

# THE NEUROSCIENCE OF IMPLICIT LEARNING

John N. Williams

University of Cambridge

First revision for *Language Learning* Special Issue (4/11/2019)

## Abstract

Over the past decades, research employing artificial grammar, sequence learning and statistical learning paradigms has flourished, not least because these methods appear to offer a window, albeit with a restricted view, on implicit learning processes underlying natural language learning. But these paradigms usually provide relatively little exposure, use meaningless stimuli, and do not even necessarily target natural language structures. So the question arises whether they engage the same brain regions as natural language. The aim of this review is to use data from brain imaging, brain stimulation, and the effects of brain damage to identify the main brain regions that show sensitivity to structural regularities in implicit learning paradigms and to consider their relationship to natural language processing and learning.

## Introduction

Implicit learning refers to the process of learning without intention, and even without awareness of what has been learned. It is regarded as a basic form of learning that makes a major contribution to the acquisition of many motor, perceptual, and cognitive skills, not least the ability to speak a first or second language. The focus here will be on implicit learning of structural regularities. Can structure-specific activity be detected in the brain during implicit learning tasks? If so, do the brain regions that are involved vary according to the nature of the regularities to be learned or the domain in which they are instantiated? Do the results of imaging studies tie in with what is known about implicit learning after various types of brain damage? As in neuroscience more generally the ultimate aim is to understand how and where specific computational functions are carried out in the brain. By studying the brain correlates of implicit learning of structural regularities of different types and in different tasks perhaps we can approach a true science of learning.

Of course, studying implicit learning of real world skills is immensely challenging, not only because of the time scales over which it occurs, but also because of the difficulty of being able to isolate learning of specific structural regularities from other features of the input, being able to prove that structural generalisations have been learned (for which a perfect record of inputs and outputs is necessary), and being able to prove that any learning is not affected by the participants' prior knowledge. Hence a number of laboratory implicit learning paradigms have been developed that provide convenient vehicles for studying the acquisition of pure structural regularities in unashamedly artificial systems. This review will focus on three such paradigms: Artificial Grammar learning (AGL), sequence learning, and statistical learning. Typically they involve short training periods of between 2 and 20 minutes (though some AGL studies have trained participants over days), the stimuli are devoid of any meaning, the underlying structural regularities are unlike any that the participants are likely to have encountered before, are generally unlike those found in natural domains like language, and the stimuli are stripped of all the other cues to structure that might occur in real world learning situations. Such tasks have certainly proved extremely useful in establishing that there is such a thing as implicit learning, in exploring its potential limits, individual differences and interactions with other learning processes. They also provide convenient vehicles for examining the brain systems involved in

implicit learning, either through brain imaging, brain stimulation, or the study of brain damaged patients. But do they have anything to do with natural language, or with natural language learning? This review aims to provide a synthesis of neuroscientific research on implicit learning in order to gauge our progress in achieving an understanding of where, and ideally how, implicit learning, in the sense defined above, happens in the brain, and how it relates to natural language processing and learning.

Coincidentally, Batterink, Paller, and Reber (2019) have recently published a review of a similar literature. The present work presents a more detailed evaluation of the evidence in order to expose inconsistencies between studies, a greater emphasis on triangulation, or in many cases lack of, across different methodologies, a closer consideration of whether studies convincingly tap into implicit learning and knowledge, and the implications of all of these for a functional interpretation of the results.

This article is not a metaanalysis of studies in the field, nor will it advance a particular theoretical perspective, or systematically lay out an agenda for future research (though indications of outstanding questions and possible avenues for research emerge along the way). Rather it is an attempt to lay out the state of the art in the field for the non-specialist, hoping to highlight, if nothing else, the complexities and inconsistencies in the evidence base, how these might be related to methodological variations, and to emphasise the problems inherent in making generalisations about localisation of brain function on the basis of the evidence reviewed. With a view to the latter, the following section briefly lays out some of the ongoing debates over the function of core brain regions in relation to language processing and non-implicit learning of natural language-like systems.

### Regions of interest

Previous reviews have considered the balance between learning-related brain changes in modality-specific and supramodal systems across different implicit and statistical learning paradigms (Batterink et al., 2019; Frost, Armstrong, Siegelman, & Christiansen, 2015; Julia & Claudia, 2018). Visual paradigms have sometimes, but not consistently, revealed learning-related changes in visual processing areas of the brain (e.g. occipital cortex, and even V1), and, as we shall see, auditory statistical

learning has revealed changes in brain areas associated with spoken language processing (e.g. left superior temporal gyrus, LSTG). In contrast, here the primary focus will be on supra-modal regions – areas that seem to perform some generic computational functions across modalities. This corresponds to the emphasis in discussions of this research in the literature, where the most prominent regions are the left inferior frontal region (LIFG), basal ganglia (shortened here to ‘BG’), and hippocampus (shortened here to ‘H’).

### *Broca’s area and adjoining regions*

Syntactic processing is strongly associated with the left inferior frontal gyrus (LIFG) which encompasses Broca’s area (i.e., Brodman’s areas (BA) 44/45, and adjoining BA 47). However, there has been much debate over the actual function of LIFG in sentence processing. Some argue that it is involved in building syntactic representations as classically conceived, i.e., as involving hierarchical structure and non-adjacent dependency relations (Friederici, 2002). Others argue that it constitutes a general “unification space” for structure building during sequential processing (Hagoort, 2013). Others argue that in fact LIFG involvement in sentence processing reflects working memory demands (Rogalsky & Hickok, 2011), or that it is primarily engaged in order to resolve syntactic or semantic ambiguity (Hsu, Jaeggi, & Novick, 2017; January, Trueswell, & Thompson-Schill, 2009; Tyler, Cheung, Devereux, & Clarke, 2013), resolve competition between alternatives (Schnur et al., 2009), or that it engages whenever general control demands increase (Novick, Trueswell, & Thompson-Schill, 2005; van de Meerendonk, Rueschemeyer, & Kolk, 2013). The domain-specificity of LIFG is also disputed. From a control perspective the involvement of LIFG in non-linguistic conflict resolution, such as in the Stroop task, has been emphasised (Hsu et al., 2017; January et al., 2009; van de Meerendonk et al., 2013). From a structure building point of view LIFG has been implicated in, amongst other things, music perception (Patel, 2003) and it has been suggested that LIFG is a “hub” that engages with other representational systems to effect either control or structure building in different domains (Hsu et al., 2017; Patel, 2003). Others have argued that there is language specialisation within the LIFG region, with syntactic structure building functions localisable to BA 44 and 45 (Fedorenko, Duncan, & Kanwisher, 2012; Hagoort, 2014), or possibly specifically BA 44, with BA 45

implicated in semantic combinatorics (Schnell et al., 2017). In the studies reviewed here, when ‘LIFG’ is referred to it can be assumed that this at least includes BA 44.

Imaging studies have explored the brain areas that are involved in adult explicit learning of natural phrase structure grammars (PSGs) under experimental conditions (Bahlmann, Schubotz, & Friederici, 2008; Hauser, Hofmann, & Opitz, 2012; Musso et al., 2003; Opitz & Friederici, 2003; Opitz & Friederici, 2004, 2007). As a shorthand these will be referred to collectively as the ‘explicit PSG’ studies. Only one of these studies used actual natural languages (Musso et al, 2003), whereas the others used phrase structure grammars (e.g. Brocanto) that generated strings of meaningless nonsense words like “aaf ploX glif rufi aak boke gum”, the structure of which corresponds to the natural language sequence determiner-noun-verb-adverb-determiner-adjective-noun. These grammars contain the equivalent of word classes – either 2 or 4 nonsense words were assigned to the ‘noun’ and ‘verb’ categories depending on the study (Opitz & Friederici, 2003; 2004). In all of these studies the participants were trained under explicit learning conditions – for example they cycled between presentation of examples and grammaticality judgment with feedback under instructions to work out the rules of the language. Such procedures encourage active hypothesis formation and testing; learning is intentional and leads to conscious knowledge. These studies have found that in grammaticality judgment tasks (GJTs) following training there is higher activation to non-grammatical (NG) than grammatical (G) items in LIFG and also adjoining areas of left ventral premotor cortex (vPMC) and the frontal operculum (FOP). Some studies have gone further and identified activity in the phylogenetically older vPMC or FOP regions with processing and learning of local, adjacent, dependencies, and activity in the more recently evolved LIFG with processing and learning of non-adjacent dependencies and hierarchical structure (Bahlmann et al., 2008; Opitz & Friederici, 2007; Wilson et al., 2015; Wilson, Marslen-Wilson, & Petkov, 2017). These studies at least tell us that, following explicit training, meaningless material that is structured in a language-like way can engage similar brain areas as natural language in the performance of a GJT. The question is whether, and for what other kinds of structures, LIFG and related regions are also engaged in implicit learning situations, and what this might tell us about the function of these regions.

