Fronto-Parietal Contributions to Phonological Processes in Successful Artificial Grammar Learning

Dariya Goranskaya[1]*, Jens Kreitewolf[1,2], Jutta L. Mueller[1,3], Angela D. Friederici[1] and Gesa Hartwigsen[1]

[1] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2] International Laboratory for Brain, Music and Sound Research, Department of Psychology, University of Montreal, Montreal, QC, Canada, [3] Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany

Sensitivity to regularities plays a crucial role in the acquisition of various linguistic features from spoken language input. Artificial grammar learning paradigms explore pattern recognition abilities in a set of structured sequences (i.e., of syllables or letters). In the present study, we investigated the functional underpinnings of learning phonological regularities in auditorily presented syllable sequences. While previous neuroimaging studies either focused on functional differences between the processing of correct vs. incorrect sequences or between different levels of sequence complexity, here the focus is on the neural foundation of the actual learning success. During functional magnetic resonance imaging (fMRI), participants were exposed to a set of syllable sequences with an underlying phonological rule system, known to ensure performance differences between participants. We expected that successful learning and rule application would require phonological segmentation and phoneme comparison. As an outcome of four alternating learning and test fMRI sessions, participants split into successful learners and non-learners. Relative to non-learners, successful learners showed increased task-related activity in a fronto-parietal network of brain areas encompassing the left lateral premotor cortex as well as bilateral superior and inferior parietal cortices during both learning and rule application. These areas were previously associated with phonological segmentation, phoneme comparison, and verbal working memory. Based on these activity patterns and the phonological strategies for rule acquisition and application, we argue that successful learning and processing of complex phonological rules in our paradigm is mediated via a fronto-parietal network for phonological processes.

Keywords: functional magnetic resonance imaging, premotor cortex, parietal cortex, phonological processes, phonological segmentation, phoneme comparison, auditory sequence processing, learning

INTRODUCTION

Successful speech processing and language learning rests on efficient processing of sequential auditory information and establishing relationships between speech elements (Hasson and Tremblay, 2015). Different elements of speech, such as phonemes, syllables, or words, are sequentially organized into speech streams, following language-specific constraints (or rules).

Converging evidence from genetic, non-human primate and cognitive neuroscience studies indicates that language and sequence learning have considerable overlap in the underlying neural mechanisms. It was thus argued that sequence learning mechanisms are important in language acquisition (Christiansen and Chater, 2015).

Artificial grammar (AG) learning paradigms provide a means to study sequence learning during language acquisition and its evolution both in human populations (children and adults) and in non-human species (e.g., primates or birds; Fitch and Friederici, 2012). One major advantage of AG learning paradigms relative to the study of natural language stimuli is that they allow for exploring the processing of rule-based regularities in a controlled way and unconfounded by semantic processes and prior knowledge (Fitch and Friederici, 2012). These paradigms usually include a learning phase and a test phase. During learning, participants are exposed to a set of syllable or letter sequences and during testing, their abilities to detect the underlying rule system are assessed (Lobina, 2014). In the test phase, participants are usually presented with novel sequences and need to classify them as correct or incorrect according to the underlying rules acquired in the learning phase (cf. Bahlmann et al., 2008). From the language acquisition perspective, AG learning represents one of a few language acquisition paradigms suitable for adult subjects. In previous studies, AG learning paradigms were mainly used as a model for syntactical and phonological processing (Uddén, 2012).

Previous neuroimaging, electrophysiological and behavioral studies used AG learning paradigms to elucidate the human capacity to detect regularities in speech sequences. Specifically, these studies investigated the processing of structured syllable sequences in healthy participants (Friederici et al., 2006; Bahlmann et al., 2008; Mueller et al., 2012; Uddén and Bahlmann, 2012). Some of the previous AG learning studies used shared articulatory features as cues for the underlying structure of the presented sequence (e.g., sequences like "di-ge-ku-to," whereby place of articulation indicates which syllable pairs belong together, namely "di-to" and "ge-ku"; Bahlmann et al., 2008; Mueller et al., 2010). It was argued that during successful AG learning in these studies, participants have to realize the pairing rule between syllables, specifically, between their starting consonants (Lobina, 2014). As soon as the relevant consonant pairs are linked by phonetic features, these features need to be rehearsed in phonological working memory to link up the different pairs.

Also, previous electrophysiological studies addressed the contribution of specific cognitive processes to language learning. These studies suggest that in the absence of semantic information during early stages of language learning, as with AG learning paradigms, attention and working memory processes are critical (De Diego-Balaguer and Lopez-Barroso, 2010).

Yet, a number of distinguishable processes are likely to contribute to successful learning of AG structures. Basic cognitive processes that are likely involved in learning include: (1) sensory or input encoding; (2) pattern extraction; (3) model building; and (4) retrieval or recognition processes (Karuza et al., 2014). Presumably, the relative involvement of these processes (and their interaction) changes in the course of a learning task. For instance, in the early phase of learning model building processes should closely interact with pattern extraction processes. In the later stages one would expect interaction of model building and recognition/retrieval processes to become more prominent. Moreover, the retrieval or recognition processes are likely to be necessary in the assessment procedure, when the outcome of learning is measured and participants need to demonstrate acquired knowledge (Karuza et al., 2014).

While most of the previous studies on AG learning focused on the outcome of learning by testing acquired rule knowledge (Friederici et al., 2006; Bahlmann et al., 2008, 2009), the neural correlates of the successful rule acquisition process itself remain largely unclear with some exceptions (Opitz and Friederici, 2004). Most studies tested processing of the AG sequences after learning and not during learning, which is likely to draw on the rule retrieval processes in the above mentioned learning framework (Karuza et al., 2014). We aimed at unraveling the functional underpinnings of successful auditory learning of phonological rules in AG sequences, focusing on the processes of pattern extraction and model building. To this end, we relied on a complex AG learning paradigm that was previously used in a behavioral study which showed reliable learning effects but also a large degree of variation in learning success (Mueller et al., 2010). In the present study, we used the known difficulty of these structures as a means to obtain variance in learning success, as we focused on the comparison between successful and unsuccessful rule learning processes. The stimuli were unconfounded by any semantic information and should thus allow for the investigation of "pure" phonological learning processes. A better understanding of the processes underlying rule extraction in AG learning paradigms might help to establish valid functional-anatomical models of rule acquisition in the healthy brain.

With respect to the underlying network for AG learning, some neuroimaging studies observed (left-lateralized or bilateral) fronto-parietal areas, including the left dorsal and ventral premotor cortex (PMC) and left inferior and superior parietal cortex (Fletcher et al., 1999; Tettamanti et al., 2002). Other studies (Friederici et al., 2006; Bahlmann et al., 2008, 2009) reported increased task-related activation in the left posterior inferior frontal gyrus (pIFG, pars opercularis) for the processing of sequences with complex as compared to simple structures and for violations of the sequence structure.

Based on the above described contribution of phonological processes to AG rule learning, we hypothesized that a successful strategy in rule learning should involve phonological processes, including phonological working memory for silent rehearsal. Specifically, successful AG learning and rule application should include search processes for specific phonetic features (i.e., consonants), phonological segmentation processes and phoneme discrimination for phoneme extraction, as well as silent rehearsal and phoneme comparison among long-distance elements. These processes should engage a left-dominant fronto-parietal network that was previously associated with phonological processes (e.g., Baddeley, 1992). If our task engages speech segmentation and phonological working memory processes, we would expect

increased task-related activity in the left lateral PMC and adjacent pIFG (e.g., Paulesu et al., 1993; Cunillera et al., 2009; Price, 2012), since this region should contribute to the active maintenance of non-meaningful verbal representations through articulatory subvocal rehearsal (Smith and Jonides, 1998) and might be crucially engaged in the first stages of learning unfamiliar new words (Baddeley et al., 1998). Additionally, we aimed to test whether the posterior IFG, which has been shown previously to be sensitive to the processing of complex AGs, would be also involved in initial stages of learning. Moreover, a posterior region in the inferior/superior parietal cortex should also contribute to phonological working memory (e.g., Kirschen et al., 2006; Romero et al., 2006), particularly to the short-term storage of information (see Buchsbaum and D'Esposito, 2008 for review). Both parietal and premotor areas were assigned to the dorsal language stream (Hickok and Poeppel, 2004, 2007) that comes into play when it is necessary to keep auditory representations in an active state during task performance (Buchsbaum et al., 2005; Aboitiz et al., 2006; Jacquemot and Scott, 2006) and contributes to novel word or phoneme learning (Rodriguez-Fornells et al., 2009).

