

On neural correlates of individual differences in novel grammar learning: an fMRI study

Olga Kepinska^{1,2,*}, Mischa de Rover^{2,3,4}, Johanneke Caspers^{1,2}, Niels O. Schiller^{1,2}

¹ Leiden University Centre for Linguistics, Postbus 9515, 2300 RA Leiden, the Netherlands

² Leiden Institute for Brain and Cognition, c/o LUMC, Postzone C2-S, P.O.Box 9600, 2300 RC Leiden, the Netherlands

³ Department of Anesthesiology, Leiden University Medical Center, Postzone P5-Q, P.O.Box 9600, 2300 RC Leiden, The Netherlands

⁴ Faculty of Social Sciences, Institute of Psychology, Clinical Psychology Unit, Pieter de la Court gebouw, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands

* Corresponding author. Tel. +31 (0)71 527 8530, e-mail: o.kepinska@hum.leidenuniv.nl

ABSTRACT. We examine the role of language analytical ability, one of the components of language aptitude - a specific ability for learning languages - during acquisition of a novel grammar. We investigated whether the neural basis of artificial grammar learning (AGL) differs between populations of highly and moderately skilled learners. Participants performed an AGL task during an fMRI scan and data from task's test phases were analysed. Highly skilled learners performed better than moderately skilled ones and engaged during the task more neural resources in the right hemisphere, i.e. in the right angular/supramarginal gyrus and superior frontal and middle frontal gyrus and in the posterior cingulate gyrus. Additional analyses investigating the temporal dynamics of brain activity during learning revealed lateralisation differences in the modulation of activity in the parietal and temporal cortex. In particular, the left angular gyrus BOLD activity was coupled with high performance on the AGL task and with a steep learning curve.

Keywords: Language learning, fMRI, Second Language Acquisition, Artificial Grammar Learning, Individual differences, Language aptitude

1. Introduction

It is a commonly observable fact that individuals learning foreign languages differ from each other both in terms of acquisition rate and the ultimate attainment of the languages. Some people are believed to have a so-called “knack” for languages, or to possess a set of special abilities, which enable them to communicate in any given foreign language successfully both faster and more proficiently than others. The question arising is what neural mechanisms are responsible for such differences between individual learners.

As language acquisition is a complex process consisting of various aspects (i.e. building up the mental lexicon, acquisition of grammatical rules, phonological regularities and pragmatic competence), capturing the neural architecture behind individual variability between learners poses important methodological challenges. Although it is possible to investigate language learning in a holistic way, employing natural language input (see e.g. Musso et al., 2003; Veroude, Norris, Shumskaya, Gullberg, & Indefrey, 2010), most neuroimaging studies on language acquisition resort to highly controllable stimuli representing only one of the facets of language learning. Acquisition of novel vocabulary items for example, is often investigated without the involvement of their morphosyntactic features (e.g. Breitenstein et al., 2005; Freundlieb et al., 2012; Hultén, Laaksonen, Vihla, Laine, & Salmelin, 2010).

The focus of this study is how new language is *ab initio* processed by the brain and how individual differences in performance are reflected in brain functionality. As a proxy for the language acquisition process, we chose to concentrate on the acquisition of new grammatical rules. We believe the grammar to be one of the most important building blocks of (second) language learning. Our aim is to capture the process of new syntax acquisition in isolation from other aspects of language learning and control for earlier exposure.

1.1. High cognitive skills for grammar learning

Within the field of Second Language Acquisition (SLA) success in language learning has been ascribed to various factors, such as learner's age, language aptitude, motivation, personality and learning style. Besides the age factor, language aptitude is the most robust predictor of achievement in a second language (L2) (Dörnyei & Skehan, 2003; Ellis, 2008; Sawyer & Ranta, 2001). This individual, relatively immutable factor, plays an important role in SLA, when language is acquired by means of instruction (i.e. in a formal setting, where it is explicitly reflected upon) (De Graaff, 1997), under incidental learning conditions (Hamrick, 2015), and naturalistically, i.e. without formal instruction (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000; Robinson, 1997; Sawyer & Ranta, 2001). Language aptitude has traditionally been operationalised by means of standardised test instruments that aim at capturing learners' abilities underlying L2 acquisition. Such tests typically consist of a number of different parts, each aiming to measure a putative separate component of the larger construct of aptitude. Most aptitude tests thus underscore its multi-componential nature (i.e. this specific talent is a combination of skills that are fairly independent from each other). Four sub-components of language aptitude are traditionally distinguished: rote learning memory, phonemic coding ability, inductive language learning ability and language analytical ability (LAA) (cf. Abrahamsson & Hyltenstam, 2008; Carroll, 1981; Dörnyei & Skehan, 2003; Ellis, 2008; Sawyer & Ranta, 2001; Skehan, 2002).

Despite the recent technological advances available for neurolinguistic research, it remains unclear how these language aptitude components can be accounted for in terms of their neural correlates (cf. Hu et al., 2013; Reiterer, Pereda, & Bhattacharya, 2011; Xiang, Dediu, Roberts, Norris, & Hagoort, 2012). With this fMRI study, we therefore wanted to gain insight into the neural mechanisms underlying language aptitude, in particular, one of its components, viz. language analytical ability. We aimed to capture the neural correlates of LAA during the process of acquisition of a novel grammar. LAA, being relevant for pattern identification during SLA which involves analysing and processing new linguistic input (Skehan, 2002), is arguably the most important of the aptitude components when it comes to grammar learning: learners with a high degree of LAA are sensitive to grammatical structure of new languages and are able to make linguistic generalisations easily. SLA research has shown that LAA plays an important role in second language acquisition in a variety of contexts, including immersion (Harley & Hart, 1997), classroom (Erlam, 2005) and lab (Yilmaz, 2012) settings.

A number of studies investigating individual differences in cognitive abilities in relation to brain function have focussed on the neural efficiency hypothesis in order to explain the mechanisms underlying high cognitive skills (Haier et al., 1988; Neubauer & Fink, 2009; Nussbaumer, Grabner, & Stern, 2015; Prat & Just, 2011; Prat, Long, & Baynes, 2007; Prat, 2011; Reichle, Carpenter, & Just, 2000). Neural efficiency is understood as using fewer mental resources, in a more focused and goal-directed way, while dealing with

demands of the task at hand (Neubauer & Fink, 2009). For example, within the domain of language abilities, Prat et al. (2007) showed that high-capacity readers (as per a reading span test), exhibited higher neural efficiency during sentence comprehension than low capacity readers.

To date, however, we are not aware of any studies investigating high skills for particular L2 learning sub-processes - such as acquisition of novel grammar rules - either corroborating or contradicting the neural efficiency hypothesis. By investigating the neural correlates of LAA during new grammar learning, we aim to contribute to understanding of neural mechanisms behind successful foreign language learning in general, as well as to the discussion on neural efficiency as the underlying mechanism behind high cognitive skills. Does neural efficiency drive successful and efficient L2 learning?

1.2. The Artificial Grammar Learning paradigm

In our approach, we employed a methodology previously used in studies investigating neural mechanisms behind the acquisition of novel grammar rules, i.e. Artificial Grammar Learning (AGL). Even though AGL paradigms do not offer a comprehensive model of language acquisition, they are often used in laboratory settings in order to gain insight in the neurobiology of syntax processing and acquisition, without the interference of semantics, phonology or pragmatics (Petersson, Folia, & Hagoort, 2012; Petersson & Hagoort, 2012; Reber, 1967) and with the advantage of being highly controllable. Also, contrary to the use of natural language stimuli, AGL excludes any interference of prior exposure. Neuroimaging investigations into the neurobiology of AGL have shown that such tasks involve the same neural resources as in case of processing and acquisition of natural languages, i.e. the left inferior frontal gyrus (LIFG) (Petersson & Hagoort, 2012). Moreover, Ettliger, Morgan-Short, Faretta-Stutenberg, & Wong (2015) have recently provided evidence for a strong relationship between performance on an artificial language learning task and L2 learning.

Although most AGL studies require an acquisition period of several days (e.g. Friederici et al., 2002; Petersson et al., 2012), on-line learning of an artificial grammar in an MRI scanner was employed in our experimental design in order to enable observation of the neural mechanisms behind the learning process in real time. Another difference between our study and traditional AGL experiments has to do with artificial grammar systems being often learnt implicitly (e.g. Petersson et al., 2012; Reber, 1967), solely on the basis of examples and without instruction or feedback. Our aim was to guide participants' attention towards discovering the grammatical rules by providing instructions to do so. Also, we wanted to include feedback as part of the learning process. Such procedure has previously been adopted in a series of experiments where the artificial language BROCANTO was used to investigate the learning mechanisms underlying grammar acquisition (Brod & Opitz, 2012; Friederici et al., 2002; Hauser, Hofmann, & Opitz, 2012; Opitz, Ferdinand, & Mecklinger, 2011; Opitz & Friederici, 2003, 2004, 2007; Opitz & Hofmann, 2015). BROCANTO studies consist of learning and test phases. During learning, participants are presented with grammatically correct sentences and are instructed to extract the underlying grammatical rules. In test phases, both grammatical and ungrammatical sentences are presented and participants are asked to give a grammaticality judgement on the sentences.

The neural architecture responsible for acquiring the BROCANTO structure has been shown to involve a dynamic interaction between left hippocampal formation and the left

inferior frontal area: during the task, activity in the hippocampus decreased as a function of time (and performance); the reverse was observed in the LIFG (Opitz & Friederici, 2003). Hauser et al. (2012) investigated the underpinnings of two types of knowledge acquired in the course of AGL: rule and similarity knowledge. They confirmed the earlier findings of Opitz & Friederici (2003) and proposed that hippocampus and right IFG support grammar learning when the acquired knowledge is based on similarity; left ventral premotor cortex was found to be responsible for rule knowledge (Hauser et al., 2012; Opitz & Friederici, 2004).

The goal of this study is then twofold: first, we want to find mechanisms responsible for processing novel grammar that are representative of individual cognitive traits measured prior to the fMRI experiment, namely the language analytical ability. Second, we are interested in the way successful learning of a novel grammar is supported by the brain and how it is represented in the neural data over time. On the basis of previous findings, we expect to observe an interaction of the hippocampal system and the prefrontal cortex when concentrating on time effect. In line with the neural efficiency hypothesis (Haier et al., 1988; Neubauer & Fink, 2009), less distributed activity networks are expected in the case of highly skilled learners, especially in the inferior frontal region.