## *Basal Ganglia*

The Basal Ganglia (BG) are a group of sub-cortical structures that include the caudate and putamen. They constitute a relatively primitive part of the brain and it was initially thought that they were primarily involved in movement control and motor learning. But more recently their role in cognitive functions, and even language, has come to the fore. It has become apparent that the BG are connected to many cortical areas. There are loops connecting putamen to motor cortex, the head of the caudate to prefrontal cortex (including BA 47, Dominey, Inui, & Hoen, 2009), and the body and tail of the caudate to visual and auditory cortex (Lim et al, 2014; Seger, 2008). With regard to function, Ullman (2001) has stressed the role of the BG in a ‘procedural’ (as opposed to ‘declarative’) system that supports automatic, and implicit, rule-based language processing and learning. Lieberman (2007) regards the BG as a general ‘sequencing engine’ that in relation to language “can form a potentially infinite number of different sentences by reordering, recombining, and modifying a finite set of words using a finite set of linguistic rules” (p. 51). This focus on subcortical structures in relation to language processing may seem surprising, and invites the idea that language was made possible by an adaptation of pre-existing motor systems (Lieberman, 2007). Given the importance of this claim it is not surprising that the precise role of the BG in language processing is debated. There is neuroimaging evidence that the BG is involved in purely syntactic, as opposed to semantic, processing (Moreno, Limousin, Dehaene, & Pallier, 2018; Moro et al., 2001). However, other researchers appear to limit its role to ancillary functions, rather than core syntactic processing. It has been argued that the BG may be particularly active when expectations are violated, a preferred interpretation suppressed, and controlled processes (in frontal cortex) need to be engaged to effect a repair, as happens in garden-path sentences for example (Mestres-Misse, Bazin, Trampel, Turner, & Kotz, 2014; Mestres-Misse, Turner, & Friederici, 2012; Sambin et al., 2012). These control functions that seem similar to those attributed to LIFG, reinforcing the idea that these regions work closely together. In the context of morphological processing and word recognition, it has been argued that the BG are only involved in “late stage” processing and that the essential syntactic and morphological computations are carried out elsewhere (Longworth, Keenan, Barker, Marslen-Wilson, & Tyler, 2005). It will be interesting to see the extent to which the BG are involved in performing implicit learning tasks given that they usually tap into

the very early phases of learning, prior to the acquisition of automaticity, and do not obviously entail reanalysis and repair processes.

### *Hippocampus*

The BG and the larger procedural system are classically contrasted with a declarative system that supports explicit learning and knowledge. The hippocampus (H) is a central component of the declarative system and is known to be essential for episodic memory – the conscious recollection of experiences (Squire, 1992). Damage to the hippocampus is associated with amnesia – the inability to recall the recent past, and more relevant here, the apparent inability to form new memories. Various kinds of implicit learning and memory are relatively intact in amnesia leading to the assumption that the declarative and procedural learning systems are dissociable. For this reason, studying amnesic performance on IL tasks is usually used as a means of assessing procedural learning without contamination from explicit knowledge.

Hippocampal functions are not typically related to structural aspects of language processing. However, studies of explicit learning of the artificial language Brocanto have found that left hippocampal activity with respect to baseline showed a decreasing trend from being greater than baseline at the start of training to being below baseline at the end (Opitz & Friederici, 2003; Opitz & Friederici, 2004). Opitz and Friederici (2004) nuance this by showing that the decreasing trend is associated with manipulations of similarity rather than rule application, and Hauser et al. (2012) show that in post-training GJT the (in this case right) hippocampus is related to similarity-, rather than rule-, based processing. The decreasing trend in hippocampal activity over training contrasts with the increasing trend in LIFG in the same studies (Opitz & Friederici, 2003; Opitz & Friederici, 2004). These studies therefore reveal the changing balance between similarity- and rule-based processing in the course of acquisition, at least as evidenced in the performance of grammaticality judgments (for related evidence based on individual differences and electrophysiology see Morgan-Short, Faretta-Stuttenberg, Brill-Scheutz, Carpenter, & Wong, 2014; Tanner, Inoue, & Osterhout, 2014; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013). This shift might suggest a competitive relationship between ‘declarative’ memory-based and ‘procedural’ rule-based systems (Ullman, 2004). Indeed there is evidence in favour of mutual exclusivity of the electrophysiological indices of these processes (in the form of P600 and N400 effects) in classroom language learning (Tanner et al., 2013). On

the other hand, using a similar artificial language training method to Opitz and Friederici (2003) an fMRI study by Kepinska, de Rover, Caspers, and Schiller (2018) found no evidence for functional connectivity between LIFG and H, implying independent contributions of these systems. Bear in mind that these studies all used explicit training paradigms. On the standard view one would not expect to find hippocampal involvement in implicit learning situations unless there is contamination from explicit memory processes.

### Triangulation

Functional magnetic resonance imaging (fMRI) reveals sites of increased blood flow in the healthy brain and is assumed to indicate processing activity. But to establish a causal connection between that activity and learning it is useful to know how alteration of the functioning of a specific region affects learning outcomes. One way of achieving this is through brain stimulation techniques such as transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS). Both of these methods induce a small electric current in the targeted brain region – in the case of TMS by using a magnet and relying on electromagnetic induction, and in the case of tDCS by actually applying a small current via electrodes to the scalp. The other way of establishing causality is by examining learning in brain damaged patients. Three types of brain damage will be considered. Damage to LIFG results in Broca’s aphasia. Patients display symptoms of agrammatism, characterised as halting and “telegraphic” production of sentences; that is, lacking grammatical morphemes. Damage to the BG results in Parkinson’s disease (PD). Dysfunction in the putamen affects movement via the loop to motor cortex, and dysfunction in the caudate affects cognitive functions via loops to other cortical regions (Hochstadt, Nakano, Lieberman, & Friedman, 2006). The BG secrete dopamine, which is necessary for motor skill learning, and so PD leads to learning impairments (Kawashima, Ueki, Kato, Ito, & Matsukawa, 2018). Damage to the hippocampus results in amnesia – the inability to consciously recall memories of events.

## Artificial grammar learning

Artificial grammar learning (AGL) experiments typically follow a methodology developed by Reber (1967, Experiment 2) – participants perform a short-term memory task on training strings, usually letters, that are generated by a finite state grammar (FSG), e.g., PVPXVPS, TXS. They are then told that the strings were in fact generated by a complex system and are asked to perform grammaticality judgments (GJs) on new strings using intuition and “gut feeling”. The FSGs used in AGL contain branching structures and recursive loops. For example, a string could start with either a P or a T. An initial T could be followed by one or more S’s, or none (TS, TSSS, or T), which would always be followed by an X (TSX, TSSSX, TX), which could either be followed by another X or a final S (e.g. TSXX, TXS). If followed by an X there could be repeating Ts, or none (TSXXTT, TSXX) always followed by a V (TSXXTTV, TSXXV), which could then either be followed by another V or P and S to end the string (TSXXTTVV, TSXXVPS). If the string had started with a P other branches and loops would be possible, and the paths from the different starting points could cross. Ungrammatical strings are formed by adding letters at inappropriate points, missing out letters, or making illegal transitions. Unsurprisingly participants are unable to report the structure of such grammars, and yet their GJs are significantly above chance. However, the extent of awareness of structure has been questioned (Shanks & St. John, 1994). Participants could be aware of some parts of the system, or of chunks (frequent bigrams and trigrams). Judgment-by-judgment measures of subjective states may provide more convincing evidence of implicit learning (Dienes & Scott, 2005) though this method has rarely been used in neuroscientific research.