MATERIALS AND METHODS

Participants

Initially, 65 healthy young volunteers (31 females) participated in the study. Four subjects had to be excluded due to technical reasons (data quality or missing behavioral responses). All of the remaining participants (31 males, 30 females; age range 20–36 years, mean age ± SD: 26.9 ± 3.81 years) were right-handed according to the Edinburgh handedness questionnaire (Oldfield, 1971). They were native German speakers with hearing levels within normal limits. Hearing levels were tested in octave steps between 250 and 8000 Hz in both ears using pure-tone audiometry. Normal hearing limits were defined as a maximum of 20 dB. They had no history of neurological or psychiatric disorders, no drug or alcohol abuse, no current pregnancy, no chronic medical disease, and no contraindications to MR-scanning. Written informed consent was collected from all participants according to the procedures approved by the local Ethics Committee of the University of Leipzig. Participants were paid after completing the experiment.

Stimuli

We used naturally spoken four-syllable sequences built according to AG rules with center-embedded structure (**Figure 1**) described in previous studies (e.g., Mueller et al., 2010). There were twelve different consonant-vowel syllables. Our AG had the structure $A_1A_2B_2B_1$ with pairwise dependencies between consonants of the syllables of two classes: A and B. Class A included syllables with voiced plosives and front vowels "e" and "i": {be, bi, ge, gi, de, di}, class B included syllables with voiceless plosives and back vowels "o" and "u": {po, pu, ko, ku, to, tu}. A and B refer to the vowel component, while indices 1 and 2 refer to the consonant component of the syllable (e.g., "be-gi-ko-pu," "be-de-tu-po"). To increase the saliency of the dependency, paired consonants (d-t,

FIGURE 1 | Stimuli structure and examples. (A) Structure of the stimuli and the rule. Four-syllable sequences were constructed according to the artificial grammar rule $A_1A_2B_2B_1$, where A and B represent two classes of consonant-vowel syllables and indices 1 and 2 represent the pairing rule. The consonants, sharing the same place of articulation (pairs "b" – "p," "g – k," "d – t"), were paired by a center-embedding rule: The first consonant was paired with the last consonant, and the second consonant was paired with the third consonant. All sequences, following this rule, were correct. In incorrect sequences, the last consonant violates the pairing rule. **(B)** Examples of correct and incorrect stimuli.

g-k, b-p) were phonetically similar (shared place of articulation). In the sequence "be-gi-ko-pu" place of articulation indicates which syllable pairs belong together, namely "be-pu" and "gi-ko". This led to a total set of 96 different four-syllable sequences, built according to the grammar rules (correct items) and 32 different four-syllable sequences with rule violations (incorrect items). Incorrect items had a violation of congruency between the first and the fourth element of the sequence (e.g., $A_1A_2B_2B_3$ sequences, like "be-gi-ku-to"). Similar sequences were previously employed to assess the learnability of complex, embedded ("syntactic") structures by using phonological cues in the input signal (Mueller et al., 2010).

Experimental Design and Procedures
Pre-scanning Training

Before the experiment, participants were familiarized with the structure and the presentation rate of the stimuli. To avoid that participants could acquire knowledge about the structure of the four-syllable sequence, we used spectrally rotated versions of the syllables (Blesser, 1972). This ensured unintelligibility, while maintaining the duration of the syllables. During presentation of each four-syllable sequence, the word "Listen" was presented on a computer screen. During the silence period between two sequences, the word "Respond" was presented and participants

FIGURE 2 | Functional magnetic resonance imaging (fMRI) paradigm. (A) Timeline of the fMRI experiment. The experiment consisted of four sessions. Each session started with a learning phase, followed by a test phase. Items were presented in blocks. **(B)** Structure of one learning block. Each learning block consisted of three items, separated by 1 s. **(C)** Structure of one test block. Each test block consisted of two items, separated by a response window of 2.5 s. After each test item participants responded with a button press, whether the test item was correct with respect to the rule or not.
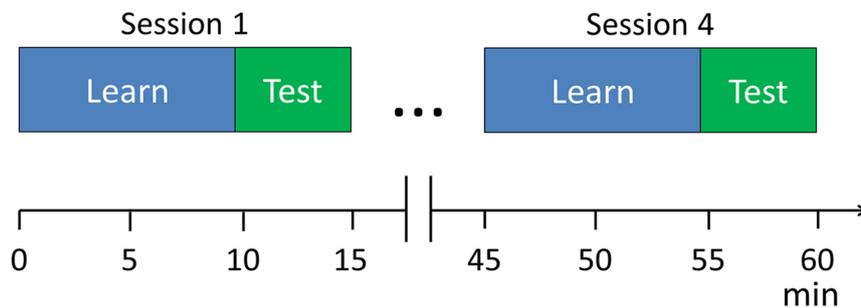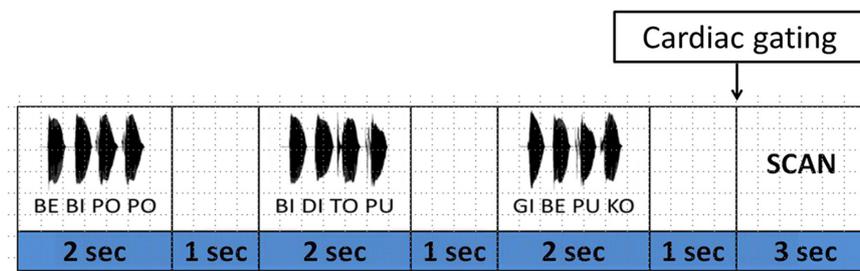
were instructed to press any response key (corresponding to "yes" or "no") before the presentation of the next sequence. The inter-trial interval was identical to that of the main experiment (i.e., 2.5 s). The aim of this familiarization procedure was to reduce the number of missing responses in the main experiment.

We also included a short familiarization session that required the participants to passively listen to two-syllable sequences A_jB_j that followed the same rule as the four-syllable sequences used in the main experiment. The practice session consisted of 12 two-syllable sequences, repeated three times in randomized order to facilitate learning (Mueller et al., 2010).

Scanning Sessions

In the main experiment participants performed four scanning sessions of a block design, each lasting approximately 15 min, with 9 min learning phase and 5 min test phase (**Figure 2**). Each session included two short visual instructions, indicating the start of the corresponding phase (i.e., learning or test phase, **Figure 2**). Participants were informed that the same rule was used in all learning phases and that all included items in learning phases were correct, while in the test phases correct and incorrect items were presented. To avoid interference with the auditorily presented stimuli, MRI volumes were acquired after each block

in an interleaved fashion (see "Magnetic Resonance Imaging section" for details). During the whole procedure a fixation cross was displayed.

In the learning phase, subjects were instructed to listen attentively and try to detect any regularities in the speech stream. Each learning phase comprised of 32 learning blocks. In each block, three four-syllable sequences were presented with an interstimulus interval of 1 s and a short pause between sequences, leading to a total number of 96 different correct sequences per session, presented in a randomized order.

In the test phase, participants performed a discrimination task requiring them to judge whether a given sequence was correct with respect to the rule they had previously learned. They were instructed to guess in case they didn't know the rule yet. In each of the 16 test blocks, two four-syllable sequences were presented with an interstimulus interval of 2.5 s, leading to a total of 32 test sequences per session (i.e., 16 correct sequences vs. 16 incorrect sequences).

After each test item, participants had to indicate via button press whether the current sequence followed the previously learned rule (binary answers: yes vs. no). Responses were collected within a time window of 4.5 s from the stimulus onset. We also included null blocks (i.e., 16 null blocks in each learning session and 8 null blocks in each test session) to avoid habituation. Participants were instructed to respond with the index and middle finger of their right hand on a custom-made two-button response box. Presentation of stimuli, recording of participants' responses and synchronization of the experiment with the MR scanner was accomplished using Cogent 2000[1].

Post-scanning Assessment

After the functional magnetic resonance imaging (fMRI) experiment, participants went through a structured interview that assessed individual strategies during rule acquisition and application as well as rule knowledge. It also covered hypotheses about the rule, attention focus (whole syllables sequences, isolated syllables, vowels, consonants) as well as application and modification of rules in the test phases. Participants were further asked to generate some examples of the correct syllable sequences, which followed the rule they were learning. General questions addressed clear stimulus presentation, motivation, fatigue, and attention drops during different phases of the experiment. Thereafter, participants underwent a short post-scanning test that required them to judge 11 test items to confirm presence or absence of rule knowledge. They were asked to verbally explain their decisions and, in the presence of rule knowledge, explicitly show how they would apply the rule on each of the test items, which included both correct and incorrect examples.

Magnetic Resonance Imaging

MRI data were obtained using a 3T Siemens Tim Trio MR scanner (Siemens Medical Systems, Erlangen, Germany). Auditory stimuli were delivered using MR-compatible headphones (MR confon GmbH, Magdeburg, Germany).

To attenuate scanner noise, participants wore flat frequency-response earplugs (ER20; Etymotic Research, Inc., Elk Grove Village, IL, USA). Prior to fMRI, sound levels were individually adjusted to a comfortable hearing level for each participant.

