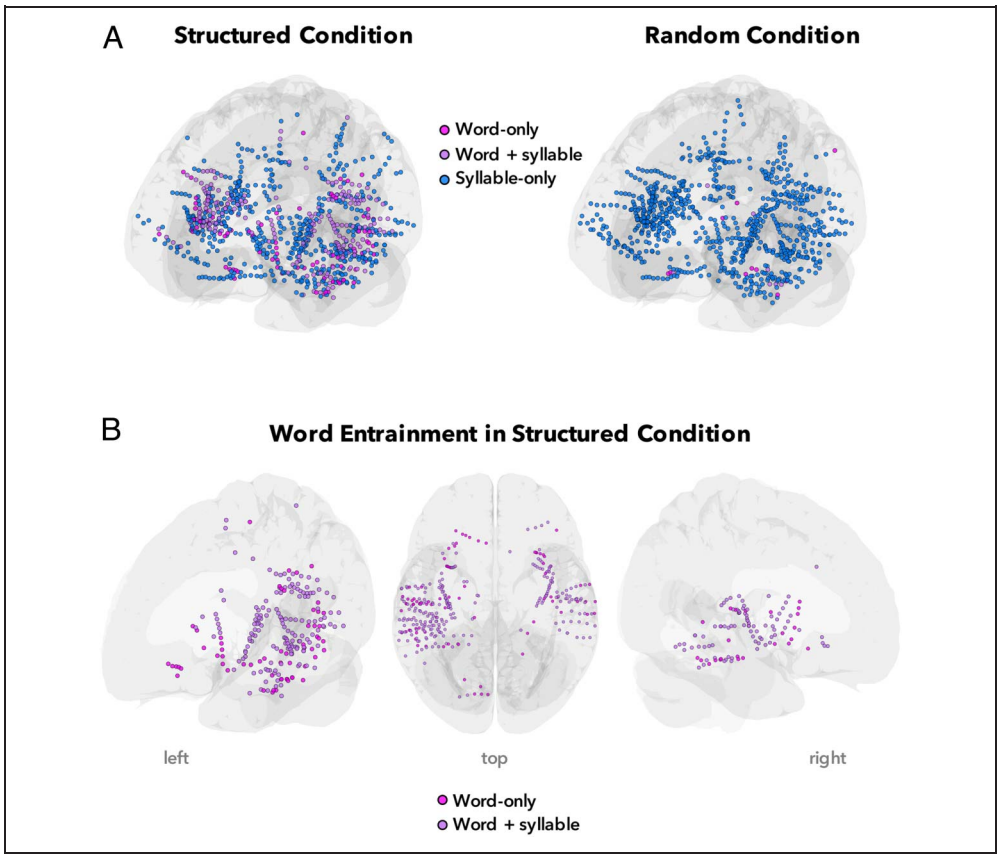

Figure 4. Neural entrainment across the brain. The images show the pooled electrode contacts across patients that showed significant entrainment to the word, word + syllable, and syllable frequencies. (A) Entrainment to the word and word + syllable frequency was observed extensively in the structured condition but not in the random condition. (B) Word entrainment in the structured condition (includes word-only and word + syllable). Areas with the highest number of entrained contacts include MTG, STG, insula, and supramarginal gyrus.

Table 2). Contrary to our hypothesis, we did not find convincing evidence of hippocampal sensitivity to the statistical regularities in the structured condition, despite dense coverage of this region (180 electrode contacts). Only eight contacts showed significant entrainment to the word frequency in the structured condition (five word-only, three word + syllable). Surprisingly, four hippocampal contacts also entrained to the word frequency in the random condition (all word-only).

LME Modeling

Our first two models (examining the effect of condition on neural entrainment at each frequency of interest) revealed that, as expected, entrainment at the word frequency was significantly higher overall in the structured condition compared with the random condition (Condition effect: $\beta = 0.86$, $SE = 0.041$, $t(4502.94) = 20.97$, $p < .001$). In contrast, entrainment at the syllable frequency was significantly lower overall in the structured condition compared with random ($\beta = -0.16$, $SE = 0.046$, $t(2260.00) = -3.52$, $p < .001$). The results from these two models align with our initial findings reported above, namely, the robust word entrainment in the structured condition and the stronger entrainment to syllable frequency in the random condition compared with the structured.

The third model, which characterized neural entrainment at the word frequency within the structured condition as a function of brain region, is summarized in Table 3. A total of 17 regions had zITC estimates significantly above zero, including the four regions identified to have greatest total number of word-entrained electrode contacts by our initial characterization above: insula, MTG, STG, and supramarginal gyrus. Other regions associated with high zITC estimates include tranverse temporal gyrus (which notably had the highest zITC mean among all regions, zITC = 5.74), inferior temporal lobe, regions within the frontal lobe (notably the pars opercularis and pars triangularis portions of the IFG), and the BG (putamen and caudate nucleus). In contrast, the hippocampus did not show above-zero entrainment to the word frequency, and interestingly, its zITC estimate was significantly below the grand mean across all regions, indicating it was actually less responsive to statistical structure than the average contact. These results fall in line with the reported findings from our initial characterization above and provide further evidence of the engagement (or lack thereof) of these brain regions in our task.

Lateralization Effects

To examine hemispheric lateralization in word frequency entrainment within the structured condition, we first conducted a chi-square test by hemisphere (left, right) in the regions identified by our initial characterization analyses to be highly responsive to the word frequency. We included

Table 2. Results from Initial Characterization (Contacts Pooled across Patients) Summarizing Neural Entrainment across Brain Regions in the Structured Condition

Region	*Word-only*	*Syllable-only*	*Word + syllable*	*Total Contacts*	*Proportion Sensitive to Word*[a]
Amygdala	2	8	0	65	0.03
Caudal anterior cingulate	0	2	0	14	0.00
Caudal middle frontal	0	2	0	14	0.00
Caudate	0	0	2	2	1.00
Cuneus	0	0	0	16	0.00
Entorhinal	0	0	0	11	0.00
Fusiform	3	12	4	56	0.13
Hippocampus	5	13	3	180	0.04
Inferior parietal	0	14	1	42	0.02
Inferior temporal	4	11	1	28	0.18
Insula	18	132	67	387	0.22
Isthmus cingulate	1	8	0	43	0.02
Lateral occipital	0	6	0	52	0.00
Lateral orbitofrontal	4	15	6	65	0.15
Lingual	2	8	1	49	0.06
Medial orbitofrontal	2	9	0	41	0.05
Middle temporal	26	116	43	390	0.18
Paracentral	1	2	1	14	0.14
Parahippocampal	0	2	0	4	0.00
Pars opercularis	2	7	1	20	0.15
Pars orbitalis	0	6	0	27	0.00
Pars triangularis	3	14	1	72	0.06
Pericalcarine	0	0	0	9	0.00
Postcentral	4	12	7	41	0.27
Posterior cingulate	0	4	1	14	0.07
Precentral	1	9	7	40	0.20
Precuneus	0	8	0	11	0.00
Putamen	4	21	8	42	0.29
Rostral anterior cingulate	0	0	0	9	0.00
Rostral middle frontal	0	5	1	26	0.04
Superior frontal	0	19	3	62	0.05
Superior parietal	0	8	1	24	0.04
Superior temporal	6	32	38	131	0.34
Supramarginal	0	19	23	66	0.35
Transverse temporal	0	1	10	11	0.91
Unclassified	4	38	22	183	0.14
Total	**92**	**563**	**252**	**2261**	

The "unclassified" brain region refers to areas of the brain (mostly white matter) that did not get classified by the algorithm into any of the specified areas in the atlas used for segmentation.

