

Limited but specific engagement of the mature language network during linguistic statistical learning

Julie M. Schneider^{1,2,*}, Terri L. Scott³, Jennifer Legault⁴, Zhenghan Qi^{2,5,6}

¹Department of Communication Sciences and Disorders, Louisiana State University, 77 Hatcher Hall, Field House Dr., Baton Rouge, LA 70803, United States

²Department of Linguistics & Cognitive Science, University of Delaware, 125 E Main St, Newark, DE 19716, United States

³Department of Communication Sciences and Disorders, Northeastern University, 360 Huntington Ave., Boston, MA, 02115, United State

⁴Department of Psychology, Elizabethtown College, One Alpha Dr, Elizabethtown, PA 17022, United States

⁵Bouvé College of Health Sciences, Northeastern University, 360 Huntington Ave, Boston, MA 02115, United States

⁶Department of Psychology, Northeastern University, 105-107 Forsyth St., Boston, MA, 02115, United States

*Corresponding author: Department of Communication Sciences and Disorders, Louisiana State University, 77 Hatcher Hall, Field House Dr., Baton Rouge, LA 70803, United States. Email: juschnei@lsu.edu

Statistical learning (SL) is the ability to detect and learn regularities from input and is foundational to language acquisition. Despite the dominant role of SL as a theoretical construct for language development, there is a lack of direct evidence supporting the shared neural substrates underlying language processing and SL. It is also not clear whether the similarities, if any, are related to linguistic processing, or statistical regularities in general. The current study tests whether the brain regions involved in natural language processing are similarly recruited during auditory, linguistic SL. Twenty-two adults performed an auditory linguistic SL task, an auditory nonlinguistic SL task, and a passive story listening task as their neural activation was monitored. Within the language network, the left posterior temporal gyrus showed sensitivity to embedded speech regularities during auditory, linguistic SL, but not auditory, nonlinguistic SL. Using a multivoxel pattern similarity analysis, we uncovered similarities between the neural representation of auditory, linguistic SL, and language processing within the left posterior temporal gyrus. No other brain regions showed similarities between linguistic SL and language comprehension, suggesting that a shared neurocomputational process for auditory SL and natural language processing within the left posterior temporal gyrus is specific to linguistic stimuli.

Key words: fMRI; individual differences; language processing; statistical learning.

Introduction

Statistical learning (SL), a robust ability to rapidly detect regularities and variabilities from inputs, is a central theoretical account for first language acquisition (Aslin and Newport 2012; Saffran et al. 1996). However, whether and how learning the embedded patterns from an artificial language engages a mature language network in the human brain remains an open question. On one hand, successful SL has been demonstrated across age groups (e.g. infants and adults), species (e.g. monkeys: Meyer and Olson 2011; rats: Toro and Trobalón 2005), modalities (e.g. nonlinguistic auditory: Gebhart et al. 2009, Saffran et al. 1999; visual: Fiser and Aslin 2002), and linguistic hierarchies (e.g. phoneme discrimination: Maye et al. 2002; syntax discrimination: Gomez and Gerken 1999, Thompson and Newport 2007), advocating for a domain-general SL mechanism. On the other hand, learning outcomes systematically vary across sensory modalities (e.g. Conway and Christiansen 2005; Emberson et al. 2011), suggesting that SL performance is constrained by domain-specific mechanisms. Recent behavioral studies emphasize individual differences in SL and hint at a domain-specific relationship between linguistic SL and complex language processing (Arnon 2019; Christiansen et al. 2012; Erickson et al. 2016; McCauley and Christiansen 2015; Schneider et al. 2020; Siegelman et al. 2018; Siegelman and Frost 2015). Indeed, the modern neurobiological framework of SL proposes a dual-system model that includes both domain-specific and domain-general learning mechanisms (Conway 2020; Frost et al. 2015). Limited studies investigating the neural mechanisms

of SL have reported the involvement of domain-general brain areas associated with executive functions, learning, and memory, as well as language-related brain regions engaged during both perception and production (Cunillera et al. 2009; Finn et al. 2013, 2018; McNealy et al. 2006; Plante et al. 2017; Smalle et al. 2022; Orpella et al. 2022; Schapiro et al. 2012). However, the exact roles of these language-related brain regions during SL remain elusive, limiting our understanding of the relationship between language and SL in the brain. It remains unclear whether the brain processes supporting newly encountered linguistic regularities are similar to those engaged during natural language processing. Elucidating the relationship between language and SL in the brain can provide key insights into the nature of domain-specific constraints on the process of SL.

Prior linguistic experiences influence SL of linguistic stimuli. Notably, adults' learning performance on word segmentation and artificial grammar learning tasks is largely constrained by both the regularities and variabilities of their native languages (e.g. Bonatti et al. 2005; Dal Ben et al. 2021; Finn and Hudson Kam 2008; Mersad and Nazzi 2011; Onnis et al. 2005; Trecca et al. 2019). Even for school-aged children with limited language experiences, it is evident that they are sensitive to rich probabilistic information in the environment that shapes their representation of language (Stärk et al. 2022). This is because language learning and language processing are not isolated processes in the real world. Humans are constantly immersed in continuous language input that is mixed with both familiar and unfamiliar patterns.

Received: June 26, 2023. Revised: March 4, 2024. Accepted: March 5, 2024

© The Author(s) 2024. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

As suggested by the predictive processing framework (e.g. [Lupyan and Clark 2015](#)), learners can leverage their prior linguistic experiences to generate predictions about the underlying structure of this continuous language input, resulting in implicit and adaptive adjustment of the existing language processing system that ultimately leads to language acquisition ([Dell and Chang 2014](#); [Elazar et al. 2022](#); [Stärk et al. 2023](#)). Thus, it is likely that the functional organization of the language processing system in mature human brains directly engages in predictive processes in a similar way as it does for auditory, linguistic SL. By examining functional similarity between natural language processing and auditory SL across domains within individual brains, the current study seeks to understand how the language network, shaped by prior linguistic experiences, is engaged during linguistic and nonlinguistic SL.

The potential similarity between SL and natural language processing has been implicated by a limited set of functional neuroimaging studies. For example, in studies of SL involving speech stimuli, the inferior frontal gyrus (IFG), superior temporal gyrus (STG), and supramarginal gyrus in the left hemisphere demonstrate sensitivity to statistical regularities ([Cunillera et al. 2009](#); [McNealy et al. 2006](#); [Karuza et al. 2013](#); [Plante et al. 2017](#); [Orpella et al. 2021, 2022](#)) and increasing activation as a function of exposure ([Plante et al. 2017](#); [McNealy et al. 2006](#); [Karuza et al. 2013](#)). Due to the anatomical similarity between these regions and regions within the core-language network, which is a collection of brain regions consistently and specifically engaged during semantic and syntactic processing during natural language comprehension ([Bautista and Wilson 2016](#); [Blank et al. 2016](#); [Julian et al. 2012](#); [Keller et al. 2001](#)), but not during processing of nonlinguistic information ([Fedorenko et al. 2011](#); [Fedorenko and Varley 2016](#)), the activation of these regions during SL is often associated with language processing, although they have never been directly compared (e.g. [Park et al. 2022](#)). Existing evidence for the association between SL and language hinges on the assumption that the neural substrates found in SL have functions related to language. Despite providing early insights into the link between SL and language processing, this practice, known as reverse inferencing, is limited by the weak functional selectivity of these anatomical labels, as well as the remarkable intersubject variabilities in the locations of activation ([Fedorenko et al. 2010](#); [Poldrack 2011](#)). The lack of direct comparison between SL and language processing is problematic as the cortical regions found in these studies can be functionally heterogeneous themselves. The only known study to investigate common neural activation during spoken language processing and SL within the same individual utilized an artificial grammar learning (AGL) task and found that similar P600 responses were elicited for structural incongruencies during spoken language and AGL. The authors interpreted this to indicate that the same neural mechanisms were recruited during syntactic processing of both language stimuli and SL of sequential patterns; however, it remains possible that both tasks are subserved by domain-general neural generators ([Christiansen et al. 2012](#)).

