Integrating When and What Information in the Left Parietal Lobe Allows Language Rule Generalization

Orpella J[1, 2, 3, 4]*, Ripollés P[4]*, Ruzzoli M[5], Amengual JL[6], Callejas A[1,7], Martinez-Alvarez A[1, 2, 3], Soto-Faraco S[5, 8], de Diego-Balaguer R[1, 2, 3, 8].

[1]Cognition and Brain Plasticity Unit, IDIBELL, L'Hospitalet de Llobregat, Spain

[2]Dept of Cognition Development and Educational Psychology, University of Barcelona, Barcelona, Spain

[3]Institute of Neuroscience, University of Barcelona, Barcelona, Spain

[4]Dept of Psychology, New York University, New York, USA

[5]Center for Brain and Cognition, Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, Barcelona, Spain

[6]Centre de Neuroscience Cognitive Marc Jeannerod, CNRS UMR 5229, Université Claude Bernard Lyon I, Bron, France

[7]Departamento de Psicología Experimental, Facultad de Psicología y Centro de Investigación Mente, Cerebro y Comportamiento, Universidad de Granada, Spain.

[8]ICREA, Barcelona, Spain

* Equal contribution

Corresponding author

Ruth de Diego-Balaguer

Faculty of Psychology

Pg. Vall d'Hebron 171

08035 Barcelona

tel. +34 934 034 768

ruth.dediego@ub.edu

Abstract

A crucial aspect when learning a language is discovering the rules that govern how words are combined in order to convey meanings. Since rules are characterized by sequential co-occurrences between elements (e.g. '***These*** cupcake***s are un***believ***able'***), tracking the statistical relationships between these elements is fundamental. However, statistical learning alone cannot fully account for the ability to create abstract rule representations that can be generalized, a paramount requirement of linguistic rules. Here, we provide evidence that, after the statistical relations between words have been extracted, the engagement of *goal-directed* attention is key to enable rule generalization. Incidental learning performance during a rule-learning task on an artificial language revealed a progressive shift from statistical learning to goal-directed attention. In addition, and consistent with the recruitment of attention, fMRI analyses of late learning stages showed left parietal activity within a broad bilateral dorsal fronto-parietal network. Critically, rTMS on participants' peak of activation within the left parietal cortex impaired their ability to generalize learned rules to a structurally analogous new language. No stimulation or rTMS on a non-relevant brain region did not have the same interfering effect on generalization. Performance on an additional attentional task showed that rTMS on the same parietal site hindered participants' ability to integrate *what* (stimulus identity) and *when* (stimulus timing) information about an expected target. The present findings suggest that learning rules from speech is a two-stage process: following statistical learning, goal-directed attention –involving left parietal regions– integrates *what* and *when* stimulus information to facilitate rapid rule generalization.

Our increasing understanding of the interplay between domain-specific and domain-general cognitive processes has gradually broadened our views on language learning. Apparently simple feats such as the learning of new words are no longer thought to result from the sole workings of a system specialized for language but are known to involve general-purpose mechanisms of statistical learning(1), memory consolidation(2), attention(3) or reward(4). Besides words, linguistic proficiency requires the learning of and, particularly, the ability to *generalize rules*, which involves the development of abstract representations of grammatical categories and an understanding of their inter-relations. While this has been the topic of much research, a fundamental question remains unresolved: what are the brain mechanisms that support language rule learning and generalization?

Given its core relevance, the mechanisms supporting rule learning have been subject to intense debate. By their very nature, language rules are characterized by sequential co-occurrences, often between non-adjacent elements (e.g. ***These*** cupcake***s are un***believ***able***). Therefore, tracking statistical relations is a key computation in this context. Yet, statistical learning alone cannot fully account for the learning of non-adjacent dependencies(5–7), and an additional role for attention has been posited as necessary(8–14). The deployment of attention has been shown to be crucial in determining what information will be learned(13) and to bias subsequent learning(15). Attention has been also proposed to underpin the ability to transfer the rules of a first language to a structurally similar second language(16,17). Indeed, diverting attention interferes with rule generalization(18). However, and despite the increasing support for the implication of attention in rule learning and generalization, there is currently no direct evidence for how attention interacts with statistical learning during the extraction of language rules, and how this interaction facilitates generalization. Moreover, the brain networks enacting this interaction for language rule learning and generalization remain largely unspecified.

In the current article, we propose that attention is essential in the process leading from the extraction of statistical regularities to the generalization of rules. In particular, we capitalized on prior research on attention to test the hypothesis that different types of attention interact with statistical learning to support language rule learning and generalization.

Research in the domain of attention posits the distinction between two networks for attention orienting(19,20): a ventral fronto-parietal network, running along perisylvian areas from the inferior parietal lobule (IPL) to the inferior frontal gyrus (IFG), involved in the automatic detection of behaviorally relevant stimuli(19,20) (i.e., *stimulus-driven attention*); and a more dorsal and bilateral fronto-parietal attention network, including the dorsolateral prefrontal cortex (dlPFC) and the intra-parietal sulcus (IPS) extending to the superior parietal lobe (SPL)(19), involved in orienting attention towards stimulus features based on internal goals (i.e., *goal-directed attention*).

An important feature of the ventral attention network is its sensitivity to input statistics, such as the probability of a cue predicting an upcoming target(21). Critically, while predicting a target at a specific spatial location correlates with right-lateralized activity, predicting *when* a target will appear elicits a left-lateralized response(22,23). Given that the relationship between rule elements in speech necessarily involves a temporal dimension, we hypothesized that the tracking of statistical relationships for rule learning will predominantly engage left-lateralized ventral network activity. Supporting this hypothesis, a variety of studies on language rule learning(24–26), as well as on statistical learning for word segmentation from speech(27), report the involvement of a left-lateralized fronto-parietal network in the early stages of exposure to a new language, highlighting an overlooked overlap between the left ventral attention network for temporal orienting and the classic language network in its fronto-parietal component. Importantly, the ventral network is also

thought to inform(20) the dorsal network for the generation and updating of internal models of the environment(21) that will ultimately guide goal-directed attention. Moreover, the right/left specialization in spatial vs. temporal domains observed in the ventral network is also observed in the dorsal network(23) for goal-directed attention.

In line with these observations, we hypothesize that rule learning is a two-stage process involving an initial statistical learning stage, recruiting the (left) ventral attention network, followed by the engagement of goal-directed attention as learning proceeds, associated with the dorsal fronto-parietal attention network. Therefore, we expected goal-directed attention to play a more prominent role at later stages of learning, and to be critical for rule generalization to new languages by providing the relevant temporal structure. Our specific prediction were that: i) the ventral attention network will be primarily activated in response to rule elements in the learning of their statistical contingency; ii) informed by this, the dorsal attention system will, in turn, direct attention towards these relevant elements of the speech stream, both in terms of their identity (i.e., *what* information; e.g. "un" predicting "able" in **un**forgett**able**) and the specific moment in time in which these elements are expected to occur (i.e., *when* information; e.g. word initial/final positions); and iii) the interaction between ventral and dorsal attention networks will be particularly critical in the *generalization* of learned rules to new instances in a language with analogous dependencies between elements. More specifically, we predicted that goal-directed attention to the time-positions of rule elements will facilitate the fast learning of new dependencies by the binding of new *what* (stimulus identity) and *when* (stimulus temporal position) information(28).

Results

In order to test our hypotheses, we followed the protocol illustrated in Figure 1. Participants were exposed to an artificial language in two different sessions: Session 1 tested rule learning, and Session 2 tested rule generalization. In all cases, the artificial language/s used comprised three-word phrases with no semantic content and an embedded rule consisting of a dependency between the initial and final elements of a phrase. Importantly, the results obtained with this kind of stimulus material are similar to those observed with real language rules(29,30) (e.g., ***is*** sing***ing, is*** play***ing***).

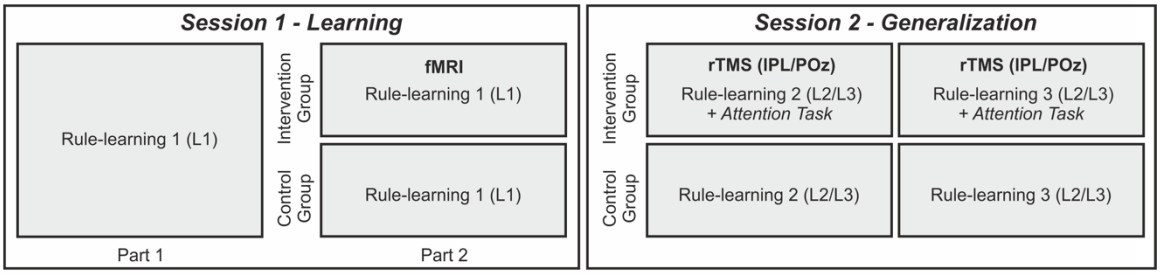

Figure 1. Schematic overview of the protocol. Sessions 1 and 2 were conducted a week apart. Session 1 assessed learning using a single language. In Session 2, generalization was assessed with two new languages that followed an analogous structure to the language learned in Session 1. Languages were counterbalanced across the protocol. rTMS intervention type (lPL/POz) was counterbalanced between participants in the Intervention group. The control group followed the exact same protocol but had no fMRI or rTMS intervention. lPL, left parietal lobe; POz, midline posterior location according to the 10-20 system electrode location. L1, L2, L3 = Languages 1-3.