The finite state grammars that are used in implicit learning research do not generate hierarchically organised, embedded, constituents of the kind that have, by some researchers, been seen as the essential characteristic of natural language for which LIFG is specialised (e.g., Friederici, Bahlmann, Heim, Schubotz, & Anwender, 2006).<sup>1</sup> Neither, unlike Brocanto, do they contain item classes. On the other hand, they do not only generate strings of adjacent dependencies – they can generate conditional dependencies as well. For example, in the grammar above, an S predicts

---

<sup>1</sup> Recursive loops might be regarded as a form of embedding, e.g. TSX, TSSX, TSSSX, but notice that only a single letter is repeated – there is no embedding of structural constituents beyond that.

an X if it followed the initial T or an S, but a single S predicts the end of the string if preceded by an X (as in TSXS or TSSXS). Hence AGs generate strings that are more complex than simple linear chains of associations between adjacent letters.

Nevertheless, in as much as they also contain adjacent dependencies then activation in the FOP or vPMC, but not LIFG, might be predicted on the basis of explicit learning paradigms.

Because GJs may be driven by item similarity as well as grammaticality ‘balanced chunk strength’ designs are sometimes used - grammatical and ungrammatical test strings are matched in terms of the frequency with which their constituent bigrams and trigrams occurred in training strings.

## Imaging

Typical fMRI studies of AG learning follow Reber’s general methodology. Note that only the grammaticality judgment task is performed inside the scanner. Hence brain activation has to be interpreted in the context of the judgment task being performed, and not as a direct measure of learning activity. An early study found that compared to a baseline task the GJT activated occipital gyri and LIFG (BA 45, 47) (Seger, Prabhakaran, Poldrack, & Gabrieli, 2000), and more recently Yang and Li (2012, implicit training condition) found greater activation in a GJT compared to baseline in 10 brain regions (many bilaterally) but including LIFG (BA 44). Other studies have also found that whilst the GJT activates a broad range of brain areas compared to a sensorimotor baseline task, the contrast between ungrammatical and grammatical test items (NG > G) reveals more restricted activations (recall that the studies of explicit learning of PSGs used the same contrast, see above). In some studies this activation difference is spread over the inferior frontal region in left and right hemispheres, BA 44, 45, 47, including the FOP (Forkstam et al, 2006; Petersson et al, 2012; Folia et al, 2014), but can also be more restricted to LIFG (BA 44, 45) (Petersson et al, 2004). Curiously, Lieberman et al (2004) report LIFG (BA 44) using the opposite G > NG contrast.<sup>2</sup> Lieberman et al. (2004) and Petersson et al. (2004) used the most similar training procedures, although there was more extensive training

---

<sup>2</sup> The reported Talarach coordinates are -44, 10, 18, which when converted to MNI are -46, 12, 17, which, for example, are very similar to the MNI coordinates for peak activations in BA 44 reported in Forkstam et al, (2006) (-45, 12, 24) and Petersson et al (2004) (-48, 16, 22).

in Petersson et al. (168 versus the 46 items in Lieberman et al.). Why this should have caused a reversal in the typical effect of grammaticality is unclear. In fact, Lieberman found no significant LIFG activations when contrasts controlled for chunk strength (e.g., G vs NG low chunk strength items), a type of analysis that is not reported in the other studies.

Where there is better agreement across studies is in the lack of LIFG sensitivity to chunk strength (Forkstam et al, 2006; Petersson et al, 2012; Folia et al, 2014; Lieberman et al, 2004) (note that Forkstam et al. (2006) found chunk strength effects in other frontal regions, just not in LIFG). This is important because it has been argued that effects of grammaticality in behavioural performance can be reduced to chunk strength if sufficiently sophisticated metrics are used (Kinder & Assmann, 2000). If this were the case then it should not be possible to find brain regions that are specifically sensitive to ‘grammaticality’ but not even relatively crude measures of chunk strength.

It is clear that activation in AGL is not confined to the FOP, and in one study using the NG > G contrast no FOP is reported at all (Petersson et al, 2004). Rather the general inferior frontal activations obtained in implicit AGL are similar to those obtained for explicit learning of grammars containing hierarchical structure and non-adjacent dependencies in that, at least across studies, they encompass both LIFG and FOP (Bahlmann et al., 2008; Opitz & Friederici, 2007). Where these studies do differ, however is with regard to lateralization, which varies considerably across AGL studies, and contrasts with the more consistent left lateralization obtained in the explicit PSG studies and in natural language syntactic processing. There is actually something of a generalisation here though – amongst the implicit AGL studies, RIFG is found when training extended over days (Forkstam et al., 2006; Petersson et al., 2012; Folia et al., 2014) but not when training occurred in a single session (Seger et al., 2000; Petersson et al., 2004; Lieberman et al., 2004; Yang & Li, 2012). Within the implicit AGL literature RIFG activity has been associated with explicit string recognition (Seger et al., 2000) and manipulations of chunk strength (Forkstam et al., 2006; Udden et al., 2008), and for explicit PSG with item similarity (Hauser et al., 2012). A recent metanalysis (Hartwigsen, Neef, Camilleri, Margulies, & Eickhoff, 2018) found RIFG involvement in a range of functions, ranging from action execution and inhibition (BA 44), spatial attention and explicit memory encoding, and cognitive control and emotional processing (BA 45) (but no language functions). It is possible

that extensive training leads to sufficiently strong chunk encoding that rejection of NG items requires greater executive control, drawing on RIFG.<sup>3</sup>

The above considerations might lead one to suspect GJT performance in the AGL studies was not entirely driven by implicit knowledge. In fact, none of the above studies used verbal report or subjective measures of awareness to establish the status of the knowledge acquired. In some studies GJT accuracy was unusually high (ranging from 73% to 87%) which is perhaps not surprising after multiple days of training (Forkstam et al., 2006; Petersson et al., 2012) or large numbers of training items on one day (Petersson et al., 2004). This might lead one to doubt that performance was based purely on implicit knowledge. Petersson et al.'s (2012) participants reported that they were only using "vague criteria" when making their judgments (behavioural data reported in Folia et al., 2008, p. 143) but their accuracy in a string generation test, though low, was strongly correlated with their grammaticality judgement accuracy, indicating a contribution of conscious knowledge. In Seger et al. (2000) performance was more in line with typical behavioural studies (58%) but this study did not use imaging contrasts that focus on grammaticality. In Yang & Li (2012) the fact that there was greater LIFG activity in a traditional implicit training condition than in an 'explicit' condition, and at a numerically lower level of accuracy (56% and 59% respectively), goes some way to showing that LIFG involvement is not necessarily associated with high accuracy, or a high likelihood of conscious knowledge.

Perhaps the best evidence in favour of the use of implicit knowledge comes from Folia & Petersson (2014) where the participants were asked to judge whether they liked or disliked the test strings, rather than judge their grammaticality. Grammatical strings were liked more than ungrammatical ones. It can be argued that performance was not contaminated by conscious knowledge of the grammar since grammaticality was strictly irrelevant to the task at hand. Indeed, unlike for the GJT task performed upon the same participants (imaging data reported in Petersson et al., 2012), there was no correlation between performance on the liking task and string generation accuracy (behavioural data for both tasks reported in Folia et al., 2008).

---

<sup>3</sup> These sets of studies also divide according to the mode of item presentation – only in Forkstam et al, Petersson et al., and Folia et al, 2014, were the training and testing strings presented one letter at a time. Why this should have induced RIFG activity is not clear, but this does illustrate how procedural variations between studies raises problems for functional interpretation.

Yet similar brain activations were obtained as for GJT – namely bilateral frontal activation of BA 44, 45, and 47 including the FOP. Overall what could be concluded about awareness is that evidence for participation of LIFG persists despite variation across studies in the likelihood of explicit knowledge contributions.

All of the above studies have looked at activation of cortical areas in relation to learning. Another approach is to look for the correlation between natural variation in brain structure and AGL performance, focusing on ‘white’ matter - the web of long-distance connections that allows communication between cortical (‘grey’ matter) regions. Modern MRI techniques (diffusion tensor tractography) can measure the ‘integrity’ of these white matter tracts (e.g. degree of myelination, density, and axon diameter). Behavioural sensitivity to grammaticality in GJT after relatively little training on an AGL (25 or 35 mins) has been shown to correlate with white matter integrity and functional connectivity in fibres emanating from left, but not right, IFG (Flöel, de Vries, Scholz, Breitenstein, and Johansen-Berg (2009), Antonenko, Meinzer, Lindenberg, Witte, and Flöel (2012)). The left lateralisation after relatively little training may be significant in light of the above discussion. Of course, here we are looking at the efficiency of connections from LIFG which is to acknowledge that performance is determined by how well it communicates with other regions. And in this case, as in the imaging studies above, the effect could in principle reflect either or both learning and judgment processes.

