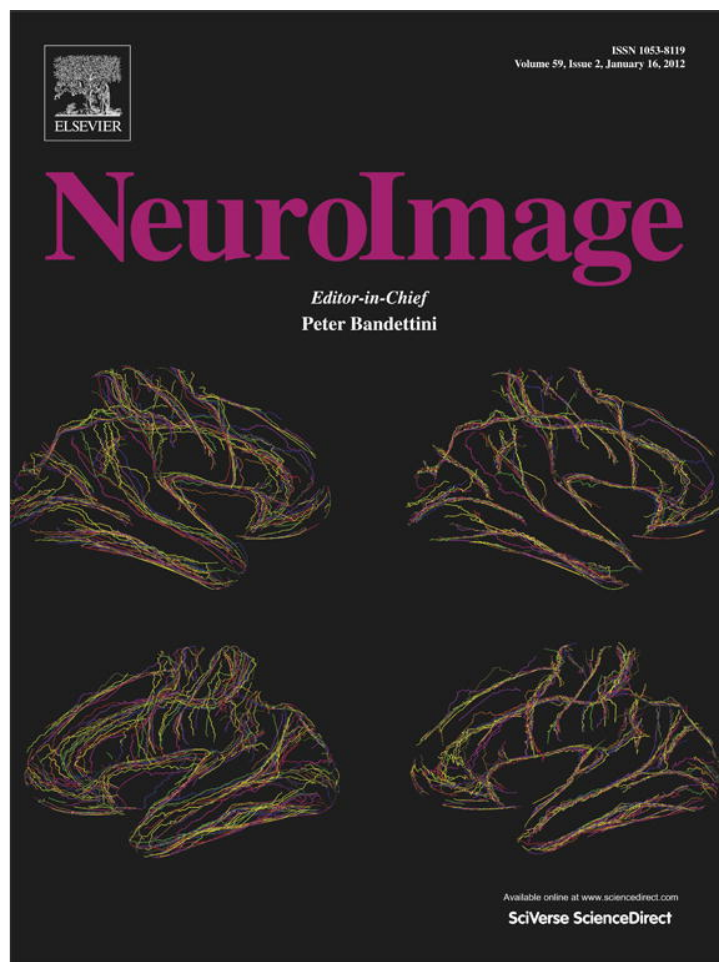


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

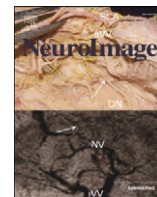
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Full Length Articles

Rule and similarity in grammar: Their interplay and individual differences in the brain

Maximilian F.A. Hauser, Juliane Hofmann, Bertram Opitz*

Experimental Neuropsychology Unit, Saarland University, Saarbrücken, Germany

ARTICLE INFO

Article history:

Accepted 6 February 2012

Available online 18 February 2012

Keywords:

AGL

IFG

Premotor

Rule

Similarity

ROC

ABSTRACT

Previous research on artificial grammar has indicated that the human ability to classify sentences or letter strings according to grammaticality relies on two types of knowledge. One is a superficial, familiarity-based understanding of a grammar the other is the knowledge of rules and critical features underlying a grammar. The fundamentally different characteristics of these systems permit an analysis of receiver-operating characteristics (ROC), which measures the extent to which each type of knowledge is used in grammaticality judgments. Furthermore, violations of a grammar can be divided into hierarchical and local violations. The present study is the first to combine the use of ROC analyses, fMRI and a grammaticality dichotomy. Based on previous neuroimaging studies, it was hypothesized that judgments based on rule knowledge, as extracted from individual ROC analyses, involve the left inferior frontal gyrus (IFG), whereas similarity would involve right IFG, as well as left hippocampal regions. With regards to violation types, it was hypothesized that hierarchical violations would recruit the opercular part of the left IFG as well as the posterior operculum, whereas local violations would bilaterally activate the premotor cortex (PMC). Results indicated that for greater reliance on rule knowledge, a ventral part of the left PMC was activated for ungrammatical items, whereas other PMC areas show a differentiated response for grammaticality for individuals less reliant on similarity. The right IFG was related to ungrammatical items as a function of similarity. Results are discussed with regards to possible error detection systems and differentiated efficiencies for respective classification strategies.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Since the first investigative ventures into the processes underlying artificial grammar learning (AGL) by Reber (1967), an increasing number of studies has sought to differentiate, identify or even challenge the existence of the main processes supporting the acquisition of a grammar. Symbols, letter strings, or artificial languages, adhering to certain grammatical rules have been used as means of surpassing the usually lengthy process of learning grammatical rules and outlining the key processes that are involved in recognizing when a symbol, letter, or word sequence adheres to, or violates the underlying rules. The Reber-paradigm (Reber, 1967), for example demonstrated that after a certain amount of exposure to letter strings formed in compliance with an underlying grammar, subjects are able to classify items according to grammaticality above chance-level. In recent years an influential view has proposed that two strategies are involved in such a grammaticality judgment task. One is a superficial process in which a mental representation, in this case the *category-exemplar* (Pothos, 2007), is compared to the target item in such a manner that the number of deviations between the two serves as an individual reference for judging an

item. This process is therefore based on knowledge of similarity. The other possible strategy is the acquisition and application of the rule-system making up the grammar itself. Being familiar with the requirements and prohibitions of a grammar allows differentiating grammatical items from their ungrammatical counterparts through critical features, such as the limited number of possible starting letters in the Reber-paradigm. Thus, the two mechanisms for grammaticality judgments can be characterized in different ways. Decisions based on similarity-knowledge are flexible, based on large parts of, or whole items, and dependent on individual thresholds, whereas rule-knowledge provides a clear definition of the grammatical status based on critical features, independent of individual thresholds.

Support for the existence of both mechanisms has been provided by several studies. On one side it was demonstrated that the ability to recognize fragments of grammatical letter strings accounts for performance on judgment tasks, which justifies a similarity based account of AGL (Perruchet and Pacteau, 1990). However, with increasing complexity of a grammatical system, a similarity based account of AGL is not sufficient to explain performance on AGL paradigms. Thus, on the other side the acquisition of rule knowledge was suggested by a study that found that participants were still able to classify items above chance when controlling for similarity-supporting features of letter-strings, such as frequency of chunks, or their position within an item (Lieberman et al., 2004). Furthermore, several studies have shown

* Corresponding author at: Experimental Neuropsychology Unit, Saarland University, Campus A2.4, 66123 Saarbrücken. Fax: +49 681/302 6516.

