The neural correlates of statistical learning in a word segmentation task: An fMRI study

Elisabeth A. Karuza[a], **Elissa L. Newport**[a,b], **Richard N. Aslin**[a,c], **Sarah J. Starling**[a], **Madalina E. Tivarus**[c,d], and **Daphne Bavelier**[a,c,e]

Elisabeth A. Karuza: ekaruza@bcs.rochester.edu; Elissa L. Newport: newport@bcs.rochester.edu; Richard N. Aslin: aslin@cvs.rochester.edu; Sarah J. Starling: sdavis@bcs.rochester.edu; Madalina E. Tivarus: mtivarus@rcbi.rochester.edu; Daphne Bavelier: daphne@cvs.rochester.edu

[a]Department of Brain and Cognitive Sciences, University of Rochester, Meliora Hall, Box 270268, Rochester, NY 14627, USA [b]Center for Brain Plasticity and Recovery, Georgetown University, Building D - 4000 Reservoir Road NW, Washington, DC 20007, USA [c]Rochester Center for Brain Imaging, University of Rochester, 430 Elmwood Ave., Medical Center Annex, Rochester, NY 14620, USA [d]Department of Imaging Sciences, University of Rochester, 110 Science Parkway, Rochester NY 14620, USA [e]FPSE, University of Geneva, Boulevard du Ponte d'Arve, 1211 Genève 4, Switzerland

Abstract

Functional magnetic resonance imaging (fMRI) was used to assess neural activation as participants learned to segment continuous streams of speech containing syllable sequences varying in their transitional probabilities. Speech streams were presented in four runs, each followed by a behavioral test to measure the extent of learning over time. Behavioral performance indicated that participants could discriminate statistically coherent sequences (words) from less coherent sequences (partwords). Individual rates of learning, defined as the difference in ratings for words and partwords, were used as predictors of neural activation to ask which brain areas showed activity associated with these measures. Results showed significant activity in the pars opercularis and pars triangularis regions of the left inferior frontal gyrus (LIFG). The relationship between these findings and prior work on the neural basis of statistical learning is discussed, and parallels to the frontal/subcortical network involved in other forms of implicit sequence learning are considered.

Keywords

fMRI; statistical learning; word segmentation; artificial language; sequence learning; Broca's area; LIFG

Corresponding author: Elisabeth A. Karuza, Department of Brain & Cognitive Sciences, Meliora Hall, RC 270268, University of Rochester, Rochester, NY 14627, USA, Phone: 1- 585-275-2508, Fax: 1-585-442-9216, ekaruza@bcs.rochester.edu.

1. Introduction

The extraction of patterns from our environment is a fundamental form of learning that enables us to develop rich representations of objects and events in our world. Sensitivity to statistical regularities in many domains is acquired by a process of distributional learning. Saffran, Newport and Aslin (1996) provided some of the first evidence of this *statistical learning* mechanism in the domain of language acquisition (see also Saffran, Aslin, & Newport, 1996). They demonstrated that infants and adults could segment word-like units from a stream of continuous speech by rapidly calculating the transitional probabilities between syllables. Statistical learning based on the predictability of patterns in sequences of elements has been shown across age groups (e.g., infants and adults), species (e.g. monkeys: Meyer & Olson, 2011; rats: Toro and Trobalón, 2005) and modalities (e.g., non-linguistic auditory: Gebhart, Newport & Aslin, 2009; Saffran, Johnson, Aslin & Newport, 1999; visual: Fiser & Aslin, 2002). In addition, a statistical learning approach to pattern extraction has been applied to levels of the linguistic hierarchy ranging from phoneme discrimination (e.g., Maye, Werker & Gerken, 2002) to syntax (Gomez & Gerken, 1999; Thompson & Newport, 2007). These experiments demonstrate that, in the absence of instructions or feedback, novel patterns embedded in temporally ordered elements can be extracted by a robust and domain-general statistical learning mechanism.

As noted by Perruchet and Pacton (2006), studies of statistical learning share several key properties with an older literature on artificial grammar learning (AGL). Classic AGL studies (e.g., Reber, 1967) used strings of letters presented simultaneously, with strings conforming to a finite-state grammar, while other studies used visual-motor sequences in a serial reaction time (SRT) task (e.g., Nissen & Bullemer, 1987). Both statistical learning and AGL studies involve implicit learning[1] of patterns that contain temporal order information (see Reber [1967] for an argument that AGL with explicit instructions may produce a different type of learning outcome). Both statistical learning and SRT studies present rapid sequences of elements, but SRT tasks assess learning through response speed while statistical learning and AGL rely on a post-exposure test of familiar vs. novel strings. Thus there are commonalities and differences that make comparisons among AGL, SRT, and statistical learning tasks of considerable interest.

In particular, there is a rich literature on the neural correlates of AGL. One of the most striking patterns that emerges across finite-state grammar processing and motor sequencing tasks is the involvement of the basal ganglia (for a review see Conway & Pisoni, 2008), particularly the caudate (Forkstam, Hagoort, Fernandez, Ingvar, & Petersson, 2006; Peigneux et al., 2000; Schendan, Searl, Melrose, & Stern, 2003; Simon, Vaidya, Howard, & Howard, 2012) and putamen (Grafton, Hazeltine, & Ivry, 1995; Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004; Schendan et al., 2003). Furthermore, both learning of artificial grammar strings and subsequent classification tasks (i.e., indicating whether a given test string is grammatical or ungrammatical) have been shown to involve prefrontal areas such as the left inferior frontal gyrus (Fletcher, Buchel, Josephs, Friston, & Dolan, 1999; Forkstam et al., 2006; Opitz & Friederici, 2003; Petersson, Folia, & Hagoort, 2012; Petersson, Forkstam, & Ingvar, 2004; Skosnik et al., 2002). Taken together, these findings suggest a modality-independent frontal/ basal ganglia circuit supporting non-declarative forms of learning (Ullman, 2004).

---

[1]In using the term *implicit learning*, we refer to a form of learning in which participants are able to extract structure from the stimuli presented to them but are unable to accurately verbalize exactly what rule or pattern they relied on in discriminating structured from unstructured test items (e.g., Reber, 1967; Turk-Browne, Scholl, Chun & Johnson, 2009). In the present study we have not tested directly whether learning is implicit, but our prior studies using the same paradigm suggest that this type of learning is largely implicit.

In contrast, there are only a handful of neuroimaging studies that have investigated statistical learning. These studies involve three potentially separate processes: (1) the storage of elements that occur during exposure, (2) the computation of one or more statistics from the element distributions, and (3) the recognition of statistically coherent (familiar) patterns after they have been learned. In many types of experimental designs using standard analyses, these processes can be difficult to distinguish. The goal of the present study is to provide insights into the second process: what are the neural substrates that mediate the computation of underlying structural information during the course of statistical learning? We chose to employ sequences of speech syllables because they have formed the basis of a substantial behavioral literature on statistical learning in the context of a word-segmentation task.

Results obtained from neuroimaging studies of statistical learning have been mixed, a fact potentially attributable to variation in the behavioral evidence of learning obtained during a scanning session. McNealy, Mazziotta & Dapretto (2006) observed increases in activation in left lateralized temporal cortices during exposure to continuous streams of speech varying in syllable-to-syllable transitional probabilities. However, they found that adult participants were unable to discriminate statistically coherent and less coherent items during a post-exposure testing phase. The authors proposed that the observed increases in neural activation were the signature of word segmentation before participants could demonstrate explicit awareness of the underlying structures. In contrast, Cunillera et al. (2009) conducted a joint ERP-fMRI study of auditory word segmentation and succeeded in obtaining statistically significant behavioral evidence of learning. Using a 2-alternative forced choice task, they found that participants could differentiate clusters of statistically coherent syllables from clusters of less coherent syllables. They saw increased activation during the exposure phase in bilateral posterior superior temporal gyrus and the superior part of the ventral premotor cortex (svPMC). Behavioral performance on the post-exposure test phase was significantly correlated with increases in activation in svPMC during the first two minutes of the exposure phase.