2. Methods

2.1. Pre-test

A language aptitude test was administered to a large group of participants ($N = 307$). We used the LLAMA Language Aptitude Test (LLAMA) (Meara, 2005), which is a computer-based test battery with automated scoring, suitable for participants with various language backgrounds. The test consists of four parts: (1) a vocabulary learning task, (2) a test of phonetic memory, (3) a test of sound-symbol correspondence and (4) a test of grammatical inferencing (LLAMA_F), being a measure of LAA. Recruitment of participants for this study was based on the scores on the LLAMA_F test.

In this test, twenty pictures are presented together with sentences in an unknown language that describe them. In the learning phase (lasting five minutes), participants are asked to discover grammatical rules (primarily concerned with agreement features) of this unknown language, and they are allowed to take notes. In the test phase, they are presented with a series of pictures, combined with two sentences and they have to decide which sentence is grammatically correct. Participants can score from 0 to 100, where 80 - 100 is defined as outstandingly good and 25 - 45 as average (Meara, 2005).

2.2. Participants

After taking the LLAMA test, forty-two healthy adults with no contra-indications for an MRI scan were invited for the second part of the study, i.e. the fMRI experiment. On the LLAMA_F test, the participants received either an “average score” (i.e. 30-50) (henceforth Average LAA), or an “outstandingly good” score (i.e. 80-100) (henceforth High LAA).

¹ Although the LLAMA manual defines “average score” as 25-45, a score of 50 was also included as average in this study. The scores are awarded at intervals of 10 and a score of 45 is not possible to obtain. Therefore, there were no participants who scored 25, either.

There were 20 Average LAA participants (16 female), age 19-39 years ($M = 23.60$ years) and 22 High LAA participants (16 female), age 19-43 years ($M = 23.18$ years). All were native speakers of Dutch, right-handed and had normal or corrected-to-normal vision.

The Medical Ethical Committee of the Leiden University Medical Centre (LUMC) (Leiden, the Netherlands) approved the protocol of the MRI experiment; behavioural testing was also conducted according to the Ethics Code of the Faculty of Humanities at Leiden University. Participants gave written informed consent prior to the experiment and were remunerated for their time.

2.3. Stimuli & Design

The stimulus material was created on the basis of the artificial grammar of BROCANTO (Brod & Opitz, 2012; Friederici et al., 2002; Hauser et al., 2012; Opitz et al., 2011; Opitz & Friederici, 2003, 2004, 2007). The AGL task was administered in the scanner and consisted of three blocks of learn and test phases, and a subsequent transfer test. The stimulus material consisted of both grammatical and ungrammatical sentences. The grammatical ones were used in the learning phases of the experiment, the test phases (and the transfer test) contained both grammatical and ungrammatical sentences. In this study, only the fMRI data acquired during the test phases are reported.

The grammar of BROCANTO follows rules found in many natural languages: a sentence (S) consists of a noun phrase (NP) and a verb phrase (VP). A version of the BROCANTO language with 8 vocabulary items was used in this experiment. Words forming a particular word class (nouns, verbs, etc.) could be distinguished by a particular vowel, e.g. ‘u’ specified a noun and ‘e’ a verb. The items were categorised into nouns (“gum”, “trul”), verbs (“pel”, “prez”), adjectives (“böke”), adverbs (“rüfi”) and determiners (“aaf”, always followed by a noun and “aak”, always followed by a modifier). The sentences contained three to eight words and had a subject-verb[-object] structure. The following possible sentence structures were included: dNv² (e.g. *aaf gum pel*), dNvm (e.g. *aaf gum pel rüfi*), DMNv (e.g. *aak böke gum pel*), dNvdN (e.g. *aaf gum pel aaf gum*), dNvDMN (e.g. *aaf gum pel aak böke gum*), dNvmDMN (e.g. *aaf gum pel rüfi aak böke gum*), dNvmdN (e.g. *aaf gum pel rüfi aaf gum*), DMNvdN (e.g. *aak böke gum pel aaf gum*), DMNvDMN (e.g. *aak böke gum pel aak böke gum*), DMNvmDMN (e.g. *aak böke gum pel rüfi aak böke gum*), DMNvm (e.g. *aak böke gum pel rüfi*) and DMNvmdN (e.g. *aak böke gum pel rüfi aaf gum*). In total, we constructed 80 sentences according to the above rules.

The ungrammatical sentences were constructed on the basis of the 80 grammatical ones. They contained syntactic violations and were created by substituting words from one category by words from a different category. The violations were either determiner-noun-agreement violations (i.e., DN instead of dN and dMN instead of DMN, e.g. **aak gum pel aaf gum* instead of *aaf gum pel aaf gum*), word class repetitions of nouns or verbs (e.g. **aaf prez pel aaf gum* instead of *aaf gum pel aaf gum*) and phrase structure violations (i.e., NP NP and NP NP VP rather than NP VP and NP VP NP, respectively, e.g. **aaf gum aaf gum pel* instead of *aaf gum pel aaf gum*). For each grammatical item there were three ungrammatical versions (according to the three violation types). From the pool of 80

² N = noun, v = verb, M = adjective, m = adverb, d = determiner (followed by a noun) and D = determiner (followed by a modifier)

grammatical and 240 ungrammatical items, we chose items for the learning and test phases of the experiment and the subsequent transfer test.

2.3.1. Presentation

The task was created and presented in E-Prime 2.0.10 software (Psychology Software Tools, 2012). Stimuli were presented on a projection screen reflected to a mirror attached to the head coil above participants' eyes. All stimuli were presented in white letters (Courier New, size 20) on an E-Prime 'black' background.

In the learning phases of the experiment, participants were instructed to discover the grammatical rules of the language. They saw forty sentences in each of the three learning phases; these were presented for 8 seconds and preceded by a fixation cross (3 seconds). Each of the three test phases included 20 samples of grammatical and 20 samples of ungrammatical sentences, presented in a random order, for 6 seconds each. The exact details of the stimulus selection algorithm and a complete list of all sentences used can be found in the Supplementary Materials.

Participants were instructed to give a grammaticality judgment by a button press within the 6 seconds of presentation of the sentence. After 6 seconds, visual feedback was provided (a green tick indicating a correct response or a red cross for wrong answers). The feedback screen was presented for 1 second. After the feedback screen, a fixation cross was presented. The duration of the fixation cross was jittered (2-6 seconds of inter-trial interval) in order to ensure that the feedback presentation would not influence the brain activation to the following sentence. Figure 1 contains an example trial from the test phase.

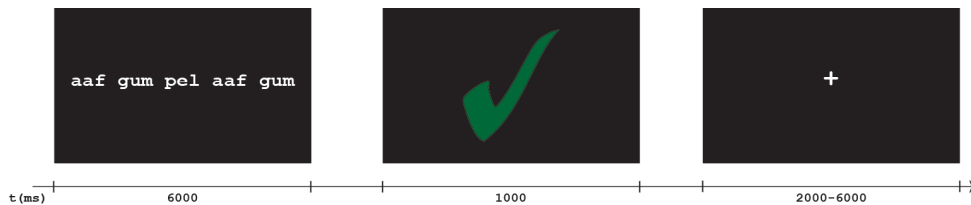


Figure 1. An example of a trial from the test phase of the AGL task: a grammatical test item is correctly classified by the participant.

Six days after the fMRI experiment, participants performed a delayed transfer test. In this test, they saw 20 grammatical and 20 ungrammatical sentences. One third of the grammatical sentences had been presented before. The order of presentation was random. The task was performed on a desktop computer or online, in 3 cases where participants could not be present in the lab six days after the fMRI scan. The online version was prepared and administered in Qualtrics (Qualtrics, 2013). In order to prevent further learning, no feedback was provided in the transfer test.

2.4. Data Acquisition

Imaging data were acquired using a Philips 3T MR-system (Best, The Netherlands) located at the Leiden University Medical Centre (LUMC) equipped with a SENSE-32 channel head coil. For each subject, changes in blood oxygen level dependence (BOLD)

were measured three times; each scan was acquired during the consecutive test phases of the AGL task. We obtained echo-planar images (EPI) using a T2*-weighted gradient echo sequence (repetition time [TR] = 2200 ms, echo time [TE] = 30 ms, matrix size: 80 x 79, 38 axial slices, 2.75 x 2.75 x 2.75 mm voxel size). EPI's were scanned parallel to the anterior–posterior commissure plane. The length of each scan sequence was 209 volumes and lasted 7.5 minutes. Anatomical imaging included a 3D gradient-echo T1-weighted sequence (TR = 9.755 ms, TE = 4.59 ms; matrix 256 x 256; voxel size: 1.2 x 1.2 x 1.2 mm; 140 slices) and a high-resolution T2-weighted image (TR = 2200 ms, TE = 30 ms; matrix 112 x 112; voxel size: 2.0 x 2.0 x 2.0 mm; 84 slices).

3. Behavioural data

3.1. Effect of LAA

The responses on the AGL task for each participant were first transformed into d' scores in order to correct for response bias (Macmillan & Creelman, 2005). The data were then analysed with the goal of establishing the learning effect and differences between High and Average LAA participants distinguished by the LLAMA_F test. Following previous studies employing similar experimental designs (Brod & Opitz, 2012; Hauser, Hofmann, & Opitz, 2012; Opitz, Ferdinand, & Mecklinger, 2011; Opitz & Friederici, 2003, 2004, 2007), a repeated measures ANOVA (alpha level = 0.05) was employed. We used SPSS version 22 (IBM SPSS, 2012). The analysis was performed with LAA as a between-subject factor (High LAA vs. Average LAA) and learning phase (first phase, second phase, last phase and transfer test) as a within-subject factor. Mauchley's test showed violations of sphericity against the factor phase, $\chi^2(5) = 21.408$, $p < .01$, therefore Greenhouse-Geisser correction for non-sphericity was used ($\epsilon = 0.769$).