E-mail address: b.opitz@mx.uni-saarland.de (B. Opitz).

that rule-knowledge is transferable to new vocabulary which would render similarity knowledge of previous vocabulary irrelevant (Beesley et al., 2010; Gomez and Schvaneveldt, 1994).

The dichotomy between similarity and rule-based grammar learning has also been investigated using analyses of receiver-operating characteristics (ROC; Kinder and Assmann, 2000). This analysis has emerged from signal detection theory and describes the ratio of the true positive rate (i.e. the sensitivity), and false positive rate ($1 - \text{specificity}$). In the case of AGL participants are asked to classify test items according to their conformity with some underlying rules, usually on a six-point confidence rating scale, ranging from *surely correct* to *surely incorrect*. The respective classification performance (i.e. true positive rate vs. false positive rate) is plotted as the confidence changes. The shape of the ROC can then be fitted by theoretical curves that were derived from the characteristics of similarity and rule knowledge. Assuming that similarity judgments are made on continuous dimension (i.e., more or less similar, cf. Pothos, 2007) the shape of the resulting model function will be a perfectly symmetrical curvilinear ROC. In contrast, rule knowledge in AGL can be described in terms of an all-or-none process or a threshold model; that is a test string either adheres to the rules somebody has learned or not (Dienes et al., 1997). Such a threshold process will generate a strictly linear ROC with a slope less than 1. If both, rule knowledge and similarity knowledge were applied in AGL, the resulting ROC is somewhere in between the two extreme positions described above, with the curvature reflecting the amount of similarity knowledge.

Kinder and Assmann (2000) have conducted a study using this paradigm, in which participants were first trained on letter-strings using the mnemonic instruction introduced by Reber (1967). After the acquisition phase the participants were informed that the learned strings adhere to some rule and that they should classify the following test items as grammatical or ungrammatical according to these rules on a six-point rating scale. Kinder and Assmann used a z-transformed version of ROC and found that their data can be fully accounted for by similarity-based processes. However, as complete knowledge of the underlying grammatical rules would lead to perfect discrimination ability, whereas similarity-based decisions would be less accurate, overall performance might be a critical aspect when investigating which knowledge type is at work in grammaticality judgments. As performance in Kinder and Assmann's (2000) experiments was rather low, it could be argued that participants were not acquainted enough with the grammar, and therefore had to rely only on knowledge of visual similarity.

In line with this argument, studies looking at the processes involved in AGL at exposure length reported an increase in rule-knowledge and a decrease in similarity-knowledge over time (Fletcher et al., 1999; Opitz and Friederici, 2004). Opitz and Friederici (2004), for example, compared the effects of switching two nouns in their artificial language BROCANTO without changing the underlying rules (i.e. similarity-knowledge) with the effects of switching a noun with its modifier (rule-knowledge). They found that changes affecting similarity-based knowledge led to an initial increase and later decrease of hippocampal activity, whilst the PMC was more and more activated with increasing understanding of rule changes. Therefore, rule knowledge may become evident only after a certain degree of familiarity with the rules of a grammar. For those reasons, a paradigm aiming to improve performance may yield new perspectives, as only higher performing individuals might employ rule knowledge. In the present study, this was accomplished by (1) using an ecologically more valid artificial language and (2) by introducing immediate feedback to the training session, in order to increase subsequent performance.

Other studies have focused on the neural correlates of grammar-related processes and provided further support for the distinction between rule and similarity knowledge as a basis for grammatical judgment tasks. For example, Lieberman et al. (2004) operationalized rule and similarity in the form of item characteristics. This was done by removing or adding several factors known to influence similarity

knowledge, thereby creating targets that were only discernable based on either similarity or rule knowledge. They reported activation in the right caudate that was associated with rule adherence, whereas medial temporal lobe activations were associated with similarity. This however, does not take into account individual differences, as it assumes that every individual has the capacity to process both, rule and similarity features. Therefore the present study seeks to identify processes related to grammaticality and knowledge types using the more fluid ROC-analysis, taking into account individual differences.

Another paradigm that identified anatomical key regions with regards to the two knowledge types was developed by Fletcher and colleagues (Fletcher et al., 1999). In a block-design, participants repeatedly viewed and judged grammatical and ungrammatical items according to their grammaticality within one block of trials. This task was performed in six blocks, each with new items, of which the grammatical ones adhered to the same grammatical structure. Functional images acquired during this task revealed that within-block performance, assumed to reflect similarity knowledge related to an increase in right middle frontal gyrus activation, whereas between-block improvement in performance, as a measure of increasing rule knowledge, was associated with an increase in activation in the left inferior frontal gyrus (IFG). The importance of this brain structure for rule knowledge has been further emphasized by studies comparing the neural correlates of processing local and hierarchical dependencies. Opitz and Friederici (2007) trained their participants on a modified version of the artificial language BROCANTO, containing both types of dependencies and found that violations of hierarchical dependencies were processed in the opercular part of the left IFG. In contrast, the premotor cortex (PMC) supports the processing of local phrase structure dependencies, an interpretation that has been supported by a study with patients suffering from a left or bilateral ventral PMC lesion (Opitz and Kotz, in press). Additionally, the anterior hippocampus was found to respond only to violations of local dependencies, which resulted from higher relational processing demands of the new relationship between visual features of a word and its position within a sentence (Opitz and Friederici, 2007, Strange et al., 2001). With regards to varying degrees of complexity of sentences, Friederici et al. (2006) have investigated the processing demands of grammatical sentences with up to two permutations permitted by the grammar, as compared to canonical sentences and ungrammatical sentences. Increasing complexity was linked to increasing activity in the left inferior pars opercularis, whilst ungrammatical items evoked increased activity in the left frontal operculum.