Examining statistical learning in the visual modality, Turk-Browne et al. (2009) offered additional support for the concept of learning without awareness (i.e., before discrimination). Across the entire exposure phase, they found that participants showed greater activation for statistically coherent relative to random shape sequences in an extensive network of areas including the striatum, medial temporal lobe, LOC, and ventral occipito-temporal cortex. A more fine-grained moving window analysis revealed differences in activation between these two conditions early on during familiarization. In analyzing the behavioral results of the posttest, they did not obtain evidence that participants could discriminate statistically coherent shape sequences from less coherent sequences (but see Turk-Browne, Scholl, Johnson & Chun, 2010 for evidence of statistical learning during a face/scene categorization task). However, performance during the first half of the test phase did show a significant learning effect. Moreover, participants' familiarity ratings of statistically coherent sequences during the test phase were shown to correlate with LIFG activation during the exposure phase, even though between-subject familiarity ratings did not differ between statistically coherent and less coherent test items.

The foregoing findings suggest some overlap in the brain areas involved in the computation of statistical regularities both within and across modalities. However, since the previous fMRI studies of segmentation show mixed behavioral evidence of statistical learning, it is challenging to compare across studies. The present experiment addresses gaps in our understanding of the neural substrates underlying statistical learning by first providing robust behavioral evidence of word segmentation. Furthermore, we assessed learning at multiple time points throughout the exposure phase, thereby providing a link between individual participants' rate of learning during the exposure phase and changes in neural

activation. Interestingly, most of the prior imaging work used increases in BOLD response within the first few minutes of exposure as an index of learning across subjects. The work of Abla, Katahira, & Okanoya (2008), on the other hand, demonstrated that participants showed considerable differences in the time course and extent of their learning. Participants were exposed to concatenated tone sequences and tested on their ability to differentiate statistically coherent and incoherent triplets. Despite a high overall accuracy score, participants could be divided into low, mid, and high learners. Event-related recordings revealed that, in the high learner group, the first member of a tonal triplet elicited the largest N100 and N400 response during the first exposure phase. In the medium learner group, this effect was found in the later phases, and for the low learner group no triplet onset effect was found. These findings clearly illustrate the importance of taking into account individual differences in learning. Correspondingly, we employed a method of analysis that could accommodate high variability across individual learning curves, assuming neither a monotonic increase across the exposure to the speech streams nor an early peak in activation. We took into account the possibility that some participants may learn quickly and then plateau or even regress in performance, while others may require more exposure before reaching peak performance. Importantly, this design allowed us to investigate the learning process as it unfolds, rather than the recognition process for items already acquired or the changes in neural responses simply due to repeated exposure to a set of stimuli. By first addressing the question of statistical learning in a word segmentation task, it will then be possible to compare our findings with the rich and developing neuroimaging literature on other forms of implicit learning. Our results support the hypothesis that regions involved in statistical learning during a word segmentation task share certain commonalities with neural circuits that are activated in other sequence learning tasks used in the AGL and SRT literatures.

2. Materials and Methods

2.1 Participants

A total of 34 undergraduates recruited from the University of Rochester completed the study. However, of that number, only 25 were included in the analyses presented here (17 female, mean age= 20.5 years, range=18–23). Two participants were excluded because they exceeded our cut-off for excess head motion (> 3.0mm in any direction), and 7 participants had to be excluded due to technical malfunctions that resulted in decreased exposure to the miniature artificial language. All participants gave written informed consent according to the protocol approved by the University of Rochester Research Subjects Review Board and were compensated $30. Participants were right-handed, native speakers of English without any known neurological problems or learning disabilities.

2.2 Stimuli and procedure

Participants completed alternating runs of exposure and testing. Prior to the start of the experiment, they were instructed to relax and listen attentively to the stimulus materials. They were also informed that they should expect tests on the familiarity of aspects of the language. In this way, we ensured that any learning that occurred during the first exposure phase would not differ fundamentally from learning during later exposure phases.

Auditory and visual materials were displayed using the presentation software DirectRT v. 2008 (Empirisoft Corporation, NY). Visual stimuli were projected onto a screen located in the rear of the magnet bore using a Titan model sx+ 3D projector (Digital Projection, Inc., GA). The screen was visible to the participants through a small mirror mounted above the eyes at an angle of 45° (viewing distance= 0.8m). Auditory stimuli were presented using pneumatic headphones (Resonance Technology Inc., CA). Participants wore earplugs to

reduce the ambient noise of the scanner and made responses using a custom-built MR-safe button box held in their right hand.

2.2.1 Exposure phase—Participants were exposed to a modified version of the miniature artificial language presented in Newport and Aslin (2004). This particular language was chosen due to the speed and robustness with which it was acquired, rendering it adaptable to a block design mode of presentation in the scanner. In previous behavioral work in a quiet environment outside of the scanner, Gebhart, Aslin and Newport (2009) showed evidence of learning this language after only 5 minutes of exposure. The speech stream contained six consonants (b, p, d, t, g, k) and 6 vowels (a, i, u, e, o, ae), which were combined to form 12 consonant-vowel syllables (pa, gi, tae, gu, te, po, da, ki, ku, bae, bu, do) and 16 trisyllabic words. An equal number of tokens of the 16 trisyllabic words were concatenated into a continuous stream with the constraint that two identical words could not occur in direct succession and that each syllable at the end of a word could only be followed by one of two different word-initial syllables. These constraints ensured that the transitional probabilities between *non-adjacent* phonetic segments (consonants) within a word were 1.0. In contrast, the transitional probabilities between *adjacent* phonetic segments (consonant   vowel or vowel   consonant) and syllables within a word were 0.5 and the transitional probabilities between non-adjacent phonetic segments and syllables across a word boundary were 0.5. The speech stream was synthesized using the female voice Victoria in MacinTalk© with a flat monotone setting so that the stream contained no pauses or prosodic cues to indicate word boundaries and were further edited using Sound Edit 16 version 2, in order to ensure that all syllable durations, both within and across words, were approximately the same.

The experiment consisted of 4 separate exposure phases each followed by a short testing phase. Before the first exposure phase, participants were instructed to relax and listen attentively to 3 different types of sound streams, each of which would be paired with a color cue indicating the type of stream being presented. The "languages" consisted of continuous streams of (1) forward speech, (2) backward speech formed by playing the recording of the forward speech stream in reverse, and (3) overlaid speech formed by layering three copies of the forward stream on top of one another with a slight temporal offset between them. Because we tested on forward and backward items, but not overlaid items, forward and backward speech streams afforded the most comparable task-relevant use of cognitive resources in this design. Therefore, for the purpose of this study, the analyses will focus on forward speech and its matched control condition of backward speech. Though the backward speech still technically contained statistical regularities, it was selected as an appropriate control because participants were largely unable to *extract* these regularities. Extensive behavioral pilot testing revealed that participants could not successfully compute the statistical relationships between syllables in the reversed stream as they did for the forward stream.

Participants were asked to listen passively during the 4 exposure phases. Stimuli during these phases were presented using a blocked design, with the ordering of the three language streams counterbalanced across participants within each exposure phase. Block duration was 30 seconds, with interleaved 15-second periods of silence paired with a black screen to serve as a baseline condition (Figure 1). Auditory fade-in and fade-out effects were applied to the beginnings and ends of each block to ensure that participants would not be able to determine the precise beginning and end of each type of stream and therefore could not make use of this potential cue to word boundaries. Within each of the 4 exposure phases, participants listened to a total of 2 minutes of each language.