The analysis revealed that the d' scores on the AGL task both among the High LAA and Average LAA participants increased over the course of the experiment (see **Error! Reference source not found.**): there was a main effect of learning phase, $F(2.308, 92.301) = 38.236$, $p < .001$, $\eta_p^2 = .489$. Furthermore, the High LAA participants performed better than the Average LAA participants which was reflected in a significant effect of LAA, $F(1, 40) = 16.762$, $p < .001$, $\eta_p^2 = .295$, and an interaction between LAA and phase, $F(2.308, 92.301) = 4.469$, $p = .011$, $\eta_p^2 = .10$.

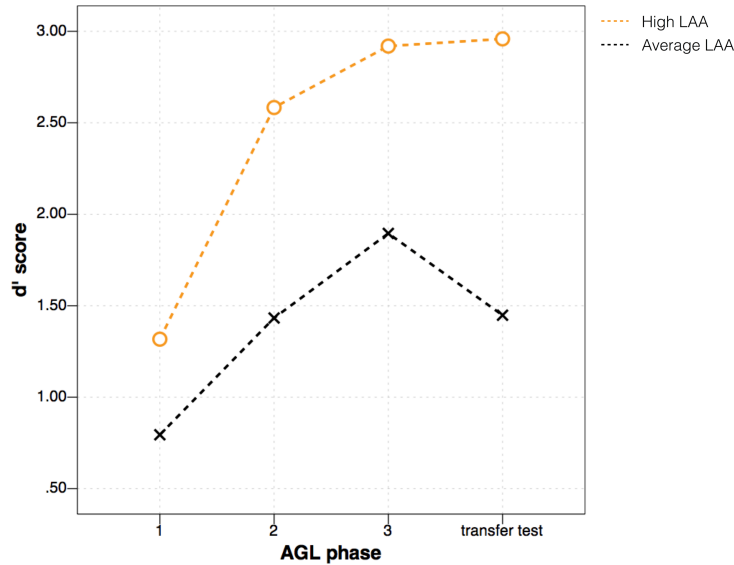


Figure 2. Performance (d' scores) across participants with High and Average LAA during the three AGL test phases and the subsequent transfer test.

3.2. Learning patterns over time

Apart from establishing whether, and to what degree (as a function of the pre-tested analytical abilities) the participants were able to acquire the grammar rules, we were interested in gaining more insight into the various ways the learning of a novel grammar proceeded in time. Individual participants exhibited various learning curves, which can arguably be coupled with different neural mechanisms of learning (cf. Karuza, Emberson, & Aslin, 2014). Therefore, we aimed to classify the behavioural AGL data sets into groups with similar learning patterns, thus taking into account the effect of time and participants' actual performance. To this end, we chose to perform a procedure enabling objective identification of relatively homogeneous groups of participants, namely a k -means cluster analysis (Aldenderfer & Blashfield, 1984, cf. Catani et al., 2007). The analysis was performed in R (R Development Core Team, 2015).

The k -means cluster analysis was run on standardized d' scores from the three AGL test blocks using 1 to 6 clusters with 1000 starts (i.e., running the analysis 1000 times, each time with a different initial clustering of the subjects and retaining the best clustering found across the 1000 analyses (see Steinley & Brusco, 2007). We determined the optimal number of clusters by using a scree-plot-like procedure (Cattell, 1966). In this procedure, the proportion explained variance per clustering solution is plotted against the number of clusters (see Figure 3) and an optimal number of clusters is determined by looking for an elbow (Thorndike, 1953) in this plot (see also Wilderjans, Ceulemans, & Meers, 2013). Looking at Figure 3, a clear elbow was found for the solution with two clusters and a smaller, but still substantial, elbow for the three cluster-solution. The two cluster-solution explained almost 60% of the variance, whereas adding a third cluster substantially increased the proportion explained variance of the solution (from 58.8% to 73.5% respectively). However, adding more clusters, which makes the solution more complex, did not result in a much better solution (i.e., 79.7%, 85.3% and 87.5% explained variance for

the four, five and six cluster-solution, respectively). In particular, in Figure 3 one can see that the increase in percentage explained variance levels off when using more than three clusters.

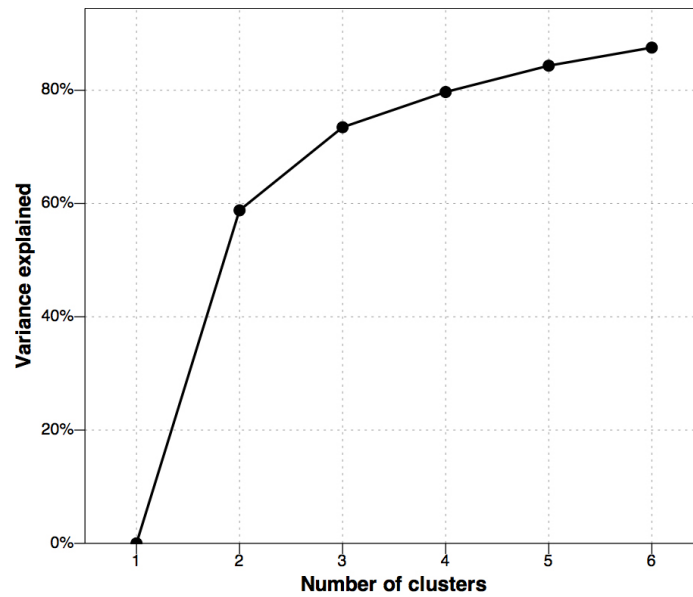


Figure 3. Graphical representation of the number of clusters versus the amount of explained variance.

The two cluster-solution classified learners into two almost equally sized groups: one group ($N = 20$) with high learners (i.e., larger d' scores in each AGL phase) and one group ($N = 22$) with low learners (i.e., lower scores in each phase). The three cluster-solution demonstrated that the cluster of high learners in fact consisted of two types of high learners: those who achieved high d' scores already in the first AGL test phase ($N = 12$), and those who had started with low scores but were able to quickly improve their performance, thus displaying a steeper increase in the obtained d' scores ($N = 9$) (see Figure 4).

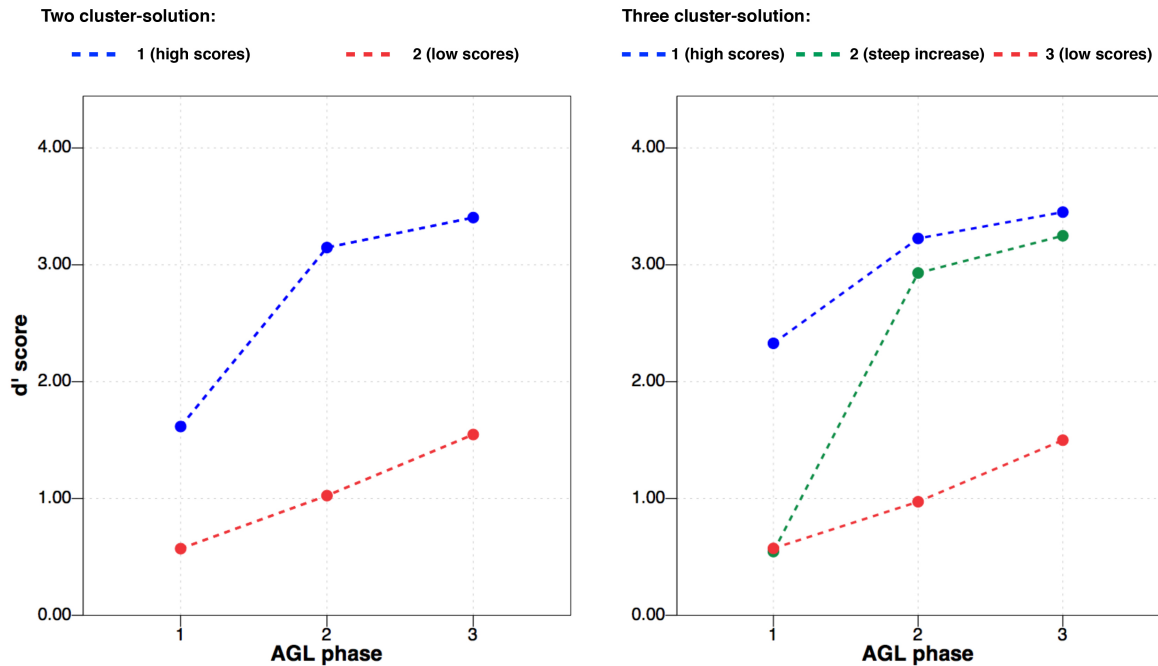


Figure 4. Results of the k -means cluster analysis on the behavioural data from the AGL task. Two solutions of the analysis are presented: the two cluster-solution on the left, and the three cluster-solution on the right. Points

When it comes to the way the original LAA scores were represented in the groups of learners determined by the analysis, the high learners, no matter whether obtaining high scores only at the end of the task, or also right from the beginning, were in both instances coupled with (on average) high analytical abilities. An independent samples t -test on the LLAMA_F scores for the high and low learners according to the two cluster-solution was significant ($t(40) = 3.39, p = .002, r = .472$); the effect of group was also significant for a three-clusters solution ($F(2, 39) = 6.38, p = .004, \eta^2 = .247$). A Games-Howell post-hoc test revealed that learners with low d' scores on the AGL task had significantly lower LAA scores ($M = 55.71, SD = 5.05$) than the “steep learners” ($M = 80.00, SD = 6.00, p = .015$), and the learners with high d' scores ($M = 79.17, SD = 6.09, p = .017$). There were no statistically significant differences between the “steep learners” and the learners with high d' scores ($p = .995$). A summary of the results of the analysis including mean d' scores per AGL task phase per group and mean scores on the pre-test and the LLAMA_F test are presented in Supplementary Table 1.

4. Imaging Data

4.1. Pre-processing

Imaging data acquired during the test phases of the AGL task were processed using FSL software Version 5.0.7 (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl; Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). Pre-processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00. The following pre-statistics processing was applied: motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002); non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian

kernel of 5 mm FWHM; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma = 50.0$ s). The functional images were registered to MNI-152 standard space (T1-standard brain averaged over 152 subjects; Montreal Neurological Institute, Montreal, QC, Canada) using a three-step registration from functional to high-resolution structural T2-image (rigid body, 6 degrees of freedom) to T1-image (rigid body, 6 degrees of freedom) to MNI-template (affine registration, 12 degrees of freedom). Registration was carried out using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001). Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). The hemodynamic response function (HRF) was computed as a double gamma function. The design matrix for each participant included grammatical and ungrammatical sentences as events of interest. Events of non-interest were not modelled. The contrasts tested for differential BOLD-response in grammaticality, i.e. for greater activity during grammatical than ungrammatical items and in ungrammaticality, i.e. for greater activity during ungrammatical than grammatical items.