Artificial grammar learning experiments have also revealed basal ganglia involvement in grammaticality judgments. A greater brain response in the caudate has been obtained for grammatical than ungrammatical strings (Forkstam et al., 2006; Petersson et al., 2012; Lieberman, Chang, Chiao, Bookheimer, and Knowlton (2004), or in the GJT compared to a baseline task (Folia et al., 2014; Petersson et al., 2012). Yang & Li (2012) found a greater caudate response in their implicit than explicit condition. Interestingly, all except Lieberman et al. (2004) used sequential string presentation in both training and grammaticality judgement, and the one study from the Petersson group that failed to find basal ganglia activity also used whole string presentation (Petersson et al., 2004). Thus, there is good evidence for basal ganglia involvement during grammaticality judgment on artificial grammar strings, although it may be particularly evident when sequential presentation is used, which would clearly have implications for inferences about function. In this regard, note also that

BG involvement is evidenced by a  $G > NG$  contrast, opposite to the  $NG > G$  contrast that reveals LIFG involvement.

With regard to hippocampal activity the evidence is less consistent. Lieberman et al. (2004) report greater hippocampal activation for high than low chunk strength test items, but Forkstam et al. (2006) found just the opposite - greater hippocampal activation for low than high chunk strength items. Again, it is not clear why opposite effects are obtained across studies making it hard to interpret what kind of process chunk strength is actually affecting. Petersson et al. (2012) and Folia et al. (2014) do not report any analyses of chunk strength effects in the hippocampus, but they do report lower hippocampal activity during the GJT and liking judgment tasks as compared to a baseline task (contrasting with greater BG activity). This 'deactivation' does not appear to be a learning-related effect, however, because Folia & Petersson (2014) report the pattern as being "very similar" on Day 1 (p. 5) even though there were no learning effects at this point (there was no GJT on day one in Petersson et al., 2012). On this evidence it could only be argued that it is the task requirements of GJT and liking judgment that cause the H to deactivate and BG to activate, perhaps reflecting a deliberate strategy of suppressing explicit memories of training items. However, connectivity analyses in Lieberman et al. (2004) revealed a negative correlation (for certain contrasts) between H and BG activity. These effects have been interpreted in terms of inherent competition between explicit item recall and rule-based processing (Batterink et al., 2019). For example, a grammatical item may invoke a stronger BG response which suppresses any tendency to explicitly recall chunk-based information using the H. This would neatly explain why chunk strength effects in behavioural data show a tendency to be weaker for grammatical than non-grammatical items (Knowlton & Squire, 1996). However, whilst this effect has been replicated in other studies using the same materials (Hendricks, Conway, & Kellogg, 2013; Kinder & Assmann, 2000; Lieberman et al., 2004), other studies using different materials have tended to find either no interaction between grammaticality and chunk strength (Forkstam et al., 2006; Petersson et al., 2012) or a greater chunk strength effect for grammatical items (Folia et al., 2008; Kürten, De Vries, Kowal, Zwitserlood, & Flöel, 2012). The inconsistency of the behavioural pattern reduces the force of the competition argument. Hence, whilst there is evidence for a dissociation between H and BG activity, whether this reflects underlying competition between systems or complementary response strategies in judgment tasks remains unclear.

It should be apparent from the above that there are inconsistencies between imaging studies, not only in relation to whether or not effects are always observed, but also in relation to the contrasts that reveal them. Given the technical challenges involved in revealing brain activity that is specifically related to such an abstract notion as ‘grammatical structure’ this is perhaps not surprising, but it does rather hinder progress towards a precise functional interpretation of what these effects mean.

The question now is, how do these imaging results tie in with evidence from brain stimulation and brain damage? Since these studies target single brain regions the discussion here will be organised in terms of region of interest.

#### LIFG: Brain stimulation and agrammatism

Brain stimulation studies provide evidence for a causal role of LIFG at least in the GJT processes following AGL, but the precise pattern of results is inconsistent between studies. Both Uddén et al. (2008) and de Vries et al. (2010) found that stimulation applied to LIFG improved correct rejection of ungrammatical items but had no effect on grammatical items. The correspondence between these studies is remarkable given the procedural differences between them - Uddén et al. applied TMS during a GJT after 5 days of training, de Vries et al applied tDCS only during a 20-minute acquisition phase (although the effects presumably persisted into the GJT). The boost to performance for NG items appears to be consistent with imaging, where in most studies LIFG activity is higher for NG than G items, suggesting that LIFG is particularly involved when a NG item has to be rejected. However, Udden, Ingvar, Hagoort, and Petersson (2017) found that TMS prior to GJT significantly reduced endorsement of grammatical items and there was no effect for ungrammatical items, precisely the opposite pattern to the earlier studies. They used a more complex ‘crossed nested dependency grammar’ containing hierarchical structure and non-adjacent dependencies, and unlike the earlier studies, verbal reports suggested that the majority of the participants had no conscious knowledge. It is not clear, however, why these differences would have caused the reversal in the pattern of results with respect to grammaticality. Once again this makes precise functional interpretation difficult.

With regard to the effect of damage to LIFG, studies have shown impaired AGL performance in agrammatics (Christiansen, Kelly, Shillcock, and Greenfield (2010), Cope et al. (2017)). Here too, though a breakdown according to

grammaticality introduces confusion – the superior performance of the controls was entirely due to better endorsement of grammatical items, and neither the patients nor the controls could reliably reject ungrammatical items (this pattern is for the complex dependencies in Cope et al). The failure to reliably reject ungrammatical items in both experiments suggests that abstract grammatical rules were not being learned by even the control groups making it unclear what the agrammatics were failing to learn. Moreover, the effect on grammatical items is inconsistent with two of the three brain stimulation studies (Uddén et al., 2008; de Vries et al., 2010).

An interesting feature of the Cope et al. (2017) study is that their FSG contained a simple adjacency – ‘pob’ was always followed by ‘jat’, with the pairing occurring, if at all, at the beginning and/or end of the string. Perhaps surprisingly, the agrammatics were impaired (though above chance) in GJT on even this simple, and one might imagine, quite salient dependency. This would not be expected if LIFG were specialised for learning complex hierarchical structures (as reflected in the other complex dependencies in this study), though it could always be argued that the learning deficit reflects damage extending into the FOP.

An alternative view of the role of LIFG in AGL comes from Schuchard and Thompson (2017). They used a phrase structure grammar to generate strings of auditory syllables and found a similar, but relatively low, level of learning in agrammatics and controls after the first training session, but significantly superior learning in the controls after a second training session. Schuchard et al argue that this difference reflects superior explicit learning strategies in the controls rather than a difference in implicit learning, an explanation that seems plausible given the possible sensitizing effect of performing grammaticality judgement tests before the second training session (this may also contribute to the control group advantage in Cope et al., 2017 since here too a repeated exposure-test cycle was employed). And once again one can raise doubts about what exactly is being learned. It is possible that the equivalent performance after the first training session reflects acquisition of adjacent dependencies, and the boosted performance for the controls reflects more abstract phrase structure-like rules acquired through explicit learning. Hence, it is hard to draw firm conclusions from studies of agrammatic learning given ambiguity in the nature of what was learned and potentially greater explicit learning in controls.

## Basal ganglia: Parkinson's disease

As we saw above the fMRI studies of AGL provided evidence of BG (caudate) involvement, at least in the grammaticality judgment phase. So it comes as some surprise that a number of studies have shown that AGL is not impaired in Parkinson's patients (Peigneux, Meulemans, Van der Linden, Salmon, & Petit, 1999; Reber & Squire, 1999; Smith, Siegert, & McDowall, 2001; Witt, Nuhman, & Deuschl, 2002b) Reber and Squire (1999); Witt et al. (2002b); Smith et al. (2001). Batterink et al. (2019) suggest that PD patients may be matching control levels of learning by relying on their intact hippocampal system. However, Peigneux et al. (1999) report equivalent grammaticality effects for control and PD patients when grammatical and ungrammatical items were matched on 12 different chunk strength measures, and for a grammar that previous research suggested resulted in no conscious knowledge. On the assumption that the H influences AGL via (possibly explicit) chunk familiarity it seems doubtful that AG learning in these PD patients could be attributed to this system.