Session 1 consisted of two parts: in Part 1, all participants were exposed to an artificial language (L1) in an incidental learning task. In Part 2, participants were re-exposed to the same (L1) language and task, but divided in two groups depending on whether they were fMRI scanned (Intervention group, N = 22) or performed the task outside of the scanner (Control group, N = 32). The Control group was included to assess task repetition effects. Session 2 was conducted one week after the initial session. This session was designed to test participants' ability to generalize the rule learned in Session 1 to new but structurally analogous artificial languages (L2 and L3; i.e., L2/L3 followed the same structure as L1 but instantiated in totally new words). In Session 2, participants in the Intervention group (top row, Fig. 1) received repetitive transcranial stimulation (rTMS) on the peak of activation within the *left parietal lobe* (lPL), or to a control site (POz) . The hotspot in IPL was individually identified from the fMRI scans of the secod part of Session 1, and was assumed to reflect goal-directed attention engaged at a later stage of learning, following Part 1. Targeting the lPL with rTMS in Session 2 was expected to hinder participants' goal-directed attention and hence the generalization of the rule learned in Session 1. The Control group (bottom row, Fig. 1) was tested with the same language tasks and with the same delays and order as the Intervention group, but without receiving rTMS. Performance from the Control group additionally informed us about the expected effects of language generalization without rTMS.

In each session, we evaluated rule learning and generalization of non-adjacent dependencies in artificial languages(13). We used two types of blocks. In rule blocks, participants were exposed to three-word phrases containing AXC-type dependencies with the first word A predicting the last word C irrespective of the intermediate word X (e.g. '*jupo* [variable word] *runi*'). In no-rule blocks, three-word phrases with no dependencies (XXX-type strings) were used. Rule and no-rule blocks were interleaved (see *Methods*). Participants were not informed about the presence of rules; their task was to detect the presence or absence of a given target word (e.g. *runi*) which always took the third position in the phrase, in both rule and no-rule blocks.

The word-monitoring task measures incidental learning of the dependencies(13) such that, as learning progresses, participants are expected to start anticipating the presence/absence of the target word (C) upon hearing the first word (A) of the phrase in the rule blocks. Learning progression, evidenced by a reaction time *gain* over trials, can be approximated as a *learning slope* via regression analysis (Fig. 2). A negative learning slope can thus be interpreted as an indication of statistical learning in the initial stages (Fig. 2; see also *Methods*). As learning accrues, we expected participants' reaction times to gradually reach a plateau, indicating that they can now use the initial element of each dependency to consistently anticipate the dependent element. Such plateau should be expressed in later learning stages as stable RTs over the course of the trials (relatively flat learning slope) for the rule blocks, that is a slope similar to no-rule blocks albeit with faster mean reaction times (RTs) in rule blocks compared to no-rule blocks. This RT advantage in rule-blocks (henceforth, *rule effect*), is calculated as the difference in the mean reaction time between rule and no-rule blocks; Fig. 2).

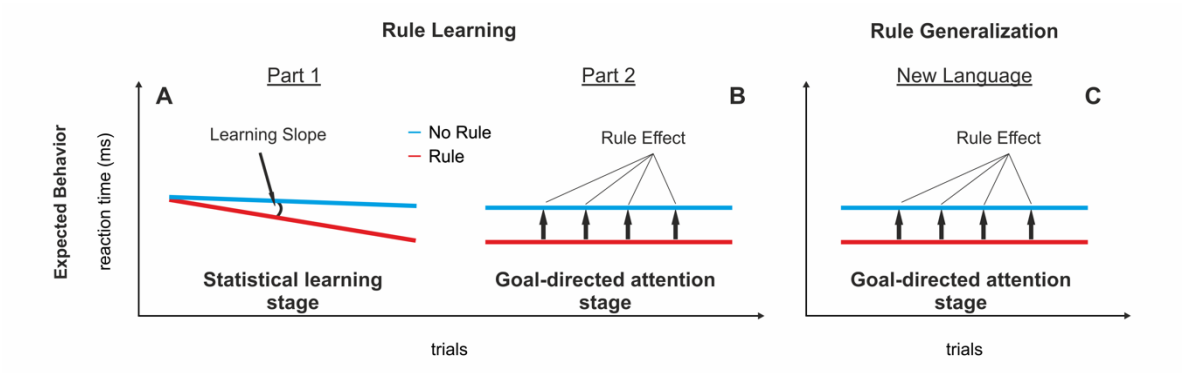

Figure 2. Hypothesized reaction time slopes for rule and no-rule blocks over repetitions of the incidental rule-learning task. For each artificial language learned, participants were exposed to blocks with rules, where the initial word determined the identity of the last word of the phrase, and no-rule blocks where the final word could not be predicted based on the first one. A. Part 1: Reflecting statistical learning, rule blocks are expected to exhibit a greater gain in reaction times across trials than no-rule blocks, as a consequence of the ability to predict the upcoming occurrence or absence of the target word. The difference between rule and no-rule slopes (Learning Slope) is thus a measure of statistical learning indicating progressive rule learning in the early stages. **B.** Part 2: If participants can benefit from previous learnings to orient attention to the initial element to consistently anticipate the final one, their reaction times for rule blocks should plateau in later learning stages and show a sustained difference compared to no-rule blocks throughout (Rule Effect, i.e., the mean difference in reaction time between rule and no-rule trials). **C.** A plateau should also be observed for participants that generalize their attentional focus on initial and final elements to a new language with the same type of dependencies (i.e. rule).

This pattern of behavior is to be expected if participants can apply the knowledge acquired during the initial stage to focus their attention in a goal-directed manner on the initial element of the phrase in order to consistently predict the appearance of the target word at the end of the phrase. Finally, a stable rule effect (i.e., a sustained difference in reaction time between rule and no-rule phrases) should also be expected in the learning of a new but structurally analogous language, given by the generalization of the learned attentional focus to the new material(16,17).

Signatures of statistical learning and attention-based behavior

During the first session, participants (N = 54, 39 women, mean age= 22.61, *sd* = 5.75) performed the incidental rule-learning task on an artificial language (e.g. L1). A short break was provided midway through the session (Session 1, see Fig. 1). Out of the main cohort, a subgroup of participants (Intervention group; N = 22, 13 women, mean age = 23.63, *sd* = 4.67) was randomly selected to undergo the rTMS intervention in Session 2. In order to obtain the appropriate coordinates for rTMS stimulation, participants in the Intervention group performed the second part of the rule-learning task in the MRI scanner (i.e. Rule-learning 1, Part 2). To assess task repetition effects, the remaining participants (Control group; N = 32, 26 women, mean age = 21.9, *sd* = 6.37) were behaviorally tested with the exact same protocol as the Intervention group but without fMRI in Session 1 or rTMS in Session 2.

Figure 3 illustrates changes in RT across trials as the task was repeated across the two parts of Session 1. A linear mixed model approach was used to obtain the *learning slopes* for the two rule learning parts of Session 1, and to compare rule vs. no-rule performance within each part of the session. In addition, *rule effects* were calculated by contrasting mean RTs in rule and no-rule blocks within each part and were assessed via paired *t*-tests (see *Methods*; four subjects were excluded from Part 1 due to missing data).

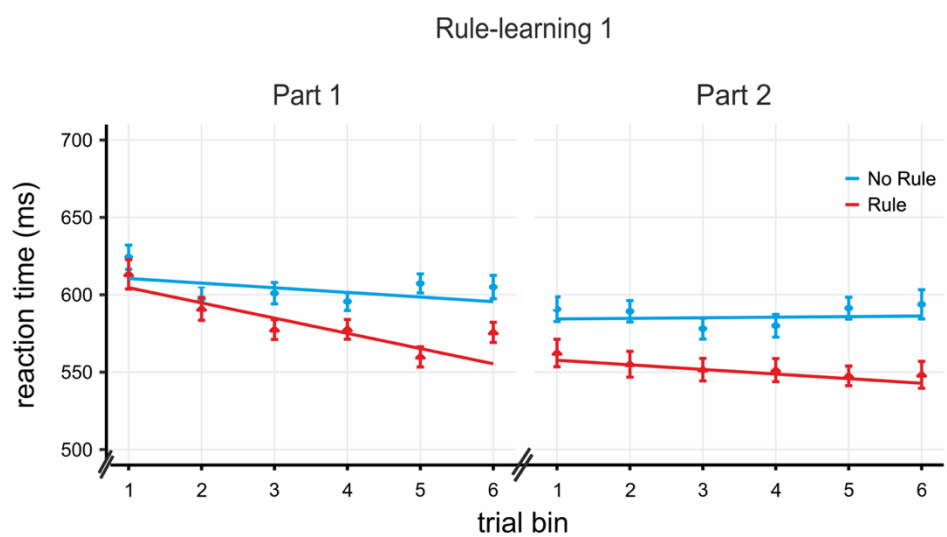

Figure 3. Incidental rule-learning task results for Session 1. Slopes for rule and no-rule blocks (N = 54) over task repetitions derived from the mixed model analysis conform to the expected pattern (Fig. 2), with a significant Learning Slope in Part 1 and a significant Rule Effect with a flat (i.e., non-significant) Learning Slope in Part 2. Actual data shown averaged into 6 trial bins with the SEM.