2.2.2 Test phase—Following each of the 4 exposure runs, participants were tested on their recognition of forward and backward 'words' and 'partwords' that occurred in the

exposure streams. Words were defined as those 3-syllable combinations with high transitional probabilities between non-adjacent phonetic segments (consonants) within the word and low transitional probabilities between consonants, vowels, and syllables across word boundaries. Partwords were statistically less consistent groupings of 3 syllables that consisted of the end of one word and the beginning of another: one pair of consonants within the partword had a low transitional probability while the other pair of consonants had transitional probabilities of 1.0. In each test, participants were presented with 16 items (4 forward words, 4 forward partwords, 4 backward words, and 4 backward partwords, in random order) and were asked to rate each one on a scale of 1 (definitely unfamiliar) to 4 (definitely familiar). Participants had 2 seconds to indicate their response on the button box. Participants were tested on both forward and backward items in order to verify that they extracted word boundaries from forward speech but not from backward speech. If participants demonstrated no rating difference between backward words and partwords, then we could confidently use backward speech as a valid control for studying learning of our forward speech stream. Although we were only interested in collecting functional imaging data during the exposure phase, we continued to scan during each of the testing phases in order to maintain similar ambient noise conditions throughout the entire experiment.

2.3 MR acquisition parameters

Data were acquired on a Siemens Trio 3T MRI scanner equipped with an 8-channel head coil. To reduce head motion and cumulative head drift, foam padding was used to secure the head and neck of participants. Thirty T2* -weighted gradient echo (GE) echo-planar axial slices were collected in an interleaved order with a TR of 2000ms (TE= 30ms, flip angle= 90°, voxel size = 4.0mm^3, FOV=256mm). Data from 277 time points were collected for each of the 4 exposure runs. Data from the 4 post-exposure tests were not included in the analyses presented here for the reasons described in section 2.2.2. At the end of the session, a high-resolution T1-weighted whole brain volume was acquired using an MPRAGE sequence (TR= 2530ms, TE= 3.39ms, flip angle= 90°, voxel size= 1.3x1.0x1.0mm^3, FOV=256mm).

2.4 Behavioral data analysis

Both reaction times and rating scores were obtained for each of the 4 tests. Data were excluded from trials in which the participants exceeded the 2-second time window to indicate a response (2.2% of 1600 total trials). A three-way repeated measures ANOVA was performed in which language (forward or backward speech), test number, and test item type (word or partword) were entered as within-subjects factors. Next, we performed planned comparisons in order to determine the source of significant effects obtained when running the ANOVA. Because these comparisons were planned, they underwent Least Significant Difference adjustment rather than multiple comparisons correction (Keppel & Wickens, 2004).

2.5 MR data analysis

2.5.1 Preprocessing—Analyses were carried out using FEAT (fMRI Expert Analysis Tool), part of the FSL software package (FSL, version 5.98, FMRIB, Oxford, UK, www.fsl.ox.a-c.uk/fsl, see also Smith et al., 2004; Woolrich et al., 2009). The first 3 volumes of each functional run were discarded to avoid the effects of any start-up magnetization transients in the data. Motion correction was then applied to each run using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002). Data from two participants were excluded from the analyses due to excessive head motion (> 3.0mm in any direction). Additional preprocessing steps included: slice timing correction, skull-stripping using the BET tool (Smith, 2002), spatial smoothing using an isotropic 3D Gaussian kernel (full-

width-half-maximum = 5mm), grand mean-based intensity normalization of all volumes by the same factor, and non-linear high pass temporal filtering to reduce low frequency artifacts. Single-subject and group-level statistical analyses were then performed.

2.5.2 Whole-brain analysis—First level statistical analysis was carried out using FILM (FMRIB's Improved Linear Model). Each type of speech (forward, backward and overlaid speech) was entered as a separate explanatory variable (EV). For each individual run, parameter estimates for forward, backward, and overlaid speech relative to baseline as well as for contrasts of interest (e.g., forward>backward) were calculated. In order to combine data across runs for each participant, these coefficients were then input into a GLM in which the random effects variance was forced to zero using FLAME (FMRIB's Local Analysis of Mixed Effects) (Beckmann, Jenkinson, & Smith, 2003; Woolrich et al., 2004). Registration to high-resolution images and the MNI-152 structural template was performed using FLIRT (Jenkinson & Smith, 2001).

At second level, the following explanatory variables were then entered for each participant: (EV1) all four runs were assigned equal weight and (EV2) delta analysis: a demeaned learning score was input for each run based on the behavioral performance of each participant on the test phase following that run. The learning score was calculated as the change in forward rating [word rating-partword rating] from the previous run: = (forward word rating-forward partword rating)$_{TestN}$ - (forward word rating- forward partword rating)$_{TestN-1}$. The learning score for the first run was calculated as the change in forward word partword rating from 0, as there would be no difference in familiarity of words and partwords prior to any speech stream exposure. This delta EV was orthogonalized with respect to the mean response EV in order to capture the *additional* effects of learning on patterns of neural activity for the contrasts forward>backward and vice versa. Z-(Gaussianized T/F statistic) images were thresholded using Z >2.3 and a corrected cluster significance threshold of p = 0.05 (Worsley, Evans, Marrett, & Neelin, 1992).

The inclusion of the delta scores as an additional covariate reveals fluctuations in neural activity predicted by individual shifts in behavioral performance across the 4 test phases. We opted to use the *change* in word/partword difference, as opposed to the absolute difference, because the delta better captures the learning process. Consider a scenario in which a participant shows a small word-partword difference on test 1, and then large differences between word and partword ratings on both of the next two tests (tests 2 and 3). This pattern of scores would suggest that most of the learning occurred during the exposure phase preceding test 2, with little learning – simply maintenance of what had already been learned – between tests 2 and 3. In this example, the learning process that occurred between tests 1 and 2 differs from the form of processing that occurred between tests 2 and 3. Because the purpose of this study was to determine the brain areas supporting the formation, not the maintenance, of structural representations, we elected to use a measure of *change* in performance that would reflect this process.

After obtaining coefficients for each participant at second level, the neural response across participants was subsequently modeled using FLAME (FMRIB's Local Analysis of Mixed Effects) stage 1 and 2 (Beckmann et al., 2003; Woolrich et al., 2004). Z-statistic images for the mean activation were thresholded using clusters determined by Z>2.3 and a corrected cluster significance threshold of p = 0.05. Because the delta effects were slightly weaker, the maps based on these coefficients were thresholded using Z>1.9 and a corrected cluster significance threshold of p = 0.05 (Worsley et al., 1992).

2.5.3 Region of interest (ROI) analysis—Based on data from prior sequence learning studies (e.g., Forkstam et al., 2006; Schendan et al., 2003), as well as the basal ganglia

activation obtained in McNealy et al. (2006) and Turk-Browne et al. (2009), we hypothesized a pattern of neural activation encompassing specific subcortical areas. Therefore, separate analyses were conducted in which pre-threshold masks were applied to the group-level activation maps for the delta analysis (uncorrected, $p = 0.05$). The masks consisted of bilateral caudate and putamen defined anatomically by the MNI structural atlas. Their application served to constrain our analysis to specific areas for which we had a strong *a priori* hypothesis of activation relevant to statistical learning.

3. Results and Discussion

3.1 Behavioral results

A three-way repeated measures ANOVA revealed a significant main effect of the test item type (word or partword) on the rating of triplet sequences ($F(1,24)=17.89$, $p<0.0001$, partial eta squared=0.43). Moreover, we found a significant interaction between the language (that is, the direction of the speech stream, forward or backward) and test item type ($F(1,24)=9.25$, $p<0.01$, partial eta squared=0.28). Planned comparisons were then carried out to determine the source of these effects, with the results indicating that for the forward stream, words were rated significantly higher than partwords for each of the four tests (test 1: mean word=2.73, mean partword=2.34, $F(1,24)= 5.25$, $p=0.03$, partial eta squared=0.18; test 2: mean word=2.99, mean partword=2.54, $F(1,24)=5.70$, $p=0.03$, partial eta squared=0.19; test 3: mean word=3.14, mean partword=2.36, $F(1,24)=18.33$ $p<0.001$, partial eta squared=0.43; test 4: mean word=2.95, mean partword=2.49, $F(1,24)=6.37$ $p=0.02$, partial eta squared=0.21). Figure 2A illustrates the mean rating differences for forward stream words and partwords. When planned comparisons were performed on *backward* items, however, no significant differences were found on three of the four tests (test 1: mean word=2.50, mean partword=$F(1,24)=0.57$, $p=0.46$, partial eta squared=0.02; test 2: mean word=2.60, mean partword=2.43, $F(1,24)=2.94$, $p=0.10$, partial eta squared=0.11; test 4: mean word=2.49, mean partword=2.42, $F(1,24)=0.63$, $p=0.44$, partial eta squared=0.03). As Figure 2B shows, a significant difference was found for the backward stream words compared to partwords for test 3 only (mean word=2.63, mean partword=2.37, $F(1,24)=5.65$, $p=0.03$, partial eta squared=0.19).