To avoid masking of the auditory stimuli by scanner noise, we used sparse temporal sampling (Hall et al., 2000; Gaab et al., 2007). Gradient-echo planar images (EPIs) were acquired in an interleaved fashion after each presentation block (whole brain coverage, 42 transverse slices in ascending order per volume; flip angle, 90°; acquisition bandwidth, 116 kHz; TR: 12 s; TE: 30 ms; TA: 2730 ms; matrix size 192 × 192; 2 mm slice thickness; 1 mm interslice gap; in-plane resolution, 3 × 3 mm; cardiac triggering). Additionally, cardiac gating was applied. On each block, after 9 s had elapsed, the scanner waited for the first heartbeat to trigger volume acquisition. Due to this, the actual repetition time (TR) was variable (mean: 12.49 s; SD: 365 ms, across all participants). In total, 300 brain volumes were acquired in four scanning sessions for each participant. Task instructions and fixation cross during auditory presentations were delivered using a LCD projector (PLC-XP50L, SANYO, Tokyo, Japan), which could be viewed via a mirror located above the head coil.

High-resolution T1-weighted MR scans of each participant were taken from the in-house database (standard MPRAGE sequences, whole brain coverage, voxel size 1 mm isotropic, matrix size 240*256*176, TR: 2300 msec, TE: 2.98 msec, TA: 6.7 min, sagittal orientation, flip angle 9°).

Statistical Analyses

Behavioral Data

For analyses of reaction times (RTs), we first excluded incorrect responses and misses from further analysis and performed outlier correction on the individual level (i.e., by excluding trials with response speed deviating more than two SD from the individual mean within each subject). Note that this resulted in an imbalance of the total number of items included for learners and non-learners. RTs were measured from the onset of the test item until button press within a time window of 4.5 s.

fMRI data

Functional images of fMRI sessions were analyzed with Statistical Parametric Mapping software (SPM12; Wellcome Trust Centre for Neuroimaging[2]; Friston et al., 2007) implemented in Matlab (release 2015). Standard preprocessing procedures comprised correction of motion-related artifacts (realignment and unwarping), coregistration of the T1 and mean EPI images in each participant, segmentation, normalization into standard stereotactic space [Montreal Neurological Institute (MNI) template], and spatial smoothing with a Gaussian kernel of 8 mm FWHM, and high-pass filtering at 128 s (Friston et al., 2007).

On the individual first level, statistical parametric maps (SPMs) were generated by modeling the evoked hemodynamic response for the two conditions (i.e., learning and test) as boxcar functions convolved with a canonical double gamma hemodynamic response function within the general linear model

---

[1]http://www.vislab.ucl.ac.uk/cogent2000.php

[2]http://www.fil.ion.ucl.ac.uk/spm/

(Friston et al., 2007), with the null blocks forming an implicit baseline. All four runs were assigned an equal weight. T-contrasts were computed for the main effect of each condition (learning and test) and for the contrast learn + test > implicit baseline.

At the second-level, between-group analyses (learners > non-learners and non-learners vs. learners) were performed with two-sample *t*-tests on the individual contrast images (learn + test > baseline, learn > baseline, test > baseline). Results were thresholded at a cluster level of $p < 0.05$, FWE corrected using an initial threshold of $p < 0.001$ (uncorrected) and a cluster threshold of $k > 250$ voxels (Friston et al., 1994).

Effect sizes for regions of interest (ROIs) were calculated as percent signal change with the rfxplot toolbox (Gläscher, 2009). ROIs were defined individually as a sphere with 3-mm radius centered on the individual subject peak, located in the vicinity of the corresponding group peak (within 6 mm) for the two main clusters from the contrast (learn + test > baseline). Neuroimaging data was visualized with MRIcron[3].

RESULTS

Subject Classification: Learners vs. Non-learners

Based on the outcome of the post-scanning assessment, participants were classified as either learners or non-learners. Participants were only classified as learners if they were able to explicitly explain the full rule (as presented in **Figure 1**) and apply it correctly to all items of the post-scanning test (i.e., 100% correct answers in classifying the examples). Learners explained if the test items were correct or incorrect according to the full rule and explicitly indicated the position of violation and the violating consonant. With respect to the learned rule, all learners reported the four-syllable structure of the sequences. They identified two parts of the syllable sequence, with vowels "e," "i" and consonants "b," "d," "g" in the first two-syllables and vowels "o," "u" and consonants "p," "t," "k" in the last two-syllables. They noted that the second and third syllable, as well as first and forth syllable were paired, leading to a pairing of "similar" consonants such as "b" and "p," "d" and "t," "g" and "k." They also reported that the structure of the last two consonants mirrored the sequence of the first two consonants in reversed order.

This classification led to a final sample of 30 learners (mean age 26.1 years; 19 females) and 31 non-learners (mean age 27.65 years; 11 females). All learners, except two, showed accuracy above 72% across test sessions. Learning success was also reflected in performance increase across test sessions, as shown in **Figure 3** (see also **Supplementary Figure S1**).

There were no between-group difference in mean RTs in any of the four sessions (Mann–Whitney test, all $W > 373$, $p > 0.05$). However, as expected, learners performed significantly better than non-learners in all sessions ($W = 282.5$, $p < 0.01$ in the first session; $W = 84.5, 52.5, 65.0$, $p < 0.001$ in all other sessions; surviving a Bonferroni–Holm correction for multiple comparisons; **Table 1**; **Figure 3**, **Supplementary Figure S1**).

---

[3]http://people.cas.sc.edu/rorden/mricron/index.html

Behavioral Strategies

In the post-scanning assessment learners reported two main strategies that were used both during learning and in the test phase after successful rule learning: a forward prediction strategy and a simple strategy. The two strategies were reported in a free manner, based on introspection (in response to the question: "Which strategies did you use during learning and test phase, respectively?"). We did not perform any quantification of response strategies since the same subject could use different strategies throughout the experiment. Notably, all learners explicitly reported the four-syllable structure of the sequences (see subject classification above).

With respect to the reported learning strategies, it should be borne in mind that our stimuli consisted of pairs of voiced and voiceless plosives between syllables 2 and 3 ("inner pair") and syllables 1 and 4 ("outer pair") (i.e., "b" was paired with "p," "d" with "t" and "g" with "k," **Figure 1**). For successful rule application, a learner needed to confirm that this pairing was present in the correct item or violated in the incorrect item. This was verbally explained and explicitly demonstrated on the test items in the post-scanning test.

Importantly, all learners reported that they consciously related and compared phonemes. Specifically, many learners reported that the structure of the last two consonants mirrored the sequence of the first two consonants in reversed order. Whenever learners stated explicitly that a specific phoneme was anticipated, we took this as evidence for a *forward prediction* strategy.

Using the *forward prediction strategy*, a learner would try to predict subsequent consonants during learning after hearing the first two consonants to find possible combinations. In the test phase, most of the learners realized that they could match the consonants in the inner pair during listening to the test items. They transformed the first consonant (e.g., "b" in "be-gi-ko-pu") into the corresponding voiceless plosive (e.g.,"p") and rehearsed the latter in working memory. During presentation of the fourth consonant, they checked if their expectations were met. For instance, if they were expecting a "p" during "be-gi-ko-pu," the response would be "correct item".

Using the *simple strategy,* a learner would try to keep the first two consonants in mind, and compare them with the following third and fourth consonant to find possible combinations. In the latest test phase after successful rule learning, a learner would keep only the first consonant in mind (e.g., "b" in "be-gi-ko-pu") and compare it with the last one ("b" and "p" in "be-gi-ko-pu," "b" and "t" in "be-gi-ko-tu") for matching.

In summary, both strategies required speech stream segmentation and phoneme matching processes as well as phonological working memory. Indeed, in the structured interview, both groups, learners and non-learners, reported silent rehearsal of the syllables or phonemes. Specifically, after successful learning, learners used silent rehearsal to keep the consonants in memory for matching. Moreover, learners reported inner visualization of the syllables, vowels or consonants during learning and rule application in test phases. Accordingly, they visualized the consonants in the test phases that followed successful rule learning. Finally, non-learners