It should be acknowledged that the present study differed from classical letter-string paradigms, not only in the way the grammar was presented, but also in informing participants about the underlying grammar prior to learning. In most other studies, the rationale for leaving participants naïve with respect to underlying rules until after the training phase was based on the implicit nature of AGL. As proposed by the declarative/procedural model of (second) language learning (Ullman, 2001) a basal ganglia-IFG network plays a crucial role in acquiring knowledge about underlying rules in typical implicit AGL paradigms (cf., Lieberman et al., 2004). This implicit system is complemented by a declarative, lexical system, largely incorporated by the medial temporal lobe, which is responsible for associative knowledge about associations between sounds or visual features and meaning of words. As Ullman (2001) argued second language learning would depend more on declarative systems, even if in first language the same linguistic forms would be processed by procedural systems. Newer work conceptualizes the declarative and procedural systems in a broader framework (see Shohamy et al., 2008, for a review). This proposal emphasizes that the basal ganglia are critical only for specific aspects of learning, namely, for gradual, incremental, feedback-based learning of associations but not for other cognitive strategies, which, nevertheless, might be important for learning. It was further suggested that the basal ganglia are specifically necessary for learning of associations, but may be less critical for mediating performance once associations have been well learned.

Instead, this final performance may be driven by representations in PFC and/or the hippocampus (Shohamy et al., 2008). In sum, the distinction between implicit and explicit knowledge of grammar is a controversial issue, as both are often closely linked (e.g., Pothos, 2007). The present study, however, was not designed to contribute to this debate, but rather investigated the neural underpinnings of superficial and abstract knowledge after having informed and explicitly trained participants.

The present study was not only designed to test the hypothesis that with increasing performance, a ROC-analysis would indeed show involvement of a rule process, but furthermore to investigate the interplay of both knowledge types with grammaticality of items. Regarding this, the hypotheses were as follows:

1. Individuals with high reliance on similarity employ right frontal areas and the hippocampus.
2. Rule reliance correlates with general left prefrontal areas, as well as PMC cortices

Material and methods

Subjects

For the experiment 17 students (7 male; mean age: 24 years, range: 19–30) were recruited and scanned using functional MRI. In all cases the first language was German. Furthermore, all participants were right-handed and had normal or corrected-to-normal vision. Written informed consent was given by all participants, after the background and possible risks of their participation were outlined. One participant was excluded from analysis, due to a malfunction of the response button during the scan.

Materials

The artificial language BROCANTO (Opitz and Friederici, 2003) was used to generate the stimulus material. In this artificial grammar system with a total vocabulary of 14 words (e.g., “trul” or “rix”) from five word categories: nouns (N), verbs (v), determiners (D, d), adjectives (M) and adverbs (m). Word categories contained four members each (except the determiner category, that contained only two words) and were identifiable by particular vowels (e.g., o, u = noun or i, e = verb). Sentences contained three to eight words, adhering to a subject-verb-[object] structure. Subject and object components of a sentence were composed of a determiner, an optional adjective and a noun. Verb phrases contained a verb and an optional adverb. An example of a correct sentence with the dNvm structure would be: *aaf gum pel riifi*. In total 400 correct sentences were formed using the artificial vocabulary, as well as 200 sentences containing violations of the artificial grammar, which was done by alternating a grammatical sentence through replacing vowels symbolizing one category with vowels of another category. Resulting from this manipulation three violation-types were generated, being the repetition of a word-category (e.g. *aaf gum pel *rix*), a violation of the determiner-noun-agreement (e.g. *aak *gum pel riifi*), or a phrase-structure-violation (e.g. *aaf gum *aaf trul pel riifi*). The reader is referred to previous studies (Opitz and Friederici, 2003, 2004) employing this artificial language for further examples and a schematic presentation of the underlying grammar system.

Experimental procedure

Participants were trained in the artificial language two days prior to the final test in the scanner. Training was done over ten blocks, each containing a training phase and a test phase. During the training phase, 20 grammatical sentences were presented on a monitor for 7 s each with the instruction to acquire the underlying grammar. The task in the test phase was to classify 20 new items (10 grammatical and 10 ungrammatical) according to their grammaticality, receiving

immediate visual feedback. The grammaticality judgment was done via six responses indicating the range from surely grammatical, to surely ungrammatical. In the final test-block, which was the only block performed in the scanner, participants were presented with 200 new items and, similarly to the tests in the training phase, were asked to classify items according to their grammaticality, although without feedback. Furthermore, a sensorimotor control task was also implemented in the test block, consisting of a forced-choice button press, related to the presentation of one of two pseudowords (BRAD or DABA).

Data acquisition

T1-weighted structural images and T2*-weighted functional images were made using a Siemens SONATA MR scanner (Erlangen, Germany) operating at 1.5 T with a standard circularly polarized whole head coil. Anatomical images of high-resolution (1 mm³ voxel size), were obtained employing a 3-D MP RAGE sequence. Changes in blood oxygen level dependence (BOLD) were measured with a gradient-echo EPI pulse sequence, using the following parameters: T_R = 1.7 s, T_E = 50 ms, flip angle = 85°, slice thickness = 4 mm, interslice gap = 1 mm, in-plane resolution = 3.5 × 3.5 mm², field of view = 224 mm, 20 axial slices parallel to anterior posterior commissure plane. A total of 1160 volumes were acquired and the initial four volumes were skipped for T1 equilibration.

Data analysis

Receiver operating characteristics

We defined true positives as the correct identification of grammatical sentences (“gr” | gr; “grammatical” answer to a grammatical item) and false positives as an “grammatical” answer to ungrammatical items (“gr” | ug). Empirical ROC points were then constructed by cumulating the mean true and false positive rates separately across levels of confidence. Thus, the first point on the ROC represents the performance for the first confidence level, i.e. surely correct/grammatical responses. This procedure was continued for each successive level of confidence ending with the surely incorrect/ungrammatical responses. To test our specific predictions regarding the contributions of similarity and rule knowledge to AGL a formal hybrid model including both processes was fitted to the empirically obtained ROC points. This model assumes similarity as a Gaussian equal-variance signal-detection process whereby the probability of accepting an item depends upon sensitivity (d' , the distance between the means of the distribution of grammatical and ungrammatical on a continuous similarity scale) and a response criterion (c_i). If performance solely relies on similarity, the probability that a grammatical item's similarity exceeds the response criterion (c_i) is then given by $P(\text{“gr”} | \text{gr})_i = \Phi(d'/2 - c_i)$ while the probability that a ungrammatical item is sufficiently similar to be incorrectly endorsed as “grammatical” is $P(\text{“gr”} | \text{ug})_i = \Phi(-d'/2 - c_i)$. The hybrid model, in addition to similarity knowledge, takes the potential contribution of rule knowledge into account. Thus, the probability of a true positive (i.e., the correct identification of grammatical sentences) estimates as $P(\text{“gr”} | \text{gr})_i = R + (1 - R) \Phi(d'/2 - c_i)$. This equation reflects the assumption that a true positive occurs when a grammatical item is endorsed either by rule knowledge [i.e., R] or is accepted as grammatical on the basis of similarity given that there is no rule knowledge [i.e., $(1 - R) \Phi(d'/2 - c_i)$]. These equations were used to fit the hybrid model to the empirically obtained ROC assuming that rule knowledge and similarity (i.e., R and d') remain constant across the ROC and only the response criterion (c_i) varies. This calculation was performed using a maximum likelihood estimation procedure described by Ogilvie and Creelman (1968) with the Excel solver. It adjusts the estimates for rule and similarity knowledge by minimizing the summed error between observed and predicted values. For comparison with the results reported by Kinder and Assmann (2000) the average ROCs are also plotted in z-space.