Although any learning from the backward stream was unexpected given our pilot results, previous work has shown that adult participants can extract certain types of regularities from non-linguistic auditory stimuli given sufficient exposure (Gebhart, Newport, et al., 2009). Note that despite showing some inconsistent sensitivity to transitional probabilities in backward speech for test 3, the rating difference between words and partwords was greater overall for forward sequences compared to backward sequences (mean forward difference = 0.52, SD = 0.68; mean backward difference = 0.10, SD = 0.21; $t(24) = 3.04$, $p < 0.01$) (Figure 2C). That is, participants may have extracted some statistical information from the backward stream, but, as revealed by the three-way ANOVA and planned comparisons, they learned significantly and substantially *better* during exposure to the forward stream. While there is extensive behavioral evidence that statistical regularities drive the acquisition of structure across domains, statistical learning is not necessarily equivalent for all types of stimuli. For example, Gebhart et al. found that, in order for successful segmentation to take place on a stream of complex and unfamiliar auditory noises, participants had to undergo a period of exposure that was 5 times longer than the exposure necessary for the segmentation of speech or music. The reduced ability of participants to segment reversed speech in the present study does not preclude the possibility that, given enough exposure, they would eventually be able to do so. Our results *do* indicate, however, that this reduced ability makes backward speech a suitable control for the rapid extraction of statistics during exposure to forward syllable streams.

3.2 Whole brain analyses

3.2.1 Forward speech streams compared to control—Before taking temporal changes and measures of learning into account, a whole brain analysis was performed in order to investigate mean differences between the forward and backward conditions. As described in section 2.5.2, activation maps across participants were created for the contrasts forward > backward and backward > forward. For the purposes of this particular analysis, participants were not divided based on the extent of their learning during the four testing points. The goal here was to determine which areas showed overall differences in activation during the presentation of forward streams relative to our control streams and to evaluate our findings with respect to prior work making use of similar contrasts (e.g., structured v. unstructured syllable streams).

Similar to previous studies contrasting statistically coherent and randomized syllable streams (Cunillera et al., 2009; McNealy et al., 2006), we obtained greater activation in the left superior temporal gyrus for the forward speech stream compared to our backward control (Table 1)[2]. This left-lateralized cluster in temporal cortex extended to portions of the middle temporal gyrus, the postcentral gyrus, and the supramarginal gyrus (Figure 4A). Some homologous regions in right temporal cortex were found as well, including the posterior portions of the middle temporal and superior temporal gyri. The right-lateralized cluster had a smaller spatial extent than its left hemisphere counterpart.

We also observed one cluster of activation that was greater in response to backward speech than forward speech. Aligning with prior work suggesting the superiority of the right hemisphere for processing of reversed speech (e.g., Kimura & Folb, 1968), this area encompassed the right angular gyrus and posterior division of the supramarginal gyrus. In addition, this right lateralized cluster overlaps in part with temporoparietal areas showing increased activation in response to unexpected or higher entropy sequences during the processing of patterned tone stimuli (Furl et al., 2011; Overath et al., 2007). While the left hemisphere areas showing greater response for forward relative to backward speech may support the processing of familiar, more predictable stimuli, the right hemisphere areas obtained with the opposite contrast may reflect neural response to unexpected, high entropy, or unusual stimuli.

3.2.2 Delta analysis: Neural correlates of learning-related changes—Though the basic comparison of forward speech streams and backward control streams was informative, it revealed little more than that there are differences in some aspect of the processing of forward and reversed streams of speech. This contrast alone cannot capture the neural substrate(s) underlying the learning process because it does not provide a link between behavioral measures of learning and fluctuations in neural activity. Assigning equal weights to each of the four exposure runs (as described in analysis 3.2.1) required the following assumptions: (1) each participant learned to the same extent in each run and (2) each participant's rate of learning was consistent across runs. However, Figure 3 illustrates the considerable amount of variability in behavioral performance between participants, shown here as the change in forward word-partword rating from one test to the next. While some participants showed evidence of learning very quickly, others did not show a jump in word-

---

[2]In addition, we obtained a significant cluster in posterior regions of the occipital lobe. This pattern seems best accounted for by the fact that we paired each stream with a different color cue. We did so to assist participants in clearly differentiating the stream types. Indeed, previous findings demonstrate that explicitly indicating to participants that they are listening to different languages facilitates learning (Gebhart, Aslin, et al., 2009). Therefore, we made use of color cues in order to obtain strong behavioral evidence of learning as exposure proceeded in our participants. These color cues were presented as a whole-field background, producing large changes in spectral intensity and luminance at the start of each stimulus condition. While we cannot discount a participation of occipital cortex in the auditory learning under study here, it is well accepted that such large visual transients would lead to the robust occipital activation reported.

partword rating difference until the third or fourth exposure phase. These results highlight the importance of taking into account individual differences during the entirety of the learning process.

The pattern of activation seen in Figure 4B was the result of relating each participant's change in learning across the four post-exposure test phases to changes in neural activity during the forward as compared to the backward exposure phase. This delta analysis resulted in a cluster of activation localized in the left inferior frontal gyrus (LIFG) pars triangularis and a small portion of the pars opercularis (peak activation at MNI coordinates x=−52, y=26, z=−6). The same delta analysis applied to the reverse contrast, backward speech> forward speech, resulted in no significant activation.

3.2.3 The role of the LIFG—Broca's area, or the posterior portion of left inferior frontal gyrus (BA 44/45), has traditionally been viewed as having a specifically linguistic function, controlling speech production or serving as the seat of syntax (e.g., Broca, 1861; Geschwind, 1970; Sakai, Hashimoto, & Homae, 2001; Santi & Grodzinsky, 2007a,b). However, recent hypotheses about the functions of the LIFG hold that it may play a more general role in the processing of linguistic and also certain types of non-linguistic materials (e.g., Hugdhal, Thomsen, & Ersland, 2006; Tillman, Janata & Bharucha, 2003). From a cognitive control perspective, a general regulatory function has been attributed to the LIFG, supporting the resolution of competing representations in phonological processing, sentence processing, and semantic retrieval (Moss et al., 2005; Novick, Trueswell & Thompson-Schill, 2005; 2010). Taking a similarly integrative approach, Petersson et al. (2012) proposed that "the left inferior frontal region is a generic on-line structured sequence processor that unifies information from various sources in an incremental and recursive manner" (p. 85). Our findings extend such interpretations of the LIFG and suggest that it functions not only as a sequence processor, but also as a substrate that drives sequence learning through the computation of statistical regularities and the formation of structural representations.

While the data obtained here can speak only to the role of the LIFG in an auditory linguistic task, neuroimaging results obtained from studies of visual and auditory non-linguistic processing raise the possibility that this area subserves domain-general statistical processing. Turk-Browne et al. (2009) observed that familiarity ratings of shape sequences were correlated with activation in LIFG. Moreover, Abla and Okanoya (2008) found a similar relationship between the segmentation of continuous tone sequences and activity in inferior frontal cortex. Participants were first trained on isolated tone triplets. Next, these statistically coherent triplets were concatenated in a continuous stream and presented in alternation with random tone sequences. Multichannel near-infrared spectroscopy recordings revealed greater changes localized near Broca's area in oxy-hemoglobin response for the structured relative to the unstructured condition. Given that similar results were obtained by Turk-Browne et al. and Alba and Okanoya in both visual and auditory segmentation studies, it is possible that LIFG operates in a modality-independent fashion. It may play a key role in the integration of sequential information, regardless of the modality of presentation. This hypothesis also fits with accounts of the neural circuits involved in implicit learning, described by many investigators as extending from portions of the basal ganglia to prefrontal areas.