A more thorough investigation of the functional overlap between SL and language processing relies on the quantitative association between two cognitive processes. Despite substantial research documenting the behavioral relationships between SL and language processing skills ([Arciuli and Simpson 2012](#); [Arciuli and Torkildsen 2012](#); [Conway et al. 2010](#); [Conway and Pisoni 2008](#); [Kirkham et al. 2002](#); [Misyak et al. 2010](#); [Qi et al. 2018](#)), only a few studies have demonstrated neurobiological evidence linking the two. For example, greater activation in the

left STG during an auditory linguistic SL task was positively associated with rapid auditory processing speed ([McNealy et al. 2006](#)). Greater neural engagement of the speech-motor regions in the frontal cortex during SL has also been associated with better SL performance ([Assaneo et al. 2019](#); [Orpella et al. 2022](#)). Evidence from populations with language and communication difficulties hint toward a similar language–SL relationship. For example, school-aged children with autism with reduced language skills showed poorer performance during linguistic, but not nonlinguistic, SL (measured by slower online learning, poorer offline pattern retrieval, and poorer overall learning performance, compared to their age-matched peers ([Hu et al. 2023](#)). In a functional MRI (fMRI) study, autistic children showed a lack of neural sensitivity to the structured speech stream compared to a random speech stream, while typically developing children showed activation at left inferior parietal lobule and putamen ([Scott-Van Zeeland et al. 2010](#)). While brain-behavior correlational approaches facilitate the interpretation of complex patterns of neural activation ([Reber 2013](#)), it is unclear whether the functional organization of SL in the brain is similar to that of language processing.

In addition to language-specific regions, many studies have highlighted the domain-generality of neural activation during SL tasks. For example, domain-general learning and memory systems, including the basal ganglia (e.g. [Finn et al. 2019](#); [Orpella et al. 2022](#)), striatum ([Orpella et al. 2021](#)), and hippocampus (e.g. [Schapiro et al. 2014, 2017](#)), are considered important for SL, though the necessity of their engagement is still greatly debated ([Covington et al. 2018](#); [Schapiro et al. 2014, 2017](#)). Intriguingly, activation at the left, right, or bilateral inferior frontal gyrus (IFG) during processing of both linguistic and nonlinguistic stimuli has been reported in most neuroimaging studies using SL paradigms ([Milne et al. 2018](#); [Karuza et al. 2013](#); [Schapiro et al. 2013](#); [Turk-Browne et al. 2009](#)). From these studies, researchers have concluded that because SL across modalities engages the LIFG, it is tapping into a domain-general executive function mechanism. Therefore, contrary to the interpretation above, where activation of the LIFG is involved in processing linguistic stimuli, activation of the LIFG has been viewed as a generic online sequence processor, which drives the computation of statistical regularities ([Pettersson et al. 2012](#); [Karuza et al. 2013](#)). These strikingly contradictory interpretations for the same anatomical region (e.g. LIFG) emphasizes how additional research is needed to establish whether or not extraction of abstract statistical patterns engages linguistic processing.

While SL tasks have been shown to activate regions often engaged during language processing, others suggest that these same regions underlie a domain-general SL mechanism. Three reasons may account for the paradox. First, as mentioned, many studies associating SL with language processing rely on correlation and reverse inferencing, oversimplifying activation within these regions that are known to be highly heterogeneous. Second, SL is highly heterogeneous across individuals. Group-level activation patterns may mask voxels that are activated at the individual level, and critical for SL. Third, activation during SL within the core-language network does not necessarily mean this region is processing regularities in an artificial language the same way as natural language processing. Taken together, there is a critical need to directly investigate whether the functional neural profiles among individual subjects during language processing and SL are similar or different.

To address this question, we utilize an auditory, linguistic SL task, auditory, nonlinguistic SL task, and a language processing

task to characterize whether auditory SL resembles the early steps of language acquisition using fMRI. The language processing task (Malik-Moraleda et al. 2022) is widely used with robust effect sizes, apparent interindividual variability, and satisfying test-retest reliability across most regions in the network (Mahowald and Fedorenko 2016). To independently and individually test the role of core-language regions in SL among individual learners, we implement a group-constrained subject-specific (GCSS) analysis (Scott et al. 2019; Julian et al. 2012; Fedorenko et al. 2010) and measure hemodynamic responses within language regions during each task. Lastly, we seek to elucidate the similarities in activation patterns between the language and SL tasks using both univariate conjunction analyses and multivariate similarity analyses. We hypothesize that if linguistic SL is constrained by an existing language system, then we should expect to see part of the mature language network actively engaged during auditory, linguistic SL. In addition, a shared functional organization of the language-related brain regions underlying auditory, linguistic SL and language processing will support a stronger version of this hypothesis.

Materials and methods

Participants

Twenty-seven adults ($M_{\text{age}} = 20.8$ years, $SD_{\text{age}} = 3.53$ years, 20 females) from the mid-Atlantic region of the United States were recruited via public flyers and websites. All participants were right-handed, monolingual English speakers, with no history of neurological disorder or developmental delay based on self-report. Participants were screened for MRI compatibility. The study was approved by the University Institutional Review Board and is in compliance with the Declaration of Helsinki. All participants provided informed consent and were compensated at a rate of \$20/hour in gift cards for their time. Due to technical issues, five participants did not yield sufficient neuroimaging data for analysis. Specifically, one participant did not complete all four runs of the auditory, SL task, two participants did not complete the language localizer task, and two participants had framewise displacement values that were significantly higher than other participants (>3.00), indicating substantial noise was present in their fMRI data. Therefore, the final sample size for this study is 22 ($M_{\text{age}} = 19.87$, $SD_{\text{age}} = 1.25$, 15 females).

Language localizer fMRI task

Stimuli

To identify brain areas sensitive to spoken language, participants underwent functional neuroimaging while they passively listened to engaging audio clips of natural speech taken from the Alice in Wonderland story. To control for low-level acoustic features of speech, these intact clips were contrasted with clips of degraded speech, which was processed to be unintelligible. The design and validation of this task is detailed in Scott et al. (2017) and Malik-Moraleda et al. (2022). Briefly, the degradation procedure consisted of creating low-pass (cutoff frequency 500 Hz) filtered copies of a set of clips of speech similar to the intact set. Then, for each clip, amplitude envelope-modulated white noise was added where the envelope was copied from the original clip. This white noise was then “softened” with low-pass filtering at pass-band frequency 8000 Hz with stop frequency of 10,000 Hz. The resulting clips are recognizable as human speech but lack lexical and phonemic content, and sound similar to an unintelligible radio

broadcast. These stimuli, their transcriptions, and a presentation script are available online. (<https://evlab.mit.edu/alice/loc/>).

Procedure

Across two separate runs, 16 blocks of intact speech and 16 blocks of degraded speech were presented. Each run lasted 5 min and 58 s. Blocks consisted of one continuous 18-s audio clip. In addition to the 16 stimulus blocks per run, five 14-s blocks of rest were included, in which participants heard no stimuli. These blocks of rest occurred every four blocks of randomly interspersed intact or degraded speech blocks. For example, a run consisted of 16 blocks (B) in this sequence: Rest B1 B2 B3 B4 Rest B5 B6 B7 B8 Rest B9 B10 B11 B12 Rest B13 B14 B15 B16 Rest (Scott et al. 2017). A fixation cross was presented on the screen continuously during the entire run. Participants were only instructed to listen attentively, and no overt responses were recorded.

Auditory SL task

A common practice in fMRI studies of SL is to compare brain responses during a familiarization phase between sequences containing regular patterns versus those containing the same stimuli, which are ordered randomly. However, implementing a random control condition often interferes with learning of structured sequences. Previous research found no evidence for behavioral learning despite neural differences between structured and random sequences (McNealy et al. 2006, 2011). We address this issue in the current study by presenting structured and random sequences containing different types of stimuli in the same run (speech syllables and non-speech tones) and counterbalancing these conditions across runs. Each participant was therefore exposed to structured speech, random speech, structured tone, and random tone sequences. Importantly though, participants were exposed to only one type of stimuli at a time, with a 6.25-s block of silence inserted between conditions. This approach has been shown to result in successful learning of structured sequences, as indicated by significantly quicker response times in the structured than the random condition (Schneider et al. 2020).