4.2. Higher level analyses

4.2.1. Effect of LAA

A multi-session and multi-subject (repeated measures – three level) analysis was conducted with the aim of detecting BOLD-response differences and modulations between participants with different degrees of LAA (High LAA vs. Average LAA). The goal of this analysis was to establish brain activations typical for participants with high and average analytical abilities, as measured prior to the experiment, during novel grammar learning. The analysis consisted of the following steps: First, mean activation maps of the three phases per subject were calculated. The three phases of the experiment were not enough for a mixed effects model, hence a fixed effects model was used, by forcing the random effects variance to zero in FLAME (fMRI's Local Analysis of Mixed Effects) (Beckmann, Jenkinson, & Smith, 2003; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004; Woolrich, 2008). Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a cluster corrected significance threshold of $p = 0.05$ (Worsley, 2001). The results of this analysis were subsequently used as input for a two-sample unpaired t-test which was carried out using FLAME stage 1 (Beckmann et al., 2003; Woolrich et al., 2004; Woolrich, 2008). Pre-threshold masking was applied and a grey matter mask was used to mask out non-grey matter regions. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a cluster-corrected significance threshold of $p = 0.05$ (Worsley, 2001).

4.2.2. Learning patterns over time

Beside exploring brain activity distinguishing highly from moderately skilled learners as determined by the LLAMA_F test, we were interested in investigating the neural architecture behind the successful grammar learning process in time. The analysis revealing the heterogeneous learning patterns in our behavioural data (see section 3.2 above) enabled us to further explore the neural underpinnings of differently realised learning curves. This approach thus facilitated an analysis integrating both behavioural responses and time with brain activity (cf. Karuza, Emberson, & Aslin, 2014).

Of interest for the further analysis of the fMRI data were the differences in activation between the first and last test phase (run) and a comparison of that effect across the

groups identified in the k -means clusters analysis. Only the first and last time point was included in the analysis in order to observe the largest contrast in terms of the increase in correct responses.

Based on the two outcomes of the k -means clusters analysis, we conducted two group analyses of the fMRI data. The first analysis was based on the two cluster-solution (a 2 x 2 between-subjects ANOVA), the second on the three cluster-solution (a 2 x 3 between-subjects ANOVA). Both analyses were conducted using FEAT Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Pre-threshold masking was applied and a grey matter mask was used to mask out non-grey matter regions. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a cluster corrected significance threshold of $p = 0.05$ (Worsley, 2001). The goal of these analyses was to examine the main effect of time (phase of the AGL task) and group (cluster) and an interaction effect between them.

4.3. Results

4.3.1. Effect of LAA

A general linear model was used in the first level fMRI analysis to test for differential BOLD-responses to grammatical and ungrammatical items. Data from the three runs were averaged per participant and subsequently a two-sample unpaired t-test was conducted in order to compare mean activations between the High and Average LAA participants. On group level, only the contrast testing for differential BOLD-response in ungrammaticality, i.e. for greater activity during ungrammatical than grammatical items, yielded significant activation clusters for both High and Average LAA participants (see Table 1 and Figure 4). Grammaticality (greater activity during the presentation of grammatical than ungrammatical items) did not produce significant activation clusters.

The greater activity for ungrammatical items among the High LAA participants (group mean) could be attributed to two clusters of activity: (1) a widespread, bilateral fronto-parieto-temporal activation, with its peak in the left supplementary motor cortex bordering on superior frontal gyrus, including superior frontal gyrus, IFG and right angular gyrus; (2) a cluster encompassing subcortical areas of the right hippocampus and right thalamus. The Average LAA participants displayed a network of activity smaller in size, but similar in topography. It included six clusters, both in the left and right hemisphere, covering the following areas: (1) bilateral frontal areas including the paracingulate gyrus, middle and superior frontal gyri; (2) left IFG and left frontal orbital cortex; (3) right temporo-parietal areas including right angular and supramarginal gyri; clusters in (4) right and (5) left temporal lobe, including middle and superior temporal gyri and (6) a left temporo-parietal cluster encompassing the angular and supramarginal gyri and stretching out to the lateral occipital cortex.

Cortical region (peak)	L/R	BA	Size (voxels)	Peak location			Z_{\max}	p -value
				X	Y	Z		
High LAA (group mean)								
(1) Supplementary Motor Cortex / Superior Frontal Gyrus	L	6	35783	-6	8	62	5.35	$p < 0.001$
(2) Hippocampus	R	28	641	26	-22	-16	3.81	$p = 0.002$
Average LAA (group mean)								
(1) Paracingulate Gyrus / Superior Frontal Gyrus	R	8	6797	0	36	40	4.86	$p < 0.001$

Cortical region (peak)	L/R	BA	Size (voxels)	Peak location			Z_{\max}	p -value
				X	Y	Z		
(2) Frontal Orbital Cortex / Inferior Frontal Gyrus, pars triangularis	L	45	2537	-48	30	-10	4.54	$p < 0.001$
(3) Angular Gyrus	R	39	1159	54	-54	32	4.25	$p < 0.001$
(4) Middle Temporal Gyrus	R	37	918	64	-46	-8	3.97	$p < 0.001$
(5) Middle Temporal Gyrus	L	21	579	-54	-42	0	3.94	$p = 0.004$
(6) Angular Gyrus / Supramarginal Gyrus	L	40	481	-54	-54	34	3.9	$p = 0.011$
Average LAA > High LAA								
-								
High LAA > Average LAA								
(1) Supramarginal Gyrus / Angular Gyrus	R	39/40	867	56	-44	30	3.7	$p < 0.001$
(2) Cingulate Gyrus, posterior division	L	23	695	0	-42	24	3.73	$p = 0.001$
(3) Superior Frontal Gyrus / Middle Frontal Gyrus	R	6	647	28	2	62	3.3	$p = 0.002$

Table 1. Brain areas exhibiting greater activity for ungrammatical than grammatical items as a function of Average LAA and High LAA group: x, y and z coordinates in MNI space, regions labelled according to Harvard-Oxford Cortical and Subcortical Structural Atlases.

The High LAA participants displayed significantly more activity than the Average LAA participants, and the difference was observed in (1) the right superior temporal lobe in the supramarginal and angular gyri, stretching to the middle temporal gyrus, (2) the left cingulate gyrus and (3) the right superior and middle frontal gyri, see Table 1 and Figure 5.

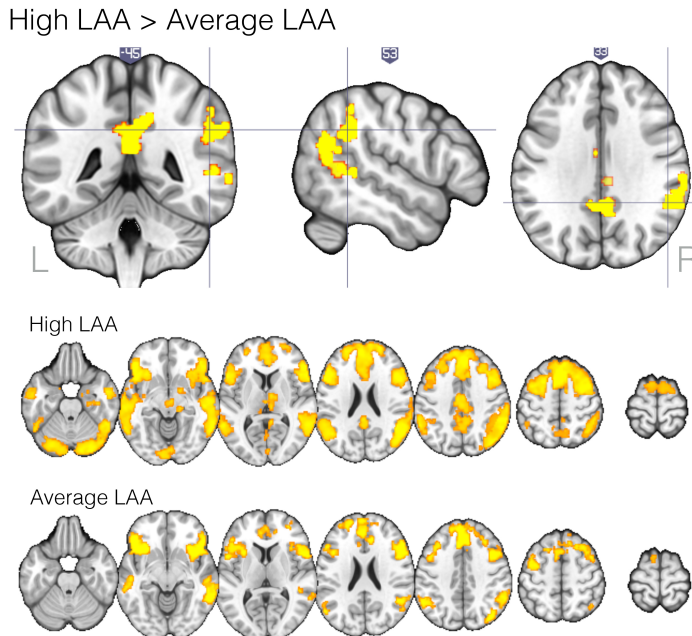


Figure 5. Brain activation in response to ungrammatical as compared to grammatical items significantly greater for the High than for the Average LAA participants (top panel) and group mean activation maps for High and Average LAA participants separately. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a cluster corrected significance threshold of $p = 0.05$ (Worsley, 2001).

4.3.2. Learning patterns over time

The second set of fMRI analyses concentrated on the different patterns of increase in correct responses on the task (as distinguished by the analysis of learning patterns in the behavioural data, see section 3.2), and two time points of the experiment: the first and last AGL task phase.

For each of the analyses two factors were distinguished in FEAT : Time (two levels: first and third phase) and Group (learners with high and low d' scores on the task as identified on the basis of the k -means cluster analysis described in section 3.2). The analysis was run twice, with the factor Group consisting of two, and three levels. Each time, we tested for the main effect of Time, main effect of Group and an interaction effect (Time x Group).

In each of the analyses, there was a significant effect of the factor Time. The factor Group yielded significant results only in the case of the 2 x 2 ANOVA analysis: there was no evidence of a group effect, or interaction between Group and Time when the groups were determined on the basis of the three cluster-solution of the k -means clusters analysis. Since the results for the factor Time are the same for both analyses, only the 2 x 2 ANOVA analysis is reported, see Table 2.

Cortical region (peak)	L/R	BA	Size (voxels)	Peak location			Z_{max}	p value
				X	Y	Z		
Time								
(1) Frontal Pole	L	9	2346	-18	54	34	4.68	$p < 0.001$
(2) Angular Gyrus / Lateral Occipital Cortex	R	39/40	1278	52	-62	38	4.43	$p < 0.001$
(3) Angular Gyrus / Lateral Occipital Cortex	L	39/40	1247	-52	-58	34	4.87	$p < 0.001$
(4) Superior Temporal Gyrus	L	21	797	-60	-30	0	3.23	$p < 0.001$
(5) Middle Temporal Gyrus	R	21	566	66	-40	-4	2.24	$p = 0.006$
Group								
(1) Angular Gyrus	L	39	540	-52	-58	30	4.17	$p = 0.008$
Interaction (Time x Group)								
-	-	-	-	-	-	-	-	-

Table 2. Brain areas exhibiting a significant differential modulation of BOLD signal for grammatical and ungrammatical items, as a function of time and performance on the AGL task (as determined by the analysis of learning patterns in the behavioural data, where two groups of learners, with high and low d' scores on the AGL task were identified); x, y and z coordinates are in MNI space, regions labelled according to Harvard-Oxford Cortical Structural Atlas.