Interestingly, and counterintuitively, PD patients do show an impairment in AGL when training consists in performing GJT with feedback (Smith & McDowall, 2006a). This is consistent with imaging and patient research showing BG involvement in 'classification learning' where trial by trial feedback is provided (Poldrack et al., 2001; Seger & Cincotta, 2005; Witt, Nuhman, & Deuschl, 2002a). Seger and Cincotta (2005) nuance this view by showing that it is the caudate head that is specifically involved in the processing of feedback information (via the loop connecting to frontal executive brain areas), and Seger (2008) claims that PD "particularly affects the head of the caudate" (p. 272). It therefore becomes possible that implicit AGL is intact because of relatively preserved function in the caudate body and tail. There are a number of problems with this view. First, even caudate head has been reported in imaging of the healthy brain in AGL (Forkstam, 2006; Lieberman et al, 2004, as classified by Seger & Cincotta, 2005). This is interesting in the context of proposed loops from caudate head to BA 47 (Dominey et al., 2009) and predicts AGL impairments in PD if caudate head is malfunctioning. Having said this, the consensus on progression of PD seems to be rather that dysfunction starts in the posterior putamen, and spreads to anterior putamen and caudate nucleus (e.g., Nurmi et al., 2001), with a constant rate of decline across all parts of the putamen and

caudate once initiated (Bruck et al., 2006). In fact, Pasquini et al. (2019) characterise the established view of PD progression in terms of “a posterior-to-anterior gradient in early PD ... with relative preservation in the head of caudate” (p. 1103, and opposite to the claim made by Seger, 2008). Although they go on to show that caudate dysfunction is detectable in about 50% of patients very early in the disease and before any symptoms are apparent, at least for those patients whose caudate remains initially unaffected, AGL should be unimpaired. However, given that Witt et al. (2002b) even found unimpaired AGL in patients with advanced stage PD when they were off medication<sup>4</sup>, it seems unlikely that preserved caudate function underlies preserved AGL in all cases.

An alternative explanation for the discrepancy between the fMRI and patient data is that all but one of the imaging studies reviewed above presented the AG strings letter by letter. In contrast the patient studies used whole string presentation. Perhaps it is only when the stimuli actually form a sequence in time that the basal ganglia are engaged (consistent with the much more robust evidence for basal ganglia involvement in perceptual and motor sequence learning, below). Yet, as ever, there is an exception – Lieberman et al. (2004) report BG activity using whole string presentation. If nothing else, this observation illustrates how procedural variations across studies can lead to confounds that hinder functional interpretation.

#### Hippocampus: Amnesia

What imaging evidence there is for MTL activity in AGL might be interpreted in terms of explicit memory for chunk/item-based information, either as a basis for GJTs (based on chunk strength effects in Lieberman et al., 2004; Forkstam et al) or as a knowledge source that needs to be suppressed in order to focus on grammaticality (based on deactivation with respect to baseline in Petersson et al., 2012 and Folia et al., 2014). One would therefore expect amnesics to perform above chance on AGL tasks, especially when grammaticality effects are measured, although possibly worse than controls depending on the extent to which controls can benefit from additional explicit item-based knowledge. Some studies have shown numerically almost

---

<sup>4</sup> The patients were diagnosed as 4.4 on the Hoehn and Yahr score (where 5 is the most severe and 0 unimpaired). Interestingly AGL performance was similar to controls across a wide range of disease severities in this study.

equivalent GJT performance in amnesics and controls: Knowlton and Squire (1994); Knowlton and Squire (1996), Experiment 1; Meulemans and Van der Linden (2003). Others have shown slightly, but not significantly, lower accuracy in amnesics: Knowlton, Ramus, and Squire (1992); Knowlton & Squire (1996) Experiment 3; Reber, Martinez, and Weintraub (2003). The studies that have used balanced chunk strength designs have found equivalent main effects of grammaticality in amnesics and controls (Knowlton & Squire, 1996, Experiment 1; Meulemans & Van der Linden, 2003).<sup>5</sup> But Knowlton & Squire (1996, Experiment 1) also found that chunk strength effects were equivalent in amnesics and controls, which is surprising given that imaging of the healthy brain has associated the H with similarity-based processing. Knowlton & Squire (1996) speculate that their amnesic chunk effects reflect implicit priming in other brain areas, whereas Lieberman et al. (2004) argue that their MTL activations reflect explicit memory processes. In support of this, deficits in chunk learning measures in amnesics have been associated with measures of conscious knowledge (Knowlton & Squire, 1996, Experiment 2; Meulemans & Van der Linden, 2003). Even so, given that Knowlton & Squire's controls could benefit from both priming and explicit memory one would have expected them to produce stronger chunk effects than amnesics. The similarity between amnesic and control performance in AGL even on chunk strength measures therefore remains remarkable.

Having said this, Channon et al. (2002) found an amnesic deficit in AGL using a more extreme contrast between high and low chunk strength test items than in previous imaging or patient studies, though the amnesics still performed above chance. It is possible then that weaker chunk manipulations are sufficient to produce imaging effects but they are not strong enough to produce an amnesic deficit. But there is a further twist – confidence measures suggested that decisions in both controls and amnesics were a reflection of implicit knowledge. This defies the common assumption that H contributions to AGL are necessarily a reflection of explicit memory.

---

<sup>5</sup> However, one should be cautious about interpreting these results as evidence of the expression of implicit knowledge in amnesics. When whole string recognition tasks have been administered (requiring discrimination between trained strings and new grammatical strings) amnesics perform significantly above chance (Knowlton & Squire, 1992; Reber et al., 2003). Though performance is still significantly worse than controls, the dissociation between GJTs and recognition is not actually as strong as would be required to argue that the two tasks tap different systems, or that GJs in amnesics are driven purely by unconscious knowledge.

## Artificial grammar learning summary

Imaging studies certainly show that LIFG, and in some cases the FOP, are involved in GJT following AGL, similar to the results obtained with more natural, hierarchical, grammars. Of specific LIFG regions, BA 44 is common across studies, which is significant given claims about the potential specialisation of this region for natural language syntax. Note, however, that the activations are less left lateralized than for natural language processing, both with respect to LIFG and the FOP. Consistent with this, Udden & Madden (2018) report a metaanalysis of 7 AGL studies, four of which have been classed as ‘explicit’ here, in which both left and right FOP emerged as the most consistently activated region using a NG > G contrast. Hence whilst it is certainly the case that the response to artificial language stimuli in general (including FSGs and hierarchical grammars) overlaps with that to natural language syntax there are also contributions of additional processes.

Interpretation of brain stimulation and agrammatic studies is more problematic, partly due to variation in the patterning of effects on grammatical and ungrammatical test items, but they do at least support the idea that LIFG is somehow involved in AGL. With regard to BG and H there is poor alignment between imaging and patient studies. Imaging provides some evidence for BG involvement in AGL, but performance is surprisingly unimpaired in PD patients. The few imaging studies that have reported H involvement have found that it is related to chunk strength rather than grammaticality (though inconsistently in terms of contrasts), yet amnesics show hardly any impairment in standard AGL even on chunk-based measures. These inconsistencies challenge a simple identification of distinct and necessary functions to individual brain areas.

Surprisingly few studies have ruled out the contribution of explicit knowledge, although the results from Floel & Petersson’s (2014) liking judgment task and Udden et al’s (2017) TMS study suggest LIFG effects even when this is likely to have been minimised. Channon et al’s (2002) amnesic study is surprising in that it shows an amnesic impairment in the expression of chunk knowledge even when responses appear to be driven by implicit knowledge, challenging the idea that hippocampal contributions are necessarily explicit, a theme that we will also see repeated below.

One must always bear in mind that all of these studies tell us about the brain areas involved in making judgements (nearly always of grammaticality) using whatever knowledge is acquired after exposure to an AGL. They do not reveal activity during learning. The research on sequence and statistical learning provides indications of how that may be achieved.