For the entire sample, a steeper *learning slope* (i.e. more negative) for rule blocks compared to no-rule blocks was observed during Part 1 ($ß_{diff}$ = -0.73, t = -3.86 $p < 0.0002$; $ß_{diff}$ is the estimate of the difference in learning slopes between the rule and no-rule conditions, see *Methods*). The significant difference indicates that learning of the statistical relations between words occurred at this stage. As expected (Fig. 2), non-significant (flatter) slopes for rule blocks were then obtained the second time the task was performed (Part 2: $ß_{diff}$ = -0.36, t = -1.81, $p > 0.07$). Crucially, the slope in the rule condition in Part 1 was significantly more pronounced than that of Part 2 ($ß_{diff}$ = 0.69, t = 3.32, $p <$ 0.001). These changes in the slope were accompanied by a significant *rule effect* in both Part 1 ($t(49)$ = 3.63, $p < 0.001$, d_{Cohen} = 0.513) and Part 2 ($t(53)$ = 4.00, $p < 0.001$, d_{Cohen} = 0.545). The rule effect was equivalent in Parts 1 and 2 (non-significant difference between the two, p = 0.649, d_{Cohen} = 0.065). The combined findings of a non-significant learning slope (significantly flatter than that for Part 1) and a significant rule effect in Part 2 supports the view of a transition from statistical learning to goal-directed attention over successive rule learning parts of Session 1, as predicted (Fig. 2). An analysis of the Intervention and Control groups separately indicated a similar pattern of results in the two groups, except that in the Intervention group some learning was still in progress in Part 2 (Fig. S1 and Fig. S2).

Functional localization of parietal areas for stimulation

With the purpose of selecting the coordinates for the rTMS intervention, we identified parietal regions engaged in the late stages of learning of Session 1 at the individual level. In particular, we localized the peak voxel with the highest activity from single-subject contrast maps for the rule vs. no-rule contrast within the left parietal lobe. The decision to stimulate the left parietal lobe was taken given the critical role this region has in temporal goal-directed attention (23). As mentioned in the introduction, we expected this process to be relevant for rule learning in later stages (Part 2, Session 1; see Fig. 2) as well as during generalization (i.e. Rule Learning Part 2 and beyond). In the cases where the rule vs. no-rule contrast showed no significant activity difference (uncorrected $p <$ 0.05), we used the contrast rule condition against the implicit baseline (N=5). In order to illustrate the variability in individual activation during rule learning, masks for each participant's activation

pattern were calculated using a $p < 0.005$ uncorrected threshold as a cut-off. These individual masks were then added to create a group overlap map. Figure S3 depicts the resulting overlap, showing voxels that were activated in at least 10 participants at the individual level during the rule-learning task. Two subjects were excluded from this analysis due to technical failure during fMRI scanning.

Exploration of the brain networks engaged during late learning stages (Part 2, Session 1)

Given the potential mixture of slow and fast learners in the Intervention group (Fig. S2), we performed an exploratory individual differences analysis of their fMRI activity. This analysis was aimed at investigating the networks related to statistical learning and goal-directed attention by correlating participants' corresponding behavioral measures (i.e. learning slope and rule effects) with the appropriate fMRI contrasts.

To identify the neural activity related to statistical learning, we derived individual *learning slopes* for rule blocks ran in the fMRI (Part 2, Session 1). These individual learning slopes were then correlated with the BOLD signal change in the same blocks (see *Methods*). Covarying with individual differences in the slope for rule blocks, this analysis revealed a ventral fronto-parietal network as related to statistical learning (see Fig. 4A). This ventral network included two large clusters of perisylvian areas in the left hemisphere including the left inferior frontal gyrus (IFG) and the left inferior parietal lobule (IPL), and a smaller cluster centered on the right IFG (Table S1).

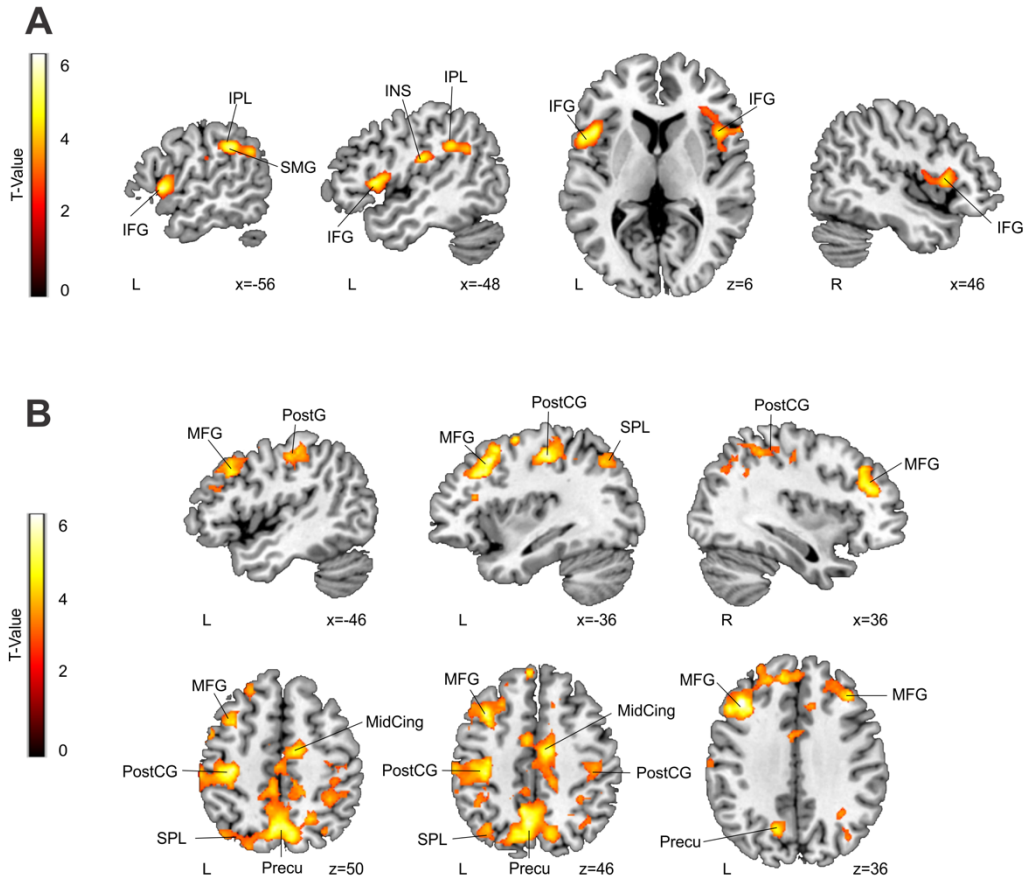

Figure 4. A. BOLD signal over a ventral fronto-parietal network in rule blocks significantly co-varies with their corresponding measure of statistical learning (Learning Slope); the more activity, the greater (i.e. more negative) the slope. **B.** BOLD signal activity over a dorsal fronto-parietal network (rule blocks – no rule blocks) co-varying with the measure of goal-directed attention (Part 2 Rule Effect – Part 1 Rule Effect); the more activity, the larger the effect. Only significant results ($p < 0.05$ FWE-corrected at the cluster level, with an additional $p < 0.001$ at the voxel level and 50 voxels of cluster extent) are shown. IFG = Inferior Frontal Gyrus; IPL = Inferior Parietal Lobule; SMG = Supra-Marginal Gyrus; INS = Insula; MFG = Middle Frontal Gyrus; PostCG = Post-Central Gyrus; SPL = Superior Parietal Lobe; MidCing = Mid-Cingulum; Precu = Precuneus. Neurological convention is used with MNI coordinates shown at the bottom right of each slice.

Since we hypothesized that goal-directed attention would increase over exposure to the same artificial language, to identify neural activity related to this process, we estimated individual measures corresponding to the change in the *rule effect* (rule vs. no-rule RTs) from Part 1 to Part 2 (i.e. Part 2 Rule Effect – Part 1 Rule Effect). We then correlated this *rule effect* increment with the rule *vs.* no-rule fMRI contrast. Individual differences in the increment of the *rule effect* from Part 1 to Part 2 covaried with activity in a more dorsal and bilateral fronto-parietal network, consistent with the engagement of goal-directed attention. This dorsal network included two large clusters centered on bilateral and predominantly superior parietal regions extending to the precuneus, as well as left middle and superior frontal gyri. In addition, there was also an involvement of the left IPL and bilateral basal ganglia (Table S2).

Note that, due to the limited number of subjects, these results can only be interpreted tentatively. However, they do hint in the direction of the hypothetized gradual transition from more ventral fronto-parietal regions for statistical learning to more dorsal fronto-parietal areas for goal-directed attention in later stages of learning.

Is the attention-related network causally related to rule generalization?

We tested the same participants (Intervention and Control groups) one week later (Session 2: Fig. 1) to test whether: (i) the left parietal lobe areas activated in the late stages of learning (functionally localizaed in Session 1) were causally related to rule *generalization*, and (ii) whether goal-directed attention for rule generalization involved the binding of *what* and *when* information. Session 2 thus assessed participants' rule learning performance on two new languages (L2 and L3) that followed the same rule structure as the language (L1) used in Session 1 (i.e. where the initial word determined the identity of the final word). Participants in the Intervention group performed the rule learning task after rTMS on the lPL (individually determined by the functional localization analysis; see *Methods*) and a control site (POz; order of stimulation site and language counterbalanced). Participants in the Control group were also tested in Session 2 with the same new languages (L2 and L3), just like the Intervention group, but without rTMS (Control group Session 2, see Fig. 1). Their performance was used to measure the effects of task repetition without rTMS.