[a] Includes word-only and word + syllable.

Table 3. Summary of Estimates for Each Region from LME Analyses

	zITC against Zero		*zITC against Grand Mean across Brain Regions*	
Region	*Estimated Marginal Mean*	*p Value*	*Beta Estimate*	*p Value*
Amygdala	0.341	*ns*	−0.507	*
Caudal anterior cingulate	0.555	*ns*	−0.293	*ns*
Caudal middle frontal	0.776	*ns*	−0.072	*ns*
Caudate	**2.887**	**	2.04	*ns*
Cuneus	0.055	*ns*	−0.793	*
Entorhinal	0.326	*ns*	−0.522	*ns*
Fusiform	**0.545**	*	−0.303	*ns*
Hippocampus	0.217	*ns*	−0.63	***
Inferior parietal	−0.266	*ns*	−1.114	***
Inferior temporal	**0.985**	**	0.137	*ns*
Insula	**1.25**	***	0.402	***
Isthmus cingulate	−0.09	*ns*	−0.938	***
Lateral occipital	−0.205	*ns*	−1.052	***
Lateral orbitofrontal	**0.487**	*	−0.361	*ns*
Lingual	−0.083	*ns*	−0.931	***
Medial orbitofrontal	0.073	*ns*	−0.775	**
Middle temporal	**0.911**	***	0.063	*ns*
Paracentral	**1.302**	**	0.455	*ns*
Parahippocampal	1.257	*ns*	0.41	*ns*
Pars opercularis	**1.279**	***	0.432	*ns*
Pars orbitalis	0.051	*ns*	−0.797	**
Pars triangularis	**0.532**	*	−0.315	*ns*
Pericalcarine	0.094	*ns*	−0.754	*ns*
Postcentral	**1.698**	***	0.85	***
Posterior cingulate	**0.916**	*	0.068	*ns*
Precentral	**1.824**	***	0.976	***
Precuneus	−0.345	*ns*	−1.193	**
Putamen	**1.207**	***	0.359	*ns*
Rostral anterior cingulate	0.375	*ns*	−0.473	*ns*
Rostral middle frontal	0.425	*ns*	−0.423	*ns*
Superior frontal	**0.777**	**	−0.071	*ns*
Superior parietal	0.066	*ns*	−0.783	*
Superior temporal	**1.924**	***	1.077	***
Supramarginal	**1.873**	***	1.025	***
Transverse temporal	**5.737**	***	4.889	***
Unclassified	**0.76**	***	–	–

When using sum contrasts in R, the estimate for the final item in the list is calculated as the negative sum of all other estimates. This value is hidden in the output, hence why the estimate for the "unclassified" region is empty. *ns* = not significant. **Bolded** values present statistically significant results.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

only the regions that had sufficient coverage and number of word-entrained electrode contacts (including word-only and word + syllable) to fit the criteria for a valid chi-square test (i.e., no expected values lower than 5 in the 2 × 2 contingency tables). The regions that fit the criteria were MTG, STG, insula, and supramarginal gyrus. The chi-square results showed significant left lateralization in STG, $\chi^2(1, n = 131) = 7.24$, $p = .007$, and supramarginal gyrus, $\chi^2(1, n = 66) = 7.70$, $p = .005$. In contrast, lateralization was not observed in MTG ($p = .573$) or insula ($p = .432$).

Next, we ran an LME model for each of these four regions separately. The models revealed that entrainment in the supramarginal gyrus was significantly stronger in the left hemisphere compared with the right ($\beta = -1.74$, $SE = 0.630$, $t(30.04) = -2.77$, $p = .010$). Similarly, the STG also showed a trend toward left lateralization, though the effect did not reach significance here ($\beta = -0.95$, $SE = 0.486$, $t(118.02) = -1.96$, $p = .053$). Furthermore, there was again no effect of laterality in MTG ($\beta = -0.14$, $SE = 0.168$, $t(386.85) = -0.84$, $p = .404$). Lastly, in contrast to the chi-square analysis, the insula showed stronger entrainment in the right hemisphere compared with the left ($\beta = 1.03$, $SE = 0.196$, $t(384.85) = 5.26$, $p < .001$). This discrepancy between the chi-square and LME results may reflect the higher sensitivity of the LME approach, as it accounts for the nesting of electrode contacts within participants and the magnitude of the zITC values (the chi-square analysis, by contrast, only incorporates the count of entrained contacts into its computation). Lastly, no major changes in lateralization were observed when the three left-handers in our patient sample were excluded.

Time Course of Neural Entrainment

To examine whether entrainment systematically increased or decreased over the course of exposure to the structured condition, we conducted one-sample *t* tests to determine whether the mean slope representing the trajectory of ITC over time at each frequency significantly differed from zero. We initially included all electrode contacts across all patients in this analysis to capture changes in entrainment that may have been present even in electrode contacts that did not reach significant ITC values when calculated across the entire 30 epochs. The *t* tests revealed that the mean slope did not significantly differ from zero for either the word frequency, $t(17) = -0.73$, $p = 0.475$, or the syllable frequency, $t(17) = -0.32$, $p = 0.756$, indicating no reliable changes in entrainment over time. Additional Bayesian one-sample *t* tests provided moderate evidence in favor of the null hypothesis for both the word frequency ($BF_{01} = 3.25$) and the syllable frequency ($BF_{01} = 3.93$). Figure 5 shows the time course of neural entrainment to the word and syllable frequencies in the structured condition, across all patients and all electrode contacts.