3.3 Word segmentation and sequence learning circuitry

3.3.1 ROI analysis: Caudate and putamen involvement in segmentation—While a whole-brain delta analysis resulted in a cluster localized to the LIFG, a more sensitive measure was needed in order to examine the possibility of basal ganglia involvement during

word segmentation. When separate pre-threshold striatal masks were applied to the group-level maps obtained from the delta analysis, activation was also revealed in bilateral caudate (peak MNI coordinates for RH: x=10, y=8, z=18, max Z= 2.300; LH: x=−20, y=16, z=10, max Z=2.733) and putamen (peak coordinates for RH: x=30, y=16, z=−2, max Z=2.605; LH: x=−20, y=16, z=10, max Z=2.733), indicating a subcortical component in addition to the prefrontal region obtained in the whole-brain delta analysis. This result suggests a potential circuit of areas underlying statistical learning and complements connectivity data showing that prefrontal cortex actually mediates activity across multiple learning and memory systems (e.g., Poldrack & Rodriguez, 2004).

3.3.2 Neural basis of modality-independent sequence learning—Given the commonalities between statistical learning tasks and other tasks that involve high demands on memory for sequential information, there is reason to consider substantial overlap in the neural structures supporting them. Ullman (2001, 2004) has posited a distinction in language between a declarative memory system involving medial temporal lobe structures and the so-called procedural memory system involving frontal, subcortical, parietal and cerebellar areas. Ullman et al. (1997) proposed that the declarative system, which is responsible for fast fact learning, is necessary for the formation of the mental lexicon. On the other end of the spectrum, the procedural system underlies domain-general cognitive abilities that cannot be accessed consciously or described explicitly. Procedural learning encompasses non-declarative skills (e.g., swinging a golf club), but it seems to be specialized for the real-time, incremental processing of rules, especially those governing relationships between any sort of sequentially presented stimuli (e.g., grammar learning or serial reaction time tasks). As previously stated, this form of learning corresponds closely to the sort of implicit processing driving statistical learning, and, indeed, the current study of word segmentation demonstrates patterns of activation that map onto frontal and subcortical structures within the proposed procedural network. Anatomical data lend support to such a network, as there are extensive fiber tracts connecting prefrontal cortex and the basal ganglia. Though many tracts project *into* the basal ganglia, the primary *output* of the basal ganglia first passes through the thalamus and then projects to prefrontal cortex (e.g., Alexander, Delong & Strick, 1986). This result also fits with the monkey physiology work of Pasupathy and Miller (2005), which showed that activity in basal ganglia was related to rapid associative learning, while slow improvements in performance over time were correlated with activity in prefrontal areas.

3.4 Conclusions

The present study has provided several advances in our understanding of the neural substrates underlying statistical learning. By presenting speech streams whose structure could be learned rapidly, even in the scanner environment, we obtained significant behavioral evidence that participants made use of statistical information in order to perform our segmentation task. We then capitalized on individual variations in the time-course of learning by assessing behavioral performance at multiple time points (after each of several repeated exposure phases). These assessments of learning enabled us to search for those regions in which neural activation was associated with *changes* in performance over the experiment, rather than examining activation in relationship to only a single outcome measure of learning after all exposure phases were completed. Thus we were able to focus on the learning process itself rather than the outcome of learning or the resulting recognition of familiar element sequences. The neuroimaging results from these whole brain analyses suggested that specific subdivisions of the left inferior frontal gyrus mediate, at least in part, statistical learning at the level of individual participants. Moreover, a region of interest analysis implicated a subcortical component encompassing areas of the basal ganglia that participate in this learning process as well.

Our findings, when taken together with prior work on the neural correlates of statistical learning, suggest that these regions, particularly the LIFG, are involved in extracting temporally ordered pattern information regardless of the modality of the input. However, because this study involves only spoken language stimuli, further neuroimaging work is needed to confirm this domain-general hypothesis. Finally, given that much of the work on statistical and other forms of implicit learning has involved the processing of sequences, additional study is needed in order to determine whether the LIFG and basal ganglia also mediate learning of element relationships that are spatially, rather than temporally, organized (e.g., Fiser & Aslin, 2005).

Acknowledgments

References

Abla D, Okanoya K. Statistical segmentation of tone sequences activates the left inferior frontal cortex: a near-infrared spectroscopy study. Neuropsychologia. 2008; 46:2787–2795. [PubMed: 18579166]

Abla D, Katahira K, Okanoya K. Online assessment of statistical learning by eventrelated potentials. Journal of Cognitive Neuroscience. 2008; 20(6):952–964. [PubMed: 18211232]

Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience. 1986; 9:357–381.

Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in fMRI. Neuroimage. 2003; 20(2):1052–1063. [PubMed: 14568475]

Broca PP. Perte de la parole, ramollissement chronique, et destruction partielle du lobe antérieur gauche du cerveau. Bulletin de la Société Anthropologique. 1861; 2:235–238.

Conway CM, Pisoni DB. Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. Annals of the New York Academy of Sciences. 2008; 1145:113–131. [PubMed: 19076393]

Cunillera T, Camara E, Toro JM, Marco-Pallares J, Sebastian-Galles N, Ortiz H, Rodriguez-Fornells A. Time course and functional neuroanatomy of speech segmentation in adults. Neuroimage. 2009; 48(3):541–553. [PubMed: 19580874]

Fiser J, Aslin RN. Encoding multi-element scenes: Statistical learning of visual feature hierarchies. Journal of Experimental Psychology: General. 2005; 134:521–537. [PubMed: 16316289]

Fiser J, Aslin RN. Statistical learning of higher-order temporal structure from visual shape sequences. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28(3):458–467.

Fletcher P, Buchel C, Josephs O, Friston K, Dolan R. Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. Cerebral Cortex. 1999; 9(2):168–178. [PubMed: 10220229]

Forkstam C, Hagoort P, Fernandez G, Ingvar M, Petersson KM. Neural correlates of artificial syntactic structure classification. Neuroimage. 2006; 32(2):956–967. [PubMed: 16757182]

Furl N, Kumar S, Alter K, Durrant S, Shawe-Taylor J, Griffiths T. Neural prediction of higher-order auditory sequence statistics. Neuroimage. 2011; 54(3):2267–2277. [PubMed: 20970510]

Gebhart AL, Aslin RN, Newport EL. Changing structures in midstream: Learning along the statistical garden path. Cognitive Science. 2009; 33(6):1087–1116. [PubMed: 20574548]

Gebhart AL, Newport EL, Aslin RN. Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. Psychonomic Bulletin & Review. 2009; 16(3):486–490. [PubMed: 19451373]

Geschwind N. The organization of language and the brain. Science. 1970; 170(3961):940–944. [PubMed: 5475022]

Gomez RL, Gerken L. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition. 1999; 70(2):109–135. [PubMed: 10349760]

Grafton ST, Hazeltine E, Ivry R. Functional mapping of sequence learning in normal humans. Journal of Cognitive Neuroscience. 1995; 7(4):497–510. [PubMed: 23961907]

Hugdahl K, Thomsen T, Ersland L. Sex differences in visuo-spatial processing: An fmri study of mental rotation. Neuropsychologia. 2006; 44(9):1575–1583. [PubMed: 16678867]

Jenkinson M, Smith S. A global optimization method for robust affine registration of brain images. Medical Image Analysis. 2001; 5(2):143–156. [PubMed: 11516708]

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17(2):825–841. [PubMed: 12377157]

Keppel, G.; Wickens, TD. Design and analysis: A researcher's handbook. Upper Saddle River, NJ: Pearson Prentice Hall; 2004.