### Sequence learning

If artificial grammar learning seems rather removed from natural language, many studies of sequence learning, on the face of it, would appear to be even more so. In the classic serial reaction time task (SRTT) the participants see a stimulus (say a cross) moving between, say, four horizontally arranged positions on the screen and they track its movement by pressing corresponding response keys. Unbeknownst to the participants the screen positions follow a repeating sequence, often a 12-item ‘second order conditional’ (SOC) sequence such as 3–2–4–1–3–4–2–3–1–2–1–4, where each number represents a screen position (Vandenberghe, Schmidt, Fery, & Cleeremans, 2006). In such sequences the position of a stimulus can only be predicted by taking into account at least the previous two positions (e.g., 2-1 predicts 4). This is why SOC’s are often referred to as tests of ‘higher order’ sequence learning, rather than of simple adjacent dependencies. Learning is evaluated by comparison with a pseudo-random baseline (a random sequence with no immediate repetitions) that is either presented after a training period or interspersed with stretches of the structured sequence during training. Sequence-specific learning, as opposed to general task practice effects, should be evident as faster response times to the structured, as compared to the random, sequence, and in fMRI studies the same contrast should reveal brain activity related to acquisition of the underlying sequence as opposed to general task effects. When structured and baseline sequences are interleaved we should be able to see learning as it happens, rather than as measured afterwards in a judgment test.

## Imaging

Imaging studies that use a whole brain analysis method have found increased activation to the fixed compared to the random sequence across surprisingly many brain areas. For example, Naismith et al. (2010, non-depressed controls) found increased activation to the fixed sequence in 19 different brain areas, including a number of frontal areas (superior and middle frontal gyrus) and subcortical structures (hippocampus, caudate, putamen, thalamus) (see also Daselaar, Rombouts, Veltman, Raaijmakers, and Jonker (2003). However, in these studies no attempt was made to determine whether sequence knowledge was conscious or not. In Willingham, Salidis, and Gabrieli (2002) the participants showed the same level of recognition for the fixed sequence as the random sequence suggesting a lack of conscious knowledge. Taking 11 regions of interest (ROIs) from previous studies they found greater activation to the fixed sequence in left prefrontal cortex (BA 46, BA10), left inferior parietal cortex, and right BG (putamen). Schendan, Searl, Melrose, and Stern (2003) also used a ROI approach and found increased activation for the fixed sequence in BG (bilateral putamen, but more prominently bilateral caudate), and prefrontal cortex, but also, most notably, the hippocampus, where activation was equivalent regardless of whether or not the participants were provided with explicit information about the sequence.<sup>6</sup> In Albouy et al. (2008) a generation task was used to establish a lack of conscious sequence knowledge (though note that this was following a block of the random sequence and so is not entirely convincing). Imaging during training showed increased activation to the fixed sequence in BG (caudate nucleus) and H, with the latter effect decreasing over training. A connectivity analysis also revealed a negative correlation between H and BG (but in the left ventral putamen). Thus, there is consistent evidence for BG activity, but also for H activity, which, surprisingly, is evident even when an attempt is made to isolate implicit knowledge.

The SRTT is underlyingly rather complex because it combines a regular perceptual sequence (the movement of the stimulus around the screen positions) with a motor sequence (a sequence of manual, or verbal, responses). This raises the question of what learning-related brain activity reflects – is it the motor sequence, or

---

<sup>6</sup> The H was not included as an ROI in Willingham et al., 2002, and it may not have shown up in Daselaar et al. (2003) due to a high threshold.

the perceptual sequence, or an integrated representation of them both? Studies that have attempted to separate these components and which have also applied stringent awareness tests to isolate implicit learning effects have found BG activity associated with both perceptual and motor components (Gheysen, Van Opstal, Roggeman, Van Waelvelde, & Fias, 2010, 2011; Rose, Haider, Salari, & Büchel, 2011). They have also revealed effects in the H, though the results are inconsistent – Rose et al. (2011) found that it was confined to perceptual learning, whereas Gheysen et al. (2010, 2011) found that it was confined to motor learning. The reasons for this discrepancy are unclear.

Studies that have targeted only perceptual sequence learning have yielded clearer results. Ling et al. (2015) examined learning of a sequence of Chinese tones (in Chinese native speakers) and found learning-related activity in bilateral hippocampus, right caudate nucleus and left superior parietal lobule. Jablonowski, Taesler, Fu, and Rose (2018) examined perceptual sequence learning of tones and found very precise learning related changes in bilateral hippocampus. Both studies used stringent tests to exclude any participants with sequence awareness. On this basis, hippocampal involvement in implicit perceptual learning looks more secure. However, all of these studies used relatively short ‘deterministic’ sequences of between 4 and 6 items, and the only study not to find hippocampal involvement in perceptual sequence learning (Gheysen et al., 2011) used a slightly more complex sequence with one non-deterministic transition (2–4–3–1–4). Could sequence simplicity explain hippocampal involvement? Probably not. Rosenthal, Andrews, Antoniadou, Kennard, and Soto (2016) used a traditional 12-item SOC in an ingenious perceptual sequence learning task in which a circle moved between 4 horizontal screen positions but prism glasses reduced participants’ conscious perception to two positions by fusing the left and right visual fields. This reduced the 4-position sequence to a 2-position sequence and hence obscured the SOC structure. In fMRI data acquired during training there was a time by condition (fixed versus random sequence) interaction in the BG (right putamen and pallidum), inferior temporal gyrus, hippocampus, and occipital (visual) areas, including V1. And in a recognition test, higher levels of activity for novel versus trained sequences (which actually looked the same to the participants) were obtained in, amongst other areas, V1 and left hippocampus. So, we see the surprising involvement of the hippocampus even when a traditional 12-item SOC sequence is used, and with what is claimed to be a

complete absence of awareness. But we also see learning-related changes in a low level visual area V1, which suggests that sequence learning-related changes can also occur within perceptual systems (see Batterink et al., 2019, for an emphasis on this point).

### Basal ganglia: Parkinson's disease

Given the repeated observation of basal ganglia involvement in a range of sequence learning tasks in healthy participants one would expect that Parkinson's Disease (PD) patients, who suffer basal ganglia dysfunction, would show a learning impairment. By and large this does indeed appear to be the case (see Clark, Lum, & Ullman, 2014 for a meta-analysis). Of course, given that PD patients have obvious movement difficulties and tremors one might not be surprised by an impairment in a task that involves rapid button-pressing. Some studies have shown reduced learning in PD when a verbal response is required (see Clark et al., 2014), whilst others have shown intact learning (Smith et al., 2001), or even the emergence of learning in a second training session after a one week interval (Smits-Bandstra & Gracco, 2015). The reason for such discrepancies is not clear. Level of medication is one possibility (Ruitenbergh, Duthoo, Santens, Notebaert, & Abrahamse, 2015) or else use of explicit learning strategies in PD patients. But the reduced learning in PD using a verbal response in at least some studies suggests that the involvement of the basal ganglia in the SRTT lies at a deeper level than mere execution of manual responses.

However, there are studies that show intact PD sequence learning under specific procedural variants. Werheid, Ziessler, Nattkemper, and von Cramon (2003) found that PD patients showed the same level of sequence learning as controls when the stimuli (letters) were presented at a central location (different letters were assigned to different response keys) whereas a comparable group of PD patients showed about half the learning effect as controls in a standard SRTT using the same sequence. The PD patients actually performed worse than the controls on tests of awareness. Smith and McDowall (2006a) report a somewhat similar result – PD patients showed unimpaired learning of either a perceptual or positional sequence, but only the controls were able to learn the combined sequence. Both groups were at chance on tests of recall and recognition of the independent perceptual and position sequences, and there was no correlation between their scores and the respective

learning effects. The conclusion is that the PD patients were able to implicitly learn picture and spatial sequences, but they were specifically impaired at combining them into an integrated stream (for a similar effect see Shin & Ivry, 2003). Smith & McDowall (2006a) suggest that the PD deficit lies in information integration rather than sequence learning as such. Yet imaging (above) shows BG involvement in simple perceptual sequence learning situations that do not impose integration demands of this kind. Hence there appears to be an inconsistency between the patient studies and imaging research – which cannot be put down to disease severity (due to deficits still being obtained under specific conditions) or to compensation from explicit learning (ruled out by tests of awareness). Once again simpler versions of the task are sufficient to show imaging effects, but not sufficient to show a patient deficit.