We expected participants' goal-directed attention to be engaged in the transfer of the rule knowledge acquired in Session 1 to structurally similar languages in the rule-learning tasks of Session 2. A mixed model analysis of their RTs indicated, as predicted, that rule-performance in the Control group in Session 2 followed the same pattern observed in Part 2 of Session 1 despite performing the task on a new language. That is, participants showed faster RTs in rule blocks compared to no-rule blocks (rule effect: $t(31) = 2.434$, $p < 0.021$, $d_{Cohen} = 430$) but no significant slope difference between the blocks ($t(31) = -1.6$, $p > 0.11$; Fig. 5). Moreover, the slope for the Control group during Session 2 was not significantly different from their slope for Part 2 in Session 1 ($t(31) = -0.84$, $p > 0.4$).

The same pattern was observed for the Intervention group after rTMS to the control site POz (rule effect: $t(19) = 2.85$, $p < 0.01$, $d_{Cohen} = 0.638$; rule *vs.* no rule slope difference: $t(19) = 0.182$, $p > 0.85$). Indeed, the rule slope under rTMS POz was comparable to that of the Control group ($t(51) = -0.79$, $p > 0.4$) confirming that stimulation to POz did not influence task performance. This pattern of results supports the idea that participants' knowledge of the rule structure acquired early on in learning (Session 1) was used to perform the task in this later phase. Thus, the behavioral plateau achieved in the later stages of Session 1 (Part 2) was maintained when a new language following the same type of rule was being learned and so transfer of the same attentional focus (pertaining to the learned rule structure) could be used in generalization.

Finally, we assessed the effects of rTMS over lPL on rule generalization performance. As can be observed in Figure 5, the pattern of performance in the Intervention group after the application of rTMS on the lPL stood in stark contrast to the pattern seen in both the Control group and in the Intervention group after rTMS to the control site POz (Session 2). First, the slope difference for rule blocks compared to no-rule blocks was significant after rTMS stimulation to the lPL ($ß_{diff} = -0.95$, $t = -3.75$, $p < 0.0002$; Fig. 5), in line with the pattern observed in the early stages of rule learning (i.e., Session 1, Part 1). The slope for rule blocks in this phase was significantly more negative than both the slope after rTMS stimulation to the control site POz ($ß_{diff} = -0.67$, $t = -2.66$, $p = 0.008$) and that of the Control group Session 2 ($ß_{diff} = -0.57$, $t = -2.49$, $p = 0.013$), suggesting that after rTMS over the lPL subjects performed the task capitalizing on statistical learning. To rule out any order effects (stimulated first/second), *order* was added as a factor to the model for both rTMS lPL and rTMS POz phases (*Methods*), yielding non-significant effects in both cases ($p > 0.1$).

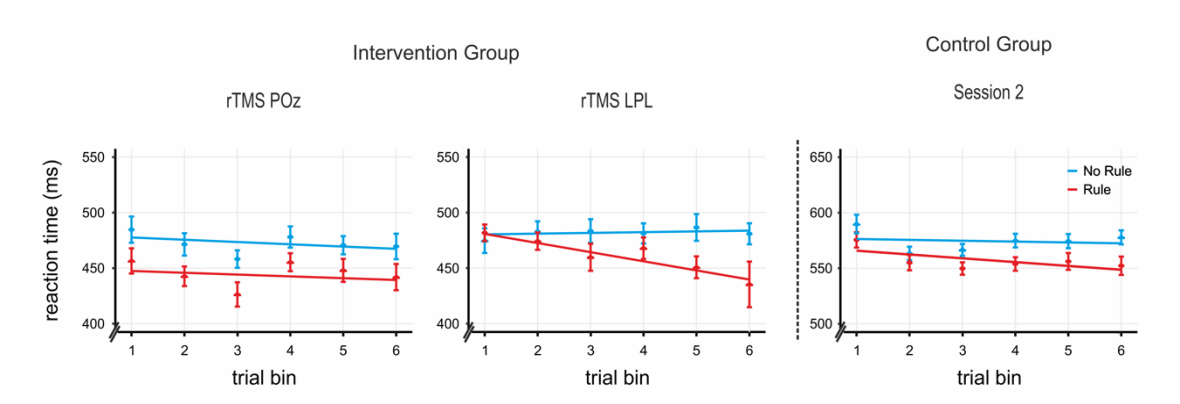

Figure 5. Incidental rule-learning task results for Session 2. Both rTMS POz (left panel) and Control Session 2 (right panel) show the expected pattern of rule learning with a significant Rule Effect and a non-significant Learning Slope, indicating that attentional focus generalized to the learning of the new languages. In contrast, a significant Learning Slope under rTMS lPL effects (center panel) suggests a return to the progressive rule-learning of early learning stages (i.e., statistical learning)**.** Actual data shown averaged into 6 trial bins with the SEM, over the slopes for rule and no-rule derived from the mixed model analysis.

The effects of rTMS on IPL on goal-oriented attention

To discern the precise role of the lPL, participants in the Intervention group performed an additional attention task specifically designed to assess the ability to use content (*what:* e.g. "un" predicting "able" in **un**forgett**able**) and temporal cues (*when:* e.g. word initial/final positions). The attention task was administered under the effects of rTMS to lPL and POz, just like the rule learning task

(Fig. 1), with the order of tasks counterbalanced between participants (Fig. 6). Participants were asked to judge whether the pitch of a target syllable at the end of a sequence of alternating syllables was higher or lower than the preceding sequence. Temporal orienting of attention (*when* information) was manipulated by presenting the sequences of syllables in either rhythmic (with isochronous inter-stimulus intervals) or non-rhythmic form (variable time interval between syllables) while keeping the overall trial length constant. Attention to the content (*what* information) was manipulated by making the first syllable of the sequence informative or uninformative of the identity of the final, target syllable. Participants were familiarized with the attention task before receiving rTMS.

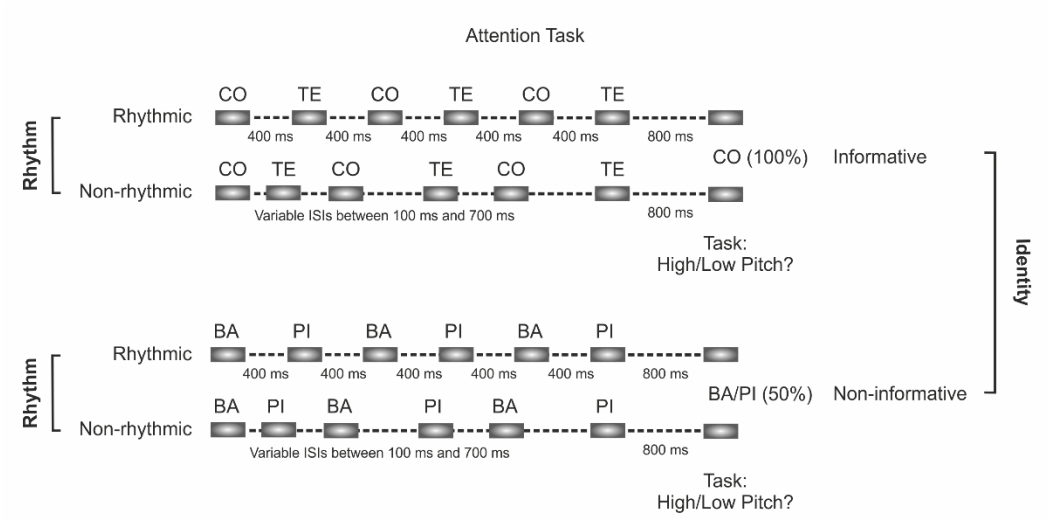

Figure 6. Illustration of the experimental design of the Attention Task. Participants had to judge the pitch of a target syllable presented after a sequence of alternating syllables. The pitch of the target syllable could be either higher or lower than that of the preceding sequence of syllables. Sequences of syllables were presented either rhythmically to engage temporal orienting (attention to *when*) or non-rhythmically (50%), with an otherwise constant trial length. At the same time, the initial syllable of each sequence could be informative or non-informative of the identity of the target syllable (50%), manipulating identity-based attention (i.e. attention to *what*).

We performed a 2x2x2 repeated measures ANOVA on reaction times to the target syllable (correct responses only), with intervention (rTMS_lPL, rTMS_POz), rhythm (rhythmic, non-rhythmic) and identity (informative, non-informative) as within-subject factors. This analysis revealed a main effect of intervention ($F(1,18) = 13.36$, $p < 0.002$, Partial $\eta^2 = 0.426$), with overall reaction times slower after rTMS to the lPL compared to rTMS POz. In order to test whether *what* and *when* cues where used differently in the two intervention conditions we analyzed the interaction between rTMS intervention, Rhythm and Identity. This interaction was significant ($F(1,18) = 7.54$, $p = 0.013$, Partial $\eta^2 = 0.295$), and not affected by the order of rTMS intervention ($F < 1$). We then proceeded to unpack the interaction by analyzing each rTMS intervention separately to observe the specific effects of *what* and *when* manipulations on each intervention. Regarding temporal attention, both under rTMS POz ($F(1,18) = 17.06$, $p < 0.001$, Partial $\eta^2 = 0.487$) and rTMS lPL ($F(1,18) = 19.42$, $p < 0.0001$, Partial $\eta^2 = 0.519$), pitch discrimination in rhythmic sequences was faster than in non-rhythmic sequences, indicating that participants benefited from the use of *when* information in both conditions. Regarding attention to identity, a facilitation in informative sequences was observed under POz stimulation ($F(1,18) = 17.49$, $p < 0.001$, Partial $\eta^2 = 0.493$) but not under rTMS on the lPL. That is, the results showed that rTMS lPL, but not on the control site ($F(1,18) = 2.58$, $p > 0.1$) interfered with the ability to use the identity of the initial syllable to anticipate the target.