fMRI data

A general linear model using a random effect model, implemented in the software package BRAINVOYAGER QX (Brain Innovation, Maastricht, The Netherlands) was used to analyze the data. During preprocessing, functional data underwent a cubic spline slice scan time correction, as well as a trilinear 3D motion correction and a two cycle temporal high-pass filter to filter out noise. Spatial smoothing was done using a 6-mm FWHM isotropic Gaussian kernel. Functional slices were then co-registered to the high-resolution whole-brain anatomical scans obtained in the beginning of the session, and were subsequently spatially transformed into stereotactic Talairach space (Talairach and Tournoux, 1988) and re-sampled to a spatial resolution of 3 x 3 x 3 mm. The hemodynamic response function (HRF) was computed as two gamma functions (onset: 0, time to response peak: 5 sec, time to undershoot peak: 15 sec). The design matrix for each participant included grammatical and ungrammatical sentences as events of interest. The sensorimotor control task was added as a predictor of no interest to the design matrix. In a first analysis contrasts tested for differential BOLD-response in grammaticality, i.e. for greater activity for grammatical than ungrammatical items. A second analysis was centered on the grammaticality x similarity/rule interaction, with the extent of a particular knowledge type being derived from the results of the ROC-analysis. Thus, the rule/similarity estimates derived from the ROC analysis were treated as a covariate in the GLM in addition to the main effect of grammaticality. This analysis sought to identify brain regions where high scores on similarity or rule estimates were associated with large differences in brain activity between grammatical and ungrammatical items.

Clusters of differential activity for grammaticality were considered significant up to a threshold of $p < .0001$ and a cluster-size of 10 contiguous voxels, whereas clusters showing differential main-effect-activity, modulated by the covariates were considered significant if they survived a threshold of $p < .001$ and a cluster-size of 5 voxels. This more liberal threshold was chosen in order to reduce the possibility of a type II error (Lieberman and Cunningham, 2009) and due to the more explorative nature of the analysis of covariates. Correlations between differential activation of grammatical and ungrammatical items averaged across activation-clusters, and respective rule or similarity estimates derived from the ROC-analysis were tested using the statistical analysis-software SPSS in a repeated-measures GLM.

Results

Behavioural

During learning accuracy in grammaticality judgments increased over blocks, as shown in Fig. 1. Comparing performance in the first two blocks with the last two blocks in an ANOVA revealed that this improvement was significant [$F(1, 15) = 37.838, P < .001, \eta_p^2 = .716$]. On average, 86% of variance in performance improvement can be explained using an exponential regression model. In order to control for a possible transfer of syntactic rules from the participants first language (German) to BROCANTO, the violation of a rule implemented only in BROCANTO (determiner-noun-agreement), was contrasted with other violation types in terms of performance increase from the first to the last two blocks. This Violation x Block interaction was not significant [$F(2, 45) = .207, P < .85$].

ROC

Receiver operating characteristics (ROC) were obtained from all participants, applying a hybrid model to the curve of true positive/false positive-ratios at different confidence levels which ranged from 1 (*surely grammatical*) to 6 (*surely ungrammatical*). The measured ROC points in probability space (Fig. 2, left) formed an asymmetrical curve and the ROC in z-space appears to be curvilinear with a slope less than 1. This indicates the involvement of both rule- and similarity-based mechanisms in learning the artificial language. This was confirmed by the parameter

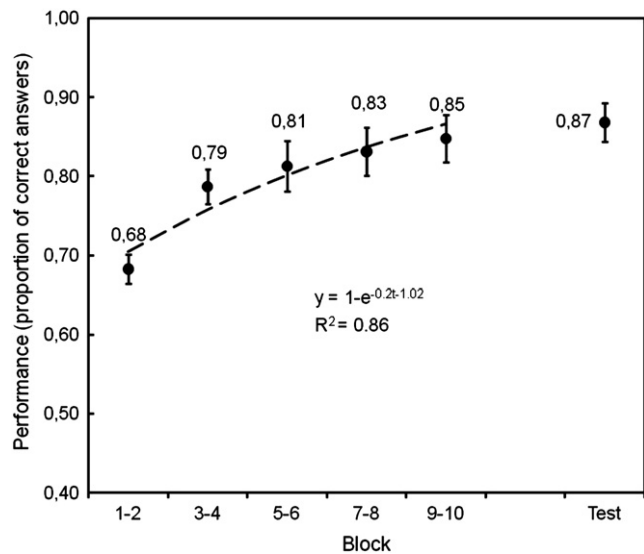


Fig. 1. Mean performance across learning blocks (averaged over two consecutive blocks) and final transfer task (with standard error bars). Significant increase in performance indicates successful learning of language, explaining variance of $R^2 = .86$.

estimates derived from the hybrid model being significantly different from zero (mean similarity estimate $S = .38, SEM = .08, t(15) = 4.44, p < .0001$ and mean rule estimate $R = .6, SEM = .06, t(15) = 8.77, p < .0001$). A marginal significant difference between the rule and the similarity estimates ($t(15) = 1.79, p < .052$, one-tailed) suggests greater reliance on rule than on similarity knowledge in grammaticality judgments.

fMRI data

Main effects of grammaticality

In a first analysis a general linear model was used to test differential BOLD responses to grammatical and ungrammatical items. This comparison between processing of grammatical and of ungrammatical items allowed a clearer picture of grammar-related processes than a baseline-comparison (i.e., comparing baseline activity with grammatical or ungrammatical items). Greater activity was observed for grammatical items (Table 1) in right frontal, cingulate, left occipital, and premotor areas, as well as the right cuneus. In contrast, ungrammatical items did not elicit any significant BOLD responses as compared to grammatical items.