Five clusters of activity were determined in the 2 x 2 FEAT analysis as being significantly modulated by the factor Time. They were localised in (1) the left frontal pole; (2) right and (3) left angular gyrus and lateral occipital cortex; (4) left and (5) right superior and middle temporal gyri. The factor Group yielded one significant activity cluster localised in the left angular and supramarginal gyri, with voxels in the lateral occipital cortex, see Figure 6. There was no significant interaction between the factors Time and Group.

FSL Featquery tool was used to further investigate the effects of Group and Time. First, we explored the lateralisation differences in the activity of parietal and temporal regions,

established to be significantly modulated by the factor Time. Four Region of Interest (ROIs) were determined: clusters of activity in the left and right temporal and parietal cortex distinguished as significantly modulated by the factor Time (see Table 2, Time). Subsequently, we extracted values for the ungrammatical > grammatical contrast from the ROI for each participant, and per each time-point and converted them into percentage signal change values. Figure 6 shows that the BOLD signal values were higher for the third AGL task phase for all ROIs, and for both groups. However, the magnitude of this effect was the strongest for the learners with high scores in the left parietal region.

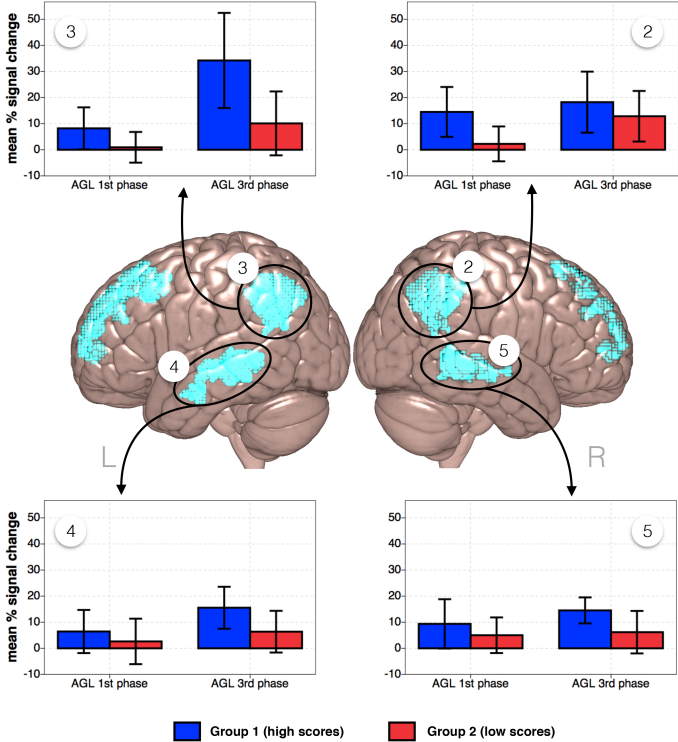


Figure 6. Time effect in the FEAT analysis investigating two time points of the experiment (first and last AGL task phase) and two groups (as determined by the analysis of learning patterns in the behavioural data): presented in the middle of the figure are the brain areas exhibiting a significant differential modulation of BOLD signal for grammatical and ungrammatical items, as a function of time, clusters 2 - 5 in Table 3. Graphs show the percentage signal change values for the ungrammatical > grammatical contrast, for each group distinguished in the analysis of learning patterns in the behavioural data, two time-points of the AGL task, and each ROI. Colours of the bars correspond to the colours of the group marking in Figure 4 (blue bars represent the mean values for learners with high scores, and red bars for learners with low scores). Error bars stand for 95% Confidence Intervals

A more detailed account for how the activity in the left parietal region reflected participants' proficiency in the AGL task was further obtained by investigating the effect of Group and extracting the contrast values from the ROI determined to be significantly modulated by the factor Group (see Table 2). After extracting the percentage signal change values for each participant, per each time-point, we correlated them with the respective d' scores. The results show that there was a significant positive correlation between the contrast values and the d' scores, but only in the 3rd phase of the AGL task ($r(40) = .144, p = .36$ and $r(40) = .400, p = .009$, for the 1st and 3rd AGL task phase respectively), see Figure 7.

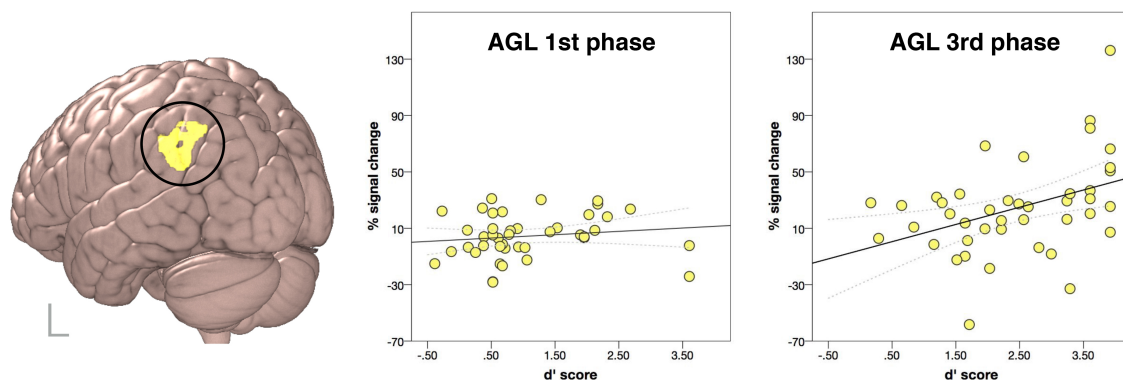


Figure 7. Group effect in the FEAT analysis investigating two time points of the experiment (first and last phase) and two groups (as determined by the analysis of learning patterns in the behavioural data): presented on the left is the brain area exhibiting a significant differential modulation of BOLD signal for grammatical and ungrammatical items, as a function of performance on the task. Graphs on the right show the percentage signal change values for the ungrammatical > grammatical contrast, in relation to the respective d' scores on the 1st and 3rd phase of the task.

Lastly, looking back at the analysis of learning patterns in the behavioural AGL task data (see Figure 4), we were interested in exploring the steep learning curve of participants who in the 1st AGL task phase obtained low scores but were able to quickly improve their performance and score highly in phase 3. We then ran additional exploratory analyses on the fMRI data of participants classified in the three groups in order to gain insight into the neural mechanism behind the differently realised learning curves. However, due to the fact that this effect was not strong enough to reach statistical significance in the 2 x 3 FEAT analysis, we wish to make only cautious observations concerning these results.

Data from two runs of the experiment of the participants classified in each group were entered in three separate paired sample t-tests. The analysis was carried out using FLAME stage 1 (Beckmann et al., 2003; Woolrich et al., 2004; Woolrich, 2008). It revealed that for the “steep learners” group, the left angular gyrus (at coordinates -48, -54, 24) exhibited more activity at the end of the task than at the beginning ($Z_{\max} = 3.14$, $p = 0.026$); no brain areas were significantly more active at the beginning of the task than at the end, see Figure 8. This effect was specific to the “steep learners” group; analyses of the other two groups of learners showed that learners with low scores exhibited a difference in activity between the third and the first phase in bilateral parietal and frontal regions; learners scoring highly from the beginning of the task showed a cluster of activity localised in the left frontal regions, see Supplementary Figure 1.

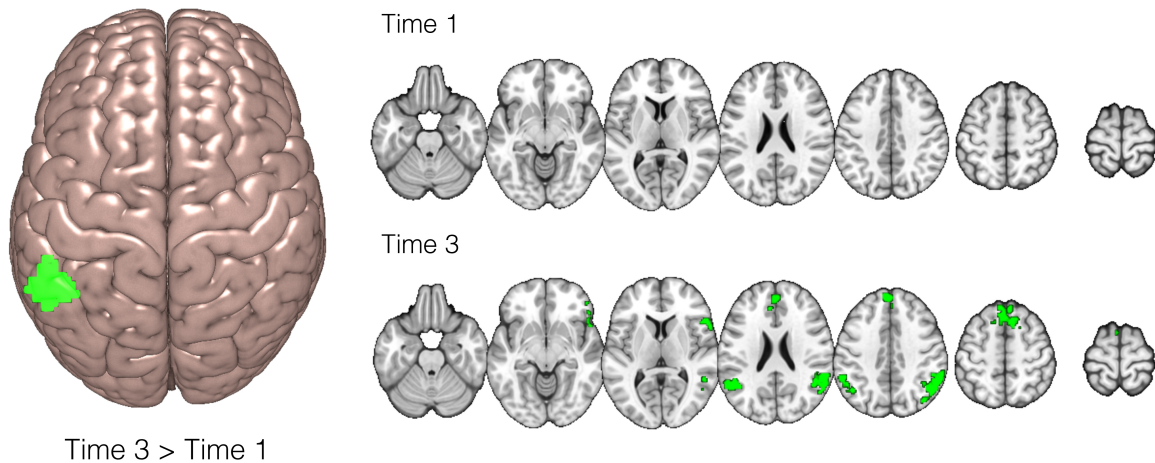


Figure 8. Brain activations in response to ungrammatical as compared to grammatical items for the group of participants determined by the analysis of learning patterns in the behavioural data as showing the most pronounced difference in scores on the AGL task between the first and the last phase. Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a cluster corrected significance threshold of $p = 0.05$ (Worsley, 2001). On the left, brain activation cluster significantly greater in phase 3 than in phase 1 of the AGL task; two mean activation maps per time point are shown on the right.

4.3.3. Summary of the Results

In summary, the obtained results indicated that participants with high language analytical abilities outperformed moderately skilled learners on the AGL task (see Figure 2). This difference observed on the behavioural level, was coupled with increased activity on the neural level among High LAA participants. The difference was localized predominantly in the right hemisphere (see Figure 5). Moreover, our design enabled a further investigation of the different ways the learning of a novel grammar proceeded over time, both behaviourally (see Figure 4), and on the neural level. We found that activity in the bilateral temporal and parietal regions increased over the course of the task and that the left parietal region displayed the biggest modulation of BOLD activity at the end of the task, especially among successful learners (see Figure 6). The amount of activity in the left angular gyrus correlated with the behavioural performance, but only in the last AGL task phase (see Figure 7). Additionally, we saw that a steep learning curve in the AGL task (starting out with low scores and quickly improving the performance) could be traced back to the modulation of BOLD activity in the left angular gyrus. Participants displaying such a pronounced difference in performance between the first and the last phase of the task, showed only one cluster of brain activation significantly greater in phase 3 than in phase 1 of the AGL task localised in the left angular gyrus (see Figure 8).