Kimura D, Folb S. Neural processing of backwards-speech sounds. Science. 1968; 161(3839):395–396. [PubMed: 5667143]

Lieberman MD, Chang GY, Chiao J, Bookheimer SY, Knowlton BJ. An event-related fmri study of artificial grammar learning in a balanced chunk strength design. Journal of Cognitive Neuroscience. 2004; 16(3):427–438. [PubMed: 15072678]

Maye J, Werker JF, Gerken L. Infant sensitivity to distributional information can affect phonetic discrimination. Cognition. 2002; 82(3):B101–B111. [PubMed: 11747867]

McNealy K, Mazziotta JC, Dapretto M. Cracking the language code: Neural mechanisms underlying speech parsing. Journal of Neuroscience. 2006; 26(29):7629–7639. [PubMed: 16855090]

Meyer T, Olson CR. Statistical learning of visual transitions in monkey inferotemporal cortex. Proceedings of the National Academy of Sciences. 2011; 108(48):19401–19406.

Moss HE, Abdallah S, Fletcher P, Bright P, Pilgrim L, Acres K, Tyler LK. Selecting among competing alternatives: Selection and retrieval in the left inferior frontal gyrus. Cerebral Cortex. 2005; 15(11):1723–1735. [PubMed: 15728742]

Newport EL, Aslin RN. Learning at a distance i. Statistical learning of nonadjacent dependencies. Cognitive Psychology. 2004; 48(2):127–162. [PubMed: 14732409]

Nissen MJ, Bullemer P. Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology. 1987; 19(1):1–32.

Novick JM, Trueswell JC, Thompson-Schill SL. Cognitive control and parsing: Reexamining the role of broca's area in sentence comprehension. Cognitive, Affective, & Behavioral Neuroscience. 2005; 5(3):263–281.

Novick JM, Trueswell JC, Thompson-Schill SL. Broca's area and language processing: Evidence for the cognitive control connection. Language and Linguistics Compass. 2010; 4(10):906–924.

Opitz B, Friederici AD. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. Neuroimage. 2003; 19(4):1730–1737. [PubMed: 12948727]

Overath T, Cusack R, Kumar S, Warren JD, von Kriegstein K, Grube M, Carlyon RP, Griffiths TD. An information theoretic characterisation of auditory encoding. PLoS Biology. 2007; 5(11):e288. [PubMed: 17958472]

Pasupathy A, Miller EK. Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature. 2005; 433(7028):873–876. [PubMed: 15729344]

Peigneux P, Maquet P, Meulemans T, Destrebecqz A, Laureys S, Degueldre C, Cleeremans A. Striatum forever, despite sequence learning variability: A random effect analysis of PET data. Human Brain Mapping. 2000; 10(4):179–194. [PubMed: 10949055]

Perruchet P, Pacton S. Implicit learning and statistical learning: One phenomenon, two approaches. Trends in Cognitive Sciences. 2006; 10(5):233–238. [PubMed: 16616590]

Petersson K-M, Folia V, Hagoort P. What artificial grammar learning reveals about the neurobiology of syntax. Brain and Language. 2012; 120(2):83–95. [PubMed: 20943261]

Petersson KM, Forkstam C, Ingvar M. Artificial syntactic violations activate Broca's region. Cognitive Science. 2004; 28(3):383–407.

Poldrack RA, Rodriguez P. How do memory systems interact? Evidence from human classification learning. Neurobiology of Learning and Memory. 2004; 82(3):324–332. [PubMed: 15464413]

Reber AS. Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior. 1967; 6:855–863.

Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science. 1996a; 274:1926–1928. [PubMed: 8943209]

Saffran JR, Newport EL, Aslin RN. Word segmentation: The role of distributional cues. Journal of Memory and Language. 1996b; 35(4):606–621.

Saffran JR, Johnson EK, Aslin RN, Newport EL. Statistical learning of tonal sequences by human infants and adults. Cognition. 1999; 70:27–52. [PubMed: 10193055]

Sakai KL, Hashimoto R, Homae F. Sentence processing in the cerebral cortex. Neuroscience Research. 2001; 39:1–10. [PubMed: 11164248]

Santi A, Grodzinsky Y. Taxing working memory with syntax: Bihemispheric modulations. Human Brain Mapping. 2007a; 28(11):1089–1097. [PubMed: 17133392]

Santi A, Grodzinsky Y. Working memory and syntax interact in Broca's area. Neuroimage. 2007b; 37(1):8–17. [PubMed: 17560794]

Schendan HE, Searl MM, Melrose RJ, Stern CE. An fmri study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron. 2003; 37(6):1013–1025. [PubMed: 12670429]

Simon JR, Vaidya CJ, Howard JH, Howard DV. The effects of aging on the neural basis of implicit associative learning in a probabilistic triplets learning task. Journal of Cognitive Neuroscience. 2012; 24(2):451–463. [PubMed: 21861675]

Skosnik PD, Mirza F, Gitelman DR, Parrish TB, Mesulam MM, Reber PJ. Neural correlates of artificial grammar learning. Neuroimage. 2002; 17(3):1306–1314. [PubMed: 12414270]

Smith SM. Fast robust automated brain extraction. Human Brain Mapping. 2002; 17(3):143–155. [PubMed: 12391568]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004; 23(Suppl 1):S208–S219. [PubMed: 15501092]

Thompson SP, Newport EL. Statistical learning of syntax: The role of transitional probability. Language Learning and Development. 2007; 3(1):1–42.

Tillmann B, Janata P, Bharucha JJ. Activation of the inferior frontal cortex in musical priming. Cognitive Brain Research. 2003; 16(2):145–161. [PubMed: 12668222]

Toro JM, Trobalon JB. Statistical computations over a speech stream in a rodent. Perception and Psychophysics. 2005; 67(5):867–875. [PubMed: 16334058]

Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK. Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. Journal of Cognitive Neuroscience. 2009; 21(10):1934–1945. [PubMed: 18823241]

Turk-Browne NB, Scholl BJ, Johnson MK, Chun MM. Implicit perceptual anticipation triggered by statistical learning. Journal of Neuroscience. 2010:11177–11187. [PubMed: 20720125]

Ullman MT, Corkin S, Coppola M, Hickok G, Growdon JH, Koroshetz WJ, Pinker S. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. Journal of Cognitive Neuroscience. 1997; 9(2):266–276. [PubMed: 23962016]

Ullman MT. A neurocognitive perspective on language: The declarative/procedural model. Nature Reviews Neuroscience. 2001; 2(10):717–726.

Ullman MT. Contributions of memory circuits to language: The declarative/procedural model. Cognition. 2004; 92(1–2):231–270. [PubMed: 15037131]

Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, Smith SM. Multilevel linear modelling for fmri group analysis using Bayesian inference. Neuroimage. 2004; 21(4):1732–1747. [PubMed: 15050594]

Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Smith SM. Bayesian analysis of neuroimaging data in FSL. Neuroimage. 2009; 45(1 Suppl):S173–S186. [PubMed: 19059349]

Worsley KJ, Evans AC, Marrett S, Neelin P. A three-dimensional statistical analysis for cbf activation studies in human brain. Journal of Cerebral Blood Flow & Metabolism. 1992; 12(6):900–918. [PubMed: 1400644]

Highlights

- fMRI was used to study the neural bases of statistical learning during word segmentation

- Participants acquired statistical regularities in a continuous syllable stream

- Analyses related learning at multiple time points to changes in neural activity