Stimuli

Syllable stimuli were constructed from twelve English consonant vowel syllables (pi/pa/pu/ti/ta/tu/di/da/du/bi/ba/bu). All syllable stimuli were made in Praat using the artificial speech synthesizer female voice (version 6.1.14; Boersma and Weenink 2020). The mean duration of each syllable was 350 ms ($SD = 40$ ms) with a stimulus onset asynchrony (SOA) set at 480 ms. Tone stimuli were constructed from 12 unique musical pure tones (F, G, D, G#, C#, B, C, F#, D#, E, A, A#) of the same duration (328 ms) within the same octave (a full chromatic scale starting from middle C). The SOA was set at 480 ms.

Participants were exposed to experimental blocks containing sequences with embedded regularities (i.e. structured blocks), sequences with no statistical regularities (i.e. random blocks), and silence (i.e. resting blocks; see Fig. 1). Structured blocks involved the presentation of sequences containing four triplets, which were created by concatenating the syllables into trisyllabic “words” (pi-tu-bi, bu-pa-da, di-ba-pu, and ta-ti-du) and the tones into four target triplets (F#DE, ABC, C#A#F, and GD#G#). Random blocks contained the same 12 stimuli as presented in the structured blocks but were ordered pseudorandomly, so that no combinations of any three stimuli were repeated more than once. A silent blank screen was presented during the resting blocks. Within each block, each triplet was repeated 8 times for a total of 32 triplets per block. Three structured blocks, three random blocks,

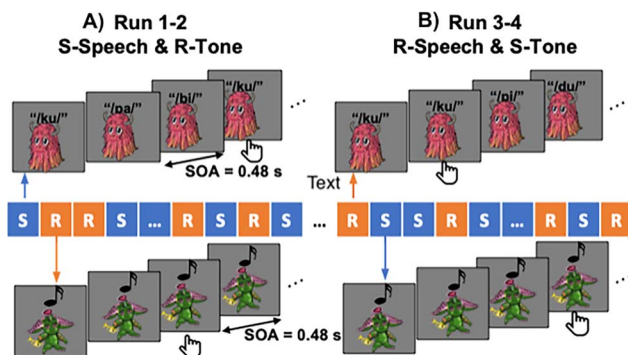


Fig. 1. Overview of auditory linguistic statistical learning fMRI task. Participants were exposed to a) two runs of structured blocks each containing a 96-syllable sequence with the embedded triplets (S-speech), intermixed with three random blocks of tone sequences (R-tone) and b) two runs of random blocks each containing a 96-syllable randomly ordered sequence (R-speech), intermixed with three structured blocks of tone sequences (S-tone). Each of these two runs always occurred consecutively. Participants were instructed to alternate between tracking a specific syllable (e.g. /ku/) and tracking a specific tone (e.g. B) in these sequences by pressing a button. Note: Only the alien images, but not the spelling of the syllables, were present on the screen.

and six resting blocks were concatenated in a pseudorandom order to create one run of auditory stimuli. The randomization was independent for each run and for each participant. Adjustment was made to ensure that no more than two blocks of the same stimulus type were adjacent to each other. To maximize the opportunities for participants to learn the embedded syllable triplet patterns, the first two experimental runs contained the syllable structured blocks intermixed with the tone random blocks and the second two experimental runs contained the syllable random blocks intermixed with the tone structured blocks.

Procedure

Participants listened to sequences of sounds and viewed a screen with an alien cartoon character in the center while performing an auditory target-detection task with a button response pad (Cambridge Research Systems). Before each block, a cue for the target was presented briefly with an auditory instruction “now listen for the sound /ku/” or “now listen for the sound /G#/#” (Fig. 1). The cue was a stimulus chosen from the last position of a triplet and was randomly assigned across each participant. Participants were instructed to press a button when they heard the target cue. Behavioral analysis of SL involved using participants’ reaction times during this cue detection task. Mean reaction time (RT) was computed as the average time it took for a participant to press the button for each target syllable. A valid button press in response to the target syllable had to occur in the time window of one stimulus before and one stimulus after the target syllable; that is, within -480 to $+960$ ms relative to the onset of the target syllable. A button press prior to the presentation of the target syllable indexed anticipation, resulting in a negative reaction time. All behavioral analyses for the current study were conducted in R (R Core Team, 2012).

Throughout the task, in each participant, the same cue was used for all the structured and random blocks of the same stimulus type. The target syllable and the target tone were randomly assigned to each participant. Participants completed four experimental runs; each lasted 6 min and 33 s. We found that the participants did not differ in their task engagement across the two

halves of the experiment (see 3.1 for analyses in target-detection behavior).

Neuroimaging data

MRI acquisition

Data were acquired on a Siemens 3-T Magnetom Prisma scanner with a 64-channel phased array head coil at the Center for Brain and Biomedical Imaging at the University of Delaware. Prior to functional imaging, whole-head, high-resolution structural images, including a T1-weighted, magnetization-prepared rapid gradient-echo anatomical volume (TR=2500 ms, TE=2.9 ms, TI=1070 ms, flip angle=8.0°, voxel resolution=1.0 mm isotropic, field of view [FOV]=256 × 256, 176 sagittal slices) and a T2-weighted anatomical volume (TR=3200 ms, TE=565 ms, flip angle=2.0°, voxel resolution=1.0 mm isotropic, FOV=256 × 256, 32 sagittal slices) were collected.

Functional data were acquired using simultaneous T2*-weighted echo-planar imaging with multislice scans acquisition (Feinberg et al. 2010; Moeller et al. 2010; Setsompop et al. 2012) with the following acquisition parameters: TR=800 ms, TE=32 ms, flip angle=61°, FOV=210 × 210 mm, and acceleration factor=6. For both SL tasks, we required 60 adjacent slices in an interleaved order with 2.5-mm slice thickness resulting in an in-plane resolution of $2.5 \times 2.5 \times 2.5$ mm³. For the language localizer task, we required 45 adjacent slices in an interleaved order with 3.0-mm slice thickness resulting in an in-plane resolution of $3.0 \times 3.0 \times 3.0$ mm³. After fMRI preprocessing, the language localizer data were resampled to the same space as the SL data using participant’s anatomical data as the reference image within FSL (FSL 5.0.9; Zhang et al. 2001). We used a nearest neighbor approach to resample, which assigns the voxel to the nearest voxel in the original space.

fMRI preprocessing

Functional and anatomical data were first converted using HeuDiConv (Halchenko et al. 2020) and then organized to cohere with the Brain Imaging Data Structure (BIDS; Gorgolewski et al. 2016). fMRIPrep 1.3.1 (Esteban et al. 2019; RRID: SCR_016216), which is based on Nipype 1.1.9 (Gorgolewski et al. 2011, 2018), was utilized for all preprocessing steps. Many internal operations of fMRIPrep use Nilearn 0.5.0 (Abraham et al. 2014), mostly within the functional processing workflow. This automated preprocessing pipeline combines methodology from AFNI (Cox and Hyde 1997), ANTs 2.2.0 (Avants et al. 2022), FreeSurfer 6.0.1 (Dale et al. 1999), FSL (FSL 5.0.9; Zhang et al. 2001), and Mindboggle (Klein et al. 2005, 2017) to provide scientifically rigorous and reproducible data for use in data analysis. fMRIPrep was first conducted on anatomical data only, described in more detail below. All functional data were then preprocessed based on these corrected T1 images.

Anatomical data preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped. Brain surfaces were reconstructed using FreeSurfer. Nonlinear spatial normalization (warping) of brain-extracted anatomical volumes to the ICBM 152 Nonlinear asymmetrical template version 2009c (Fonov et al. 2011) was performed. Brain tissue segmentation of cerebrospinal fluid, white matter, and gray matter was performed on the brain-extracted T1w.