#### Hippocampus: Amnesia

As we saw above, some fMRI studies of the standard SRT task do not report hippocampal activation, whereas two studies do (Naismith et al., 2010; Schendan et al., 2003). Can studies of SRT learning in amnesics shed any more light on the issue? Early studies using SOC sequences found no significant difference in reaction time slow-downs to sequence violations in amnesics and controls (Curran, 1997; Reber & Squire, 1994; Reber & Squire, 1998; Van Tilborg, Kessels, Kruijt, Wester, & Hulstijn, 2011), although in a fine-grained analysis of transition times Curran (1997) did reveal significant group differences at specific sequence positions. Interestingly, even this detailed analysis did not reveal any amnesic deficit for a ‘first order’ conditional sequence in which each element uniquely predicted the successor. This suggests that sequence complexity might be an important variable.<sup>7</sup> Other studies suggest that an amnesic deficit can be revealed when variability is introduced into the sequence. Vandenberghe et al. (2006) compared learning of a standard ‘deterministic’ SOC with a ‘non-deterministic’ version in which there was a 20% probability that a transition would be drawn from another sequence. Amnesics and controls showed equivalent performance on the deterministic version, but only the controls showed any learning of the non-deterministic version. Measures of conscious knowledge

---

<sup>7</sup> Similarly, in an animal model of the SRT task, (Ergorul & Eichenbaum, 2006) found that complexity in terms of sequence length and conditional structure (first or second order) affected learning rate more for rats with hippocampal lesions than controls.

(using the process dissociation procedure) suggested that learning had been implicit in both groups. Shanks, Channon, Wilkinson, and Curran (2006) also found no sequence-specific learning in amnesic patients using a method for generating non-deterministic SOCs similar to Vandenberg et al. (2006). In fact, it is worth noting that the studies that show no amnesic impairment for SOCs exposed the participants to purely structured material throughout training (Reber & Squire, 1994, 1998; Van Tilborg et al., 2011), whereas Curran (1997) interleaved sections of structured and random material, as do imaging studies. Hence it is possible that dependence on the hippocampus is affected by the extent to which the structural regularity has to be extracted from noisy input.

#### LIFG: Agrammatism and TMS

Given the lack of imaging evidence for LIFG involvement in sequence learning it is not surprising that the few studies that have examined agrammatics have found no impairment (Goschke, Friederici, Kotz, & van Kampen, 2001; Schuchard & Thompson, 2014). Interestingly, Schuchard and Thompson (2014) found that agrammatics were worse than controls when trained under explicit conditions, consistent with their argument that Broca's area involvement in AGL might reflect explicit learning or expression of explicit knowledge (Schuchard & Thompson, 2017).

However, a TMS study on healthy participants fails to conform. Clerget, Poncin, Fadiga, and Olivier (2012) showed that TMS applied to LIFG (BA 44) obliterated learning in a standard SRTT [check]. The fact that a control group showed learning effects without conscious knowledge (as assessed by a sequence generation test) suggests that the results were not because TMS was interfering with explicit learning. One reason for these divergent results might lie in the use of an unusually long 20-item repeating sequence which was created in such a way as to encourage chunking (though note that detailed RT analyses revealed that although a chunking pattern could indeed be discerned it was equivalent in both groups). Alamia et al. (2016) provide corroborating evidence that Broca's area is involved in chunking operations, this time in a perceptual/conceptual sequence where response times at the transitions between the highest level chunks were specifically slowed by application of TMS to BA44. Although their task required explicit learning of the sequence through trial and error, and hence unlike Clerget et al. (2012) did not tap implicit

learning, this research does point to LIFG involvement when hierarchically organised sequences are employed.

### Sequence learning summary

Imaging studies reveal learning-related increases in activity in BG and hippocampus for sequence learning, though when motor and perceptual sequence learning are separated inconsistencies between what are on the face of it similar studies emerge. There is good evidence that these effects occur without sequence awareness, ruling out explicit learning as an explanation of hippocampal activity. Sequence learning shows impairments in PD, but there is evidence that it is not impaired when the demands of information integration are reduced, a finding that contradicts imaging evidence for BG involvement in perceptual sequence learning. Amnesics show remarkable preservation of sequence learning in light of the imaging evidence, but do appear to show impairments when variability is introduced. There is no imaging evidence for LIFG involvement (and not even of the FOP), and agrammatic sequence learning is intact for standard SOC sequences, but there is evidence for LIFG involvement when longer sequences that invite chunking are used. Once again, we see poor alignment between imaging and patient studies, but a recurring theme is that specific areas become more necessary when complexity is introduced.

### Statistical learning

The statistical learning paradigm offers a convenient vehicle for studying learning-related changes in brain activity. We shall be concerned here only with statistical learning studies that examine the segmentation of units from continuous streams, which is where most of the imaging work has been carried out. Participants are exposed to a stream of meaningless syllables (e.g. pa-bi-ku-go-la-tu-da-ro-pi-pa-bi-ku ..), or in visual versions passively observe a stream of nonsense shapes. This stream is in fact composed of a small set of recurring but randomly ordered syllable or shape triplets (pa-bi-ku, go-la-tu, da-ro-pi). The only cue to the boundaries between the triplets is the transition probabilities between the syllables or shapes. Statistical

learning is a form of sequence learning akin to the kinds of perceptual sequence learning described above. The difference lies primarily in the nature of what is learned. Statistical learning of this type simply depends on learning adjacent dependencies between syllables, but these are obscured by the random concatenation of triplets. ‘Word’-like or ‘event’-like structures are emergent from the stream. These studies look at extremely rapid learning, occurring after 2 or 3 minutes of exposure, and sometimes with behavioural evidence of learning after less than a minute.

## Imaging

In the tradition of statistical learning research there are no tests of conscious knowledge – it is assumed that the statistical computations underlying learning are themselves unconscious even if the products of learning are not. Indeed, imaging studies have generally given participants instructions to work out the words in the language in preparation for a test (Cunillera et al., 2009; Karuza et al., 2013; López-Barroso et al., 2013; Plante et al., 2015) and some have used repeating cycles of exposure and test to provide a behavioural index of the learning trajectory (Karuza et al., 2013; Plante et al., 2015). Such procedures are likely to encourage explicit learning strategies, and hence it cannot be assumed that the only learning-related processes going on are related to the passive computation of transition probabilities. Only a study by McNealy, Mazziotta, and Dapretto (2006) seems to approximate an implicit learning procedure – participants were scanned whilst being exposed to a combination of structured or random syllable streams with no instruction other than to just listen.

Whilst just listening to auditory syllable streams activates what McNealy et al. (2006) describe as “a large-scale bilateral neural network” (p. 7633) specific measures of learning-related effects reveal more localised activation – but much depends on the measure of learning-related activity that is adopted. Taking the difference between structured and random syllable streams (structured > random) has revealed effects in superior temporal gyrus (STG), largely left lateralized, but the time course varies across studies – increasing over 2 mins of exposure (McNealy et al., 2006), but decreasing from being significant over the first 2 min exposure to being non-significant in the second 2 minutes (Cunillera et al., 2009). Combining these very

similar studies we might conclude that learning effects increase over the first two minutes and then disappear, suggesting that LSTG activity reflects computation of TPs, which might tail off once representations are established. The problem with this idea is that behavioural accuracy, whilst being at chance in McNealy et al. after 2 mins, was far from perfect in Cunillera et al. after 4 mins (68%). Also, Plante et al. (2015) found evidence for an increasing effect in LSTG over more than 8 mins of exposure with respect to a low transition probability baseline. However, this study is not directly comparable to the previous two in that they used a more naturalistic method (real two-syllable Norwegian words in sentence contexts), 3 cycles of exposure and test (which may have stimulated explicit learning), and an Independent Components Analysis method. Hence the time course of LSTG activity, though potentially informative, remains unclear. Plante et al also report a response in the LIFG in the first two blocks, but because this disappears in the third, despite the behavioural effect being greatest at that point, they conclude that it is unlikely to reflect computation of transition probabilities. The effect in the BG (right caudate) only appeared in Block 1 and hence its role in learning is not clear.

Studies have also examined correlations between the structured > random difference and behavioural performance, which one might think would give the most illuminating picture of learning-related activity. But no clear localisation emerges across studies. McNealy et al. (2006) report a correlation with LSTG (which at least is consistent with their structured > random comparison), Cunillera et al. (2009) report a correlation with premotor cortex (BA 6, for activity in the first 2 minutes, which they attribute to rehearsal in working memory). Plante et al. (2015) report correlations with a variety of temporal, frontal and parietal regions, but not specifically LSTG, LIFG, or BG. Karuza et al. (2013) developed a fine-grained measure relating change in accuracy over a particular block to brain activity in that block. On a whole brain analysis they only found correlations with LIFG, and also with BG on a regions of interest analysis with reduced threshold.