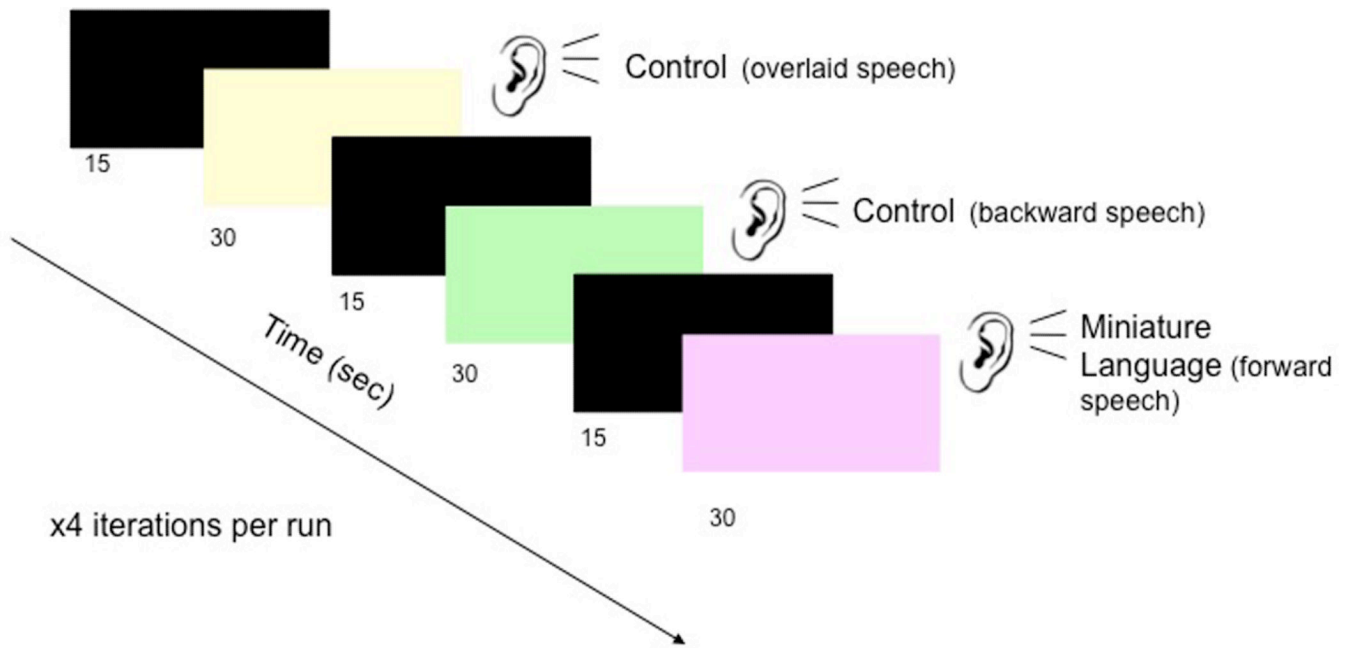- Results showed the central involvement of the LIFG in this type of learning