Rule and similarity

To further differentiate the main effect according to individual similarity or rule knowledge, individual estimates of rule and similarity knowledge obtained from the ROC analysis were entered into the general linear model as covariates for the main contrast of grammaticality. Similarity knowledge significantly modulated the main grammaticality effect in the right inferior frontal gyrus and bilaterally in the premotor cortex (Fig. 3) encompassing also parts of the left insula (Table 2). Furthermore, the same modulation of the main effect of grammaticality by similarity knowledge was observed in the inferior parietal lobule, right fusiform gyrus and right hippocampus (Fig. 3). As the focus in the present study lies on the involvement of premotor, frontal and hippocampal areas in superficial systems, only respective activation patterns are depicted in greater detail (Figs. 3 and 4). In all these brain areas, higher similarity estimates were linked to a larger difference in brain activity between ungrammatical and grammatical items.

Rule knowledge modulated brain activity elicited by grammatical and ungrammatical items only in the ventral premotor cortex (Table 2 and Fig. 4). This modulation is characterized by an increase in BOLD response for ungrammatical items over grammatical ones when participants achieved higher scores on rule knowledge.

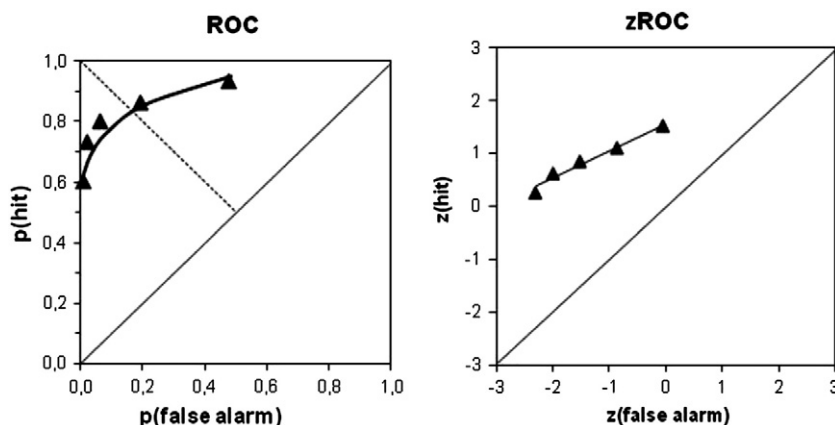


Fig. 2. Left: Mean ROC-curve produced from hit vs. false alarm-ratios accumulated at different confidence levels. Right: Mean z-score of hit vs. false alarm increase at different confidence levels.

Discussion

The present study set out to investigate the neural correlates of rule and similarity knowledge, as the basis of the grammaticality classification of sentences formed according to the artificial language BRO-CANTO. Rule knowledge is based on an understanding of the abstract grammar system underlying the artificial language, whilst similarity knowledge requires mental representations built on superficial features of grammatical sentences. Furthermore, it was hypothesized that a ROC analysis would provide estimates of rule as well as similarity reliance on an individual level. The present behavioral data produced ROC-curves that were in accord with a hybrid model, implying the use of both knowledge types. This result complements the findings of Kinder and Assmann (2000), as they found a signal detection model, thus a function assuming only the acquisition of similarity knowledge, to comply with their data. The fact that the present data provided evidence for the acquisition of rule knowledge in addition to similarity knowledge could be a result of the higher performance level in the present study. This increase in performance may represent a more in-depth acquisition of the grammar or rule knowledge caused by the immediate feedback during training. This is in line with a recent proposal that learning the rules requires explicit feedback (e.g., Ashby and O'Brien, 2005; Opitz et al., 2011). A recent study on L2 learning compared native and non-native speakers after brief exposure to correct Italian sentences without any feedback (Mueller et al., 2009). From the diverging pattern of brain responses between native and non-native speakers it can be inferred that non-native speakers did not acquire an abstract representation of the underlying syntactic rules after mere exposure to simple Italian sentences suggesting that feedback is necessary for the acquisition of a grammatical rule set.

More importantly however, neuroanatomical correlates have been investigated as a function of individual rule and similarity knowledge. In the hippocampus, the difference in BOLD activity for grammatical and ungrammatical items was modulated by the amount of similarity

knowledge as indicated by an interaction between similarity estimates derived from the ROC analysis and the differential brain activity. High similarity estimates were related to a greater sensitivity to ungrammatical items. This finding is in a good agreement with previous studies reporting a decreasing activity in the left posterior hippocampus with longer exposure to BRO-CANTO (Opitz and Friederici, 2003). This decrease in activity was argued to represent decrease in importance of relational processing (Opitz and Friederici, 2004, 2007). Similarly to the declarative component in Ullman's (2004) model, it was proposed that the hippocampus supported the learning of specific word combinations (rather than word-class combinations) on a superficial level, a strategy which is believed to be of importance in initial stages of grammar learning. Although this observation was based on exposure time, the present data could be interpreted in a similar manner, given that the different similarity estimates of the ROC analysis reflect the extent of similarity based strategies. The greater sensitivity to ungrammatical items could be an attempt to integrate the novel combinations of words into an already learned pattern. In support of this idea, Forkstam et al. (2006) reported that the right hippocampus responded to items with low associative chunk strength, meaning that a probabilistic approach would require more effort in classifying items.

In addition to the hippocampus, activations in the right inferior frontal gyrus (IFG) and superior right and middle left premotor cortex (PMC) were modulated by the extent of individual similarity knowledge. In all regions greater similarity knowledge was associated with an increased signal for ungrammatical over grammatical items. The observation of similarity-related activity in the right IFG is partially in accord with the interpretation of Fletcher et al. (1999), who proposed, after employing a within/between block paradigm, that the right middle frontal gyrus supports similarity knowledge. It should be noted however, that in the present study the right frontal involvement was not purely task related, as in the Fletcher et al. (1999) study but rather stems from an interaction between similarity knowledge and grammaticality of items (Fig. 3 second from top). Taking into account that similarity based judgments rely on the superficial comparison between the target item and a mentally stored representation, the greater activity for ungrammatical items linked with high similarity estimates implies that the right IFG serves as an error detection system, responding to mismatches between target and mental representation. Although this is similar to the proposed mechanisms supported by the hippocampus, the hippocampal response is based on novel word combinations, whereas the IFG is assumed to respond to mismatches between target item and a vague, but general mental representation of grammatical items. Such an interpretation of superficial mismatch processing is further supported by previous findings that violations in phonotactic judgment tasks, which are unrelated to any form of grammar, result in an activation of right prefrontal areas, as compared to syntactic

Table 1
Brain areas exhibiting greater activity for grammatical than ungrammatical items on a cluster level.