FIGURE 3 | Behavioral performance (accuracy and reaction time) for learners and non-learners across sessions. (A) Accuracy as percentage of correct answers across all trials in each session (mean percent correct ± SEM). **(B)** Reaction time across all correct trials in each session in milliseconds (mean ± SEM). Statistical significance (*p* < 0.01) is marked by *.
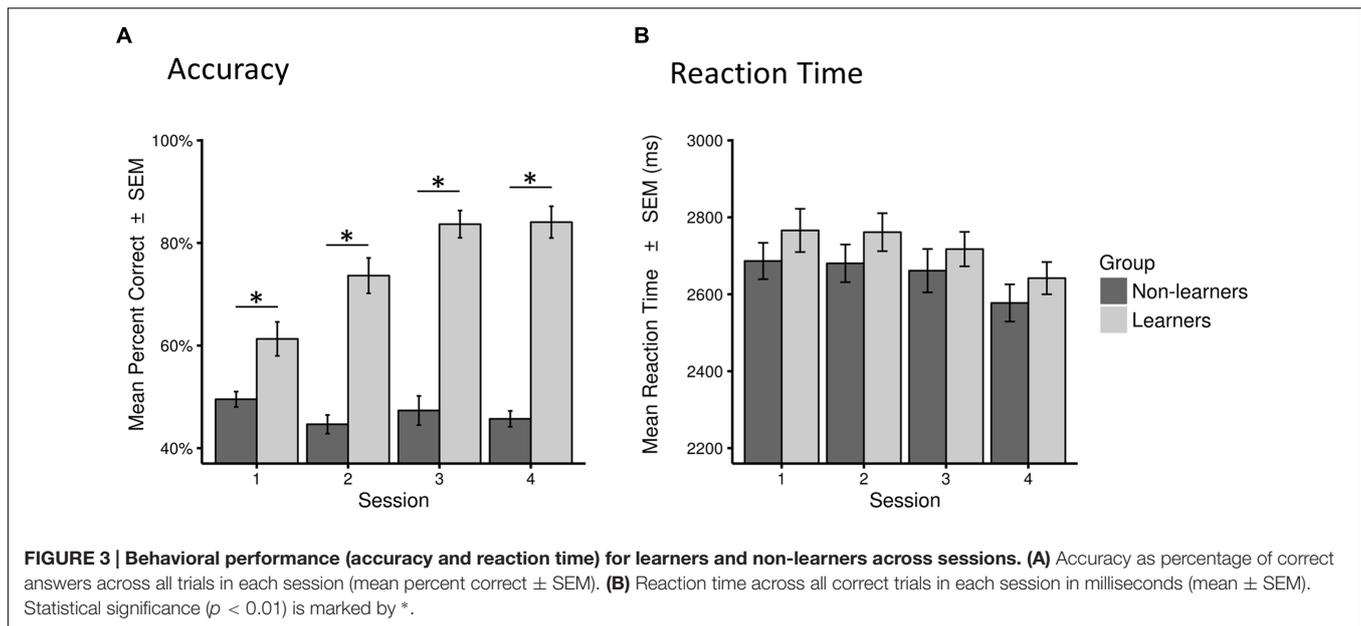
reported difficulties in ignoring distracting information (e.g., vowels, which are irrelevant) during learning.

fMRI Results

Since we were mainly interested in the difference of task-related activity in successful learners vs. non-learners, we investigated second-level contrasts directly addressing group comparison (i.e., learners > non-learners, and additionally non-learners > learners).

To identify brain regions that are involved in both successful learning and application of AG rules, we first compared the global learn + test > baseline contrast between learners and non-learners using a second-level, two-sample *t*-test. Overall, learners showed stronger activity in a large network of frontal and parietal areas (**Figure 4A**; **Table 2**) than non-learners.

The frontal cluster was located mainly in the left lateral PMC (peak at x, y, $z = -54$, 10, 40), partly extending to the middle, superior and inferior frontal gyri (pars triangularis and pars opercularis) as well as primary motor cortex. Increased neural activity was also found in the right PMC (peak at x, y, $z = 26$, -8, 52). The parietal clusters included significant activation in inferior and superior parietal lobes (IPL/SPL) bilaterally (peaks

at x, y, $z = -34$, -44, 52; x, y, $z = 34$, -48, 48), left precuneus (peak at x, y, $z = -8$, -66, 54), bilateral postcentral gyrus (PcG) and SMG (peaks at x, y, $z = -42$, -36, 46; x, y, $z = 36$, -32, 44). Effect sizes for the peaks of the two major clusters in PMC and IPL/SPL in each run/group are summarized in **Figure 4B**.

To identify specific brain regions involved in either learning or application of the rule, we further computed contrasts for each condition (i.e., learn > baseline and test > baseline) between groups. In the learning > baseline contrast, learners displayed stronger activation in the left lateral PMC (i.e., at the border between left dorsal and ventral PMC; peak at x, y, $z = -56$, 12, 38) and bilateral IPL/SPL (peaks at x, y, $z = -30$, -48, 44; x, y, $z = 32$, -48, 48; **Figure 5A**; **Table 2**). In the test > baseline contrast, task-related activity also engaged left PMC (peak at x, y, $z = -62$, -22, 36) and left parietal areas, including precuneus (peak at x, y, $z = -8$, -68, 54), PcG and SMG (peak at x, y, $z = -62$, -22, 36), as well as IPL/SPL (peak at x, y, $z = -34$, -44, 52; **Figure 5B**; **Table 2**).

We did not find any significant activation in the second-level contrasts comparing non-learners vs. learners in either learn or test sessions, probably due to a strong variability of strategies in non-learners.

TABLE 1 | Behavioral data for learners and non-learners.

	Session 1	Session 2	Session 3	Session 4
Mean accuracy (%) ± SEM				
Learners	62% ± 3.4%	74% ± 3.5%	84% ± 2.7%	84% ± 3.2%
Non-learners	50% ± 1.5%	45% ± 1.9%	47% ± 2.9%	46% ± 1.6%
Effect size (Cohen's d)	0.83	1.87	2.40	3.11
Mean reaction time (ms) ± SEM				
Learners	2765.93 ± 56.22	2761.40 ± 49.24	2717.38 ± 44.78	2641.90 ± 41.89
Non-learners	2686.46 ± 47.21	2680.45 ± 78.91	2661.36 ± 56.39	2577.43 ± 48.31
Effect size (Cohen's d)	0.28	0.22	0.20	0.46
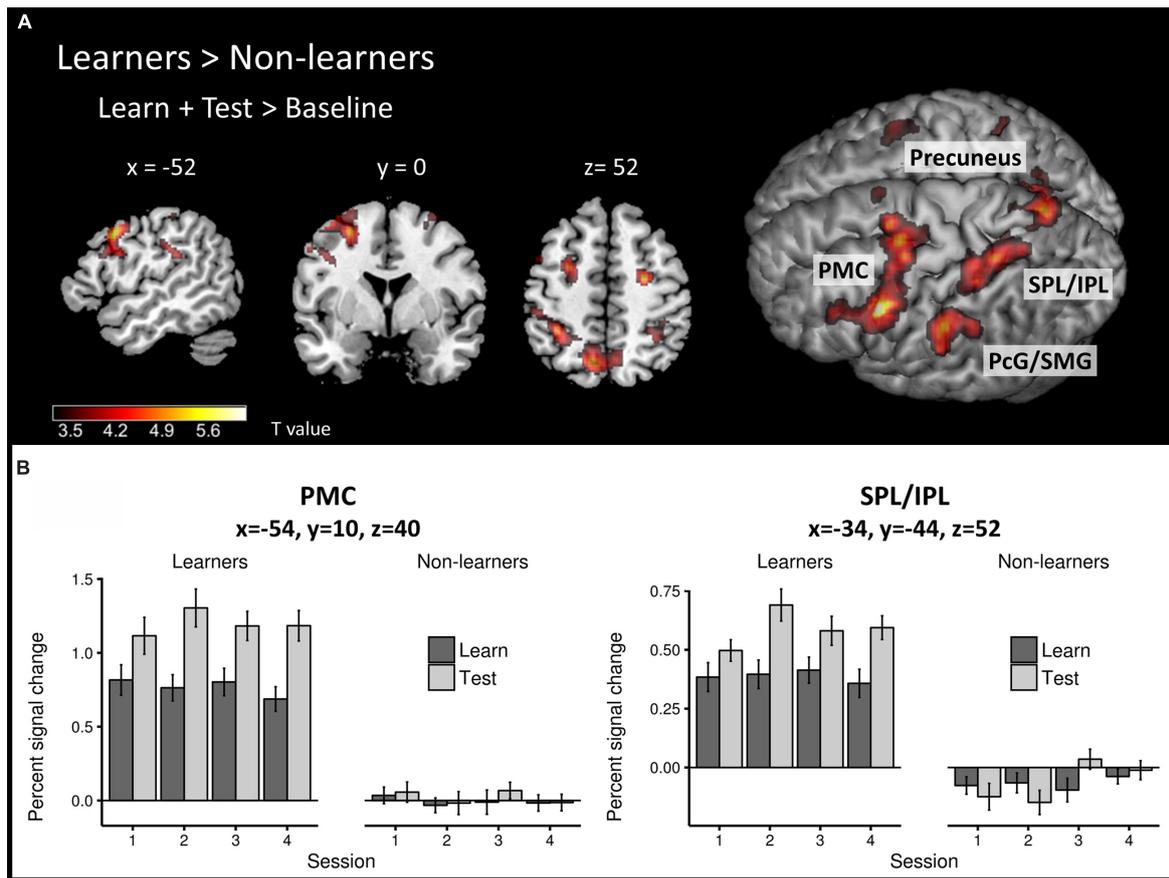
FIGURE 4 | Task-related neural activity across learning and test phases. (A) Between-group comparison (learners > non-learners) for the learn+test > baseline contrast (two-sample *t*-test; FWE corrected, *p* < 0.05). (B) Percent signal change in the learn and test conditions across sessions at PMC and SPL/IPL peaks in each group. PMC – premotor cortex, SPL/IPL – superior parietal lobe/inferior parietal lobe, PcG – postcentral gyrus, SMG – supramarginal gyrus.