Nevertheless, these effects were in turn modulated by a significant interaction between Rhythm and Identity in both rTMS interventions (rTMS POz: $F(1,18) = 6.00$, $p < 0.025$, Partial $\eta^2 = 0.250$; rTMS lPL: $F(1,18) = 5.99$, $p < 0.025$, Partial $\eta^2 = 0.250$). Interestingly these interactions between *what* and *when* effects took opposite directions in each intervention (Fig. 7). Under stimulation to the control site (rTMS POz), *what* and *when* cues appeared to show a facilitatory effect: that is, both cues could be used in conjunction. The benefit of informative cues over non-informative cues was significant in rhythmic sequences ($t(19) = -5.52$, $p < 0.0001$, $d_{Cohen} = 0.596$) but only marginal in non-rhythmic sequences ($t(19) = -1.93$, $p < 0.07$, $d_{Cohen} = 0.231$). Rhythm effects when the identity of the initial syllable was informative ($t(19) = -4.09$, $p < 0.001$, $d_{Cohen} = 0.412$) were also greater than in non-informative sequences ($t(19) = -1.56$, $p > 0.13$, $d_{Cohen} = 0.155$). In contrast, after rTMS to the lPL, the facilitatory effect was not observed: only Rhythm effects were observed when sequences were non-informative ($t(19) = -4.45$, $p < 0.0001$, $d_{Cohen} = 0.503$), and Identity effects when the sequences were non-rhythmic ($t(19) = -2.81$, $p < 0.011$, $d_{Cohen} = 0.313$). Again, order of intervention did not significantly impact on the interaction of interest in none of the intervention conditions (Identity x Rhythm x Order: rTMS POz: $F < 1$; rTMS lPL: $F(1,18) = 2.61$, $p > 0.1$).

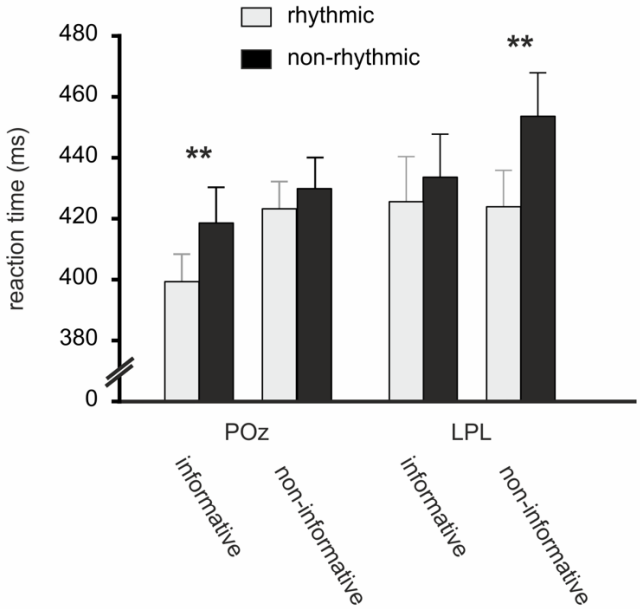

Figure 7. Attention task results. Black and grey shaded bars represent mean RTs for rhythmic and non-rhythmic conditions, respectively, in the informative and non-informative trials for the phases with rTMS on POz and lPL (as specified). $^{**}p < 0.0001$.

The rTMS lPL intervention therefore appeared to impact the interaction between the two forms of attention, that is, on the joint use of *both when* and *what* cues to optimize performance in the pitch discrimination task. In sum, while rhythm and information about future content seem to act jointly to decrease RTs for discrimination, this benefit was lost under rTMS at lPL, where either content or rhythm but *not both concurrently* could be used to improve task performance.

Discussion

In this study, we provide converging behavioral and neuroimaging data in favor of the hypothesis that successful rule learning is a two-stage process. In the early exposure to an artificial language,

participants showed a gradual decrease in response times for rule compared to no-rule phrases consistent with the progressive learning of the statistical relationship between rule elements (Fig. 3, left panel). In the latter part of exposure, and supported by a dorsal fronto-parietal network (Fig. 4, and Fig. S2), a sustained advantage in response times for rule phrases over no-rule phrases (Fig. 3, right panel) suggests instead the engagement of goal-directed attention to focus on the elements pertaining to the rule at this stage. This advantage in response times was maintained for new languages with the same kind of rule. This is consistent with the hypothesis that goal-directed attention in later learning stages facilitates the abstraction of the underlying rule, allowing participants to generalize their acquired knowledge regarding the location (i.e. initial-final) of the elements of the rule to new material (Fig. 5; Control group; and Intervention group, left panel). Moreover, rTMS to a key node of the dorsal network (the lPL) interfered with rule generalization (Fig. 5, Intervention group, right panel), pointing to a further role for goal-directed attention in this stage. Finally, by means of an additional attention task, we were also able to specify the mechanism underlying generalization as carried out by the dorsal network. In particular, rTMS to the lPL hindered participants' ability to integrate stimulus temporal and identity information, that is, information about *what* element is expected, and *when* (Fig. 7), suggesting that rule generalization involves the integration of such features.

Here, by contrasting reaction times over trials for rule and no-rule phrases in an orthogonal word-monitoring task(13), we show that rule learning is a two-stage process involving: i) the progressive learning of statistical contingencies, followed by ii) the engagement of goal-directed attention for rule extraction and generalization. The use of this *online* measure of learning thus enabled us to tap into processes that parallel those arising in natural language learning, where rule knowledge accrues from successive encounters with a particular grammatical rule. Our results -replicated across different sessions and cohorts- further speak for the reliability of this measure of *online* learning, and add substantially to the knowledge gained by the use of test measures acquired after language exposure (i.e. *offline*), which are blind to the learning dynamics(31,32).

We hypothesize that, while still tied to specific content (e.g. the particular words forming a dependency), the ability to direct attention to particular moments in time at later learning stages leads to the emergence of abstract knowledge regarding classes of words (i.e. categories of words that can occur at specific positions in a phrase), which facilitates generalization (*cf.*(33)). Note that category knowledge implies the notion of time-bound category *slots* and *placeholders*, relating to how specific content is apted to take a particular slot in time. What is further required for generalization is the transfer of the learned temporal structure (i.e., the attentional focus) to the new material, as it occurs in the acquisition of a second language(16). That is, generalization simply involves the reverse process of filling in the category slots with the specific word-forms of the new language.

The two-stage account of rule learning evidenced at the behavioral level is also supported by the observed shift in the cortical networks involved. Approximated as a learning slope, statistical learning in the early stages correlated with brain activity in a left-lateralized *ventral* fronto-parietal network (Fig. 4A). In contrast, later-arising sustained rule effects (mean reaction time differences between rule and no rule phrases) engaged a more *dorsal* and bilateral fronto-parietal network (Fig. 4B) related to goal-directed attention(19). Although these fMRI analyses should be taken with caution due to the limited sample size, activity in the reported areas is consistent with the existing literature on both language learning (24,26,27,34–38) and attention(39).

Regarding language learning, several studies have previously implicated regions of the ventral network, such as the inferior frontal gyrus and parietal lobule, in statistical word(36–38) and rule learning(24,26,27,34,35). Regarding attention, on the other hand, the same network has been shown to respond to statistical regularities in the environment (cue-target contingencies(21)), and it is classically associated with the detection of behaviorally-relevant stimuli (i.e. stimulus-driven attention(20)). In our work, the engagement of this ventral network could thus reflect the detection of rule elements (predictive and predicted), which become relevant by virtue of their statistical relationship. Moreover, a left-lateralization of the ventral network has been shown for stimulus-driven temporal orienting(23). This is consistent with our reported left-lateralized activity in this network in response to rules instantiated in *speech*, which necessarily imposes a temporal dimension.

In contrast to the ventral network, the increment in *rule effect* arising in later learning stages was related to activity in a more dorsal and bilateral fronto-parietal network. This dorsal network has been traditionally associated with the goal-directed orienting of attention, in both the spatial(19) and temporal(23) domains. We further confirmed the causal implication of this network in rule generalization by the interference of its left parietal node, identified for each individual during later learning. Indeed, rTMS to this region hindered participants' ability to use acquired rule knowledge for the fast learning of new dependencies. Interestingly, while the same participants stimulated on a non-relevant cortical area (POz) showed the expected generalization behavior, in line with the Control group, participants under the effects of rTMS lPL were seen to return to progressive rule-learning, suggesting that left parietal stimulation did not hinder statistical learning. In other words, rTMS on the lPL did not impede the computation and use of statistical and temporal order information for the gradual learning of the rules. This further speaks for the relative independence of the two mechanisms, congruent with their putative support by the two reported cortical networks.

Finally, we devised an attention task to clarify the mechanisms underlying rule generalization. In particular, we measured participants' ability to use temporal and stimulus identity cues (i.e. information about *when* and *what*, respectively) under the same rTMS effects. Our results show that, although participants could still benefit from either *what* or *when* -type cues, rTMS on the left parietal cortex interfered with the use of both kinds of cues *in conjunction*. This suggests that the binding of stimulus temporal and identity information is required for successful generalization, and is a characteristic of goal-directed attention in latter learning stages.