Cortical region	BA	Size (voxels)	Peak location			F-value (cluster level)	p-value
			X	Y	z		
Right Middle Frontal Gyrus	6	10	28	-14	45	32.405	p<0.001
Right Cuneus	7	388	21	-74	32	34.357	p<0.001
Right Posterior Cingulate	30	275	22	-59	8	34.943	p<0.001
Cingulate Gyrus	24	10	5	4	33	19.247	p<0.005
Left Middle Occipital Gyrus	37	174	-38	-68	2	58.341	p<0.001
Left Premotor Cortex	6	22	-53	-8	39	31.313	p<0.001

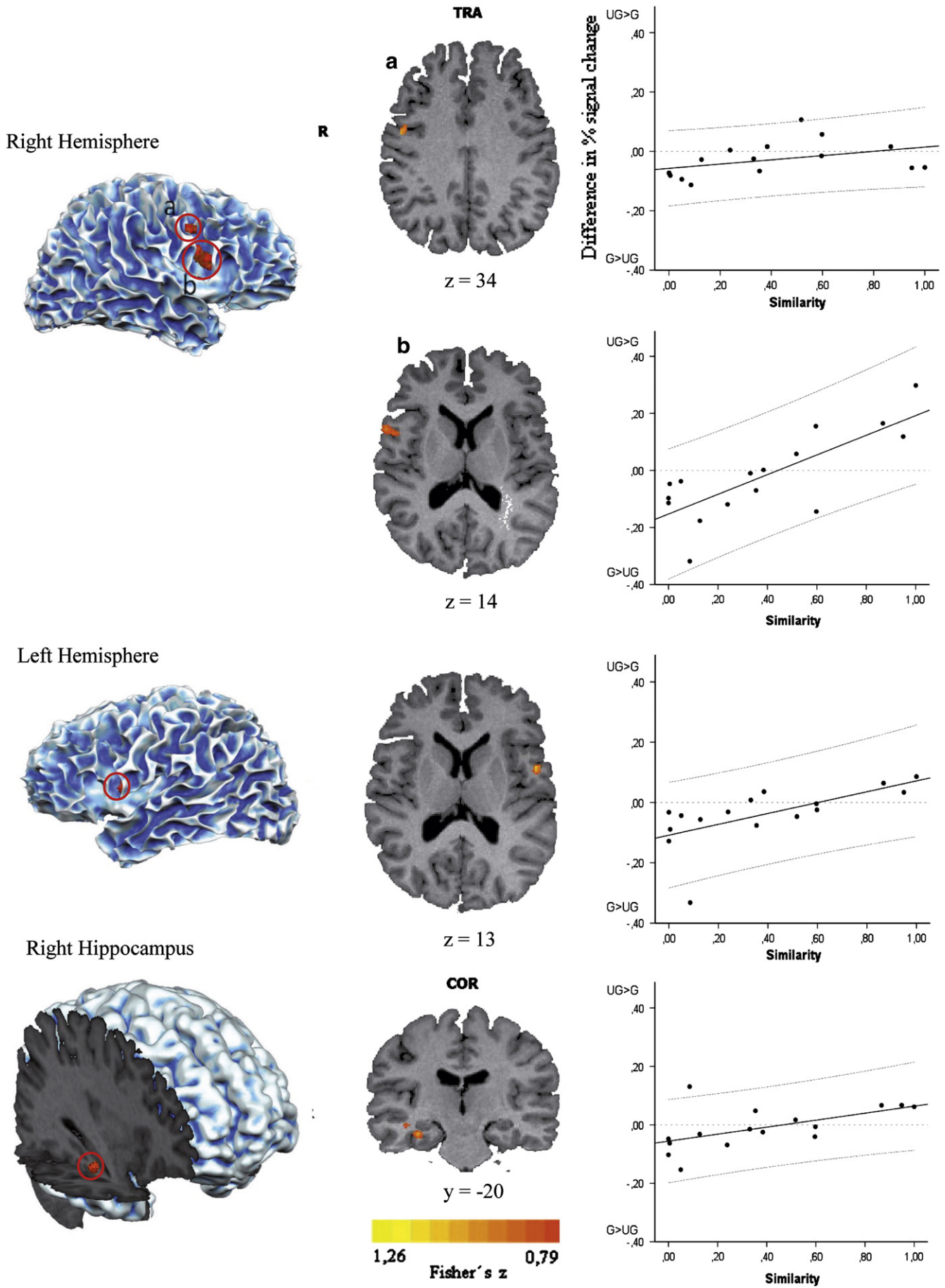


Table 2

Brain areas that exhibit a significant differential modulation of BOLD signal for grammatical and ungrammatical items as a function of similarity or rule knowledge (areas only found to interact with covariates on a voxel-level are marked with an asterisk).

Cortical region	BA	Siz (voxels)	Peak location			F-value	p-value (cluster level)
			X	y	z		
Similarity × Grammaticality							
Right Inferior Frontal Gyrus	44	12	59	10	14	9.507	p<0.01
		192	50	5	19	10.688	p<0.01
Right Inferior Parietal Lobe	40	11	53	−50	36	6.736	p<0.05
Right Premotor Cortex	6	47	48	0	34	5.628	p<0.05
Right Fusiform Gyrus	20	12	38	−39	−15	5.488	p<0.05
Hippocampus		9	32	−20	−15	6.549	p<0.05
Left Insula	13	10	−37	19	2	5.456	p<0.05
		16	−37	−2	10	8.49	p<0.05
Left Fusiform Gyrus*	20	10	−41	−28	−27		
Left Precentral Gyrus	6	42	−53	3	13	5.147	p<0.05
Rule × Grammaticality							
Left Superior Temp. gyrus*	38	35	−36	13	−30		
Left Precentral Gyrus	6	6	−46	−8	9	18.648	p<0.005

judgment tasks and input/output related processes (Indefrey et al., 2001). Furthermore, studies on episodic (Aggleton and Brown, 1999) and long-term memory (Simons and Spiers, 2003) have demonstrated that prefrontal cortices receive input from hippocampal areas. Thus, it is conceivable that the IFG serves as a general, bottom-up classification facility for similarity-strategies, based on information received from hippocampal areas.