DISCUSSION

In this study, we investigated the functional underpinnings of successful learning of a complex phonological rule in an AG learning paradigm. Our main finding was that successful learning and rule application was associated with increased neural activity in a fronto-parietal network that encompassed left lateral premotor and prefrontal areas as well as bilateral regions in the IPL/SPL. These areas have been previously associated with learning and phonological processing (Tettamanti et al., 2002; Liebenthal et al., 2013). In the learning phase, the contribution of the IPL/SPL regions was more bilaterally distributed while the test phase showed a more left-lateralized parietal activation pattern. Notably, we did not find evidence for a contribution of left posterior IFG to AG learning. However, during rule application in the test phase, task-related premotor activity clearly extended into left IFG (i.e., pars opercularis and triangularis) and neighboring inferior frontal sulcus (IFS). This might indicate that the left IFG is not needed to support AG learning with our paradigm, but rather serves rule representation, as suggested by a previous AG learning study (Opitz and Friederici, 2004). That study used an AG that mimicked natural language rules

and reported a shift of neural activation from the hippocampus in early learning stages to the posterior IFG when abstract rule representations were built. Our IFG/IFS cluster was located close to the activation pattern of another previous study that associated IFS activation with memory-related aspects during complex sentence processing (Makuuchi et al., 2009). This supports our claim of a key contribution of verbal working memory processes to our AG learning paradigm (see below).

In the following sections, we will discuss the cognitive processes associated with successful AG learning and the role of fronto-parietal regions in these processes. These data support our hypothesis that our task mainly required phonological processes and strongly engages phoneme comparison and verbal working memory.

The Role of Phonological Processes in AG Learning: Contributions of Premotor Regions

With respect to the processes associated with phonological rule learning, our task required speech stream segmentation and phoneme comparison. It was previously argued that solving AG

TABLE 2 | Local activation maxima in Montreal Neurological Institute (MNI) coordinates for the group comparison learners > non-learners.

Contrast	Location	Cluster size	Cluster FWE-corrected	p (unc)	T-value	Z-value	x	y	z
Learn + Test > Baseline	Left premotor cortex	1320	0	0	6.21	5.42	−54	10	40
					5.05	4.58	−26	0	56
					4.82	4.41	−26	6	70
	Right premotor cortex	275	0.032	0.004	5.09	4.61	26	−8	52
					3.76	3.54	30	2	68
	Left postcentral gyrus, supramarginal gyrus	345	0.014	0.002	5.08	4.61	−62	−20	36
					4.49	4.15	−66	−20	24
					4.22	3.93	−56	−34	26
	Left precuneus, superior/inferior parietal lobe, postcentral gyrus	1350	0	0	5.01	4.56	−8	−66	54
					4.79	4.38	−34	−44	52
					4.73	4.33	−42	−36	46
	Right postcentral gyrus, supramarginal gyrus	300	0.024	0.003	3.94	3.7	36	−32	44
					3.94	3.7	34	−48	48
					3.79	3.57	30	−44	56
Learn > Baseline	Left premotor cortex	272	0.032	0.004	5.12	4.64	−56	12	38
					4.53	4.18	−52	6	28
					3.56	3.38	−42	4	40
	Right superior/inferior parietal lobe	523	0.002	0	4.61	4.25	32	−48	48
					4.27	3.97	36	−48	58
					4.11	3.84	46	−40	44
	Left superior/inferior parietal lobe	414	0.005	0.001	4.12	3.85	−30	−48	44
					4.08	3.81	−42	−38	46
					3.91	3.67	−34	−58	60
Test > Baseline	Left postcentral gyrus, supramarginal gyrus	262	0.036	0.005	5.71	5.07	−62	−22	36
					3.67	3.46	−52	−34	26
	Left premotor cortex	1096	0	0	5.63	5.02	−54	10	40
					4.73	4.34	−24	4	70
					4.54	4.19	−26	0	54
	Left postcentral gyrus, supramarginal gyrus	281	0.028	0.004	4.73	4.34	−34	−44	52
					4.15	3.87	−44	−34	62
					4.13	3.85	−42	−36	44
	Left precuneus, superior/inferior parietal lobe	523	0.002	0	4.59	4.23	−8	−68	54
	Right cerebellum (lobule VI)	301	0.022	0.003	4.55	4.2	26	−64	−26
					4.22	3.93	30	−56	−32

rules for syllable sequence structures can be done solely by phoneme matching, without the necessity to build hierarchies (Lobina, 2014).

Indeed, successful learners in our study reported the application of these strategies. Specifically, many of our successful learners relied on phoneme manipulation in the learning phase of our study (e.g., by transforming "b" into "p" in the forward prediction strategy described above). Other processes that might have been involved in learning our AG rule include discrimination of initial consonants of the syllables and phoneme monitoring. In accordance with our observation of increased task-related activation of lateral premotor areas (dorsal and ventral PMC) during successful AG learning, several previous neuroimaging studies demonstrated a contribution of the left PMC to speech segmentation and categorization tasks (Burton

et al., 2000; Sato et al., 2009) as well as phoneme or syllable perception (Wilson et al., 2004; Pulvermueller et al., 2006) and production (Bohland and Guenther, 2006; Peeva et al., 2010; Hartwigsen et al., 2013).

A number of neuroimaging and non-invasive brain stimulation studies demonstrated that the left lateral ventral PMC is engaged in phoneme categorization (Burton et al., 2000; Wilson et al., 2004; Meister et al., 2007; Chang et al., 2011; Alho et al., 2012; Chevillet et al., 2013; Krieger-Redwood et al., 2013; Du et al., 2014). Burton et al. (2000) observed increased activity of left ventral PMC and adjacent inferior frontal gyrus during a phoneme discrimination task, when participants were asked to discriminate initial or final consonants in pairs of CVC syllables (e.g., "fat"–"tid," "dip"–"ten"), compared to a pitch or a tone discrimination task. Moreover, Pulvermueller
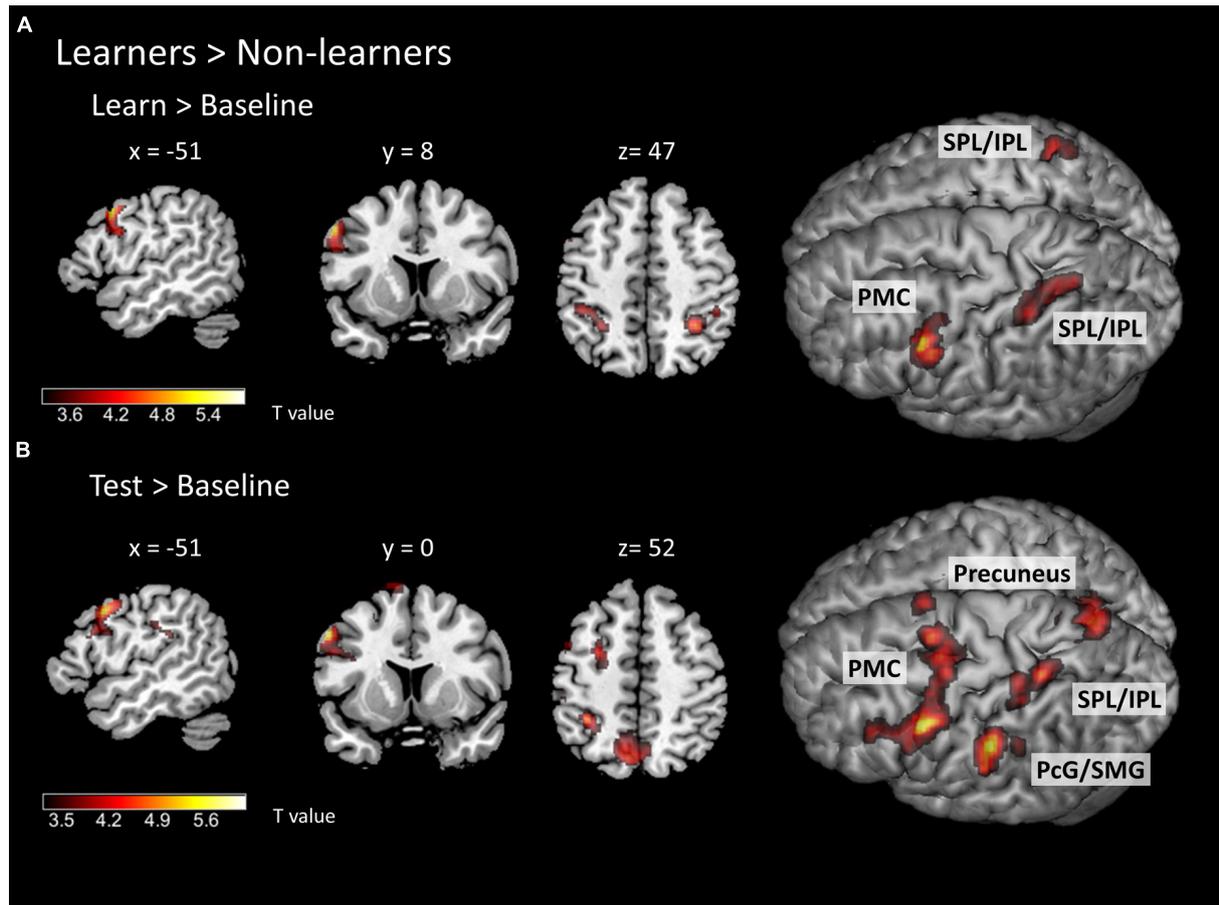
FIGURE 5 | Task-related neural activity during learning and test phases. Between-group comparison (learners > non-learners) for **(A)** learn > baseline contrast and **(B)** test > baseline contrast (two-sample *t*-test; FWE corrected at a threshold of *p* < 0.05). PMC – premotor cortex, SPL/IPL – superior parietal lobe/inferior parietal lobe, PcG – postcentral gyrus, SMG – supramarginal gyrus.