5. Discussion and Conclusions

The present study set out to investigate the neural correlates of individual differences in Artificial Grammar Learning. We wanted to examine both the correlates of cognitive abilities measured independently from the fMRI task, i.e. the language analytical abilities, and task-specific activations representing the learning process over time.

Our results suggest that individuals learning an artificial grammar who were chosen on the basis of high or average score on a language aptitude test, not only differ when it comes to behavioural data on an AGL task, but also on the level of brain activity. Both groups of

participants showed an increase in the amount of correct responses, but at the end of the task only the High LAA participants were on average able to correctly classify more than 90% of the items.

In the analyses of the imaging data, only the contrast testing for differential BOLD-response in ungrammaticality, i.e. greater activity during ungrammatical than grammatical items, yielded significant activation clusters for both High and Average LAA participants. Comparable findings illustrating response to violations of newly acquired grammatical rules were reported by e.g. Forkstam, Hagoort, Fernandez, Ingvar, & Petersson (2006) and Petersson, Forkstam, & Ingvar (2004) and interpreted (Petersson et al., 2004) in terms of a model of learning through negative evidence (cf. Karuza et al., 2014). In our analyses of imaging data, we aimed at investigating such process of learning and pinpointing the mechanisms responsible for individual differences.

We showed that participants with high analytical abilities, contrary to our expectations, formulated on the basis of the neural efficiency hypothesis, did not exhibit less distributed activity networks when compared to the moderately skilled ones. We believe this result to be coupled with the task's complexity and clear differences in performance between the High and Average LAA participants. As argued by Neubauer & Fink (2009), neural efficiency can be most clearly applicable to tasks where high and low ability individuals perform on a similar level (see also Nussbaumer et al., 2015). However, the two groups did not perform on the same level, most probably due to the task being more challenging than expected and the training in the paradigm being rather short (45 minutes). According to Neubauer & Fink (2009), in cases where higher ability individuals perform much better than lower ability individuals, the high performance is coupled with a stronger brain activation, and the neural efficiency hypothesis is not applicable. In other words, “[w]hen it comes to difficult tasks it seems that brighter individuals [...] invest more cortical effort than individuals with lower ability may do” (p. 1014). In particular, the difference between the High and Average LAA participants lay in the High LAA participants displaying more activity in the brain areas traditionally described as belonging to the language network (cf. Hagoort, 2014; Joseph, 1982), but localised in the right hemisphere, i.e. the right angular/supramarginal gyrus and superior frontal and middle frontal gyrus and in the posterior cingulate gyrus.

Although language processing has been repeatedly reported to activate a left-hemisphere focused network (Gernsbacher & Kaschak, 2003), and despite the notion present in the literature that suppression of contralateral activity can benefit language performance (Antonenko, Meinzer, Lindenberg, Witte, & Flöel, 2012; Schäfer et al., 2012; Thiel et al., 2006), our study showed a bilateral network of regions involved in the AGL task. The engagement of regions in the right hemisphere, in our view, might reflect the difference between proficient language processing and language learning, in a similar way to which bilateral language-related activity is found in young children during L1 processing (cf. e.g. Everts et al., 2009; Holland et al., 2001; Ressel, Wilke, Lidzba, Lutzenberger, & Krägeloh-Mann, 2008). Successful acquisition of new grammatical information in adults might thus be governed by parallel engagement of additional neural resources. Engagement of broader networks seems advantageous for the ultimate outcome of the learning process and the activity in the right homologue areas to the classically left-lateralised language regions might cause the learning to be more efficient.

In addition to examining the correlates of pre-tested language analytical ability, in an analysis integrating behavioural responses and time with brain activity, we found that

modulation of angular gyrus activity could be tied to differences in participants' performance on the task. Angular gyrus, middle temporal gyrus (both bilaterally) and the frontal pole were modulated when we investigated the effect of time. Furthermore, effort was made to identify the mechanisms involved in the progress of the learning process. By correlating the BOLD contrast values extracted from the left angular gyrus with the *d'* scores on respective AGL task phases, we believe to have underscored the importance of this region for successful learning of a novel grammar and rule acquisition.

The behavioural data on the AGL task increased steeply over the course of the experiment among a sub-group of the participants. A learning curve of such shape has been previously interpreted within the framework of Skill Acquisition Theory as typical for proceduralisation of newly acquired knowledge (DeKeyser, 1997, 2015) As argued by DeKeyser, L2 knowledge typically develops along three stages: declarative, procedural, and automatic. Declarative knowledge, acquired through observation and memorisation can turn into procedural knowledge rather quickly, and at fairly early stages of learning, especially among high-aptitude adult learners³. One of the clearest manifestations of such a shift is a decrease in error rates, most noticeably present among participants classified as “steep learners” in our data set (see Figure 4). DeKeyser's skill acquisition perspective on L2 development is also compatible with Ullman's (2004) declarative/procedural model, proposing distinct neural circuits for each type of knowledge, with - among others - the parietal cortex making up the procedural memory system⁴. By comparing two runs of the experiment among participants with the steepest learning curve on the task, we then believe to have tapped into the process of proceduralisation of the acquired knowledge. However, as pointed out in section 0 above, this particular set of findings should be seen as a trend and investigated further in future studies in order for the results to be inarguably robust.

Support for our interpretations comes nevertheless from recent studies establishing similar qualitative differences in the neural substrates of second language grammar processing associated with learning. In their study, Morgan-Short and colleagues (2015) employed an adapted version of BROCANTO, BROCANTO2, with the aim of investigating how an implicit learning context may mediate the neural representation of L2. The authors showed some learners to be able to quickly rely on neural circuits associated with L1 grammar and procedural memory (in this case, the LIFG and basal ganglia). Other learners increasingly used neural circuits possibly tied to extralinguistic processing (e.g., the left middle frontal gyrus). Moreover, the behavioural analysis of their data provided evidence for relationships between declarative and procedural learning abilities and L2 development (Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter, & Wong, 2014). The authors established a role for declarative memory at early, and for procedural memory at later stages of syntactic development. That early grammar learning relies on

³ We limit our interpretation of the results to the notion of “proceduralisation”, since according to DeKeyser (2015) automaticity can be achieved much later in L2 development: the shift from declarative to procedural knowledge is “followed by a much slower process of automatization of procedural knowledge” (p. 96) and depends on a large amount of practice.

⁴ An alternative, but congruent interpretations of L2 knowledge development has been recently offered by Opitz & Hofmann (2015) who present rule- and similarity-based learning mechanisms as the basis of AGL. In our understanding, the similarity-based learning can be interpreted on a par with the declarative knowledge in DeKeyser's and Ullman's account of L2 learning; the rule-based learning, with the procedural knowledge.

declarative memory and later, on procedural, has also been recently established by Tagarelli (2014). Comparable individual differences in recruitment of neural resources were also shown in natural L2 processing: early stage L2 learners were observed to progress through distinct stages of learning, with a varying rate of progression across learners (Tanner, McLaughlin, Herschensohn, & Osterhout, 2013).

Successful acquisition of a new grammar appears to highly depend upon the engagement of the parietal cortex and in particular, the angular and supramarginal gyri. Due to its location at the intersection of parietal, occipital and temporal lobes, the angular gyrus has been characterised as a cross-modal integrative hub bringing together information from different modalities and processing subsystems (see Seghier, 2013 for review). Moreover, already in the early 1900s, the left lower parietal region was delineated as the “language talent area” by the German neurologist Pötzl (1925), who pointed it out to be the site underlying the exceptionality of the brains of multilinguals (Della Rosa et al., 2013). His findings seem to converge with recent neuroimaging investigations: the inferior parietal cortex has been shown to play a crucial role in complex language functions and memory processes, phonological representation, semantic integration, and second language vocabulary learning (Della Rosa et al., 2013; Li, Legault, & Litcofsky, 2014; Mechelli et al., 2004; Yang, Gates, Molenaar, & Li, 2015). More specifically, O’Connor, Han, & Dobbins (2010) have found the angular and supramarginal gyri to act as markers of violations in memory expectations, where a violation is a sign of contradiction between retrieval outcomes and expectations, and argued for the importance of both superior and inferior aspects of the lateral parietal regions for noting a lack of correspondence between what is anticipated from memory retrieval, and what is unusual. The inferior parietal cortex has furthermore been argued to form part of the attention and memory systems in that it is part of a “bottom-up” attentional subsystem facilitating the automatic allocation of attention to task-relevant information (Ciaramelli, Grady, Levine, Ween, & Moscovitch, 2010), particularly to retrieved memories (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008; cf. Seghier, 2013). More recently, in the effort to explore the computational characteristics of sentence processing regions, the angular gyrus has been implicated in the combinatorial network participating in the construction of complex meaning (Poeppel, 2014).

In the current study, right angular and supramarginal gyri activity differentiated highly from moderately skilled learners who were presented with grammatical and ungrammatical sentences of a newly acquired artificial language. The regions were more active during presentation of ungrammatical sentences among the High LAA than in the Average LAA participants, pointing to the fact that High LAA learners reacted more strongly to violations in sentences in a newly acquired language than the Average LAA learners, either due to allocating more attention to the stimuli or by having formed better knowledge representations of the acquired material, or both. Since the materials used in this study did not have a semantic component, the activations of angular gyrus present in our data suggest that its role in the construction of complex meaning (as proposed by Poeppel, 2014) might be predominantly combinatorial in nature. Moreover, as pointed out above, modulation of activity in the left angular gyrus seems to be tied to the proceduralisation process of new rules - rules determining ways in which a number of unknown words can be combined and form sentences.