In contrast to the IFG, activation of the PMC could be interpreted as reflecting the prediction of upcoming words in a sentence. Schubotz (2007) proposed that the PMC serves as a predictive system, not only for human-like actions, but also for abstract sequences such as music or sea-waves, which cannot be mirrored by any human movement. This predictive role has also been suggested for the processing of grammatical rule systems (Bahlmann et al., 2009). In the present study, the abstract level of violations appeared to be of importance also in superficial strategies, implying the PMC as a further contributor to the processes in the right IFG. Thus, similarity strategies in this particular study might build on different levels of abstractness.

Lieberman et al. (2004) have already pointed out the possibility of a competitive element between accessibility of grammaticality and associative chunk strength in items. They found that in items with high chunk strength, grammaticality had less of an impact than in low chunk strength items, suggesting that at an equal accessibility, superficial features are favored. Although they did not assess to what extent their participants preferred, or had acquired any of the two knowledge types, such a competition between rule and similarity may also apply to the present similarity-reliant individuals. As long as knowledge of rule-structure is not evolved and efficient enough to act as the governing process in the classification task, it may work in parallel to superficial processes, which are accommodated in hippocampal structures. As individuals with little knowledge of abstract features will process the rule level of items less efficiently, these abstract processes will fail to compete with the faster, more efficient similarity-based processes. Therefore, the operating characteristics observed in these participants will reflect a tendency towards similarity strategies, regardless of abstract processes working in parallel. Nonetheless, the attempt at processing underlying rules results in an activation of the rule-based PMC areas also for individuals with greater reliance on superficial features.

In contrast, the key role maintained by PMC in abstract strategies becomes evident looking at the activation patterns found in the left

ventral PMC, where functional response to grammaticality interacted with rule scores. This interaction was characterized by a greater activation for ungrammatical items in individuals with high rule scores, whereas low rule scores were linked to a greater activation in the grammatical condition. For greater rule reliance, this implies certain automaticity in the processing of grammatical items, whilst the detection of violations leads to further processing due to integration or prediction difficulties (Forkstam et al., 2006).

This suggested double-role of the PMC in similarity and rule processes may be attributed to a differentiation between different types of dependencies of words within a sentence. Previous studies suggested that the PMC is crucially involved in the processing of so called local or adjacent dependencies that can be fully specified by transition probabilities between neighboring elements in a sequence. Functional imaging studies (Bahlmann et al., 2009; Opitz and Friederici, 2007) as well as studies on patients with lesions centered in the PMC (Opitz and Kotz, in press) consistently demonstrated the involvement of the PMC in processing violations of such adjacent dependencies. In contrast, hierarchical structures, characterized by long-distance dependencies, have been shown to depend on left inferior frontal areas. Bahlmann et al. (2009) found that the PMC operates as a common basis for both grammar types, suggesting that it acts as a very basic rule-processor, with the previously mentioned task of predicting abstract events. With regards to the present PMC activations, the predictive role this area fulfills would regard adjacent, local sequences. Superficial features of such sequences, relevant for similarity strategies, yet based on adjacent local rules, have been defined in previous studies (Perruchet and Pacteau, 1990; Lieberman et al., 2004), as chunks or fragments. In studies employing the Reber-paradigm, these consist of mentally stored letter-groups (bigrams or trigrams), whilst in the present study these would contain small word-sequences. With increasing exposure to the grammar, size and complexity of these fragments increases to the point that they can be unified into rule knowledge (Dulany et al., 1984; Forkstam et al., 2006; Pothos, 2007). This exposure-dependent increase in size of predictable fragments may represent the competition-efficiency discussed previously with regards to similarity-strategies. As long as the PMC is only able to process shorter abstract sequences, processing entire sentences on a rule-level is reduced in efficiency, thus slower and unable to compete with processing speed of hippocampal and right frontal structures, leading to a similarity based strategy. In contrast, if such

Fig. 3. Left and middle: Activation clusters of grammaticality × similarity interaction for right PMC ($z = 34$), right IFG ($z = 14$), left PMC ($z = 13$) and right hippocampus (bottom). Right: Correlation between similarity-estimates and difference in BOLD signal between ungrammatical (UG) and grammatical (G) items for the respective activation.

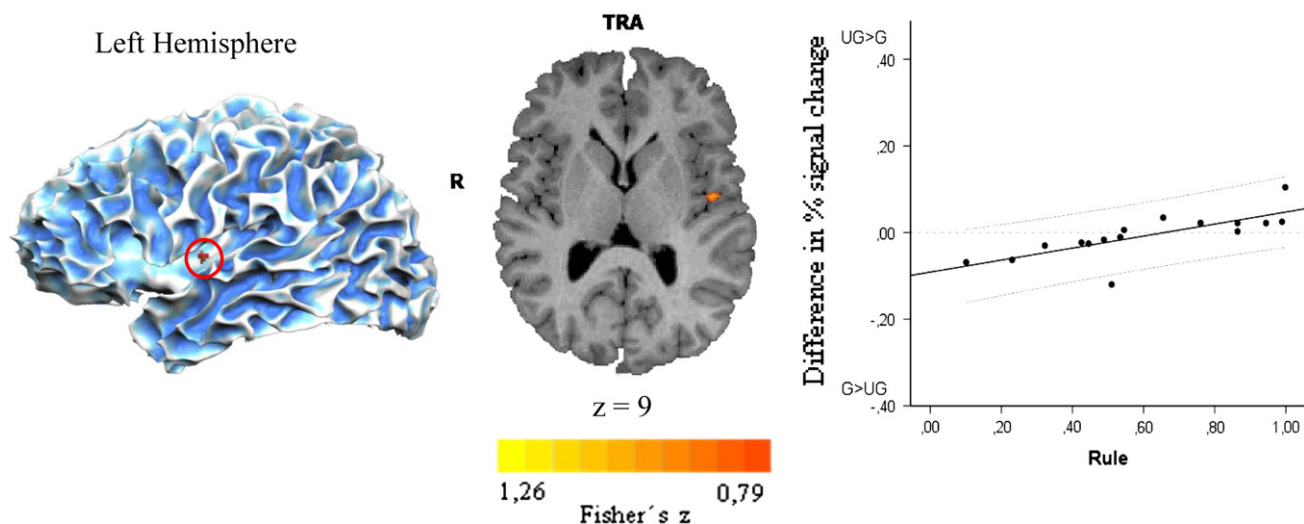


Fig. 4. Activation of left ventral premotor cortex and interaction with rule knowledge.

fragments become longer and more predictable on an abstract level, a rule-based strategy may take over, outcompeting superficial processing structures, resulting in more rule-like operating characteristics.