et al. (2006) showed a somatotopic activation in the left PMC during phoneme discrimination, for both subvocal production and passive listening tasks. The task specificity of this premotor activation was further shown by Krieger-Redwood et al. (2013), who used focal disruption of premotor activity induced by transcranial magnetic stimulation to demonstrate that the PMC is causally relevant for phoneme judgements, but not for speech comprehension. Together, the results of the previous and present studies suggest that the left lateral PMC is a key node for phoneme comparison, a process that is, among others, required during successful learning of speech sequences.

In a similar vein, another transcranial magnetic stimulation study by Sato et al. (2009) revealed a causal contribution of left PMC (overlapping our premotor cluster) to phonological segmentation. To successfully identify and apply the rule in our study, learners also used phonological segmentation to extract the consonants from a syllable sequence and compare the respective consonants on the corresponding positions. This process required basic acoustic analysis for phoneme identification, and verbal storage to discriminate phonemic contrasts when comparing syllables/consonants.

Premotor regions were also consistently associated with the planning and execution of speech gestures in previous neuroimaging studies (Gracco et al., 2005; Bohland and Guenther, 2006; Sörös et al., 2006). Accordingly, it was suggested that the recruitment of auditory-motor transformation and articulatory-based representations during phonological processing depends on the use of phonemic segmentation and working memory demands (Démonet et al., 1994; Zatorre et al., 1996; Hickok and Poeppel, 2004, 2007; Burton and Small, 2006). It was further argued that one specific strategy for performing short-term maintenance is phonological rehearsal by using inner speech (Herwig et al., 2003; see also Price, 2012 for review). This is well in line with the reported strategy from the successful learners in our study who used silent articulatory rehearsal of syllables or phonemes to compare specific consonants during learning. Note that to find the rule, our participants were specifically required to identify the pairing between voiced and unvoiced consonants according to the place of articulation.

Moreover, Chevillet et al. (2013) demonstrated that the left PMC (again overlapping with our cluster) was selectively activated during phoneme categorization but not acoustic

phonetic tasks when subjects listened to a place-of-articulation continuum between the syllables "da" and "ga". In that study, phoneme category selectivity in the PMC correlated with explicit phoneme categorization performance, suggesting that premotor recruitment accounted for performance on phoneme categorization tasks. This supports our hypothesis that successful learning requires distinguishing phonetic categories and matching corresponding phonemes.

The Role of Phonological Processes in AG Learning: Contributions of Parietal Regions

Aside from the observed increases in premotor activity during successful learning and application of AG rules, we also found a strong upregulation of inferior and superior parietal areas (IPL/SPL) during both learning and test phase. Several previous studies have shown that left IPL regions and premotor areas jointly contribute to phonological tasks (Liebenthal et al., 2013), for example during the rehearsal of verbal sequences (Koelsch et al., 2009) and support sensorimotor integration (Poldrack et al., 1999; Buchsbaum et al., 2011; Liebenthal et al., 2013). The IPL was assigned a key role as an auditory-motor interface (Hickok and Poeppel, 2007, Rauschecker and Scott, 2009). Consistent with that view, left IPL activation was reported during phonemic categorization tasks, with the level of activity being related to the individual categorization ability (Jacquemot et al., 2003; Raizada and Poldrack, 2007; Desai et al., 2008; Liebenthal et al., 2013) and the learning of new phonemic categories (Kilian-Hütten et al., 2011). A meta-analysis by Turkeltaub and Branch Coslett (2010) also found significant activation likelihood in left IPL/SMG during categorical phoneme perception when subjects were required to focus on the differences between phoneme categories. This cluster overlaps with the location of the observed parietal activation in the learning phase in our study. These authors claimed that although the exact role of the IPL in categorical perception of speech sounds remains unclear, one explanation would be that increased IPL activity during phoneme perception might be related to phonological working memory processes (Baddeley, 2003a,b; Jacquemot and Scott, 2006; Buchsbaum and D'Esposito, 2008). As an alternative explanation, it was argued that the left IPL might serve as a sensorimotor sketchpad for distributing predictive information between motor and perceptual areas during speech perception and production (Rauschecker and Scott, 2009). The left IPL might also play a more domain general role in consolidating continuous features of percepts or concepts into categories and comparing stimuli during discrimination tasks (Turkeltaub and Branch Coslett, 2010).

Interestingly, the observed IPL/SPL activation in our study was more bilaterally distributed in the learning phase and more left-lateralized in the test phase. This might reflect different strategies or increased cognitive load during learning compared to application of the rule. Indeed, it was suggested that right hemispheric IPL/SPL regions might support their left-hemispheric homologs under demanding

task conditions (Nakai and Sakai, 2014). Learning efficiency in second or artificial language learning paradigms was also associated with bilateral or right-hemispheric parietal networks (Kepinska et al., 2016; Prat et al., 2016). Moreover, bilateral contribution of parietal areas was associated with increased attention demands due to higher task difficulty during pseudoword vs. real word processing (Newman and Twieg, 2001), which might also have contributed to the observed bilateral parietal activation in the learning phase of our study.

Notably, in the test but not learning phase, additional left parietal activation was observed at the border between left postcentral and supramarginal gyrus as well as in the left precuneus. The left SMG has previously been associated with phonological working memory processes (Kirschen et al., 2006; Romero et al., 2006; Hartwigsen et al., 2010; Deschamps et al., 2014). This might indicate that during the test phase, our subjects relied more on phonological working memory processes to recall and apply the learned rule. Postcentral activation, on the other hand, was previously associated with motor-speech processes in tasks that required overt articulation (e.g., pseudoword repetition; Lotze et al., 2000; Dogil et al., 2002; Riecker et al., 2005; Peschke et al., 2009), or articulation without phonation (Pulvermueller et al., 2006). It was suggested that the left somatosensory cortex plays a crucial role in speech-motor control, forming part of a somatosensory feedback system (Guenther, 2006). Finally, left precuneus was associated with correct grammaticality decisions in another AG learning paradigm in a previous study (Skosnik et al., 2002) and might be related to retrieval success in working memory demanding tasks (Konishi et al., 2000; von Zerssen et al., 2001). Consequently, our observation of a strong upregulation of these regions in the test but not learning phase might indicate that subjects focused on phoneme manipulation in the learning, but relied stronger on working memory and rehearsal processes in the test phase. The upregulation of the precuneus might also point toward an engagement of mental imagery processes during the test phase since this region has been identified as a core node for visuo-spatial imagery previously (for review, Cavanna and Trimble, 2006). Moreover, activation of precuneus has been reported in previous fMRI studies on number comparisons or arithmetic calculations (Pinel et al., 2001; Nakai and Sakai, 2014).

We also found increased activity in the right cerebellum for learners vs. non-learners in the test phase. A contribution of the cerebellum is observed in many neuroimaging and electrophysiological studies in the cognitive and language domain (Marvel and Desmond, 2010; Price, 2012; De Smet et al., 2013). It was argued that the cerebellum plays an important role in the prediction of outcome associated with sensory input or actions (De Smet et al., 2013). Moreover, this area was also suggested to be involved in verbal working memory (Marvel and Desmond, 2010), potentially reflecting pre-articulatory or 'internal speech' processes. More specifically, the superior/lateral cerebellum (lobule VI) was associated with the encoding phase in covert speech and verbal working memory tasks. Together, the

previous and present results indicate that the (right) cerebellum contributes to verbal working memory processes during speech processing that are necessary for successful encoding and rule application.

Overall, we found a contribution of fronto-parietal areas during both the learning and test phase in successful learners, indicating that a fronto-parietal network orchestrates successful AG processing. The stronger engagement of phonological processes in learners compared with non-learners in our study might reflect a more consistent application of the rule during and after successful learning. Particularly, non-learners might have lost attention or were engaged in wrong strategies such as focusing on intonation or rhythm, or simply relying on gut feeling or passive listening.