Since it was possible that we did not observe any changes in neural entrainment due to the inclusion of all electrode contacts (even those that did not show significant entrainment to either of our frequencies of interest in our main analyses), we extracted only the contacts that were responsive to our task and repeated the analysis. More specifically, we separately grouped the electrode contacts exhibiting each of the three entrainment profiles described above (word-only, word + syllable, syllable-only) and calculated the linear slopes for each group. Once again, the one-sample *t* tests comparing slope values to zero revealed no significant change over time for any group or frequency (all *p* values > .05). Bayesian one-sample *t* tests yielded evidence in favor of the null hypothesis for both the word frequency (word-only contacts: $BF_{01} = 2.74$ word + syllable contacts: $BF_{01} = 2.43$; syllable-only contacts: $BF_{01} = 3.91$) and the syllable frequency (word-only contacts: $BF_{01} = 1.30$; word + syllable contacts: $BF_{01} = 3.45$; syllable-only contacts: $BF_{01} = 2.23$).

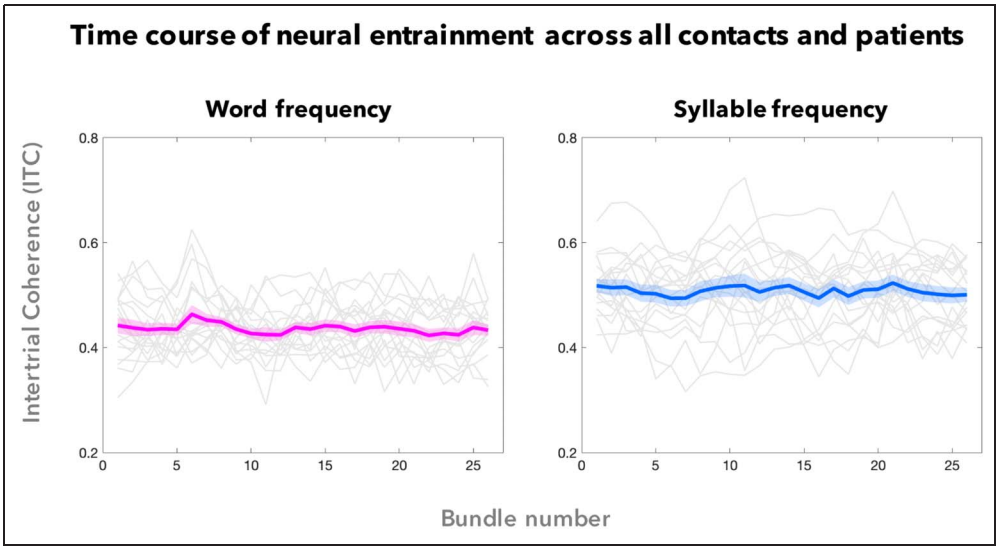

Figure 5. Time course of neural entrainment in the structured condition across all 2261 electrode contacts. ITC for each frequency (word and syllable) is plotted over 26 sliding time windows (or bundles). Thick lines show the average trajectory across all patients. Thin lines show the average trajectory for each individual patient. One-sample *t* tests comparing slope values to zero revealed no significant change in neural entrainment over time for either frequency (*p* values > .05). Shaded areas indicate *SEM*.

DISCUSSION

Our study investigated linguistic auditory SL in a group of patients with epilepsy that underwent implantation of intracranial depth electrodes for presurgical investigation. We measured neural entrainment at the frequency of individual syllables and trisyllabic words, with entrainment to the trisyllabic words serving as a neural marker of SL. We also obtained behavioral measures of SL through implicit and explicit tasks in the same individuals. Behaviorally, we found evidence of SL at the group level in both implicit and explicit tasks. At the neural level, we found strong entrainment to the word frequency in the structured condition, but not in the random condition. These word entrainment effects were observed primarily in temporal regions (MTG, STG, and Heschl's gyrus), insula, and supramarginal gyrus, with some notable entrainment also present in BG (putamen and caudate) and frontal regions (notably pars opercularis and pars triangularis). Furthermore, we observed evidence of right lateralization in the insula and left lateralization in the supramarginal gyrus and STG. Lastly, contrary to our initial hypothesis, we did not find convincing evidence of hippocampal involvement in tracking regularities from speech through neural entrainment.

Processing of Statistical Regularities in Modality-specific Regions

Our finding that primary auditory (Heschl's gyrus) and associative auditory regions (MTG, STG) are among the most highly responsive regions to the statistical word structures falls largely in line with previous studies implicating modality-specific regions in SL. For instance, fMRI studies using auditory stimuli have consistently found engagement of the STG in SL (Karuza et al., 2013; Cunillera et al., 2009; McNealy et al., 2006), along with other parts of the temporal lobe, including MTG (Karuza et al., 2013; McNealy et al., 2006). In the visual domain, fMRI studies have found that different regions within the occipital cortex are involved in processing visual statistical regularities (Karuza et al., 2017; Turk-Browne et al., 2009). Recent studies using neural entrainment as a measure of SL have further supported these findings. A magnetoencephalography (MEG) study using tone stimuli reported neural entrainment effects in STG and supramarginal gyrus (Moser et al., 2021), which are two of the most responsive regions to words in the present study. Furthermore, an iEEG study using both auditory and visual stimuli observed entrainment to trisyllabic words in auditory temporal cortex and entrainment to image pairs in occipital cortex (Henin et al., 2021). Another iEEG study used visual stimuli organized into regularities at exemplar and categorical levels and found entrainment in the visual cortex to statistical regularities at both levels (Sherman et al., 2023). The common empirical finding that modality-specific sensory cortices support SL is congruent with a theoretical account in which similar computations (here, as reflected in neural entrainment) are applied across domains but are ultimately constrained by modality-specific neurocircuitry (Frost, Armstrong, Siegelman, & Christiansen, 2015).

Our finding of left lateralization in STG and supramarginal gyrus is consistent with previous fMRI studies of linguistic auditory SL. For instance, engagement of the left STG was reported by Cunillera et al. (2009) and Karuza et al. (2013). Although McNealy et al. (2006) found bilateral STG activation during the exposure phase of their task, they found that only left STG activation was correlated with post-exposure behavioral performance. Furthermore, left supramarginal gyrus activation was found by both McNealy et al. (2006) and Karuza et al. (2013). The STG and supramarginal gyrus are both part of a well-described canonical language network that is lateralized to the left hemisphere in most individuals (Fedorenko et al., 2024). Interestingly, an fMRI study that used auditory tones instead of syllables found activation of STG but not supramarginal gyrus (Abla & Okanoya, 2008), while an MEG study that also used tones did find engagement of supramarginal gyrus, but this response was bilateral with a larger cluster in the right hemisphere (Moser et al., 2021). Furthermore, another recent study observed that the left posterior temporal gyrus was engaged in both natural language processing and linguistic auditory SL, but not in nonlinguistic auditory SL (Schneider et al., 2024). Therefore, in line with Frost et al.'s (2015) account of modality and stimulus-specific constraints on SL, it is likely that linguistic auditory SL engages regions within the language network that are not necessarily recruited by nonlinguistic auditory stimuli.