On a more general note, our results are consistent with the dual-process hypothesis of rule learning from speech that we recently proposed in the developmental domain(14). Specifically, we hypothesized that the maturation of the dorsal fronto-parietal network facilitates the exploitation of knowledge acquired through statistical learning, via attentional mechanisms. Hence, the gradual acquisition of statistical contingencies via incidental learning will allow children after their second year of life to form abstract, long-lasting rule-based knowledge. The results of the present study suggest a similar dual-process mechanism in adult learners when facing a new language.

In conclusion, our data supports the view of two distinct stages in rule-learning, predominantly characterized by the engagement of stimulus-driven and goal-directed attention, respectively. Importantly, we have established a causal link between goal-directed attention as engaged in later learning stages and the abstraction and posterior generalization of language rules. In considering the precise implication and role of attentional mechanisms, therefore, this view of rule-learning integrates seemingly irreconcilable experimental findings and theoretical proposals(5,6,8,18,40).

Materials and Methods

Participants

The study included one group of participants (N = 54, 39 women, mean = 22.61, *sd* = 5.75) who completed the whole protocol (Session 1 and Session 2) under different conditions. A subgroup of 22 right-handed participants (Intervention group: 13 women, mean age = 23.63, *sd* = 4.67), underwent an fMRI session (part 2, Session 1) and rTMS (Session 2). The remaining 32 right-handed participants formed the Control group (26 women, mean age = 21.9, *sd* = 6.37) who was exposed to the artificial languages the same number of times as the Intervention group but without fMRI acquisition or rTMS. The Control group was crucial to characterize the course of events under repeated exposure to the same language task. All participants were native Spanish speakers and had no history of neurological or auditory problems. Participants in the Intervention group were screened for compatibility with fMRI and rTMS procedures(41). The study was approved by the ethics committee of the Universitat de Barcelona, the Universitat Pompeu Fabra (CIEC Parc de Mar) and the ERC ethics scientific office, and was conducted in accordance with the Declaration of Helsinki. Participants were remunerated and completed a written informed consent.

In each part and session of the experiment, several participants within each group were excluded for technical reasons or after being identified as outliers within their group. We considered participants as outliers when their learning slopes were 2 SD above or below the mean for their group and part. In Session 1, two participants in Part 1 (final sample N=30) and one participant in Part 2 (final sample N=31) of the Control group were identified as outliers. For the Intervention group in Session 1, five participants were excluded from Part 1 (one outlier and four for technical reasons; final sample N=17) and two were identified as outliers during Part 2 (fMRI; N=19). In Session 2, none of the participants in the control group were identified as outliers (final sample N=32), while two participants from the Intervention group were excluded (final sample N=20) because they were stimulated > 2 mm from the appropriate coordinate within the lPL during rTMS (see Repetitive TMS protocol section below).

Experimental design

The overall experiment (main text Fig. 1) comprised two sessions that took place approximately a week apart from one another. In Session 1, participants performed an artificial language incidental rule-learning task twice (Part 1, behavioral only for both the Intervention and Control groups; Part 2, behavioral for the Control group and during fMRI scanning for the Intervention one), followed by an offline recognition test (see *Supporting Information* for further details on the offline recognition test). The artificial language (L1) and its non-adjacent dependencies were the same for both Part 1 and 2 of Session 1. For the Intervention group, the fMRI data was used to identify relevant activity within the lPL during rule learning in order to select appropriate coordinates for rTMS stimulation. For the Intervention group only, during Session 2, 1Hz rTMS (15 minutes off-line) was used to stimulate the individually determined max peak of activation within the left parietal lobe (rTMS lPL phase) or a task-irrelevant brain area (POz electrode location; rTMS POz phase). The order between the rTMS lPL phase and the rTMS POz phase was counter-balanced across participants. After rTMS application, participants performed two tasks: a language task similar to that in Session 1; and an attention task, which tested for the effects of rTMS (lPL/POz) on the goal-directed use of content and temporal cues (i.e., *what* and *when* information). Importantly, the total duration of the attention plus rule learning tasks was kept within 30 minutes so as to ensure post rTMS performance under the estimated rTMS effects (see section on Repetitive TMS protocol

below). During Session 2, the Control group completed the same attention and language tasks as the Intervention group but without rTMS stimulation. The languages used for the language rule-learning task during Session 2 were different from the one used for Session 1 (a total of three differentiated artificial languages were thus used: one for Session 1 and two for Session 2, counterbalanced; see Rule Learning Task section below).

Rule Learning Task

Three different artificial languages counting twenty-eight bi-syllabic (consonant-vowel-consonant-vowel) nonsense words each were created. One language (L1) was used in Session 1, the remaining two languages (L2 and L3) in Session 2 (main text, Fig. 1). The order of languages was counterbalanced between sessions and participants.

Twenty-eight words were created for each language stream and synthesized using Mbrola speech synthesizer software(42) by concatenating diphones from the Spanish male database (http://tcts.fpms.ac.be/synthesis/) at 16 KHz. Words (385 ms), were combined using Adobe Audition® software to form three-word phrases with 100 ms gaps between word. Phrase stimuli were presented via Presentation® software (Neurobehavioral Systems) through appropriate headphones and at a volume level adjusted for the participant.

A total of 96 rule phrases and 96 filler phrases were used in this task. Rule phrases conformed to an AXC structure whereby the initial word (A) always predicted the final word (C), while the middle word (X) was variable. Two different A_C dependencies (A1_C1 and A2_C2) were created out of 4 words from the total word pool. The remaining 24 served as middle (X) elements for each of the two A_C dependencies. The transitional probability was always 1 between A and C elements, 0.04 between A and X, and 0.5 between X and C. Half of the filler phrases consisted in the combination of 3 of the 24 X elements and so took the form XXX, with the only constraints that each X had an equal probability to appear in each position but could never appear twice in the same phrase. The combination of two XX elements (following the same constraints as for XXX) and C1 (C2) formed the remaining fillers. Note that these set of fillers, therefore, conformed to an XXC structure where, in contrast to the rule block, C elements could not be predicted on the basis of previous elements. Since two blocks of rule and no rule phrases were presented, a total of 96 rule phrases (48x2) and 96 filler phrases (48x2) were used over the entire task.

Participants were presented with the randomized 96 rule and 96 filler phrases in four alternated rule and no-rule blocks, with the order of blocks counterbalanced between participants. In the fMRI version of the task, data was acquired in two runs, including a block of rule and no-rule each (counterbalanced). A short break was given between the second and third blocks (i.e., between runs in the fMRI). A single offline recognition test was issued after the fourth block (see Supporting Information). In order to obtain a measure of incidental rule-learning, participants performed a cover word-monitoring task. Specifically, they were instructed to detect, as fast and accurately as possible via a button press, the presence or absence of a given target word, which was always one of the C elements (C1 or C2, counterbalanced). A given target word remained constant for each participant throughout the experiment and was displayed in the middle of the screen at all times for reference during the blocks. Participants were not informed about the presence of rules. Inter-trial interval was jittered, using pseudo-random values between 1000 and 3000 ms for optimal fMRI acquisition, and fixed at 500 ms in the remaining phases. A maximum of 1000 ms after the end of a given phrase was allowed for participants to respond before the next trial started. Reaction times (RTs) were calculated from the onset the last word in the phrase until button press. Performance in

interleaved rule and no rule blocks was jointly analyzed by concatenating blocks of a same kind. Only correct response trials with RTs within mean ± 2sd were included for the analysis (mean over groups and conditions of 5.57 ± 3.55% of total trials removed; note that the rejection rate for the comparisons of interest was similar between groups: 4.79 ± 1.36% for the rTMS IPL phase and 4.81 ± 1.61% for the rTMS POz phase in the Intervention group and 5.73 ± 3.44% for the rTMS phase in the Control group).

We reasoned that if incidental rule-learning occurs over exposure in the rule block, participants' gradual ability to predict the appearance or non-appearance of a target word C_j on the basis of the identity of the initial word A_j should be reflected in a RT gain (i.e., faster RTs) over trials within the blocks (see section Linear Mixed Models analysis). We also expected an overall RT advantage over target words in the no-rule blocks (*rule effect*), where prediction is possible (no prediction can be made during no-rule blocks). Participants' *rule effect* for the different parts/sessions was calculated as the mean RT difference between no-rule and rule trials.

Linear Mixed Models analysis

In order to assess online rule learning in the Rule Learning Task within each experimental session, we used a linear mixed model approach to fit learning slopes that reflect reaction time gains over trials for the rule and no-rule conditions. Analyses were performed using the *lme4*(43) and lmerTest(44) packages as implemented in the R statistical language (R Core Team, 2012). Our basic model included reaction time (*rt*) and *trial* as continuous variables, no-rule/rule (*NR_R*) and target/no target (*TNT*) as 2-level factors, and *subject* as a factor with as many levels as participants.

$$rt \sim NR_R + trial + NR_R*trial + TNT + (1|subject) \quad (1)$$

$$rt = ß0\ (\text{intercept}) + ß1(NR_R) + ß2(trial) + ß3(TNT) + ß4(NR_R*trial) \quad (2)$$

$$rt \sim TNT + (1+NR_R*trial|suje) \quad (3)$$

As detailed in (1), which shows the specified model, *NR_R*, *trial* and their interaction were introduced as fixed effects terms. *TNT* was also included here albeit as a predictor of no interest. To allow for a different intercept per participant and so account for inter-subject variability in basal response speed, *subject* was introduced as a random effect. The algebraic expression of the fixed effects part of the model is given in (2). Note that, in this this model, ß4 (*NR_R*trial*) represents an estimate of the difference in learning slopes between the rule and no-rule conditions and can thus be used as a de-trended learning slope estimate for the Rule condition. For the sake of clarity, we have referred to this estimate as $ß_{diff}$. A statistically significant negative $ß_{diff}$ value therefore indicates that online rule learning effectively took place in that experimental phase over and beyond any reaction time gain that may be attributed to within-phase practice effects. Individual slopes were estimated for correlations via a random slopes model specifying the interaction term in the random effects part (3).