Functional data preprocessing

For each of the functional runs per subject (four auditory, linguistic SL runs, four auditory, nonlinguistic SL runs, and two language localizer runs), the following preprocessing was performed. First, a brain-masked functional reference volume was generated using a custom methodology of fMRIPrep. A deformation field to correct for susceptibility distortions was estimated based on a coregistered field map, using a custom workflow of fMRIPrep derived from D. Greve's `epidewarp.fsl` script (<http://www.nmr.mgh.harvard.edu/~greve/fbirt/b0/epidewarp.fsl>) and further improvements from HCP Pipelines (Glasser et al. 2013). Based on the estimated field distortion, a distortion-corrected BOLD reference was calculated for a more accurate coregistration with the anatomical reference. The BOLD reference was then used to compute the linear transformation between the subject's functional data and anatomical volume using boundary-based registration (bbregister; Greve and Fischl 2009). Coregistration was configured with 9 degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters for each volume in the time series were estimated with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters). Time-series data were then high-pass filtered and underwent slice-timing correction, recommended for multiband data with short TRs (Parker and Razlighi 2019). Head-motion and field distortion corrections were then applied to the time series with a single, composite transformation. The BOLD time series were resampled to MNI 152.

First-level statistical analysis was carried out using FEAT (FMRI Expert Analysis Tool; Woolrich et al. 2001). For the language localizer task, we computed contrast images between the intact and the degraded speech conditions for each functional run and each participant, and then we averaged the contrast images across the two runs (see [Supplementary Analyses](#) for results of this univariate analysis). For both the auditory, linguistic and auditory, nonlinguistic SL tasks, for each individual run, parameter estimates for structured and random speech relative to baseline were calculated. Then, these contrast images were averaged across runs for each experimental condition, such that in both SL tasks, structured versus rest and random versus rest contrasts, and in the language localizer task, intact versus degraded and degraded versus intact, were generated for each participant. To directly examine the effect of SL and to take run order difference into account, we also computed the structured-random contrast for the syllable and tone stimuli separately by computing the differences between the contrast images of task and rest blocks within each run: structured speech versus rest compared to random speech versus rest and structured tones versus rest compared to random tones versus rest. Parameter estimates for each contrast of interest were constructed. Group-level means for each contrast of interest were computed using FLAME (FMRIB's Local Analysis of Mixed Effects; Beckmann et al. 2003).

GCSS analyses

Since traditional group-level approaches can result in the loss of meaningful individual variability, we performed GCSS analyses designed to take into account intersubject variability (Scott et al. 2019; Julian et al. 2012; Fedorenko et al. 2010). With these analyses, we defined probabilistic regions of interest ("parcels"), which we used to constrain our individual subject analyses. Specifically, we used the updated language parcels provided in Malik-Moraleda et al. (2022) to measure patterns of similarity between tasks within individuals. From within each language parcel, the

contrasts between structured speech during SL and rest, random speech during SL and rest, structured tones during SL and rest, random tones during SL and rest, intact sentence processing and rest, and degraded sentence processing and rest were extracted and compared (Results section "Subject-specific language regions are engaged during auditory SL").

Multivoxel pattern analyses

To ensure thorough investigation of neural patterns shared across auditory SL and language processing, we used local pattern similarity analysis. This technique is designed to identify brain regions that support similar functions during different tasks, regardless of the exact level of activation (Scott and Perrachione 2019). For this analysis, unsmoothed functional contrast maps from each task are compared. These maps have undergone the same preprocessing and first-level analysis as the data discussed so far, including warping to common MNI space, except that these maps have not been spatially smoothed and thresholded. To first determine whether the pattern of activity in each language parcel reflects similar task engagement during linguistic SL and language processing, we computed Pearson correlation coefficients between the structured > random linguistic SL, structured > random nonlinguistic SL, and intact > degraded language processing contrasts across all voxels in each parcel, within individual subjects. We assessed the significance of these correlations across our participants under a null hypothesis in which unrelated patterns of activity had a correlation of zero (Results section "Linguistic SL and language processing show representational similarity in the left Posterior Temporal Gyrus").

Results

Auditory SL performance

To ensure participants attended to and learned the embedded triplets in the speech stream, we measured response reaction time to both structured and random speech sequences. We predicted that participants would respond faster to target syllables when listening to structured sequences (structured speech) as compared to random sequences (random speech). One participant was removed from the speech RT analysis, as they did not have enough valid key presses (<6 trials) during the exposure phase ($M_{\text{age}} = 19.90$ years, $SD_{\text{age}} = 1.26$ years, number of females = 14). Consistent with what we have reported in the complete sample in Schneider et al. (2020), the average RT to the target syllable across the entire exposure phase was significantly quicker in the structured ($M = 643.29$ ms, $SD = 83.27$ ms) than the random condition ($M = 670.07$ ms, $SD = 55.41$ ms; $t(19) = -1.91$, one-tailed $P = 0.04$). We also predicted that participants would respond faster to target tones when embedded in statistical regularities (structured tones) as compared to random sequences (random tones). Four participants were removed from the tone RT analysis, as they did not have enough valid key presses (<6 trials) during the exposure phase ($M_{\text{age}} = 19.67$ years, $SD_{\text{age}} = 0.91$ years, number of females = 13). The average RT to the target tone across the entire exposure phase was significantly quicker in the structured ($M = 515.52$ ms, $SD = 74.64$ ms) than the random condition ($M = 567.89$ ms, $SD = 78.03$ ms; $t(15) = -2.63$, one-tailed $P = 0.009$).

Since structured and random conditions within each stimulus type were presented at different runs, one possible confounding factor was task engagement. Therefore, we also computed A' values as this is a measure of participants' sensitivity for target stimulus based on hit and false alarm rates (Aaronson and Watts 1987; Grier 1971; Pallier 2002). An A' near 1.0 indicates good

discriminability for target over nontarget stimuli, while a value near 0.5 indicates chance performance. The mean A' values for each condition were as follows: structured speech condition 0.90 ($SD=0.10$), random speech 0.89 ($SD=0.05$), structured tone 0.80 ($SD=0.14$), and random tone 0.76 ($SD=0.14$). There were no significant differences between structured and random sequences in the speech ($t(20)=0.32$, $P=0.76$) and tone ($t(20)=-0.82$, $P=0.42$) conditions, indicating that participants were equally engaged in both conditions, and the RT difference between structured and random conditions was not due to differences in hit and false alarm rates.

Subject-specific language regions are engaged during auditory SL

Univariate conjunction analyses between auditory, linguistic SL and language processing suggest functional overlap at the group level (Supplementary Analyses). To test whether language regions are engaged during SL in individual participants, we used a publicly available set of updated language parcels from Fedorenko et al.'s (2010) language localizer task. These parcels were created by contrasting sentences > nonwords conditions across 220 participants and their right-hemisphere homologues. We chose to use these parcels rather than the parcels created from our own sample and language localizer task for the following reasons. First, the results will facilitate comparison with other existing studies using the same parcels. Second, in a study using the same "Alice-in-Wonderland" task across 45 languages, the language parcels created from Intact > Degraded Speech contrast were nearly identical to the original set (Malik-Moraleda et al. 2022).

We next asked whether subject-specific language functional region of interest (fROI), defined by the most activated voxels within each language parcel during the "Alice-in-Wonderland" task (Intact vs. Degraded contrast), are engaged during auditory, linguistic SL. To statistically evaluate this functional overlap, we extracted the mean activation during linguistic SL for each condition (structured vs. rest and random vs. rest) for each language fROI. This results in a measure of how sensitive each language parcel is to the linguistic SL contrast for each participant.

Listening to structured sequences of speech syllables resulted in significantly greater hemodynamic responses than listening to random sequences of speech syllables in the left posterior temporal gyrus ($t(39.16)=1.70$, one-tailed $P=0.048$). No other language regions showed sensitivity to differences in statistical regularities of speech (P 's > 0.05; see Supplementary Table 1 for full statistics).

To clarify whether activation of the left posterior temporal gyrus during SL is domain specific or domain general, we next extracted the mean activation during tone SL for each condition (structured vs. rest and random vs. rest). Listening to structured sequences of tone syllables did not result in a significantly greater hemodynamic response than listening to random tone sequences in the left posterior temporal gyrus ($P>0.05$; see Supplementary Table 2 for full statistics). No other language regions showed sensitivity to differences in statistical regularities of tones (P 's > 0.05; see Supplementary Table 1 for full statistics).