Entrainment beyond Auditory Regions

Our results also align with previous studies reporting engagement of the IFG during SL. Although our initial characterization showed limited involvement of the IFG (pars triangularis, pars orbitalis, and pars opercularis in Table 2) in terms of the raw proportion of word-entrained electrodes, our more sensitive LME approach revealed that the pars opercularis and pars triangularis portions of the IFG show significant entrainment at the word frequency. The IFG, which contains the canonical Broca's area (Keller, Crow, Foundas, Amunts, & Roberts, 2009), has been associated with several different language functions, including semantic and phonological processing (Liakakis, Nickel, & Seitz, 2011), complex syntactic processing (Friederici & Gierhan, 2013), and pre-articulatory processing (Flinker et al., 2015). Within the context of SL, the IFG has been found to be involved in both linguistic auditory (Henin et al., 2021; Karuza et al., 2013; McNealy et al., 2006) and nonlinguistic auditory (Abla & Okanoya, 2008) studies, suggesting that it plays a domain-general role in learning regularities across modalities.

Furthermore, our LME models also revealed significant entrainment in regions within the BG, notably the putamen. This finding converges with previous fMRI studies

that have reported engagement of the putamen in both visual (Karuza et al., 2017; Turk-Browne et al., 2009) and auditory SL (Karuza et al., 2013; McNealy et al., 2006). Beyond SL, the putamen has been implicated in many other types of implicit learning, including probabilistic implicit sequence learning (Wilkinson & Jahanshahi, 2007) and implicit contextual learning (van Asselen et al., 2009). Therefore, due to the similarities between SL and other forms of implicit learning (Batterink et al., 2019; Christiansen, 2019; Perruchet & Pacton, 2006), it is not surprising that the putamen was engaged in our task as well.

Interestingly, we found robust entrainment at the word frequency within the insula, with stronger entrainment in the right hemisphere (as shown by our LME approach). This region is functionally heterogeneous, with functions ranging from somatosensory to cognitive (Uddin, Nomi, Hébert-Seropian, Ghaziri, & Boucher, 2017). At the cognitive level, the insula's role in the salience network is well established (Uddin, 2015), but other studies have linked it to specific language functions as well. For instance, in a meta-analysis, Oh, Duerden, and Pang (2014) found that bilateral regions within the anterior insula were activated in tasks composed of linguistic reception and expression (whole words, syntax, morphology, or pragmatics). Furthermore, bilateral insular activation was also found in perception and production of speech stimuli (phonetic, syllabic, or nonword), with speech production showing more left-hemisphere dominance (Oh et al., 2014). Moreover, an earlier study found that the left precentral insular gyrus is involved in coordinating articulation (Dronkers, 1996). Thus, it is possible that the insula was responsive to our task due to its general role in language and speech processing. However, a recent iEEG study of visual SL also reported neural entrainment to visual regularities in the insula (Sherman et al., 2023), so this leaves open the possibility that the insula has a role in SL across sensory modalities that is independent of any linguistic process. Obtaining more clarity on the role of the insula in our task (and its potential right lateralization) will require further consideration of the functional differentiation within this structure in future work.

Three Neural Entrainment Profiles

Across the electrode contacts that were responsive to our task, we found three distinct entrainment profiles: entrainment to the word frequency only, to the syllable frequency only, and to both the word and the syllable frequencies. These discrete profiles of entrainment cannot be observed through scalp EEG, which reflects the summation of activity across millions of neurons distributed over the cortex, and can only be revealed with iEEG due to its far superior spatial resolution. The existence of word-only entrainment in a subset of electrode contacts suggests that some brain regions may be involved in processing primarily higher-order statistical units, without directly processing the lower-level input. In contrast, other regions support processing of lower-level sensory input but are not involved in discovering the higher-level regularities. Finally, a third category of regions appear to be involved in processing both the sensory input as well as the higher-level regularities. Henin et al. (2021) suggested that mainly sensory regions (i.e., STG) showed the third (syllable + word) profile, while higher-order, domain-general regions (i.e., IFG and anterior temporal lobe) had a higher incidence of entrainment restricted to the frequency of the statistical regularities (word-only). These findings are interpreted as a functional and anatomical processing hierarchy, wherein sensory brain regions are involved in early, lower-level processing, and domain-general regions engage in later, higher-order processing of the statistical regularities.

Our results provide some support for this account. Similar to Henin and colleagues, electrode contacts within primary auditory regions (STG and Heschl's gyrus) mostly showed syllable-only and word + syllable entrainment, with few contacts entraining to the word frequency only (see Table 2). Furthermore, the supramarginal gyrus, which has been implicated in early phonological processing (Sliwinska, Khadilkar, Campbell-Ratcliffe, Quevenco, & Devlin, 2012; Hartwigsen et al., 2010), also showed syllable-only and word + syllable entrainment, without any contacts showing word-only entrainment. We observed word-only entrainment mostly in insula and MTG, which are not among the regions described by Henin et al. (2021) as showing a word-only entrainment profile, but which also both play a role in higher-level cognitive processes, including language processing (e.g., Yu, Song, & Liu, 2022; Oh et al., 2014; Visser, Jefferies, Embleton, & Lambon Ralph, 2012). Lastly, we were unable to examine entrainment effects in the anterior temporal lobe—a region that was highlighted by Henin and colleagues as being especially sensitive to statistical regularities—due to insufficient coverage of this area in our sample (see Supplemental Information A).

The Role of the Hippocampus in SL

In contrast with our initial hypothesis, we did not find convincing evidence of hippocampal involvement in our task. Only eight electrode contacts (4.4%), within only four patients, showed significant entrainment to the word frequency following FDR correction. This number was similar to that observed in the random condition, leading us to interpret even this minimal involvement with caution. Furthermore, our LME analyses revealed that the hippocampus actually showed significantly weaker entrainment to the word frequency compared with the average entrainment estimate across brain regions. The lack of word-level entrainment in the hippocampus is unlikely to be due to insufficient electrode coverage in this structure. In fact, the hippocampus was one of the regions with the densest overall coverage across our patient sample (180 electrode contacts).

There are a number of possible factors that may account for why hippocampal entrainment to the word frequency was largely absent. First, it is possible that the role of the hippocampus in SL may be specific to the visual modality. fMRI studies that have reported hippocampal engagement in SL have mainly used visual nonlinguistic stimuli (Ellis et al., 2021; Sherman & Turk-Browne, 2020; Karuza et al., 2017; Turk-Browne et al., 2009). To our knowledge, fMRI studies on SL using auditory stimuli, either linguistic or nonlinguistic, have generally not reported hippocampal activation (Schneider et al., 2024; Karuza et al., 2013; Cunillera et al., 2009; McNealy et al., 2006).