Figure 1.
Exposure phase presented in a blocked design. In each of 4 runs, participants listened to 2 minutes each of forward, backward and overlaid speech streams. Each type of speech stream was paired with a unique color cue. 30-second blocks of language exposure were interleaved with 15-second periods of silence. The total duration of each run was approximately 8 minutes and 30 seconds.
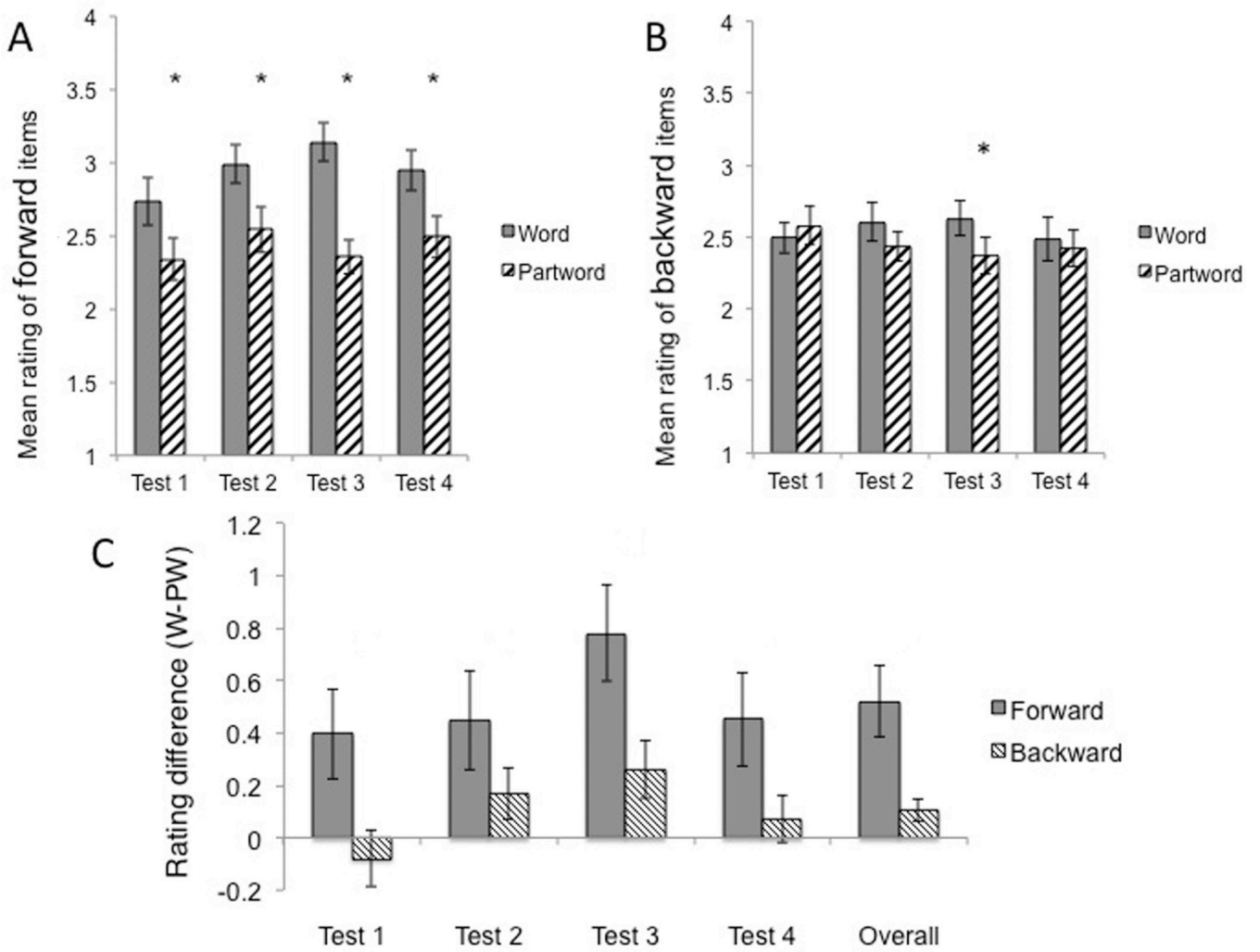
Figure 2.
Differences in word and partword rating for test items presented as forward speech and as backward speech. (A) Forward words were rated as significantly more familiar than forward partwords on each test. (B) Backward words were rated significantly higher than backward partwords only on test 3. (C) The rating difference between words and partwords presented in forward speech was significantly higher than the rating difference for items presented in backward speech.
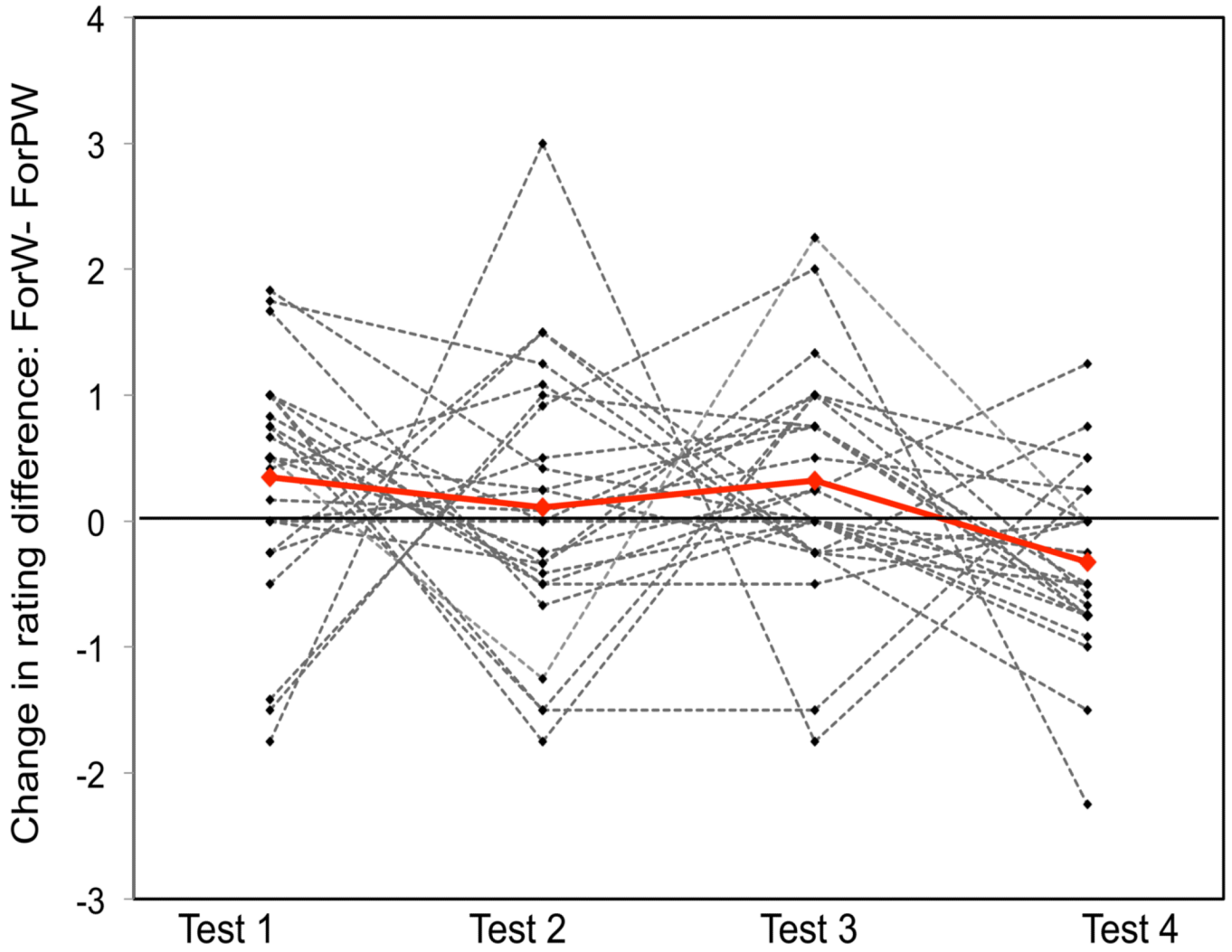
Figure 3.
Variability in learning. On each test, participants' delta-learning scores across tests spanned a wide range (measured as change in forward W-PW rating for each test as compared to the test preceding it). The mean change in rating for each test is indicated in red.
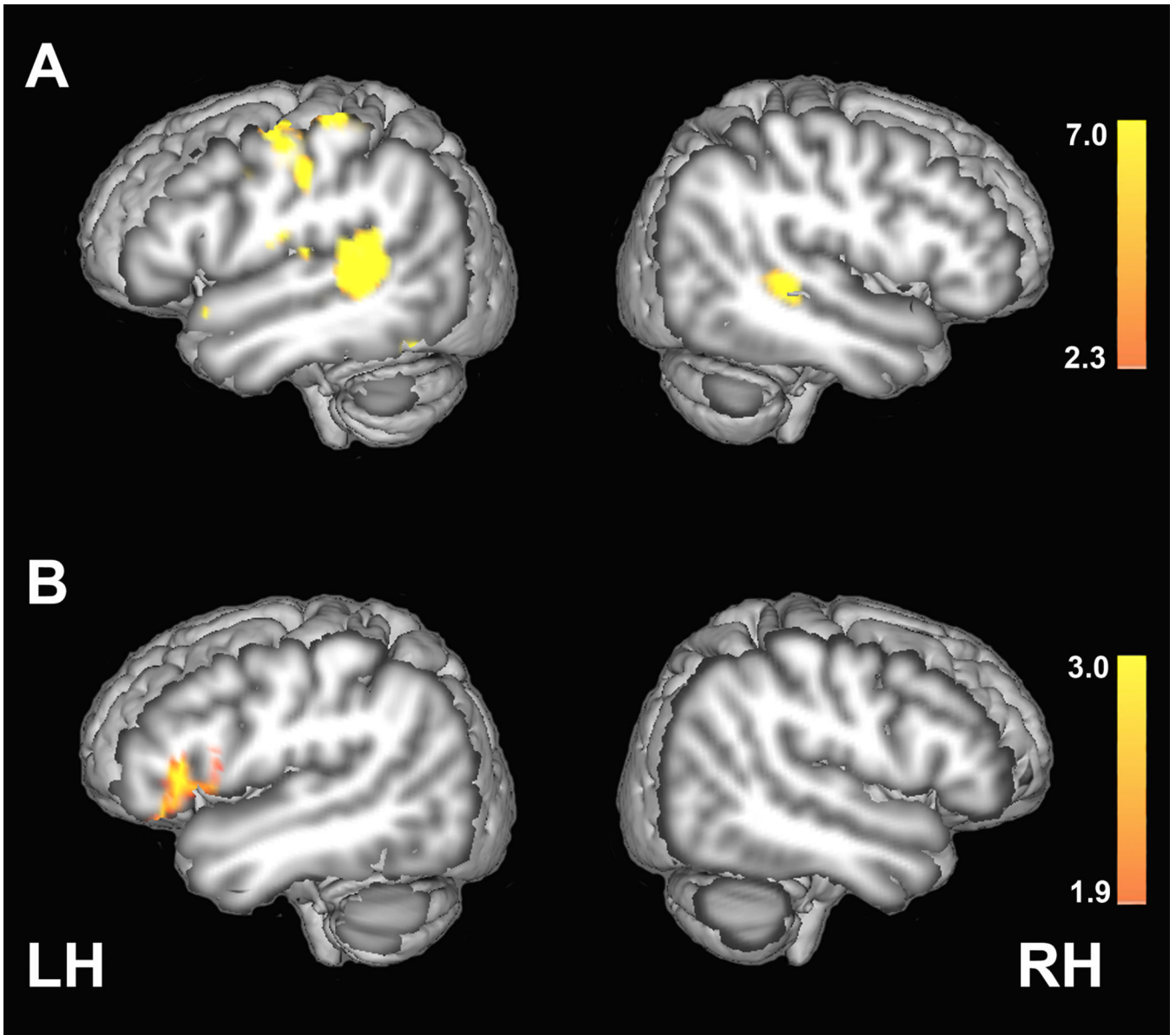
Figure 4.
Results of whole brain analysis with sagittal sections removed (A) Mean activation in response to forward > backward continuous speech streams (Z>2.3, p<0.05) (B) Neural changes related specifically to learning in the left inferior frontal gyrus as revealed by a delta analysis performed on the contrast forward >backward (Z>1.9, p<0.05).

Table 1

Location of activation peaks in MNI coordinates for mean differences between conditions: forward speech > backward speech and backward speech> forward speech

Contrast	Cluster	Extent (voxels)	Region	x	y	z	Z max
Forward > backward							
	1	19382					
			R Intracalcarine cortex	6	−86	6	6.68
			R Occipital pole	10	−92	0	6.22
			R Lingual gyrus	8	−78	−4	6.06
	2	4085					
			L Superior temporal gyrus, posterior division	−64	−12	4	4.96
			L Middle temporal gyrus, temporo-occipital part	−46	−46	8	4.32
			L Postcentral gyrus	−62	−12	42	4.24
			L Supramarginal gyrus, posterior division	−52	−48	10	4.17
	3	450					
			R Planum temporale	58	−34	14	3.65
			R Middle temporal gyrus, posterior division	64	−20	−4	3.54
			R Superior temporal gyrus, posterior division	48	−34	2	3.43
Backward > forward							
	1	812					
			R Angular gyrus	62	−48	44	3.85
			R Supramarginal gyrus, posterior division	62	−36	52	3.66