Attention task

In addition to the Rule Learning Task, during Session 2, participants in the Intervention group also performed an attention task with the aim of better characterizing the impact of rTMS lPL and rTMS POz. Note that all tasks were completed within 30 minutes in order to ensure post rTMS performance under the estimated rTMS effects for both tasks. Half of the subjects performed the language task first and then the Attention Task while the other half took the reverse order under each of the rTMS interventions (see main text Fig 1).

In the Attention Task (main text Fig. 6), participants were asked to make a pitch discrimination judgment (higher or lower) on an isolated syllable that followed a sequence of alternating syllables. A pool of 200 ms-long syllables at three different pitch heights (249, 440 and 554 Hz) was created from four original tokens ('ba', 'co', 'pi' and 'te') previously recorded by a female native Spanish speaker. All syllables were normalized for homogeneous output volume using ADOBE AUDITION (Adobe Systems Software Ireland Ltd.). Variants at 440 Hz were used to build sequences of six alternating stimuli consisting of either the syllables '/ba' and '/pi' or the syllables '/co' and '/te'. A higher (554 Hz) or lower pitch (440 Hz) syllable was then attached 800 ms following the end of each sequence and represented the target syllable upon which the pitch discrimination was to be judged. Sequences could be either informative, the first syllable matching the target syllable in identity, or non-informative, the identity of the target syllable could not be anticipated (50% probability) from the first syllable. The time interval between the syllables of a given sequence was also manipulated to be either constant (fixed at 400 ms; rhythmic condition) or pseudo-random (ranging between 100-700 ms; mean 400 ms; non-rhythmic condition). The overall design, therefore, conformed to a standard 2x2 orthogonal design with factors Identity (informative/non-informative) and Rhythm (rhythmic/non-rhythmic).

Participants were instructed to judge whether the target syllable was either higher or lower in pitch compared to the preceding syllables in the sequence, responding as fast as possible while avoiding errors. Participants were also informed about the Identity factor and they were encouraged to use informative sequences to anticipate target appearance. No reference to sequence rhythm was made. Each trial began with a fixation cross of variable duration (500-1000 ms) followed by the syllable sequence and target syllable. A maximum of 2000 ms was allowed for response after target presentation. An inter-trial interval of 1900 ms followed responses, after which the next trial started.

Participants ran a set of 16 practice trials (with only rhythmic sequences) in which speed and accuracy feedback was given after each trial. This was included in order to ensure proper understanding of the task. A familiarization session was then completed, consisting of three blocks of 32 sequences each. The Attention Task were completed twice, after rTMS lPL and rTMS POz. Each run consisted of two blocks of 32 sequences each. Experimental conditions were manipulated within blocks, pseudo-randomly. Only correct responses were included for the analysis. Outlier trials per condition on the basis of a mean ± 2.5 *sd* criterion were also removed from further analysis (rTMS POz: 6.7% of trials; rTMS lPL: 6.3% of trials). A repeated measures ANOVA was used to assess the effects of the factors [within subject factors: Intervention (rTMS lPL, rTMS POz), Rhythm (rhythmic, non-rhythmic), Identity (informative/non-informative); and between subject factor: Intervention Order]. An ANOVA for each intervention type was performed, with interactions therein explored via paired-samples *t*-tests.

fMRI acquisition and apparatus

Functional T2*-weighted images were collected using a General Electric MRI 3T scanner and a gradient echo-planar imaging sequence to measure blood oxygenation level dependent (BOLD)

contrast over the whole brain [repetition time (TR) 2000 ms, echo time (TE) 30 ms; 35 slices acquired in ascending interleaved order; slice-thickness: 3 mm with a 0.3 mm gap, 64 × 64 matrix, in plane resolution = 3 × 3 mm; flip angle, 90 º]. Two main fMRI runs with 325 images were acquired for the rule-learning task. Structural images were collected using the MPRAGE-sequence with the following parameters: TR = 7.3 ms, TE/-TI = 2.6, flip angle = 90°, FOV = 256 × 256 × 160 mm, spatial resolution = 1mm^3/voxel. Auditory stimuli were presented using an amplifier (Sensimetrics^{TM}, Malden, MA, USA) attached to the earphones, and visual stimuli were presented using MRI-compatible goggles.

fMRI preprocessing and regression analysis

Data were preprocessed using Statistical Parameter Mapping software (SPM8, Wellcome Trust Centre for Neuroimaging, University College, London, UK, www.fil.ion.ucl.ac.uk/spm/). The two functional runs were first realigned and a mean image of all the EPIs was created. This mean image was then co-registered to the T1 image, which was then segmented into grey and white matter by means of the Unified Segmentation algorithm(45). The resulting normalization parameters were applied to normalize the whole functional set to the MNI space. Finally, functional EPI volumes were re-sampled into 2×2×2 mm voxels and spatially smoothed with an 8 mm FWHM kernel.

An event-related design matrix was specified using the canonical hemodynamic response function. Trial onsets were modeled at the moment of the presentation of the first word in a sentence. Two main conditions were modeled: rule and no-rule. A temporal 1st order modulator and a parametric modulator (we included, for each trial, its reaction time) were added to each condition. Data were high-pass filtered (to a maximum of 1/90 Hz) and serial autocorrelations were estimated using an autoregressive (AR(1)) model. Remaining motion effects were minimized by also including the estimated movement parameters in the model. First-level contrasts were specified for all participants for the main conditions against the implicit baseline. In addition a rule vs. no-rule contrast was also calculated at the individual level. First level contrast images were fed to two second level one sample models with one covariate each to calculate correlations between brain activity and behavior: rule vs implicit baseline contrasts were correlated with individual learning slopes and rule vs no-rule contrasts were correlated with the increment in the rule effect in Part 2 over Part 1.

Group results are reported at a $p < 0.05$ corrected threshold at the cluster level with 50 voxels of cluster extent, with an additional uncorrected $p < 0.005$ threshold at the voxel level. Anatomical and cytoarchitectonical areas were identified using the Automated Anatomical Labelling Atlas(46) (Tzourio-Mazoyer et al., 2002) included in the xjView toolbox (http://www.alivelearn.net/xjview8/).

Repetitive TMS protocol

Transcranial magnetic stimulation was delivered through a figure eight coil attached to a standard MagStim Rapid 2 stimulator (Maximum stimulation output 2.2T). The rTMS was delivered off-line, 15 minutes before the Rule Learning and the Attention tasks, at a frequency of 1Hz and an intensity of the 60% of the maximum stimulation output. This protocol is known to decrease the excitability in the cortical regions beneath the coil position(47) for an average duration of 30 minutes(48) after the end of the stimulation period.

Participants performed two rTMS conditions corresponding to: i) the target stimulation site (obtained from the peak activation of the BOLD signal on the lPL; rTMS lPL phase); and ii) the

control stimulation site (POz electrode location according to the 10/20 EEG international system; rTMS POz phase). rTMS was applied within an average of 3 mm from that peak.

The Control group did not receive rTMS but followed the same behavioral protocol. Since no rTMS was applied, data from L2 and L3 sessions were merged for this group.

When applied onto the lPL, the handle of the coil was angled at 45º away from the midline. In contrast, the coil was placed in a vertical position (with the coil handle pointing backwards) for stimulation at POz. Stimulation sites were identified on subject's scalp using the SofTaxic navigator system (EMS, Italy) according to the T1 image and a marker of the individual target location, using the coordinates of the functional localizer (as described above). Mean stimulation MNI coordinates located at the lPL were x = -47.85 ± 6.91; y = -39.6 ± 8.62; z = 48.8 ± 5.52 (see Table S3 for individual coordinates). SofTaxic navigator system (EMS Italy) was also used for coil maintenance during the rTMS application (allowed error < 2 mm in 3D space). Two subjects were stimulated beyond 3 mm away from the peak of activation and were thus excluded from rTMS analyses.

Author Contributions

R. d. D.-B., J. L. A and A. C. designed the research. J. O., P. R., M. R., J. L. A., A. C., S. S.-F., and R. d. D.-B. wrote the manuscript. M. R., J. L. A., A. M.-A. acquired the data. J.O., P. R. analyzed the data.

Acknowledgements

This work was supported by the European Research Council grant ERC-StG-313841 (R. d. D.-B.), the European Research Council (Proof of Concept, ERC, 727595 (S. S.-F.), and the Juan de la Cierva Post-Doctorate Fellowship (JCI-2012-12335, Ministerio de Economia y Competividad (M. R.)

References

1. Saffran JR, Aslin RN, Newport EL. Statistical Learning by 8-Month-Old Infants. Science (80-) [Internet]. 1996 [cited 2019 May 7];274(5294):1926–8. Available from: http://linguistics.berkeley.edu/~kjohnson/ling290e/saffran_et_al_1996.pdf

2. Davis MH, Gaskell MG. A complementary systems account of word learning: neural and behavioural evidence. Philos Trans R Soc Lond B Biol Sci. 2009;364(1536):3773–800.