Another possibility is that the hippocampus is not critically involved in the learning process itself, but rather in retrieving learned regularities to generate predictions of upcoming stimuli. Previous SL studies using visual stimuli have indeed reported evidence of prediction in the hippocampus (Sherman et al., 2022; Sherman & Turk-Browne, 2020), but these findings remain to be confirmed in auditory SL.

A final possibility is that neural entrainment is an insensitive index of SL within the hippocampus, at least with the word presentation frequency used in the current study. Henin et al. (2021) used a similar word presentation rate in their work and, like us, did not observe entrainment effects in the hippocampus. However, after performing a representational similarity analysis, they found that the hippocampus uniquely represented the identity of the word units, as shown by similar neural activity patterns to syllables that belonged to the same word. Therefore, it is possible that the hippocampus plays a role in SL even for auditory linguistic stimuli, but that its computations do not involve tight phase-locking of neural activity to the incoming stimuli and, as a consequence, would not be well captured through neural entrainment measures.

The Time Course of Neural Entrainment

We did not observe a clear pattern of change in entrainment to the word frequency over the course of exposure to the structured stream. This was true when all electrode contacts were included and even when we included only the responsive electrode contacts in the analysis. This result is at odds with previous EEG studies using a similar sliding time window analysis, as these studies have generally shown that entrainment to the word frequency across electrodes tends to gradually increase over time, while entrainment to the syllable frequency remains relatively stable throughout exposure (Moreau et al., 2022; Smalle, Daikoku, Szmalec, Duyck, & Möttönen, 2022; Batterink & Choi, 2021; Ordin, Polyanskaya, Soto, & Molinaro, 2020; Batterink & Paller, 2017). However, another iEEG study also conducted this analysis on their data and, like us, found no discernible change in entrainment over the course of exposure in either of their two experimental conditions (Sherman et al., 2023).

Although we do not have a definitive explanation for these results, they could be due to a few factors. First, the electrical activity measured by scalp EEG differs from that recorded by intracranial electrodes (Mercier et al., 2022). Namely, EEG measures the activity of millions of neurons distributed over the cortex, while the electrodes used in iEEG measure more localized activity. In the case of sEEG, depth electrodes measure electrical activity directly from deep regions within the brain. Thus, it is possible that these deeper regions have different learning trajectories than the more superficial regions, and these differences in trajectories are obscured when averaging electrode contacts across the brain. It is also possible that the sliding time window analysis is not the best analytic approach for iEEG data, as it may be too coarse to account for the unique entrainment profiles observed in this type of data and the disparate coverage across the brain that is inherent to any iEEG study. Future studies that account for differences in learning trajectory between regions and/or type of entrainment (e.g., word-only, word + syllable, syllable-only) could provide a more fine-grained view of neural entrainment changes over time across the brain.

Robust Behavioral Evidence for SL

Overall, we found robust evidence of SL on the target detection task. Not only did patients show strong facilitation to predictable syllables at the group level, as evidenced by faster RTs as a function of syllable position, but 11 out of 17 patients had reliable evidence of learning at the individual level, representing a relatively stringent statistical criterion. Moreover, nearly the entire sample (16/17) showed numerically faster RTs to more predictable syllables. Overall, we interpret our results from the target detection task as robust evidence of implicit SL in our sample. Thus, our data provide evidence of SL not just at the level of neural entrainment but also in behavior.

In contrast, evidence of learning in the familiarity rating task is somewhat weaker. While as a group the patients showed significantly higher familiarity ratings for words compared with the two foil types (part-words and nonwords), only 11 out of the 18 patients rated words numerically higher than part-words and nonwords. The small number of trials on this task precluded us from testing for significance on this task at the individual level—as we did for the target detection task—so we do not have information on significant learning at the individual patient level. The less robust evidence for learning on this task compared with the target detection task could potentially be explained by the epilepsy diagnosis of our participants. It is known that patients with epilepsy often suffer from declarative memory deficits, in particular in association with temporal lobe focus (Tramoni-Negre, Lambert, Bartolomei, & Felician, 2017), among other cognitive impairments (Holmes, 2015). Since the familiarity rating task is thought to engage mostly explicit memory mechanisms, it is not surprising that our participants showed less

robust performance on this task (effect size of word type on ratings: $\eta^2 = .362$) compared with the target detection task (effect size of syllable position on RT: $\eta^2 = .581$).[3] However, we note that differences in task reliability may also contribute to differential performance on these two tasks (Christiansen, 2019; Isbilen, McCauley, Kidd, & Christiansen, 2017).

A recent study that used our stimuli and the same tasks on an amnesic patient with a dentate gyrus lesion yielded similar results—namely, the patient had intact performance on the target detection task while his performance on the familiarity rating task was impaired (Wang, Rosenbaum, et al., 2023). Furthermore, the familiarity rating task may be more demanding even for healthy individuals due to the explicit memory judgments that it requires. For instance, Batterink et al. (2015) found that the target detection task revealed learning in a greater proportion of healthy adults compared with a two-alternative forced-choice task (which, like the familiarity rating task, is a measure of explicit learning).

Limitations

As with any human iEEG study, one of our fundamental limitations was the heterogeneous electrode coverage across patients, which prevents us from making meaningful comparisons between participants. Moreover, we did not include the patients' neuropsychological profiles or medication intake as variables in our analyses. This is especially limiting in our lateralization analyses, as we included both right-handed ($n = 15$) and left-handed ($n = 3$) patients in our study. It has been shown that the incidence of atypical language network lateralization is higher in left-handers compared with right-handers—although the majority of left-handers still have typical left-lateralized language (Mazoyer et al., 2014; Szaflarski et al., 2002; Pujol, Deus, Losilla, & Capdevila, 1999). Lastly, another intrinsic limitation of any iEEG study is the fact that electrodes are implanted in brain regions estimated to contain the seizure onset and propagation zones (Mercier et al., 2022). Although we tested patients at a time window in which epileptic spikes were thought to be minimal by the clinical staff, our iEEG data were not inspected by a neurologist to remove the individual electrode contacts that had epileptic activity (if any). Therefore, it is possible that some epileptic activity may have contaminated our signal, though we note that any such effects would work against our main reported entrainment findings.

Conclusion

Overall, we found strong evidence of the engagement of auditory and language-specific brain regions in tracking the statistical regularities within speech-based SL. This is consistent with previous fMRI and iEEG studies, which point to a major role of sensory cortices in SL. Furthermore, we did not find convincing evidence of hippocampal engagement with our neural entrainment measure, though future alternative analyses could reveal crucial information about processing of statistical regularities in this region. Lastly, participants in the current study showed more robust learning on our implicit measure as compared with our explicit measure of SL. The current results highlight the important contributions of modality-specific brain regions to auditory speech-based SL and call for caution in generalizing between findings from SL tasks in different domains and modalities.