Another neural marker of high analytical abilities in our data was the posterior cingulate gyrus, which has previously been shown to be modulated by working memory tasks (Greicius, Krasnow, Reiss, & Menon, 2003), binding semantic representations on a multi-modal level (Jessen et al., 1999), to be implicated in successful episodic memory retrieval

(Cabeza & Nyberg, 2000) and recently in memory consolidation of complex events (Bird, Keidel, Ing, Horner, & Burgess, 2015). Morgan-Short et al. (2015) found the activity in left cingulate gyrus to increase over time during the AGL task and suggested increased engagement of attentional processes as an interpretation for this finding. The posterior cingulate activations in our study might help understand the nature of the skill in question: LAA seems to depend on good working memory, attention abilities and episodic memory and future studies employing neuropsychological testing and SLA diagnostic instruments might be able to elucidate the way these constructs interact with each other.

Interestingly, the previous findings concerning Artificial Grammar Learning and the interactions of the hippocampal system and the prefrontal cortex (Opitz & Friederici, 2003) were only partially replicated in this study. Hippocampus activity was only observed among the High LAA participants and was not part of the significant activation clusters revealed by the Time by Group analysis. The role of the hippocampus in learning language rules and its involvement in language analytical abilities remains unclear on the basis of present results.

In terms of its neural correlates, success in foreign language learning has so far been linked to, among others, an anatomically larger primary auditory cortex (Golestani, Molko, Dehaene, LeBihan, & Pallier, 2007), increased white matter integrity (Flöel, de Vries, Scholz, Breitenstein, & Johansen-Berg, 2009; Loui, Li, & Schlaug, 2011), differences in the functionality of the hippocampus (Breitenstein et al., 2005), as well as differences in structural (Xiang et al., 2012) and functional connectivity (Reiterer et al., 2011; Veroude et al., 2010) of the brain (see also Erard, 2012). With this study, we aimed at extending these results and gaining more insights into the neural mechanisms underlying individual differences in second language acquisition, with particular focus on novel grammar learning. We have shown that activity of the higher order associative brain areas (i.e. angular gyrus, cingulate gyrus) plays a crucial role in successful acquisition of a new grammar. In particular, we found that the activity of the right angular and supramarginal gyri during learning of a new grammar might be seen as a marker of high analytical skills, and the left angular gyrus as a site marking progress in acquiring language-like rules. In other words, our results point to the conclusion that engagement of right parietal regions is crucial for efficient learning, but in order to process the newly acquired rules proficiently, the language specialised left-hemisphere activations must be present. The sooner they occur, the better the task performance.

The present study has at least two theoretical implications. First it brings nuance to the neural efficiency hypothesis by showing that its predictions could not be borne out in a demanding AGL task focussing on initial stages of learning. Predictions of neural efficiency are yet to be tested in future studies investigating L2 learning in more advanced stages, or in contexts where learners' behavioural performance is equal, but the underlying skills vary. Secondly, the present study suggests that the neural mechanisms supporting L2 syntax abilities differ as a function of proficiency in the task. Whether these mechanisms can unequivocally be coupled with distinct stages of knowledge development (i.e. declarative, procedural and automatic as proposed by DeKeyser, 2015), or engagement of distinct memory systems (i.e. declarative and procedural, as put forward by Ullman, 2004), should be investigated further. A conclusive test for the predictions of the Skill Acquisition Theory and the neural underpinnings of development of L2 knowledge along the declarative, procedural, and automatic stages would in any case require a longitudinal design tracking learners over a period of L2 acquisition considerably longer than the 45 minutes offered by the present design.

Further questions deriving from this research concern inter-individual differences at the network level, such as differences in connectivity networks active during similar tasks. Also, research should try to elucidate the nature of structural correlates of LAA and explore whether the involvement of the right hemisphere in the task employed in this study, could also be characterised by weaker lateralization of e.g. the arcuate fasciculus (cf. Catani et al., 2007; López-Barroso et al., 2013). Other subcomponents of language aptitude await further neurolinguistic investigations. Furthermore, it remains an open question whether results obtained by this or any other study on individual differences in SLA, such as language aptitude levels, are also due to individual's personal experiences, or only their innate qualities.

6. Acknowledgements

This research was supported by the Leiden University Centre for Linguistics and the NWO Graduate Programme. The funding sources had no involvement in the preparation of the manuscript, or in the decision to submit the article for publication. We thank Bahar Azizi, Ferdi Van De Kamp and Bobby Ruijgrok for their help in data acquisition, and Bertram Opitz for advice on creation of the stimulus material. We are grateful to Tom F. Wilderjans for advice on the statistical analysis and to two anonymous reviewers for their insightful feedback on an earlier version of the manuscript.

7. References

- Abrahamsson, N., & Hyltenstam, K. (2008). The Robustness of Aptitude Effects in Near-Native Second Language Acquisition. *Studies in Second Language Acquisition*, 30(4), 481. <http://doi.org/10.1017/S027226310808073X>
- Aldenderfer, M., & Blashfield, R. (1984). *Cluster Analysis*. Newbury Park, CA, CA: Sage Publications.
- Antonenko, D., Meinzer, M., Lindenberg, R., Witte, A. V., & Flöel, A. (2012). Grammar learning in older adults is linked to white matter microstructure and functional connectivity. *NeuroImage*, 62(3), 1667–74. <http://doi.org/10.1016/j.neuroimage.2012.05.074>
- Beckmann, C., Jenkinson, M., & Smith, S. M. (2003). General multi-level linear modelling for group analysis in fMRI. *Neuroimage*, 20(2), 1052–1063.
- Bird, C. M., Keidel, J. L., Ing, L. P., Horner, A. J., & Burgess, N. (2015). Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex. *Journal of Neuroscience*, 35(43), 14426–14434. <http://doi.org/10.1523/JNEUROSCI.1774-15.2015>
- Breitenstein, C., Jansen, A., Deppe, M., Foerster, A.-F., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage*, 25(3), 958–68. <http://doi.org/10.1016/j.neuroimage.2004.12.019>
- Brod, G., & Opitz, B. (2012). Does it really matter? Separating the effects of musical training on syntax acquisition. *Frontiers in Psychology*, 3(December), 543. <http://doi.org/10.3389/fpsyg.2012.00543>

- Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). Parietal Cortex and Episodic Memory: An Attentional Account. *Nature Reviews. Neuroscience*, *9*(8), 613–625. <http://doi.org/10.1038/nrn2459>
- Cabeza, R., & Nyberg, L. (2000). Imaging Cognition II: An Empirical Review of 275 PET and fMRI Studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47. <http://doi.org/10.1162/08989290051137585>
- Carroll, J. B. (1981). Twenty-five years of research in foreign language aptitude. In K. Diller (Ed.), *Individual Differences and Universals in Language Learning Aptitude* (pp. 83–118). Newbury House.
- Catani, M., Allin, M. P. G., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., & Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proc. Natl. Acad. Sci. U. S. A.*, *104*(43), 17163–8. <http://doi.org/10.1073/pnas.0702116104>
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245–276.
- Ciaramelli, E., Grady, C., Levine, B., Ween, J., & Moscovitch, M. (2010). Top-down and bottom-up attention to memory are dissociated in posterior parietal cortex: neuroimaging and neuropsychological evidence. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(14), 4943–4956. <http://doi.org/10.1523/JNEUROSCI.1209-09.2010>
- De Graaff, R. (1997). The Esperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, *19*, 249–276.
- DeKeyser, R. M. (1997). Beyond Explicit Rule Learning. *Studies in Second Language Acquisition*, *19*(02), 195–221.
- DeKeyser, R. M. (2000). The Robustness Of Critical Period Effects In Second Language Acquisition. *Studies in Second Language Acquisition*, *22*, 499–533.
- DeKeyser, R. M. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (Second, pp. 97–114). New York: Routledge.
- Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., & Abutalebi, J. (2013). A neural interactive location for multilingual talent. *Cortex: a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(2), 605–8. <http://doi.org/10.1016/j.cortex.2012.12.001>
- Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. Oxford: Blackwell Publishing.
- Ellis, R. (2008). *The Study of Second Language Acquisition* (2nd Ed.). Oxford: Oxford University Press.
- Erard, M. (2012). *Babel No More: The Search for the World's Most Extraordinary Language Learners*. Free Press.
- Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*(2), 147–172.

- Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. C. M. (2015). The Relationship Between Artificial and Second Language Learning. *Cognitive Science*. <http://doi.org/10.1111/cogs.12257>
- Everts, R., Lidzba, K., Wilke, M., Kiefer, C., Mordasini, M., Schroth, G., ... Steinlin, M. (2009). Strengthening of laterality of verbal and visuospatial functions during childhood and adolescence. *Human Brain Mapping*, *30*(2), 473–83. <http://doi.org/10.1002/hbm.20523>
- Flöel, A., de Vries, M. H., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *NeuroImage*, *47*(4), 1974–81. <http://doi.org/10.1016/j.neuroimage.2009.05.046>
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage*, *32*(2), 956–967. <http://doi.org/10.1016/j.neuroimage.2006.03.057>
- Freundlieb, N., Ridder, V., Dobel, C., Enriquez-Geppert, S., Baumgaertner, A., Zwisler, P., ... Liuzzi, G. (2012). Associative vocabulary learning: development and testing of two paradigms for the (re-) acquisition of action- and object-related words. *PLoS One*, *7*(6), e37033. <http://doi.org/10.1371/journal.pone.0037033>
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(1), 529–534.
- Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging Studies of Language Production and Comprehension. *Annual Review of Psychology*, *54*(1), 91–114. <http://doi.org/10.1146/annurev.psych.54.101601.145128>
- Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, *17*(3), 575–582.
- Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(1), 253–258. <http://doi.org/10.1073/pnas.0135058100>
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28C*, 136–141. <http://doi.org/10.1016/j.conb.2014.07.013>
- Haier, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., ... Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence*.
- Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9–15. <http://doi.org/10.1016/j.lindif.2015.10.003>
- Harley, B., & Hart, D. (1997). Language Aptitude And Second Language Proficiency In Classroom Learners Of Different Starting Ages. *Studies in Second Language Acquisition*, *19*(3).