To determine whether the condition differences during auditory linguistic SL elicited a response that is similar in magnitude to that produced during language processing, we also examined hemodynamic responses during listening to intact and degraded sentences during the language localizer task within the left posterior temporal gyrus, and all other language parcels. Listening to intact speech resulted in a significantly greater hemodynamic response than listening to degraded speech in the left posterior temporal gyrus ($t(36.85)=4.04$, one-tailed $P<0.001$). All parcels,

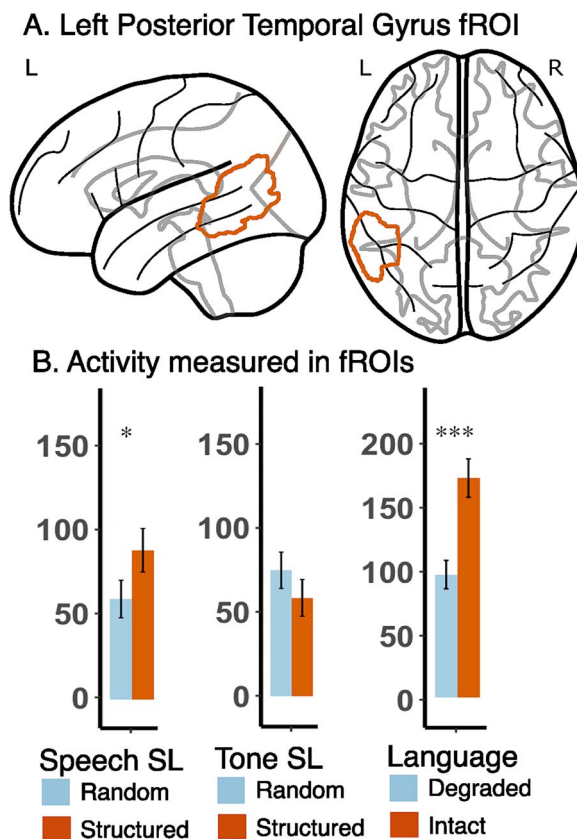


Fig. 2. Mean activation during processing of structured speech and random speech, structured tone and random tone, and intact and degraded speech within the left posterior temporal gyrus. A) Greater activation for structured speech sequences compared to random speech sequences was observed in the left posterior temporal gyrus. B) Mean activity (b) for random (blue) and structured (orange) speech sequences, random (blue) and structured (orange) tone sequences, and degraded (blue) and intact (orange) language processing within the left posterior temporal gyrus. Error bars represent the within-subject standard error. Asterisks represent one-tailed P values: * $P<0.05$, *** $P<0.001$.

with the exception of the bilateral angular gyrus, and the right IFG and MFG, also demonstrated significant differences between intact and degraded speech (see Supplementary Table 3 for full statistics). Average hemodynamic responses within the left posterior temporal gyrus across tasks are presented in Fig. 2.

Due to the imbalanced distribution of gender in the current sample (15 female, 7 male), we also examined whether there were gender differences within the left posterior temporal gyrus using a one-way ANOVA predicting differences in BOLD activation (structured-random) during auditory SL between males and females. We found no difference between males and females in the activation of the left posterior temporal gyrus during speech SL ($F(1,19)=0.16$, $P=0.70$) or tone SL ($F(1,19)=0.66$, $P=0.43$).

Linguistic SL and language processing show representational similarity in the left posterior temporal gyrus

Univariate conjunction analysis applies a voxel-wise activation threshold and searches for voxels consistently coactivated by both tasks at the group level. This approach does not necessarily indicate similarities within individuals across both tasks. Our analyses in language parcels revealed greater activation for structured speech sequences compared to random speech sequences.

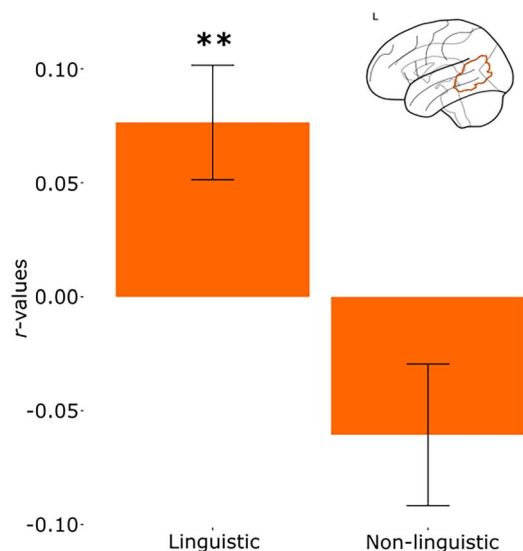


Fig. 3. Representational similarity between language and auditory, linguistic SL within the left posterior temporal gyrus, which was sensitive to differences in statistical regularities of speech during structured versus random sequences. Cross-task correlation values between language and auditory, linguistic SL within the left posterior temporal gyrus significantly differed from zero. There was a lack of representational similarity between language and auditory, nonlinguistic SL in this same region, and cross-task correlation values between these tasks did not significantly differ from zero. Error bars represent the within-subject standard error. Note: ** $P < 0.01$.

These findings confirm that learning embedded speech regularities engages language region; however, voxels which are most activated within a given parcel might differ across tasks. To further test whether underlying neurocomputational processes are similar between the two tasks, we turned to multivoxel pattern analyses for answers. To determine whether the multivoxel functional organization within the language parcels was similar across tasks, we masked each task's contrast map using the language parcel shown to be sensitive to linguistic SL (left posterior temporal gyrus) and extracted the contrast value from each voxel within the whole parcel for each task for each participant (language: intact vs. degraded; linguistic/nonlinguistic SL: structured vs. random). Next, we computed Pearson correlation coefficients between the language localizer task and each SL task across all voxels in each parcel for each participant. There was a positive, significant cross-task correlation between the language localizer and linguistic SL tasks ($t(21) = 3.04$, $P = 0.006$, compared to zero). In contrast, there was a nonsignificant anticorrelation between the language localizer and nonlinguistic SL tasks ($t(21) = -1.95$, $P = 0.06$). These results are visualized in Fig. 3. When computing Pearson correlation coefficients between auditory linguistic and nonlinguistic SL, we found a significant, negative correlation ($t(21) = -3.51$, $P = 0.002$). These findings suggest that the left posterior temporal gyrus is not only specifically engaged during auditory linguistic SL, but also presents similar multivoxel activation pattern as natural language processing.

Discussion

The current study investigates the relationship between auditory SL and natural language processing by testing whether these tasks share similar neural substrates in adults. Processing of speech streams with embedded patterns, as opposed to random syllable

sequences, activated a similar brain region, the left posterior temporal gyrus, as was activated during natural language processing. Given this region was not activated during auditory, nonlinguistic SL, it appears that the left posterior temporal gyrus may underlie domain-specific properties relevant for both linguistic SL and higher-level language processing. Furthermore, our multivoxel analysis revealed that patterns within the left posterior temporal gyrus language parcel are associated across natural language processing and auditory, linguistic SL, suggesting similarities in the underlying neurocomputation of linguistic SL and that of natural language comprehension in this region. Such similarity was absent between nonlinguistic SL and natural language comprehension. Despite our findings that the left posterior temporal gyrus supports both language comprehension and auditory, linguistic SL of an artificial language, this was the only region identified as relevant for both tasks among the 12 regions of interest. Taken together, the current findings suggest that while the left posterior temporal gyrus is associated with domain-specific properties relevant for learning and using language in adults, the lack of task similarity in other language regions suggests that critical differences exist between language learning and processing in mature language learners.