Conclusion

In summary, the present study demonstrated that, given sufficient prior training, an analysis of receiver-operating characteristics can reveal individual differences in use of knowledge types. Furthermore, the neural correlates of similarity knowledge were proposed to be composed of the hippocampus supporting the processing of specific word-combinations, feeding into comparison processes of superficial word features to mental exemplars in the right IFG. In contrast, rule structures in the present study were related to the PMC processing dependencies of a local nature. More specifically, it was proposed that rule and similarity knowledge work in parallel and compete in processing-efficiency, leading to an initial superiority of similarity-based classification, and a subsequent dominance of rule-based processes, once a critical amount of abstract knowledge of adjacent dependencies was acquired.

References

- Aggleton, J.P., Brown, M.W., 1999. Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behav. Brain Sci.* 22, 425–489.
- Ashby, F.G., O'Brien, J.B., 2005. Category learning and multiple memory systems. *Trends Cogn. Sci.* 9, 83–89.
- Bahlmann, J., Schubotz, R.I., Mueller, J.L., Koester, D., Friederici, A.D., 2009. Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res.* 1298, 161–170.
- Beesley, T., Wills, A.J., Le Pelley, M.E., 2010. Syntactic transfer in artificial grammar learning. *Psychon. Bull. Rev.* 17, 122–128.
- Dienes, Z., Kurz, A., Bernhaupt, R., Perner, J., 1997. Application of implicit knowledge: Deterministic or probabilistic? *Psychol. Belg.* 37, 89–113.
- Dulany, D.E., Carlson, R.A., Dewey, G.I., 1984. A case of syntactical learning and judgment: How conscious and how abstract? *J. Exp. Psychol. Gen.* 113, 541–555.
- Fletcher, P., Buechel, C., Josephs, O., Friston, K., Dolan, R., 1999. Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cereb. Cortex* 9, 168–178.
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., Petersson, K.M., 2006. Neural correlates of artificial syntactic structure classification. *Neuroimage* 32, 956–967.
- Friederici, A.D., Fiebach, C.J., Schlesewsky, M., Bornkessel, I.D., von Cramon, D.Y., 2006. Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb. Cortex* 16, 1709–1717.
- Gomez, R.L., Schvaneveldt, R.W., 1994. What is learned from artificial grammars? Transfer tests of simple association. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 396–410.
- Indefrey, P., Hagoort, P., Herzog, H., Seitz, R.J., Brown, C.M., 2001. Syntactic processing in left prefrontal cortex is independent of lexical meaning. *Neuroimage* 14, 546–555.
- Kinder, A., Assmann, A., 2000. Learning artificial grammars: No evidence for the acquisition of rules. *Mem. Cognit.* 28, 1321–1332.
- Lieberman, M.D., Cunningham, W.A., 2009. Type I and Type II error concerns in fMRI research: re-balancing the scale. *Soc. Cogn. Affect. Neurosci.* 4, 423–428.
- Lieberman, M.D., Chang, G.Y., Chiao, J., Bookheimer, S.Y., Knowlton, B.J., 2004. An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *J. Cogn. Neurosci.* 16, 427–438.
- Mueller, J.L., Oberecker, R., Friederici, A.D., 2009. Syntactic learning by mere exposure—an ERP study in adult learners. *BMC Neurosci.* 10, 89. doi:10.1186/1471-2202-10-89.
- Ogilvie, J.C., Creelman, C.D., 1968. Maximum-likelihood estimation of receiver operating characteristic curve parameters. *J. Math. Psychol.* 5, 377–391.
- Opitz, B., Friederici, A.D., 2003. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *Neuroimage* 19, 1730–1737.
- Opitz, B., Friederici, A.D., 2004. Brain correlates of language learning: The neuronal dissociation of rule-based versus similarity-based learning. *J. Neurosci.* 24, 8436–8440.
- Opitz, B., Friederici, A.D., 2007. Neural basis for processing sequential and hierarchical syntactic structures. *Hum. Brain Mapp.* 28, 585–592.
- Opitz, B. & Kotz, S.A. (in press). Ventral premotor cortex lesions disrupt learning of sequential grammatical structures. *Cortex*. doi:10.1016/j.cortex.2011.02.013.
- Opitz, B., Ferdinand, N.K., Mecklinger, A., 2011. Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Front. Hum. Neurosci.* 5, 8. doi:10.3389/fnhum.2011.00008.
- Perruchet, P., Pacteau, C., 1990. Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *J. Exp. Psychol. Gen.* 119, 264–275.
- Pothos, E.M., 2007. Theories of artificial grammar learning. *Psychol. Bull.* 133, 227–244.
- Reber, A.S., 1967. Implicit learning of artificial grammars. *J. Verbal Learn. Verbal Behav.* 6, 855–863.
- Schubotz, R.I., 2007. Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11, 211–218.
- Shohamy, D., Myers, C.E., Kalanithi, J., Gluck, M.A., 2008. Basal ganglia and dopamine contributions to probabilistic category learning. *Neurosci. Biobehav. Rev.* 32, 219–236.
- Simons, J.S., Spiers, H.J., 2003. Prefrontal and medial temporal lobe interactions in long-term memory. *Nat. Rev. Neurosci.* 4, 637–648.
- Strange, B.A., Henson, R.N.A., Friston, K.J., Dolan, R.J., 2001. Anterior prefrontal cortex mediates rule learning in humans. *Cereb. Cortex* 11, 1040–1046.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publishers, Inc., New York.
- Ullman, M.T., 2001. The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Biling. Lang. Cogn.* 4, 105–122.
- Ullman, M.T., 2004. Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270.