Corresponding authors: Stefan Köhler or Laura J. Batterink, Western Institute for Neuroscience, University of Western Ontario, London, Ontario, Canada, e-mail: stefank@uwo.ca or dherrer3@uwo.ca.

Data Availabitiliy Statement

The experimental paradigm is available on the Open Science Framework (https://osf.io/evydn/files/osfstorage). We do not have permission to share patient data. Supplemental Material can be accessed on this article's homepage: https://doi.org/10.1162/JOCN.a.2411.

Author Contributions

Daniela Herrera-Chaves: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing—Original draft. Greydon Gilmore: Data curation; Methodology; Software. Mohamad Abbass: Formal analysis; Visualization. Lyle Muller: Resources. Ana Suller-Marti: Resources; Supervision; Writing—Review & editing. Seyed M. Mirsattari: Resources. Stefan Köhler: Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing—Review & editing. Laura Batterink: Conceptualization; Funding acquisition; Methodology; Resources; Supervision; Writing—Review & editing.

Funding Information

Natural Sciences and Engineering Research Council of Canada (https://dx.doi.org/10.13039/501100000038), grant number: RGPIN-2018- 05770 to Stefan Köhler, grant number: RGPIN-2019-05132 to Laura J. Batterink.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited

were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

Notes

1. One of the part-words contained the two syllables from an original word, but these two syllables were presented in the wrong order (*rusuti* instead of *rutisu*).
2. Note that we did not include electrode as a nested factor here as there was only one observation per electrode contact in this subset of the data.
3. It is worthwhile to note that effect sizes for the familiarity rating task and target detection task are similar to one another in healthy participants, as shown in a reanalysis of Batterink and Paller (2017). Effect size of word type on ratings: $\eta^2 = .534$; effect size of syllable position on RT: $\eta^2 = .580$.

REFERENCES

Abla, D., & Okanoya, K. (2008). Statistical segmentation of tone sequences activates the left inferior frontal cortex: A near-infrared spectroscopy study. *Neuropsychologia*, *46*, 2787–2795. https://doi.org/10.1016/j.neuropsychologia.2008.05.012, PubMed: 18579166

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*, 464–473. https://doi.org/10.1111/j.1467-7687.2009.00937.x, PubMed: 21477186

Arnon, I., Kirby, S., Allen, J. A., Garrigue, C., Carroll, E. L., & Garland, E. C. (2025). Whale song shows language-like statistical structure. *Science*, *387*, 649–653. https://doi.org/10.1126/science.adq7055, PubMed: 39913578

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Batterink, L. (2020). Syllables in sync form a link: Neural phase-locking reflects word knowledge during language learning. *Journal of Cognitive Neuroscience*, *32*, 1735–1748. https://doi.org/10.1162/jocn_a_01581, PubMed: 32427066

Batterink, L. J., & Choi, D. (2021). Optimizing steady-state responses to index statistical learning: Response to Benjamin and colleagues. *Cortex*, *142*, 379–388. https://doi.org/10.1016/j.cortex.2021.06.008, PubMed: 34321154

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. https://doi.org/10.1016/j.cortex.2017.02.004, PubMed: 28324696

Batterink, L. J., & Paller, K. A. (2019). Statistical learning of speech regularities can occur outside the focus of attention. *Cortex*, *115*, 56–71. https://doi.org/10.1016/j.cortex.2019.01.013, PubMed: 30771622

Batterink, L. J., Paller, K. A., & Reber, P. J. (2019). Understanding the neural bases of implicit and statistical learning. *Topics in Cognitive Science*, *11*, 482–503. https://doi.org/10.1111/tops.12420, PubMed: 30942536

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004, PubMed: 26034344

Bogaerts, L., Frost, R., & Christiansen, M. H. (2020). Integrating statistical learning into cognitive science. *Journal of Memory and Language*, *115*, 104167. https://doi.org/10.1016/j.jml.2020.104167

Boros, M., Magyari, L., Török, D., Bozsik, A., Deme, A., & Andics, A. (2021). Neural processes underlying statistical learning for speech segmentation in dogs. *Current Biology*, *31*, 5512–5521. https://doi.org/10.1016/j.cub.2021.10.017, PubMed: 34717832

Buiatti, M., Pena, M., & Dehaenelambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, *44*, 509–519. https://doi.org/10.1016/j.neuroimage.2008.09.015, PubMed: 18929668

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). Preverbal infants discover statistical word patterns at similar rates as adults: Evidence from neural entrainment. *Psychological Science*, *31*, 1161–1173. https://doi.org/10.1177/0956797620933237, PubMed: 32865487

Christiansen, M. H. (2019). Implicit statistical learning: A tale of two literatures. *Topics in Cognitive Science*, *11*, 468–481. https://doi.org/10.1111/tops.12332, PubMed: 29630770

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 24–39. https://doi.org/10.1037/0278-7393.31.1.24, PubMed: 15641902

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The necessity of the hippocampus for statistical learning. *Journal of Cognitive Neuroscience*, *30*, 680–697. https://doi.org/10.1162/jocn_a_01228, PubMed: 29308986

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., et al. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, *48*, 541–553. https://doi.org/10.1016/j.neuroimage.2009.06.069, PubMed: 19580874

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*, 159–161. https://doi.org/10.1038/384159a0, PubMed: 8906789

Ellis, C. T., Skalaban, L. J., Yates, T. S., Bejjanki, V. R., Córdova, N. I., & Turk-Browne, N. B. (2021). Evidence of hippocampal learning in human infants. *Current Biology*, *31*, 3358–3364. https://doi.org/10.1016/j.cub.2021.04.072, PubMed: 34022155

Fedorenko, E., Ivanova, A. A., & Regev, T. I. (2024). The language network as a natural kind within the broader landscape of the human brain. *Nature Reviews Neuroscience*, *25*, 289–312. https://doi.org/10.1038/s41583-024-00802-4, PubMed: 38609551

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*, 499–504. https://doi.org/10.1111/1467-9280.00392, PubMed: 11760138

Fiser, J., & Aslin, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 458–467. https://doi.org/10.1037/0278-7393.28.3.458, PubMed: 12018498

Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 15822–15826. https://doi.org/10.1073/pnas.232472899, PubMed: 12429858

Flinker, A., Korzeniewska, A., Shestyuk, A. Y., Franaszczuk, P. J., Dronkers, N. F., Knight, R. T., et al. (2015). Redefining the role of Broca's area in speech. *Proceedings of the National Academy of Sciences, U.S.A.*, *112*, 2871–2875. https://doi.org/10.1073/pnas.1414491112, PubMed: 25730850