3. Smith LB. Learning How to Learn Words: An Associative Crane. In: Becoming a Word Learner: A Debate on Lexical Acquisition. 2010.

4. Ripolles P. The role of reward in word learning and its implications for second language acquisition.

5. Gómez RL. VARIABILITY AND DETECTION OF INVARIANT STRUCTURE. Psychol Sci [Internet]. 2002 May;13(5):431–6. Available from: http://doi.wiley.com/10.1002/da.20417

6. Peña M, Bonatti LLL, Nespor M, Mehler J. Signal-driven computations in speech processing. Science (80-) [Internet]. 2002;298(5593):604–7. Available from: http://www.sciencemag.org/content/298/5593/604.abstract%5Cnhttp://www.sciencemag.org/content/298/5593/604.short

7. Newport EL, Aslin RN. Learning at a distance I. Statistical learning of non-adjacent dependencies. Cogn Psychol. 2004;48(2):127–62.

8. Perruchet P, Pacton S. Implicit learning and statistical learning: one phenomenon, two approaches. Trends Cogn Sci. 2006;10(5):233–8.

9. Pacton S, Sobaco A, Perruchet P. Is an attention-based associative account of adjacent and nonadjacent dependency learning valid? Acta Psychol (Amst). 2015;157(1):195–9.

10. De Diego Balaguer R, Toro JM, Rodriguez-Fornells A, Bachoud-Lévi AC. Different neurophysiological mechanisms underlying word and rule extraction from speech. PLoS One. 2007;2(11).

11. de Diego-Balaguer R, Fuentemilla L, Rodriguez-Fornells A. Brain Dynamics Sustaining Rapid Rule Extraction from Speech. J Cogn Neurosci [Internet]. 2011;23(10):3105–20. Available from: http://www.mitpressjournals.org/doi/10.1162/jocn.2011.21636

12. Rodriguez-Fornells A, Cunillera T, Mestres-Misse A, de Diego-Balaguer R. Neurophysiological mechanisms involved in language learning in adults. Philos Trans R Soc B Biol Sci [Internet]. 2009;364(1536):3711–35. Available from: http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2009.0130

13. López-Barroso D, Cucurell D, Rodríguez-Fornells A, de Diego-Balaguer R. Attentional effects on rule extraction and consolidation from speech. Cognition. 2016;152:61–9.

14. De Diego-Balaguer R, Martinez-Alvarez A, Pons F. Temporal Attention as a Scaffold for Language Development. Front Psychol [Internet]. 2016;7(February):1–15. Available from: http://journal.frontiersin.org/Article/10.3389/fpsyg.2016.00044/abstract

15. Jiang Y, Chun MM. Selective attention modulates implicit learning. Q J Exp Psychol. 2010;54(4):1105–24.

16. Ellis NC. Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. Appl Linguist. 2006;27(2):164–94.

17. Robinson P. Language Awareness and Multilingualism. Cenoz J et al., editor. 2017;1–10. Available from: http://link.springer.com/10.1007/978-3-319-02325-0

18. Toro JM, Sinnett S, Soto-Faraco S. Speech segmentation by statistical learning depends on attention. Cognition. 2005;97(2):25–34.

19. Corbetta M, Shulman GL. Control of Goal-Directed and Stimulus-Driven Attention in the Brain. Nat Rev Neurosci [Internet]. 2002;3(3):215–29. Available from: http://www.nature.com/doifinder/10.1038/nrn755

20. Corbetta M, Patel G, Shulman GL. The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron. 2008;58(3):306–24.

21. Macaluso E, Doricchi F. Attention and predictions: control of spatial attention beyond the endogenous-exogenous dichotomy. Front Hum Neurosci [Internet]. 2013;7(October):685. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3800774&tool=pmcentrez&rendertype=abstract

22. Coull JT, Nobre a C. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. J Neurosci [Internet]. 1998;18(18):7426–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9736662

23. Coull JT, Cotti J, Vidal F. Differential roles for parietal and frontal cortices in fixed versus evolving temporal expectations: Dissociating prior from posterior temporal probabilities with fMRI. Neuroimage. 2016;141:40–51.

24. Opitz B, Friederici AD. Brain Correlates of Language Learning: The Neuronal Dissociation of Rule-

Based versus Similarity-Based Learning. J Neurosci [Internet]. 2004;24(39):8436–40. Available from: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2220-04.2004

25. Bahlmann J, Schubotz RI, Friederici AD. Hierarchical artificial grammar processing engages Broca's area. Neuroimage. 2008;42(2):525–34.

26. Goranskaya D, Kreitewolf J, Mueller JL, Friederici AD. Fronto-Parietal Contributions to Phonological Processes in Successful Artificial Grammar Learning. Front Hum Neurosci. 2016;10:1–14.

27. Karuza EA, Newport EL, Aslin RN, Starling SJ, Tivarus ME, Bavelier D. The neural correlates of statistical learning in a word segmentation task: An fMRI study. BRAIN Lang. 2013;127(1):1–21.

28. Treisman A. How the deployment of attention determines what we see [Internet]. [cited 2019 May 8]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1832144/pdf/nihms-8070.pdf

29. Pelucchi, B., Hay, J.F. and Saffran J. Statistical Learning in a Natural Language by 8-Month-Old. Child Dev. 2009;80(3):674–85.

30. Friederici AD, Mueller JL, Oberecker R. Precursors to natural grammar learning: Preliminary evidence from 4-month-old infants. PLoS One. 2011;6(3).

31. Karuza EA, Emberson LL, Aslin RN. Combining fMRI and Behavioral Measures to Examine the Process of Human Learning. Neurobiol Learn Mem. 2014;1–31.

32. Siegelman N, Bogaerts L, Kronenfeld O, Frost R. Redefining "Learning" in Statistical Learning: What Does an Online Measure Reveal About the Assimilation of Visual Regularities? Cogn Sci. 2017;1–36.

33. Endress AD, Bonatti LL. Rapid learning of syllable classes from a perceptually continuous speech stream. Cognition. 2006;105(2):247–99.

34. Opitz B, Friederici AD. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. Neuroimage. 2003;19(4):1730–7.

35. Tettamanti M, Alkadhi H, Moro A, Perani D, Kollias S, Weniger D. Neural correlates for the acquisition of natural language syntax. Neuroimage. 2002;17(2):700–9.

36. Cunillera T, Càmara E, Toro JM, Marco-Pallares J, Sebastián-Galles N, Ortiz H, et al. Time course and functional neuroanatomy of speech segmentation in adults. Neuroimage [Internet]. 2009;48(3):541–53. Available from: http://dx.doi.org/10.1016/j.neuroimage.2009.06.069

37. McNealy K, Mazziotta JC, Dapretto M. Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing. J Neurosci [Internet]. 2006;26(29):7629–39. Available from: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5501-05.2006

38. López-Barroso D, Catani M, Ripollés P, Dell'Acqua F, Rodríguez-Fornells A, de Diego-Balaguer R, et al. Word learning is mediated by the left arcuate fasciculus. Proc Natl Acad Sci U S A [Internet]. 2013;110(32):13168–73. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3740909&tool=pmcentrez&rendertype=abstract

39. Corbetta M, Shulman GL. Control of Goal-Directed and Stimulus-Driven Attention in the Brain. Nat Rev Neurosci [Internet]. 2002;3(3):215–29. Available from: http://www.nature.com/doifinder/10.1038/nrn755

40. Endress AD, Mehler J. Primitive computations in speech processing. Q J Exp Psychol. 2009;62(11):2187–209.

41. Rossi S, Hallet M, Rossini PM, Pascual-Leone A. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin

Neurophysiol [Internet]. 2009;120(12):2008–39. Available from: http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/dltr2011&section=17

42. Dutoit T, Pagel V, Pierret N, Bataille F, van der Vrecken O. The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes. Proceeding Fourth Int Conf Spok Lang Process ICSLP '96 [Internet]. 1996;3(Icslp 96):1393–6. Available from: http://ieeexplore.ieee.org/document/607874/

43. Bates D, Mächler M, Bolker BM, Walker SC. Fitting Linear Mixed-Effects Models Using lme4. Journalo f Stat Softw. 2015;67(1):1–47.

44. Kuznetsova A, Brockhoff PB, Christensen RHB. **lmerTest** Package: Tests in Linear Mixed Effects Models. J Stat Softw [Internet]. 2017;82(13). Available from: http://www.jstatsoft.org/v82/i13/

45. Ashburner J, Friston KJ. Unified segmentation. Neuroimage. 2005;26(3):839–51.

46. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002;15(1):273–89.

47. Muellbacher W, Ziemann U, Boroojerdi B, Hallett M. Effects of low-frequency transcranial magnetic stimulation on motor excitability and basic motor behavior [Internet]. [cited 2019 May 7]. Available from: www.elsevier.com/locate/clinph

48. Valero-Cabré A, Amengual JL, Stengel C, Pascual-Leone A, Coubard OA. Transcranial magnetic stimulation in basic and clinical neuroscience: A comprehensive review of fundamental principles and novel insights. Neurosci Biobehav Rev [Internet]. 2017;83(July):381–404. Available from: https://doi.org/10.1016/j.neubiorev.2017.10.006

49. Macmillan NA, Kaplan HL. Detection Theory Analysis of Group Data. Estimating Sensitivity From Average Hit and False-Alarm Rates. Psychol Bull. 1985;98(1):185–99.