Our results contribute to the ongoing debate surrounding domain-general versus domain-specific nature of SL processes. Our subject-specific univariate analysis determined that the left posterior temporal gyrus, part of the frontotemporal core-language network (Fedorenko and Thompson-Schill 2014; Price 2010), is sensitive to embedded regularities in a stream of meaningless syllables, but not tones. Importantly, our multivariate pattern similarity analysis adds to these results by demonstrating that the neurocomputational processes supporting language processing and auditory, linguistic SL share certain degrees of similarity within the left posterior temporal gyrus. Based on these findings, we believe that the left posterior temporal gyrus is functionally relevant for learning auditory linguistic regularities, but not statistical regularities in general. While SL has historically been viewed as involving a set of domain-general neurobiological mechanisms for learning, representation, and processing that detect and encode a wide range of distributional properties within different modalities or types of input (Saffran et al. 2007), more recent accounts suggest that SL is not represented by a unitary learning system but, rather, by separate neural networks in different cortical areas (e.g. visual, auditory, and somatosensory cortex; Frost et al. 2015). In the current study, we add to this modern theoretical account by demonstrating that the process of encoding an internal representation follows constraints that are determined by the specific properties of the input being processed. As a result, the outcomes of computations occurring in the left posterior temporal gyrus, while not necessarily modality specific, are invoking similar sets of computational principles and some shared neural architecture to process and learn language. It is important to note that processing of structured speech regularities resulted in a similar BOLD response magnitude to degraded speech, which was substantially lower than intact speech during natural language comprehension. The degraded speech, despite being of low intelligibility, maintains the sequential regularities of speech acoustic cues. Previous findings suggest an association between implicit SL and speech perception under degraded listening conditions (Conway et al. 2010). Similar activation magnitude across structured artificial language and degraded natural language in the left posterior temporal gyrus is potentially attributed to the co-occurring patterns embedded in both conditions. Because this region is also reliably activated

during syntactic and semantic processing of natural languages (Fedorenko et al. 2020; Snijders et al. 2009), it is not surprising that activation magnitude in the intact speech condition was higher than linguistic SL.

In addition to these important confirmatory results, our study contributes to the recent emerging psycholinguistic frameworks aiming to unify language processing and acquisition. For example, part of the P-chain framework (Chang and Dell 2014) proposes implicit learning as the mechanism for how a mature language system adapts to changing inputs during processing. Specifically, linguistic experiences throughout childhood are known to shape both SL ability (Conway 2020; Kuhl 2004; Onnis and Thiessen 2013; Potter et al. 2017; Poulin-Charronnat et al. 2017; Siegelman et al. 2018) and language-related brain regions (Li et al. 2014; Pliatsikas et al. 2020; Wong et al. 2016). These same contributions of linguistic experience in shaping SL are also observed in adulthood: adults' ability to learn trisyllabic nonsense words in artificial speech is associated with greater sensitivity to high-frequency trigrams in natural language (Isbilen et al. 2022). Thus, increased familiarity with linguistic inputs and language proficiency throughout development might in turn facilitate SL in the linguistic domain. Despite our research demonstrating cross-task activation in the left posterior temporal gyrus, the directionality of this relationship remains unknown, making it unclear whether the current findings are attributed to the language network being constrained by prior language experiences, in turn, constraining SL processes, or vice versa. Taking a longitudinal, developmental approach by comparing children and adults in their recruitment of the language network during SL is necessary to test whether the maturity of the language system is associated with greater or lesser statistical sensitivity of an artificial language.

Despite the left posterior temporal gyrus demonstrating coactivation during both auditory, linguistic SL and language processing, our findings indicate that most other language regions are not robustly engaged during SL. This may seem surprising given the established role of the left IFG in processing both linguistic and nonlinguistic regularities (Milne et al. 2018; Karuza et al. 2013; Schapiro et al. 2013; Turk-Browne et al. 2009) as well as language learning (Ventura-Campos et al. 2013; Folia et al. 2010; Myers and Swan 2012; Ishkhanyan et al. 2020; Friederici 2009; Hagoort 2005; Heim et al. 2009; for review, see Qi and Legault 2020). However, there has been no empirical evidence associating SL and language-specific left IFG within individual participants. Given the anatomical adjacency of domain-general and language-specific frontal regions, it is likely that different subregions of the left IFG are engaged in linguistic SL and language processing. Indeed, our approaches, relying on rigorous functional localization techniques, did not find evidence supporting such an SL-language association in other regions than the left posterior temporal gyrus. Lastly, the current findings are based on a sample of 22 participants. Although 96% of highly cited experimental fMRI studies have a single group of participants and median sample size of 12 (Szucs and Ioannidis 2020), replications with larger sample sizes are necessary to verify these findings especially for the null results.

To further investigate whether human brain processes statistical regularities in artificial language similarly as in natural language, future research may consider implementing a task that requires retrieving existing statistical regularities from language knowledge (e.g. multiword chunking task in Isbilen et al. 2022). Similarly, altering the language localizer task utilized may uncover particular cognitive components shared between language processing and SL, as the language parcels identified

in the current study are biased toward holistic comprehension (Fedorenko et al. 2020). Specifically, it is important to consider how the neural mechanisms responsible for SL are linked to different elements of language processing, such as auditory perception, speech perception, and speech-motor functions that can be involved during passive listening (e.g. Assaneo et al. 2019; Orpella et al. 2022). Although participants were instructed to attend to stimuli in both tasks, only the SL tasks required voluntary attention to the target syllable/tone in the form of a target-detection task. While concentration of attention during exposure to statistical regularities in speech has been shown to not strongly influence the perceptual component of SL, it may differentially engage learning-related processes, compared to a passive SL task (Schneider et al. 2022; Batterink and Paller 2019). Therefore, the current SL findings may be task dependent in that different neural regions may be more or less activated in tasks where cognitive load is reduced. It is also important to note that we examined the process of learning across the entire exposure period, which encompasses various learning rates that might be substantially different across individual participants. For example, López-Barroso et al. (2015) revealed that the connectivity strength of the auditory-premotor network (one of three dorsal language-related networks relevant for SL) varied as a function of individual learning rates. While beyond the scope of the current paper, future research should consider how activation of the frontotemporal core-language network during SL may vary based on the speed by which individuals extract linguistic regularities.

Taken together, our findings provide evidence that auditory linguistic SL recruits subregions within the left posterior temporal gyrus that are specifically engaged for natural language comprehension. Moreover, multivoxel pattern analyses suggest that similar neural processes are undertaken by these subject-specific left posterior temporal gyrus regions during natural language comprehension and auditory linguistic SL. However, we did not find commonly activated brain regions outside of this region across the entire language network. Our study is the first to specifically examine whether and how SL engages a core-language network in adults. Future research should be careful to not draw assumptions about the similarity between SL and language processing on the basis of activation in a given region alone, as our findings indicate that there are substantial regional differences between tasks when subject-specific approaches are used and multivoxel patterns are examined at the individual level.

Acknowledgments

The data collection was conducted at the University of Delaware. We thank the faculty and staff at the Center for Biological and Brain Imaging at the University of Delaware for their critical support for our neuroimaging data collection: Keith Schneider, James Hoffman, John Christopher, Ibrahim Malik, and Trevor Wigal. We thank Patrick McMahon for IT support. We thank Violet Kozloff and An Nguyen for their assistance in stimulus construction and programming for the experiment, and Yi-Lun Weng and Anqi Hu for their contribution in data collection, organization, and behavioral data analysis. We thank Ev Fedorenko for generously sharing the materials of the language localizer task. We thank Morten Christiansen, Noam Siegelman, Laura Batterink, Casey Lew Williams, Alexis Pérez-Bellido, and Kelly Hay Yee Chan for providing helpful comments.

Author contributions

Julie Schneider (Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing—original draft, Writing—review & editing), Terri L. Scott (Formal analysis, Resources, Writing—review & editing), Jennifer Legault (Methodology, Writing—review & editing), and Zhenghan Qi (Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Writing—review & editing).

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

This work was supported in part by the National Institute on Deafness and Other Communication Disorders (NIDCD) Grant R21DC010576 to Z.Q. The first author's time was supported by the National Science Foundation Directorate for Social, Behavioral & Economic Sciences under Grant 1911462 to J.M.S.

Conflict of interest statement: None declared.

Data and code availability

Preprocessed fMRI data, analysis scripts used to analyze fMRI data, and statistical analyses conducted in R can all be found online (https://github.com/juliagoolia28/manuscripts/tree/master/langloc_sl).