- Hauser, M. F. A. a, Hofmann, J., & Opitz, B. (2012). Rule and similarity in grammar: Their interplay and individual differences in the brain. *NeuroImage*, *60*(4), 2019–2026. <http://doi.org/10.1016/j.neuroimage.2012.02.016>
- Holland, S. K., Plante, E., Weber Byars, A., Strawsburg, R. H., Schmithorst, V. J., & Ball, W. S. (2001). Normal fMRI Brain Activation Patterns in Children Performing a Verb Generation Task. *NeuroImage*, *14*(4), 837–843. <http://doi.org/10.1006/nimg.2001.0875>
- Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: behavioural predictors and neural substrates. *Brain and Language*, *127*(3), 366–76. <http://doi.org/10.1016/j.bandl.2012.11.006>
- Hultén, A., Laaksonen, H., Vihla, M., Laine, M., & Salmelin, R. (2010). Modulation of brain activity after learning predicts long-term memory for words. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(45), 15160–4. <http://doi.org/10.1523/JNEUROSCI.1278-10.2010>
- IBM SPSS. (2012). IBM SPSS. *IBM Software Business Analytics*.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. M. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*, 825–841.
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156.
- Jessen, F., Erb, M., Klose, U., Lotze, M., Grodd, W., & Heun, R. (1999). Activation of human language processing brain regions after the presentation of random letter strings demonstrated with event-related functional magnetic resonance imaging. *Neuroscience Letters*, *270*(1), 13–16. [http://doi.org/10.1016/S0304-3940\(99\)00453-X](http://doi.org/10.1016/S0304-3940(99)00453-X)
- Joseph, R. (1982). The neuropsychology of development hemispheric laterality, limbic language, and the origin of thought. *Journal of Clinical Psychology*, *38*(1), 4–33.
- Karuza, E. A., Emberson, L. L., & Aslin, R. N. (2014). Combining fMRI and behavioral measures to examine the process of human learning. *Neurobiology of Learning and Memory*, *109*(September), 193–206. <http://doi.org/10.1016/j.nlm.2013.09.012>
- Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, (June), 1–24. <http://doi.org/10.1016/j.cortex.2014.05.001>
- López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(32), 13168–73. <http://doi.org/10.1073/pnas.1301696110>
- Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage*, *55*(2), 500–507. <http://doi.org/10.1016/j.neuroimage.2010.12.022>

- Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide. A User's Guide*. Psychology Press.
- Meara, P. M. (2005). *Llama Language Aptitude Tests*. Swansea: Lognostics.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757.
- Morgan-Short, K., Deng, Z., Brill-Schuetz, K. A., Faretta-Stutenberg, M., Wong, P. C. M., & Wong, F. C. K. (2015). A View Of The Neural Representation Of Second Language Syntax Through Artificial Language Learning Under Implicit Contexts Of Exposure. *Studies in Second Language Acquisition*, *37*(02), 383–419.
<http://doi.org/10.1017/S0272263115000030>
- Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, *17*(01), 56–72.
<http://doi.org/10.1017/S1366728912000715>
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, *6*(7), 774–81.
<http://doi.org/10.1038/nn1077>
- Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience and Biobehavioral Reviews*, *33*(7), 1004–23.
<http://doi.org/10.1016/j.neubiorev.2009.04.001>
- Nussbaumer, D., Grabner, R. H., & Stern, E. (2015). Neural efficiency in working memory tasks: The impact of task demand. *Intelligence*, *50*, 196–208.
<http://doi.org/10.1016/j.intell.2015.04.004>
- O'Connor, A. R., Han, S., & Dobbins, I. G. (2010). The inferior parietal lobule and recognition memory: expectancy violation or successful retrieval? *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(8), 2924–2934.
- Opitz, B., Ferdinand, N. K., & Mecklinger, A. (2011). Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience*, *5*(February), 8. <http://doi.org/10.3389/fnhum.2011.00008>
- Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, *19*(4), 1730–1737.
[http://doi.org/10.1016/S1053-8119\(03\)00170-8](http://doi.org/10.1016/S1053-8119(03)00170-8)
- Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity-based learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(39), 8436–40.
<http://doi.org/10.1523/JNEUROSCI.2220-04.2004>
- Opitz, B., & Friederici, A. D. (2007). Neural basis of processing sequential and hierarchical syntactic structures. *Human Brain Mapping*, *28*(7), 585–92.
<http://doi.org/10.1002/hbm.20287>

- Opitz, B., & Hofmann, J. (2015). Concurrence of rule- and similarity-based mechanisms in artificial grammar learning. *Cognitive Psychology*, *77*, 77–99. <http://doi.org/10.1016/j.cogpsych.2015.02.003>
- Petersson, K. M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, *120*(2), 83–95. <http://doi.org/10.1016/j.bandl.2010.08.003>
- Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, *28*(3), 383–407. <http://doi.org/10.1016/j.cogsci.2003.12.003>
- Petersson, K. M., & Hagoort, P. (2012). The neurobiology of syntax: beyond string sets. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1971–1983.
- Poeppel, D. (2014). The Neurobiology of Language. In G. Marcus & J. Freeman (Eds.), *The Future of the Brain: Essays by the World's Leading Neuroscientists* (pp. 139–148). Princeton University Press.
- Pötzl, O. (1925). Über die parietal bedingte Aphasie und ihren Einfluss auf das Sprechen mehrerer Sprachen. *Zeitschrift Für Die Gesamte Neurologie Und Psychiatrie*, *96*, 100–1124.
- Prat, C. S. (2011). The brain basis of individual differences in language comprehension abilities. *Linguistics and Language Compass*.
- Prat, C. S., & Just, M. A. (2011). Exploring the neural dynamics underpinning individual differences in sentence comprehension. *Cerebral Cortex (New York, N.Y. : 1991)*, *21*(8), 1747–60. <http://doi.org/10.1093/cercor/bhq241>
- Prat, C. S., Keller, T. A., & Just, M. A. (2007). Individual differences in sentence comprehension: a functional magnetic resonance imaging investigation of syntactic and lexical processing demands. *Journal of Cognitive Neuroscience*, *19*(12), 1950–63. <http://doi.org/10.1162/jocn.2007.19.12.1950>
- Prat, C. S., Long, D. L., & Baynes, K. (2007). The representation of discourse in the two hemispheres: an individual differences investigation. *Brain and Language*, *100*(3), 283–94. <http://doi.org/10.1016/j.bandl.2006.11.002>
- Psychology Software Tools. (2012). E-Prime. Pittsburgh, PA.
- Qualtrics. (2013). Qualtrics. Provo, Utah.
- R Development Core Team. (2015). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*. <http://doi.org/10.1007/978-3-540-74686-7>
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*.
- Reichle, E. D., Carpenter, P. A., & Just, M. A. (2000). The neural bases of strategy and skill in sentence-picture verification. *Cognitive Psychology*, *40*(4), 261–95. <http://doi.org/10.1006/cogp.2000.0733>

- Reiterer, S., Pereda, E., & Bhattacharya, J. (2011). On a Possible Relationship between Linguistic Expertise and EEG Gamma Band Phase Synchrony. *Frontiers in Psychology*, 2(November), 334. <http://doi.org/10.3389/fpsyg.2011.00334>
- Ressel, V., Wilke, M., Lidzba, K., Lutzenberger, W., & Krägeloh-Mann, I. (2008). Increases in language lateralization in normal children as observed using magnetoencephalography. *Brain and Language*, 106(3), 167–76. <http://doi.org/10.1016/j.bandl.2008.01.004>
- Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47(1), 45–99.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 319–353). Cambridge University Press.
- Schäfer, K., Blankenburg, F., Kupers, R., Grüner, J. M., Law, I., Lauritzen, M., & Larsson, H. B. W. (2012). Negative BOLD signal changes in ipsilateral primary somatosensory cortex are associated with perfusion decreases and behavioral evidence for functional inhibition. *NeuroImage*, 59(4), 3119–27. <http://doi.org/10.1016/j.neuroimage.2011.11.085>
- Seghier, M. L. (2013). The angular gyrus: multiple functions and multiple subdivisions. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 19(1), 43–61. <http://doi.org/10.1177/1073858412440596>
- Skehan, P. (2002). Theorizing and updating aptitude. In *Individual Differences and Instructed Language Learning* (pp. 69–93).
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155.
- Steinley, D., & Brusco, M. J. (2007). Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques. *Journal of Classification*, 24(1), 99–121. <http://doi.org/10.1007/s00357-007-0003-0>
- Tagarelli, K. M. (2014). The Neurocognition of Adult Second Language Learning: An fMRI Study.
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, 16(02), 367–382. <http://doi.org/10.1017/S1366728912000302>
- Thiel, A., Schumacher, B., Wienhard, K., Gairing, S., Kracht, L. W., Wagner, R., ... Heiss, W.-D. (2006). Direct demonstration of transcallosal disinhibition in language networks. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 26(9), 1122–7. <http://doi.org/10.1038/sj.jcbfm.9600350>
- Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, 18(4), 267–276. <http://doi.org/10.1007/BF02289263>
- Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative / procedural model, 92, 231–270. <http://doi.org/10.1016/j.cognition.2003.10.008>

- Veroude, K., Norris, D. G., Shumskaya, E., Gullberg, M., & Indefrey, P. (2010). Functional connectivity between brain regions involved in learning words of a new language. *Brain and Language*, *113*(1), 21–7. <http://doi.org/10.1016/j.bandl.2009.12.005>
- Wilderjans, T. F., Ceulemans, E., & Meers, K. (2013). CHull: a generic convex-hull-based model selection method. *Behavior Research Methods*, *45*(1), 1–15. <http://doi.org/10.3758/s13428-012-0238-5>
- Woolrich, M. W. (2008). Robust group analysis using outlier inference. *NeuroImage*, *41*(2), 286–301.
- Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747.
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage*, *14*(6), 1370–1386.
- Worsley, K. J. (2001). Statistical analysis of activation images. In P. Jezzard, P. M. Matthews, & S. M. Smith (Eds.), *Functional MRI: An Introduction to Methods*. OUP.
- Xiang, H., Dediu, D., Roberts, L., Norris, D. G., & Hagoort, P. (2012). Language Aptitude , Working Memory , and IQ in the Perisylvian Language Network. *Language Learning*, *62*(September), 110–130.
- Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2015). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics*, *33*, 29–49. <http://doi.org/10.1016/j.jneuroling.2014.09.004>
- Yilmaz, Y. (2012). Relative Effects of Explicit and Implicit Feedback: The Role of Working Memory Capacity and Language Analytic Ability. *Applied Linguistics*, *34*(3), 344–368. <http://doi.org/10.1093/applin/ams044>