With respect to the interaction between both regions, it was previously argued that at least during verbal working memory tasks, left PMC provides phonological information to parietal areas (Herwig et al., 2003). Accordingly, these authors suggest that the concept of phonological storage can be regarded as a premotor-mediated top-down activation of internal representations of the memorized items in parietal regions for later recognition. However, it should be borne in mind that our task cannot solely be explained by verbal working memory processes, since it required explicit phonological manipulation (i.e., phonological segmentation and phoneme comparison). We argue that explicit phonological manipulation requires more than just temporary storing of the stimulus or its parts (constituent phonemes). Indeed, left IPL regions were previously associated with phonological manipulation processes. For instance, Peschke et al. (2012) reported increased IPL activity during phonological segmental manipulation as compared to prosodic manipulation. Moreover, it was suggested that activation of the premotor-parietal network during rehearsal tasks might represent formation and maintenance of sensorimotor codes that contributes not only to verbal processing, but also to a number of (motor) sequencing tasks (Jobard et al., 2003; Koelsch et al., 2009) and might thus indicate more domain general processing.

CONCLUSION

In the present study, we investigated the functional underpinnings of successful learning and application of complex phonological rules, implemented in an AG learning paradigm with auditory presented structured syllable sequences. The observed fronto-parietal network comprised left premotor areas and bilateral superior/inferior parietal cortex. This network together with the reported strategies provides strong evidence for a core contribution of phonological processes, specifically phonological segmentation, phoneme comparison and inner rehearsal, as well as verbal working memory to successful AG learning. We might further speculate that successful AG learning might not require a specific contribution of the posterior IFG. However, this region comes into play during rule application in the test phase, further supporting the notion of a key role of the IFG in rule representation. Further studies might disentangle the exact role of the observed brain regions in AG learning and application, specifics of the time-course changes in activity and the interaction between these processes.

AUTHOR CONTRIBUTIONS

DG, JK, JM, and AF designed the experiment. DG and JK conducted the experiment. DG, JK, and GH performed the data analysis and interpretation. All authors contributed to the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnhum. 2016.00551/full#supplementary-material

FIGURE S1 | Boxplots for behavioral performance (accuracy and reaction time) for learners and non-learners across sessions. (A) Accuracy as percentage of correct answers across all trials in each session. **(B)** Reaction time across all correct trials in each session in milliseconds.

REFERENCES

Aboitiz, F., García, R. R., Bosman, C., and Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain Lang.* 98, 40–56. doi: 10.1016/j.bandl.2006.01.006

Alho, J., Sato, M., Sams, M., Schwartz, J.-L., Tiitinen, H., and Jääskeläinen, I. P. (2012). Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. *Neuroimage* 60, 1937–1946. doi: 10.1016/j.neuroimage.2012.02.011

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Baddeley, A. (2003a). Working memory and language: an overview. *J. Commun. Disord.* 36, 189–208. doi: 10.1016/S0021-9924(03)00019-4

Baddeley, A. (2003b). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201

Baddeley, A., Gathercole, S., and Papagno, C. (1998). The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173. doi: 10.1037/0033-295X.105.1.158

Bahlmann, J., Schubotz, R. I., and Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *Neuroimage* 42, 525–534. doi: 10.1016/j.neuroimage.2008.04.249

Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., and Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res.* 1298, 161–170. doi: 10.1016/j.brainres.2009.08.017

Blesser, B. (1972). Speech perception under conditions of spectral transformation. I. Phonetic characteristics. *J. Speech Hear. Res* 15, 5–41. doi: 10.1044/jshr.1501.05

Bohland, J. W., and Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage* 32, 821–841. doi: 10.1016/j.neuroimage.2006.04.173

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., et al. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory – An aggregate analysis of lesion and fMRI data. *Brain Lang.* 119, 119–128. doi: 10.1016/j.bandl.2010.12.001

Buchsbaum, B. R., and D'Esposito, M. (2008). The search for the phonological store: from loop to convolution. *J. Cogn. Neurosci.* 20, 762–778. doi: 10.1162/jocn.2008.20501

Buchsbaum, B. R., Olsen, R. K., Koch, P., and Berman, K. F. (2005). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* 48, 687–697. doi: 10.1016/j.neuron.2005.09.029

Burton, M. W., and Small, S. L. (2006). Functional neuroanatomy of segmenting speech and nonspeech. *Cortex* 42, 644–651. doi: 10.1016/S0010-9452(08)70400-3

Burton, M. W., Small, S. L., and Blumstein, S. E. (2000). The role of segmentation in phonological processing: an fMRI investigation. *J. Cogn. Neurosci.* 12, 679–690. doi: 10.1162/089892900562309

Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. doi: 10.1093/brain/awl004

Chang, E. F., Edwards, E., Nagarajan, S. S., Fogelson, N., Dalal, S. S., Canolty, R. T., et al. (2011). Cortical spatio-temporal dynamics underlying phonological target detection in humans. *J. Cogn. Neurosci.* 23, 1437–1446. doi: 10.1162/jocn.2010.21466

Chevillet, M. A., Jiang, X., Rauschecker, J. P., and Riesenhuber, M. (2013). Automatic phoneme category selectivity in the dorsal auditory stream. *J. Neurosci.* 33, 5208–5215. doi: 10.1523/JNEUROSCI.1870-12.2013

Christiansen, M. H., and Chater, N. (2015). The language faculty that wasn't: a usage-based account of natural language recursion. *Front. Psychol.* 6:1182. doi: 10.3389/fpsyg.2015.01182

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

De Diego-Balaguer, R., and Lopez-Barroso, D. (2010). Cognitive and neural mechanisms sustaining rule learning from speech. *Lang. Learn.* 60, 151–187. doi: 10.1111/j.1467-9922.2010.00605.x

De Smet, H. J., Paquier, P., Verhoeven, J., and Mariën, P. (2013). The cerebellum: its role in language and related cognitive and affective functions. *Brain Lang.* 127, 334–342. doi: 10.1016/j.bandl.2012.11.001

Démonet, J. F., Price, C., Wise, R., and Frackowiak, R. S. (1994). A PET study of cognitive strategies in normal subjects during language tasks.Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain* 117(Pt. 4), 671–682. doi: 10.1093/brain/117.4.671

Desai, R., Liebenthal, E., Waldron, E., and Binder, J. R. (2008). Left posterior temporal regions are sensitive to auditory categorization. *J. Cogn. Neurosci.* 20, 1174–1188. doi: 10.1162/jocn.2008.20081

Deschamps, I., Baum, S. R., and Gracco, V. L. (2014). On the role of the supramarginal gyrus in phonological processing and verbal working memory: evidence from rTMS studies. *Neuropsychologia* 53, 39–46. doi: 10.1016/j.neuropsychologia.2013.10.015

Dogil, G., Ackermann, H., Grodd, W., Haider, H., Kamp, H., Mayer, J., et al. (2002). The speaking brain: a tutorial introduction to fMRI experiments in the production of speech, prosody and syntax. *J. Neurolinguistics* 15, 59–90. doi: 10.1016/S0911-6044(00)00021-X

Du, Y., Buchsbaum, B. R., Grady, C. L., and Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc. Natl. Acad. Sci. U.S.A* 111, 7126–7131. doi: 10.1073/pnas.1318738111

Fitch, W. T., and Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1933–1955. doi: 10.1098/rstb.2012.0103

Fletcher, P., Büchel, C., Josephs, O., Friston, K., and Raymond, D. (1999). Learning-related Neuronal Responses in Prefrontal Cortex Studied with Functional Neuroimaging. *Cereb. Cortex* 9, 168–178. doi:10.1093/cercor/9.2.168.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., and Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2458–2463. doi: 10.1073/pnas.0509389103

Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T. E., and Penny, W. D. (eds). (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images.* Cambridge, MA: Academic Press.

Friston, K. J., Worsley, K. J., Frackowiak, R. S., Mazziotta, J. C., and Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Hum. Brain Mapp.* 1, 210–220. doi: 10.1002/hbm.460010306

Gaab, N., Gabrieli, J. D. E., and Glover, G. H. (2007). Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. *Hum. Brain Mapp.* 28, 721–732. doi: 10.1002/hbm.20299

Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics* 7, 73–82. doi: 10.1007/s12021-008-9042-x

Gracco, V. L., Tremblay, P., and Pike, B. (2005). Imaging speech production using fMRI. *Neuroimage* 26, 294–301. doi: 10.1016/j.neuroimage.2005.01.033

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *J. Commun. Disord.* 39, 350–365. doi: 10.1016/j.jcomdis.2006.06.013

Hall, D. A., Summerfield, A. Q., Gonçalves, M. S., Foster, J. R., Palmer, A. R., and Bowtell, R. W. (2000). Time-course of the auditory BOLD response to scanner noise. *Magn. Reson. Med.* 43, 601–606. doi: 10.1002/(SICI)1522-2594(200004)43:4<601::AID-MRM16>3.3.CO;2-I

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., and Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16494–16499. doi: 10.1073/pnas.1008121107

Hartwigsen, G., Saur, D., Price, C. J., Baumgaertner, A., Ulmer, S., and Siebner, H. R. (2013). Increased facilitatory connectivity from the Pre-SMA to the left dorsal premotor cortex during pseudoword repetition. *J. Cogn. Neurosci.* 25, 580–594. doi: 10.1162/jocn_a_00342

Hasson, U., and Tremblay, P. (2015). "Neurobiology of auditory statistical information processing," in *The Neurobiology of Language,* eds S. L. Small and G. Hickok (Amsterdam: Elsevier).