References

- Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, Gramfort A, Thirion B, Varoquaux G. Machine learning for neuroimaging with scikit-learn. *Front Neuroinform*. 2014;8(FEB):14. <https://doi.org/10.3389/FNINF.2014.00014/FULL>.
- Arciuli J, Simpson IC. Statistical learning is related to reading ability in children and adults. *Cogn Sci*. 2012;36(2):286–304.
- Arciuli J, Torkildsen JVK. Advancing our understanding of the link between statistical learning and language acquisition: the need for longitudinal data. *Front Psychol*. 2012;3:324.
- Arnon I. Statistical learning, implicit learning, and first language acquisition: a critical evaluation of two developmental predictions. *Top Cogn Sci*. 2019;11(3):504–519.
- Aslin RN, Newport EL. What statistical learning can and can't tell us about language acquisition. *Infant Pathways to Language: Methods, Models, and Research Directions*. Psychology Press. 2012:15–30. <https://doi.org/10.4324/9780203809907>.
- Assaneo M, Ripollés P, Orpella J, Lin W, de Diego-Balaguer R, Poeppel D. Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nat Neurosci*. 2019;22(4):627–632 <https://www.nature.com/articles/s41593-019-0353-z>.
- Avants B, Tustison NJ, Song G. Advanced normalization tools: V1.0. *Insight J*. 2022:1–35.
- Batterink LJ, Paller KA. Statistical learning of speech regularities can occur outside the focus of attention. *Cortex*. 2019;115:56–71.
- Bautista A, Wilson SM. Neural responses to grammatically and lexically degraded speech. *Lang Cogn Neurosci*. 2016;31(4):567–574.
- Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in fMRI. *NeuroImage*. 2003;20(2):1052–1063.
- Blank I, Balewski Z, Mahowald K, Fedorenko E. Syntactic processing is distributed across the language system. *NeuroImage*. 2016;127:307–323.
- Boersma P, Weenink D. Praat: doing phonetics by computer [computer program]. Version. 2020:6(1):14.
- Christiansen MH, Conway CM, Onnis L. Similar neural correlates for language and sequential learning: evidence from event-related brain potentials. *Language and Cognitive Processes*. 2012;27(2):231–256. <https://doi.org/10.1080/01690965.2011.606666>.
- Conway C. How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neurosci Biobehav Rev* https://www.sciencedirect.com/science/article/pii/S0149763419307122?casa_token=ob7zkHwEDkUAAAAA:zjvk-YNMnzsGodZrsyl9crOAK1kTYL1Qhas5ARMxYuUJZ4Jt-wp8jYlzX1WUTJqllWlunWg1y08. 2020;112:279–299.
- Conway CM, Bauernschmidt A, Huang SS, Pisoni DB. Implicit statistical learning in language processing: word predictability is the key. *Cognition*. 2010;114(3):356–371.
- Conway CM, Christiansen MH. Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J Exp Psychol Learn Mem Cogn*. 2005;31(1):24–39.
- Conway CM, Pisoni DB. Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Ann N Y Acad Sci*. 2008;1145(1):113–131.
- Cox RW, Hyde JS. Software tools for analysis and visualization of fMRI data. *NMR Biomed*. 1997;10(4–5):171–178.
- Cunillera T, Càmarà E, Toro JM, Marco-Pallares J, Sebastián-Galles N, Ortiz H, Pujol J, Rodríguez-Fornells A. Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*. 2009;48(3):541–553.
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. *NeuroImage*. 1999;9(2):179–194.
- Dell GS, Chang F. The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philos Trans R Soc Lond Ser B Biol Sci*. 2014;369(1634):20120394.
- Elazar A, Alhama RG, Bogaerts L, Siegelman N, Baus C, Frost R. When the “tabula” is anything but “rasa”: what determines performance in the auditory statistical learning task? *Wiley Online Library*. 2022;46(2):13102.
- Emberson LL, Conway CM, Christiansen MH. Timing is everything: changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Q J Exp Psychol*. 2011;64(5):1021–1040.
- Erickson LC, Kaschak MP, Thiessen ED, Berry CAS. Individual differences in statistical learning: conceptual and measurement issues. *Collabra*. 2016;2(1):14. <https://doi.org/10.1525/collabra.41>.
- Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods*. 2019;16(1):111–116.
- Fedorenko E, Behr MK, Kanwisher N. Functional specificity for high-level linguistic processing in the human brain. *Proc Natl Acad Sci USA*. 2011;108(39):16428–16433.
- Fedorenko E, Blank IA, Siegelman M, Mineroff Z. Lack of selectivity for syntax relative to word meanings throughout the language network. *Cognition*. 2020;203:104348.
- Fedorenko E, Hsieh PJ, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N. New method for fMRI investigations of language:

- defining ROIs functionally in individual subjects. *J Neurophysiol.* 2010;104(2):1177–1194.
- Fedorenko E, Thompson-Schill SL. Reworking the language network. *Trends Cogn Sci.* 2014;18(3):120–126.
- Fedorenko E, Varley R. Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Ann N Y Acad Sci.* 2016;1369(1):132–153.
- Finn AS, Hudson Kam CL, Ettlinger M, Vytlačil J, D'Esposito M. Learning language with the wrong neural scaffolding: the cost of neural commitment to sounds. *Front Syst Neurosci.* 2013;7(November):85.
- Finn AS, Kharitonova M, Holtby N, Sheridan MA. Prefrontal and hippocampal structure predict statistical learning ability in early childhood. *J Cogn Neurosci.* 2018;31(1):126–137.
- Fiser J, Aslin RN. Statistical learning of higher-order temporal structure from visual shape sequences. *J Exp Psychol Learn Mem Cogn.* 2002;28(3):458–467.
- Folia V, Uddén J, De Vries M, Forkstam C, Petersson KM. Artificial language learning in adults and children. *Lang Learn.* 2010;60(s2):188–220.
- Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage.* 2011;54(1):313–327.
- Friederici AD. Pathways to language: fiber tracts in the human brain. *Trends Cogn Sci.* 2009;13(4):175–181.
- Frost R, Armstrong BC, Siegelman N, Christiansen MH. Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn Sci.* 2015;19(3):117–125 Elsevier Ltd.
- Gebhart A, Newport EL, Aslin RN. Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychon Bull Rev.* 2009;16(3):486–490.
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, et al. The minimal preprocessing pipelines for the human connectome project. *NeuroImage.* 2013;80:105–124.
- Gomez RL, Gerken L. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition.* 1999;70(2):109–135.
- Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Front Neuroinform.* 2011;5:13. <https://doi.org/10.3389/fninf.2011.00013>.
- Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data.* 2016;3(1):1–9.
- Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage.* 2009;48(1):63–72.
- Hagoort P. On Broca, brain, and binding: a new framework. *Trends Cogn Sci.* 2005;9(9):416–423.
- Heim S, Eickhoff SB, Friederici AD, Amunts K. Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Struct Funct.* 2009;213(4–5):441–456.
- Hu A, Kozloff V, Owen Van Horne A, Chugani D, Qi Z. Dissociation between linguistic and nonlinguistic statistical learning in children with autism. *J Autism Dev Disord.* 2023;1–16. <https://doi.org/10.1007/S10803-023-05902-1>.
- Isbilen ES, McCauley SM, Christiansen MH. Individual differences in artificial and natural language statistical learning. *Cognition.* 2022;225:105123. <https://doi.org/10.1016/J.COGNITION.2022.105123>.
- Ishkhanyan B, Michel Lange V, Boye K, Mogensen J, Karabanov A, Hartwigsen G, Siebner HR. Anterior and posterior left inferior frontal gyrus contribute to the implementation of grammatical determiners during language production. *Front Psychol.* 2020;11:685.
- Julian JB, Fedorenko E, Webster J, Kanwisher N. An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage.* 2012;60(4):2357–2364.
- Karuz EA, Newport EL, Aslin RN, Starling SJ, Tivarus ME, Bavelier D. The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 2013;127(1):46–54.
- Keller TA, Carpenter PA, Just MA. The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing. *Cereb Cortex.* 2001;11(3):223–237.
- Kirkham NZ, Slemmer JA, Johnson SP. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition.* 2002;83(2):B35–B42.
- Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, Rossa B, Reuter M, Chaibub Neto E, et al. Mindboggling morphometry of human brains. *PLoS Comput Biol.* 2017;13(2):e1005350.
- Klein A, Mensh B, Ghosh S, Tourville J, Hirsch J. Mindboggle: automated brain labeling with multiple atlases. *BMC Med Imaging.* 2005;5(1):1–14. <https://doi.org/10.1186/1471-2342-5-7>.
- Kuhl PK. Early language acquisition: cracking the speech code. In: *Nat Rev Neurosci.* 2004;5(11):831–843.
- Li P, Legault J, Litcofsky KA. Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex.* 2014;58:301–324.
- López-Barroso D, Ripollés P, Marco-Pallarés J, Mohammadi B, Münte TF, Bachoud-Lévi A-C, Rodríguez-Fornells A, de Diego-Balaguer R. Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *NeuroImage.* 2015;110(0):182–193. <https://doi.org/10.1016/j.neuroimage.2014.12.085>.
- Lupyan G, Clark A. Words and the world: predictive coding and the language-perception-cognition interface. *Curr Dir Psychol Sci.* 2015;24(4):279–284.
- Mahowald K, Fedorenko E. Reliable individual-level neural markers of high-level language processing: a necessary precursor for relating neural variability to behavioral and genetic variability. *NeuroImage.* 2016;139:74–93.
- Malik-Moraleda S, Ayyash D, Gallée J, Affourtit J, Hoffman M, Mineroff Z. An investigation across 45 languages and 12 language families reveals a universal language network. *Nat Neurosci.* 2022;25(25):1014–1019. <https://www.nature.com/articles/s41593-022-01114-5>.
- Maye J, Werker JF, Gerken LA. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition.* 2002;82(3):B101–B111.
- McCauley S, Christiansen M. Individual differences in chunking ability predict on-line sentence processing. *CogSci.* 2015: <https://csl-lab.psych.cornell.edu/files/2021/02/2015-mc-cogsci.pdf>.
- McNealy K, Mazziotta JC, Dapretto M. Cracking the language code: neural mechanisms underlying speech parsing. *J Neurosci.* 2006;26(29):7629–7639.
- McNealy K, Mazziotta JC, Dapretto M. Age and experience shape developmental changes in the neural basis of language-related learning. *Dev Sci.* 2011;14(6):1261–1282.
- Meyer T, Olson CR. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc Natl Acad Sci USA.* 2011;108(48):19401–19406.
- Misyak JB, Christiansen MH, Tomblin JB. On-line individual differences in statistical learning predict language processing. *Front Psychol.* 2010;1:31.