Fló, A., Brusini, P., Macagno, F., Nespor, M., Mehler, J., & Ferry, A. L. (2019). Newborns are sensitive to multiple cues for word

segmentation in continuous speech. *Developmental Science*, *22*, e12802. https://doi.org/10.1111/desc.12802, PubMed: 30681763
Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., & Collins, D. L. (2011). Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*, *54*, 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033, PubMed: 20656036
Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5
Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage. https://doi.org/10.32614/CRAN.package.car
Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*, 250–254. https://doi.org/10.1016/j.conb.2012.10.002, PubMed: 23146876
Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*, 1128–1153. https://doi.org/10.1037/bul0000210, PubMed: 31580089
Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*, 117–125. https://doi.org/10.1016/j.tics.2014.12.010, PubMed: 25631249
Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 16494–16499. https://doi.org/10.1073/pnas.1008121107, PubMed: 20807747
Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., et al. (2021). Learning hierarchical sequence representations across human cortex and hippocampus. *Science Advances*, *7*, eabc4530. https://doi.org/10.1126/sciadv.abc4530, PubMed: 33608265
Holmes, G. L. (2015). Cognitive impairment in epilepsy: The role of network abnormalities. *Epileptic Disorders*, *17*, 101–116. https://doi.org/10.1684/epd.2015.0739, PubMed: 25905906
Isbilen, E. S., & Christiansen, M. H. (2022). Statistical learning of language: A meta-analysis into 25 years of research. *Cognitive Science*, *46*, e13198. https://doi.org/10.1111/cogs.13198, PubMed: 36121309
Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2017). Testing statistical learning implicitly: A novel chunk-based measure of statistical learning. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, *39*, 564–569. https://escholarship.org/uc/item/8jd7m3df
Karuza, E. A., Emberson, L. L., Roser, M. E., Cole, D., Aslin, R. N., & Fiser, J. (2017). Neural signatures of spatial statistical learning: Characterizing the extraction of structure from complex visual scenes. *Journal of Cognitive Neuroscience*, *29*, 1963–1976. https://doi.org/10.1162/jocn_a_01182, PubMed: 28850297
Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*, 46–54. https://doi.org/10.1016/j.bandl.2012.11.007, PubMed: 23312790
Keller, S. S., Crow, T., Foundas, A., Amunts, K., & Roberts, N. (2009). Broca's area: Nomenclature, anatomy, typology and asymmetry. *Brain and Language*, *109*, 29–48. https://doi.org/10.1016/j.bandl.2008.11.005, PubMed: 19155059
Lenth, R. V. (2025). *Emmeans: Estimated marginal means, aka least-squares means*. R package Version 11.2. https://doi.org/10.32614/CRAN.package.emmeans
Liakakis, G., Nickel, J., & Seitz, R. J. (2011). Diversity of the inferior frontal gyrus—A meta-analysis of neuroimaging studies. *Behavioural Brain Research*, *225*, 341–347. https://doi.org/10.1016/j.bbr.2011.06.022, PubMed: 21729721
Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., et al. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS One*, *9*, e101165. https://doi.org/10.1371/journal.pone.0101165, PubMed: 24977417
McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *Journal of Neuroscience*, *26*, 7629–7639. https://doi.org/10.1523/JNEUROSCI.5501-05.2006, PubMed: 16855090
McNealy, K., Mazziotta, J. C., & Dapretto, M. (2010). The neural basis of speech parsing in children and adults. *Developmental Science*, *13*, 385–406. https://doi.org/10.1111/j.1467-7687.2009.00895.x, PubMed: 20136936
Mercier, M. R., Dubarry, A.-S., Tadel, F., Avanzini, P., Axmacher, N., Cellier, D., et al. (2022). Advances in human intracranial electroencephalography research, guidelines and good practices. *Neuroimage*, *260*, 119438. https://doi.org/10.1016/j.neuroimage.2022.119438, PubMed: 35792291
Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., et al. (2010). Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine*, *98*, 278–284. https://doi.org/10.1016/j.cmpb.2009.09.002, PubMed: 19818524
Moreau, C. N., Joanisse, M. F., Mulgrew, J., & Batterink, L. J. (2022). No statistical learning advantage in children over adults: Evidence from behaviour and neural entrainment. *Developmental Cognitive Neuroscience*, *57*, 101154. https://doi.org/10.1016/j.dcn.2022.101154, PubMed: 36155415
Moser, J., Batterink, L., Li Hegner, Y., Schleger, F., Braun, C., Paller, K. A., et al. (2021). Dynamics of nonlinguistic statistical learning: From neural entrainment to the emergence of explicit knowledge. *Neuroimage*, *240*, 118378. https://doi.org/10.1016/j.neuroimage.2021.118378, PubMed: 34246769
Narizzano, M., Arnulfo, G., Ricci, S., Toselli, B., Tisdall, M., Canessa, A., et al. (2017). SEEG assistant: A 3DSlicer extension to support epilepsy surgery. *BMC Bioinformatics*, *18*, 124. https://doi.org/10.1186/s12859-017-1545-8, PubMed: 28231759
Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, *23*, 913–926. https://doi.org/10.1016/j.tics.2019.08.004, PubMed: 31606386
Oh, A., Duerden, E. G., & Pang, E. W. (2014). The role of the insula in speech and language processing. *Brain and Language*, *135*, 96–103. https://doi.org/10.1016/j.bandl.2014.06.003, PubMed: 25016092
Ordin, M., Polyanskaya, L., Soto, D., & Molinaro, N. (2020). Electrophysiology of statistical learning: Exploring the online learning process and offline learning product. *European Journal of Neuroscience*, *51*, 2008–2022. https://doi.org/10.1111/ejn.14657, PubMed: 31872926
Orpella, J., Assaneo, M. F., Ripollés, P., Noejovich, L., López-Barroso, D., Diego-Balaguer, R. D., et al. (2022). Differential activation of a frontoparietal network explains population-level differences in statistical learning from speech. *PLoS Biology*, *20*, e3001712. https://doi.org/10.1371/journal.pbio.3001712, PubMed: 35793349
Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: Experiments in

behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y, PubMed: 30734206
Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233–238. https://doi.org/10.1016/j.tics.2006.03.006, PubMed: 16616590
Petersson, K.-M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, *120*, 83–95. https://doi.org/10.1016/j.bandl.2010.08.003, PubMed: 20943261
Pujol, J., Deus, J., Losilla, J. M., & Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology*, *52*, 1038–1043. https://doi.org/10.1212/wnl.52.5.1038, PubMed: 10102425
Ramos-Escobar, N., Mercier, M., Trébuchon-Fonséca, A., Rodriguez-Fornells, A., François, C., & Schön, D. (2022). Hippocampal and auditory contributions to speech segmentation. *Cortex*, *150*, 1–11. https://doi.org/10.1016/j.cortex.2022.01.017, PubMed: 35305505
Ringer, H., Sammler, D., & Daikoku, T. (2024). *Neural tracking of auditory statistical regularities is reduced in adults with dyslexia*. https://doi.org/10.1101/2024.08.26.609678
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. https://doi.org/10.1126/science.274.5294.1926, PubMed: 8943209
Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4, PubMed: 10193055
Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621. https://doi.org/10.1006/jmla.1996.0032
Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105. https://doi.org/10.1111/j.1467-9280.1997.tb00690.x
Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, *26*, 1736–1747. https://doi.org/10.1162/jocn_a_00578, PubMed: 24456393
Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, *22*, 1622–1627. https://doi.org/10.1016/j.cub.2012.06.056, PubMed: 22885059
Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*, 486–492. https://doi.org/10.1038/nn.3331, PubMed: 23416451
Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *372*, 20160049. https://doi.org/10.1098/rstb.2016.0049, PubMed: 27872368
Schneider, J. M., Scott, T. L., Legault, J., & Qi, Z. (2024). Limited but specific engagement of the mature language network during linguistic statistical learning. *Cerebral Cortex*, *34*, bhae123. https://doi.org/10.1093/cercor/bhae123, PubMed: 38566510
Sherman, B. E., Aljishi, A., Graves, K. N., Quraishi, I. H., Sivaraju, A., Damisah, E. C., et al. (2023). Intracranial entrainment reveals statistical learning across levels of abstraction. *Journal of Cognitive Neuroscience*, *35*, 1312–1328. https://doi.org/10.1162/jocn_a_02012, PubMed: 37262357
Sherman, B. E., Graves, K. N., Huberdeau, D. M., Quraishi, I. H., Damisah, E. C., & Turk-Browne, N. B. (2022). Temporal dynamics of competition between statistical learning and episodic memory in intracranial recordings of human visual cortex. *Journal of Neuroscience*, *42*, 9053–9068. https://doi.org/10.1523/JNEUROSCI.0708-22.2022, PubMed: 36344264
Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, *32*, 15–20. https://doi.org/10.1016/j.cobeha.2020.01.015, PubMed: 32258249
Sherman, B. E., & Turk-Browne, N. B. (2020). Statistical prediction of the future impairs episodic encoding of the present. *Proceedings of the National Academy of Sciences, U.S.A.*, *117*, 22760–22770. https://doi.org/10.1073/pnas.2013291117, PubMed: 32859755
Sjuls, G. S., Harvei, N. N., & Vulchanova, M. D. (2024). The relationship between neural phase entrainment and statistical word-learning: A scoping review. *Psychonomic Bulletin & Review*, *31*, 1399–1419. https://doi.org/10.3758/s13423-023-02425-9, PubMed: 38062317
Sliwinska, M. W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., & Devlin, J. T. (2012). Early and sustained Supramarginal gyrus contributions to phonological processing. *Frontiers in Psychology*, *3*, 161. https://doi.org/10.3389/fpsyg.2012.00161, PubMed: 22654779
Smalle, E. H. M., Daikoku, T., Szmalec, A., Duyck, W., & Möttönen, R. (2022). Unlocking adults' implicit statistical learning by cognitive depletion. *Proceedings of the National Academy of Sciences, U.S.A.*, *119*, e2026011119. https://doi.org/10.1073/pnas.2026011119, PubMed: 34983868
Szaflarski, J. P., Binder, J. R., Possing, E. T., McKiernan, K. A., Ward, B. D., & Hammeke, T. A. (2002). Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology*, *59*, 238–244. https://doi.org/10.1212/WNL.59.2.238, PubMed: 12136064
Tramoni-Negre, E., Lambert, I., Bartolomei, F., & Felician, O. (2017). Long-term memory deficits in temporal lobe epilepsy. *Revue Neurologique*, *173*, 490–497. https://doi.org/10.1016/j.neurol.2017.06.011, PubMed: 28838789
Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*, 1934–1945. https://doi.org/10.1162/jocn.2009.21131, PubMed: 18823241
Tustison, N. J., Cook, P. A., Holbrook, A. J., Johnson, H. J., Muschelli, J., Devenyi, G. A., et al. (2021). The ANTsX ecosystem for quantitative biological and medical imaging. *Scientific Reports*, *11*, 9068. https://doi.org/10.1038/s41598-021-87564-6, PubMed: 33907199
Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, *16*, 55–61. https://doi.org/10.1038/nrn3857, PubMed: 25406711
Uddin, L. Q., Nomi, J. S., Hébert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and function of the human insula. *Journal of Clinical Neurophysiology*, *34*, 300–306. https://doi.org/10.1097/WNP.0000000000000377, PubMed: 28644199
van Asselen, M., Almeida, I., Andre, R., Januário, C., Freire Gonçalves, A., & Castelo-Branco, M. (2009). The role of the basal ganglia in implicit contextual learning: A study of

Parkinson's disease. *Neuropsychologia*, *47*, 1269–1273. https://doi.org/10.1016/j.neuropsychologia.2009.01.008, PubMed: 19428390

Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: Distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience*, *24*, 1766–1778. https://doi.org/10.1162/jocn_a_00244, PubMed: 22621260

Wang, H. S., Köhler, S., & Batterink, L. J. (2023). Separate but not independent: Behavioral pattern separation and statistical learning are differentially affected by aging. *Cognition*, *239*, 105564. https://doi.org/10.1016/j.cognition.2023.105564, PubMed: 37467624

Wang, H. S., Rosenbaum, R. S., Baker, S., Lauzon, C., Batterink, L. J., & Köhler, S. (2023). Dentate gyrus integrity is necessary for behavioral pattern separation but not statistical learning. *Journal of Cognitive Neuroscience*, *35*, 900–917. https://doi.org/10.1162/jocn_a_01981, PubMed: 36877071

Wilkinson, L., & Jahanshahi, M. (2007). The striatum and probabilistic implicit sequence learning. *Brain Research*, *1137*, 117–130. https://doi.org/10.1016/j.brainres.2006.12.051, PubMed: 17229409

Yu, M., Song, Y., & Liu, J. (2022). The posterior middle temporal gyrus serves as a hub in syntactic comprehension: A model on the syntactic neural network. *Brain and Language*, *232*, 105162. https://doi.org/10.1016/j.bandl.2022.105162, PubMed: 35908340

Zhang, M., Riecke, L., & Bonte, M. (2021). Neurophysiological tracking of speech-structure learning in typical and dyslexic readers. *Neuropsychologia*, *158*, 107889. https://doi.org/10.1016/j.neuropsychologia.2021.107889, PubMed: 33991561