Herwig, U., Abler, B., Schönfeldt-Lecuona, C., Wunderlich, A., Grothe, J., Spitzer, M., et al. (2003). Verbal storage in a premotor–parietal network: evidence from fMRI-guided magnetic stimulation. *Neuroimage* 20, 1032–1041. doi: 10.1016/S1053-8119(03)00368-9

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., and Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J. Neurosci.* 23, 9541–9546.

Jacquemot, C., and Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends Cogn. Sci.* 10, 480–486. doi: 10.1016/j.tics.2006.09.002

Jobard, G., Crivello, F., and Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage* 20, 693–712. doi: 10.1016/S1053-8119(03)00343-4

Karuza, E. A., Emberson, L. L., and Aslin, R. N. (2014). Combining fMRI and behavioral measures to examine the process of human learning. *Neurobiol. Learn. Mem.* 109, 193–206. doi: 10.1016/j.nlm.2013.09.012

Kepinska, O., de Rover, M., Caspers, J., and Schiller, N. O. (2016). On neural correlates of individual differences in novel grammar learning: an fMRI study. *Neuropsychologia* doi: 10.1016/j.neuropsychologia.2016.06.014 [Epub ahead of print].

Kilian-Hütten, N., Vroomen, J., and Formisano, E. (2011). Brain activation during audiovisual exposure anticipates future perception of ambiguous speech. *Neuroimage* 57, 1601–1607. doi: 10.1016/j.neuroimage.2011.05.043

Kirschen, M. P., Davis-Ratner, M. S., Jerde, T. E., Schraedley-Desmond, P., and Desmond, J. E. (2006). Enhancement of phonological memory following transcranial magnetic stimulation (TMS). *Behav. Neurol.* 17, 187–194.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., and Gruber, O. (2009). Functional architecture of verbal and tonal working memory: an FMRI study. *Hum. Brain Mapp.* 30, 859–873. doi: 10.1002/hbm.20550

Konishi, S., Wheeler, M. E., Donaldson, D. I., and Buckner, R. L. (2000). Neural correlates of episodic retrieval success. *Neuroimage* 12, 276–286. doi: 10.1006/nimg.2000.0614

Krieger-Redwood, K., Gaskell, M. G., Lindsay, S., and Jefferies, E. (2013). The selective role of premotor cortex in speech perception: a contribution to phoneme judgements but not speech comprehension. *J. Cogn. Neurosci.* 25, 2179–2188. doi: 10.1162/jocn_a_00463

Liebenthal, E., Sabri, M., Beardsley, S. A., Mangalathu-Arumana, J., and Desai, A. (2013). Neural dynamics of phonological processing in the dorsal auditory stream. *J. Neurosci.* 33, 15414–15424. doi: 10.1523/JNEUROSCI.1511-13.2013

Lobina, D. J. (2014). Probing recursion. *Cogn. Process.* 15, 435–450. doi: 10.1007/s10339-014-0619-z

Lotze, M., Seggewies, G., Erb, M., Grodd, W., and Birbaumer, N. (2000). The representation of articulation in the primary sensorimotor cortex. *Neuroreport* 11, 2985–2989. doi: 10.1097/00001756-200009110-00032

Makuuchi, M., Bahlmann, J., Anwander, A., and Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8362–8367. doi: 10.1073/pnas.0810928106

Marvel, C. L., and Desmond, J. E. (2010). Functional topography of the cerebellum in verbal working memory. *Neuropsychol. Rev.* 20, 271–279. doi: 10.1007/s11065-010-9137-7

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., and Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Curr. Biol.* 17, 1692–1696. doi: 10.1016/j.cub.2007.08.064

Mueller, J. L., Bahlmann, J., and Friederici, A. D. (2010). Learnability of embedded syntactic structures depends on prosodic cues. *Cogn. Sci.* 34, 338–349. doi: 10.1111/j.1551-6709.2009.01093.x

Mueller, J. L., Friederici, A. D., and Mannel, C. (2012). Auditory perception at the root of language learning. *Proc. Natl. Acad. Sci. U.S.A.* 109, 15953–15958. doi: 10.1073/pnas.1204319109

Nakai, T., and Sakai, K. L. (2014). Neural mechanisms underlying the computation of hierarchical tree structures in mathematics. *PLoS ONE* 9:e111439. doi: 10.1371/journal.pone.0111439

Newman, S. D., and Twieg, D. (2001). Differences in auditory processing of words and pseudowords: an fMRI study. *Hum. Brain Mapp.* 14, 39–47. doi: 10.1002/hbm.1040

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Opitz, B., and Friederici, A. D. (2004). Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity-based learning. *J. Neurosci.* 24, 8436–8440. doi: 10.1523/JNEUROSCI.2220-04.2004

Paulesu, E., Frith, C. D., and Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature* 362, 342–345. doi: 10.1038/362342a0

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J.-L., Nazarian, B., et al. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage* 50, 626–638. doi: 10.1016/j.neuroimage.2009.12.065

Peschke, C., Ziegler, W., Eisenberger, J., and Baumgaertner, A. (2012). Phonological manipulation between speech perception and production activates a parieto-frontal circuit. *Neuroimage* 59, 788–799. doi: 10.1016/j.neuroimage.2011.07.025

Peschke, C., Ziegler, W., Kappes, J., and Baumgaertner, A. (2009). Auditory–motor integration during fast repetition: the neuronal correlates of shadowing. *Neuroimage* 47, 392–402. doi: 10.1016/j.neuroimage.2009.03.061

Pinel, P., Dehaene, S., Rivière, D., and Le Bihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage* 14, 1013–1026. doi: 10.1006/nimg.2001.0913

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10, 15–35. doi: 10.1006/nimg.1999.0441

Prat, C. S., Yamasaki, B. L., Kluender, R. A., and Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain Lang.* 15, 44–50. doi: 10.1016/j.bandl.2016.04.007

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062

Pulvermueller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7865–7870. doi: 10.1073/pnas.0509989103

Raizada, R. D. S., and Poldrack, R. A. (2007). Selective amplification of stimulus differences during categorical processing of speech. *Neuron* 56, 726–740. doi: 10.1016/j.neuron.2007.11.001

Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331

Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., et al. (2005). fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology* 64, 700–706. doi: 10.1212/01.WNL.0000152156.90779.89

Rodriguez-Fornells, A., Cunillera, T., Mestres-Missé, A., and de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3711–3735. doi: 10.1098/rstb.2009.0130

Romero, L., Walsh, V., and Papagno, C. (2006). The neural correlates of phonological short-term memory: a repetitive transcranial magnetic stimulation study. *J. Cogn. Neurosci.* 18, 1147–1155. doi: 10.1162/jocn.2006.18.7.1147

Sato, M., Tremblay, P., and Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain Lang.* 111, 1–7. doi: 10.1016/j.bandl.2009.03.002

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., and Reber, P. J. (2002). Neural correlates of artificial grammar learning. *Neuroimage* 17, 1306–1314. doi: 10.1006/nimg.2002.1291

Smith, E. E., and Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12061–12068. doi: 10.1073/pnas.95.20.12061

Sörös, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., and Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *Neuroimage* 32, 376–387. doi: 10.1016/j.neuroimage.2006.02.046

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., and Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *Neuroimage* 17, 700–709. doi: 10.1006/nimg.2002.1201

Turkeltaub, P. E., and Branch Coslett, H. (2010). Localization of sublexical speech perception components. *Brain Lang.* 114, 1–15. doi: 10.1016/j.bandl.2010.03.008

Uddén, J. (2012). *Language as Structured Sequences: A Causal Role of Broca's Region in Sequence Processing*. Ph.D. Thesis, Karolinska Institutet, Stockholm.

Uddén, J., and Bahlmann, J. (2012). A rostro-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2023–2032. doi: 10.1098/rstb.2012.0009

von Zerssen, G. C., Mecklinger, A., Opitz, B., and von Cramon, D. Y. (2001). Conscious recollection and illusory recognition: an event-related fMRI study. *Eur. J. Neurosci.* 13, 2148–2156. doi: 10.1046/j.0953-816x.2001.01589.x

Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702. doi: 10.1038/nn1263

Zatorre, R. J., Meyer, E., Gjedde, A., and Evans, A. C. (1996). PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cereb. Cortex* 6, 21–30. doi: 10.1093/cercor/6.1.21