- Myers EB, Swan K. Effects of category learning on neural sensitivity to non-native phonetic categories. *J Cogn Neurosci*. 2012;24(8):1695–1708.
- Onnis L, Thiessen E. Language experience changes subsequent learning. *Cognition*. 2013;126(2):268–284.
- Orpella J, Assaneo MF, Ripollés P, Noeovich L, López-Barroso D, de Diego-Balaguer R, Poeppel D. Differential activation of a frontoparietal network explains population-level differences in statistical learning from speech. *PLoS Biol*. 2022;20(7):e3001712. <https://doi.org/10.1371/JOURNAL.PBIO.3001712>.
- Orpella J, Mas-Herrero E, Ripollés P, Marco-Pallarés J, de Diego-Balaguer R. Language statistical learning responds to reinforcement learning principles rooted in the striatum. *PLoS Biol*. 2021;19(9):e3001119. <https://doi.org/10.1371/JOURNAL.PBIO.3001119>.
- Parker DB, Razlighi QR. The benefit of slice timing correction in common fMRI preprocessing pipelines. *Front Neurosci*. 2019;13(JUL):821.
- Petersson KM, Folia V, Hagoort P. What artificial grammar learning reveals about the neurobiology of syntax. *Brain Lang*. 2012;120(2):83–95.
- Plante E, Patterson D, Sandoval M, Vance CJ, Asbjørnsen AE. An fMRI study of implicit language learning in developmental language impairment. *NeuroImage Clin*. 2017;14:277–285.
- Pliatsikas C, DeLuca V, Voits T. The many shades of bilingualism: language experiences modulate adaptations in brain structure. *Lang Learn*. 2020;70(S2):133–149.
- Poldrack RA. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*. 2011;72(5):692–697.
- Potter CE, Wang T, Saffran JR. Second language experience facilitates statistical learning of novel linguistic materials. *Cogn Sci*. 2017;41(S4):913–927.
- Poulin-Charronnat B, Perruchet P, Tillmann B, Peereman R. Familiar units prevail over statistical cues in word segmentation. *Psychol Res*. 2017;81(5):990–1003.
- Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. In *Annals of the New York Academy of Sciences* (Vol. 1191, pp. 62–88). Malden, MA: Blackwell Publishing Inc., 1.
- Qi Z, Legault J. Neural hemispheric organization in successful adult language learning: is left always right? *Psychol Learn Motiv*. 2020;72:119–163.
- Qi Z, Sanchez Araujo Y, Georgan WC, Gabrieli JD, Arciuli J. Hearing matters more than seeing: a cross-modality study of statistical learning and reading ability. *Sci Stud Read*. 2018;23(1):101–115.
- Reber PJ. The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia*. 2013;51(10):2026–2042.
- Saffran JR, Thiessen ED, Hoff E, Shatz M. Domain-general learning capacities. In: *Blackwell Handbook of Language Development*, Malden, MA; 2007. pp. 68–86.
- Saffran JR, Johnson EK, Aslin RN, Newport EL. Statistical learning of tone sequences by human infants and adults. *Cognition*. 1999;70(1):27–52.
- Saffran JR, Newport EL, Aslin RN. Word segmentation: the role of distributional cues. *J Mem Lang*. 1996;35(4):606–621.
- Schapiro AC, Kustner LV, Turk-Browne NB. Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr Biol*. 2012;22(17):1622–1627.
- Schneider JM, Hu A, Legault J, Qi Z. Measuring statistical learning across modalities and domains in school-aged children via an online platform and neuroimaging techniques. *J Vis Exp*. 2020;2020(160):1–21.
- Schneider JM, Weng YL, Hu A, Qi Z. Linking the neural basis of distributional statistical learning with transitional statistical learning: the paradox of attention. *Neuropsychologia*. 2022;172:108284.
- Scott TL, Gallée J, Fedorenko E. A new fun and robust version of an fMRI localizer for the frontotemporal language system. *Cogn Neurosci*. 2017;8(3):167–176.
- Scott-Van Zeeland AA, McNealy K, Wang AT, Sigman M, Bookheimer SY, Dapretto M. No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biol Psychiatry*. 2010;68(4):345–351.
- Siegelman N, Bogaerts L, Elazar A, Arciuli J, Frost R. Linguistic entrenchment: prior knowledge impacts statistical learning performance. *Cognition*. 2018;177:198–213.
- Siegelman N, Frost R. Statistical learning as an individual ability: theoretical perspectives and empirical evidence. *J Mem Lang*. 2015;81:105–120.
- Smalle EHM, Daikoku T, Szmalec A, Duyck W, Onen RM. Unlocking adults' implicit statistical learning by cognitive depletion. *Proc Natl Acad Sci USA*. 2022;119(2):e2026011119.
- Snijders TM, Vosse T, Kempen G, Van Berkum JJA, Petersson KM, Hagoort P. Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cereb Cortex* (New York, N.Y.: 1991). 2009;19(7):1493–1503.
- Stärk K, Kidd E, Frost R. The effect of children's prior knowledge and language abilities on their statistical learning. *Appl Psycholinguist*. 2022;43(5):1045–1071.
- Stärk K, Kidd E, Learning RF-L. Close encounters of the word kind: attested distributional information boosts statistical learning. *Wiley Online Library*. 2023;73(2):341–373.
- Szucs D, Ioannidis JP. Sample size evolution in neuroimaging research: an evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. *NeuroImage*. 2020;221:117164.
- Thompson SP, Newport EL. Statistical learning of syntax: the role of transitional probability. *Lang Learn Dev*. 2007;3(1):1–42.
- Toro JM, Trobalón JB. Statistical computations over a speech stream in a rodent. *Percept Psychophys*. 2005;67(5):867–875.
- Ventura-Campos N, Sanjuán A, González J, Palomar-García MÁ, Rodríguez-Pujadas A, Sebastián-Gallés N, Ávila C. Spontaneous brain activity predicts learning ability of foreign sounds. *J Neurosci*. 2013;33(22):9295–9305.
- Wong B, Yin B, O'Brien B. Neurolinguistics: structure, function, and connectivity in the bilingual brain. *Biomed Res Int*. 2016;2016:1–22.
- Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage*. 2001;14(6):1370–1386.
- Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging*. 2001;20(1):45–57.