It is difficult to draw conclusions from the above. Changes over time with respect to control streams point to effects in LSTG, but the time course is inconsistent across studies. Brain-behaviour correlations are only consistent the overall structured > random contrast in the case of McNealy et al. (2006) (the most implicit) whereas other studies find evidence of BG, LIFG, and even premotor involvement, though the

use of repeated exposure-test cycles may have encouraged explicit learning (involving working memory) in these cases.

There is some evidence of LIFG involvement in other tasks relating to processing of already-segmented units. McNealy et al. (2006) found this for a word > part-word contrast whilst the participants passively listened to isolated words or part-words between the exposure and behavioural test phases. Abla and Okanoya (2008) found effects on LIFG and LSTG when listening to streams of pre-learned tone triplets. And in a study of statistical learning from continuous syllable streams in which words were separated by 25 millisecond pauses to provide a prosodic cue to word boundaries, López-Barroso et al. (2013) found correlations between word learning and functional connectivity in the left arcuate fasciculus – the tract between LIFG (and premotor cortex) and STG. Hence there is evidence for LIFG involvement of processing of learned, or already segmented, units, but evidence for involvement in statistics-driven learning processes comes largely from Karuza et al. (2013) who used the least implicit of the procedures, but possibly the most sensitive measure of learning-related effects.

Experiments have developed visual analogues of the standard auditory statistical learning paradigm involving triplets of “glyph” nonsense shapes concatenated into continuous streams. Turk-Browne, Scholl, Chun, and Johnson (2009) found greater activation to structured than random streams in the right caudate after just 32 seconds of exposure to each stream. And by around 48 seconds greater activation was obtained in the right hippocampus and right STG. These rapid learning effects were obtained despite the fact that after just over 3 minutes of exposure triplet familiarity was not significantly above chance (56%). In a post-experiment debriefing only 2 out of the 16 participants reported any awareness of pairings (but not actual triplets) and there was no relationship between subjective confidence and accuracy. Neither was there any correlation between familiarity and H or BG activation (this being confined to left frontal cortex). Therefore, it seems unlikely that the hippocampal activity after 48 seconds was a reflection of explicit processes.<sup>8</sup>

---

<sup>8</sup> Batterink et al. (2019) note that when only participants who were at or below chance on the familiarity test were excluded BG, but not H, activation was obtained, which they argue suggests that the H effects are a reflection of explicit memory. However, this seems rather a severe criterion given that it excludes people who were above chance purely by chance and people for whom familiarity judgements may have been driven by implicit knowledge. In any case, it could be argued that in relation to explicit memory processes, verbal report is the most relevant measure.

Subsequent research by the Turk-Browne group, and most notably by Schapiro, has revealed that items that tend to occur together evoke more similar neural responses, as measured at the level of individual voxels, in the hippocampus (Schapiro, Kustner, & Turk-Browne, 2012), again under conditions under which no participants reported awareness of the associations, although in this case they were able to discriminate between strongly and weakly associated pairs at an above chance level. Hence, in these studies, at least if verbal report is taken as the primary indicator of awareness, there appears to be H involvement in implicit visual statistical learning.

Subsequent studies examined systems in which patterns of pairwise association would lead to the emergence of ‘community structure’ (similar to how grammatical and semantic classes might emerge from distributional analysis of word co-occurrences). Items from the same community were found to evoke more similar voxel patterns in LIFG, insula, left ATL, and left STG (Schapiro, Rogers, Cordova, Turk-Browne, & Botvinick, 2013), and in subsequent analyses also in the hippocampus when selected as an ROI, and in the connectivity between the hippocampus and LIFG (Schapiro, Turk-Browne, Norman, & Botvinick, 2016) suggesting that these two areas were working in concert to represent the community structure of the input. Although it is not clear to what extent these effects are associated with conscious knowledge these studies are significant in showing how it is possible to actually detect where structural relationships are encoded in the brain, which it turns out is in more than one place.

### Hippocampus: Amnesia

Given the above claims about the role of the hippocampus one would expect statistical learning in amnesics to be seriously impaired. One study has found that a single patient performed at chance and significantly below the mean of a group of 28 controls on a set of statistical learning tasks involving syllables, shapes, scenes, and tones (Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014). A larger scale replication involving 4 patients and 14 controls per task found that the patients were significantly above chance (58%) but significantly worse than the controls (at 72%) (Covington, Brown-Schmidt, and Duff (2018). However, as noted by Batterink et al. (2019), given the evident variability in control group performance on statistical

learning tasks, and the relatively small numbers of participants involved, caution should be exercised in concluding that the patient deficit is due specifically to their amnesia. Additionally, it may be that control performance is boosted by hippocampus-dependent conscious knowledge (some control participants scored over 90% in both studies). It would be interesting to see whether amnesics and controls differ when measures of implicit knowledge are used. Nevertheless, there is some evidence for preserved statistical learning in amnesia, contrasting with the emphasis on the hippocampus in at least the visual statistical learning imaging studies.

#### LIFG and LSTG: Broca's and Wernicke's aphasia

Peñaloza et al. (2015) examined auditory statistical learning in heterogeneous groups of stroke patients. The patients performed significantly worse than matched controls on a forced choice discrimination test following 5.2 minutes of exposure, although their performance was significantly above chance. However, only 4 patients out of 14 performed significantly above chance at an individual level, all of them mild fluent (Wernicke type) aphasics with either parietal or temporal lesions ( $n = 2$  each). Strikingly patients with anterior lesions (Broca's type aphasics) did not achieve above chance performance. There was also a general correlation between verbal working memory performance and statistical learning ( $r = 0.655$ ). Hence these data point to an association between frontal (possibly LIFG) damage and a lack of statistical learning, but one that is impossible to distinguish from a verbal working memory impairment.

#### Statistical learning summary

The weight of evidence from auditory statistical learning of syllable sequences appears on the face of it to point to a role of the LSTG. This is to be expected given the known role of this region in auditory processing and word learning (Hickok & Poeppel, 2007). However, there are inconsistencies - the time course of the effect is uncertain, it is strange how the LSTG only shows up in brain-behaviour correlations in one study (McNealy et al., 2006), and in the one study of aphasics it was the Wernicke's patients who appeared to be least impaired. Brain-behaviour correlations are inconsistent across studies, but there is some evidence for BG and LIFG involvement using possibly the most sensitive method (Karuza et al., 2013). When

evaluating the relative contribution of the LSTG and LIFG much depends on which measure of learning-related brain changes one emphasises. Variation in procedures across studies also raises problems of interpretation. To what extent do the results reflect intentional learning processes encouraged by instructions to learn words, and repeating cycles of exposure and test? This is particularly problematic for interpretation of LIFG activity given that this region is implicated more generally in phonological processing (Flinker et al., 2015; Sahin, Pinker, Cash, Schomer, & Halgren, 2009) and phonological short-term memory (Papagno et al., 2017). Note also the correlation between auditory statistical learning and verbal working memory in Peñaloza et al.'s (2015) study of aphasics. There is better evidence for LIFG involvement in processing of already learned units, or learning when additional prosodic cues are provided. With regard to hippocampal involvement, there is some evidence for impaired auditory and visual statistical learning in amnesia, and yet only imaging studies of visual statistical learning have reported hippocampal activation, at least so far. These effects appeared to occur in the absence of awareness of the relevant structure.

## Discussion

Given the inconsistencies between different methodologies (imaging and patient studies in particular), and inconsistencies amongst studies within the same methodology (particularly amongst imaging studies) then it is clearly difficult to draw generalisations across this literature. In relation to imaging, there is ample evidence of learning-related brain responses to structural regularities that participants have acquired without intention and without awareness. Some of the best evidence for this comes from perceptual sequence learning. Including studies that are less scrupulous about establishing lack of awareness, then it appears that there is good imaging evidence for LIFG/FOP in relation to AGL, almost no evidence for it in sequence learning, and some evidence for it in statistical learning. BG involvement is most widely attested in sequence learning, less so in statistical learning and AGL. There is some evidence for H involvement in all tasks, even in situations where concerted efforts have been made to rule out the contribution of conscious knowledge.

However, in only two cases is there good alignment between imaging and patient/stimulation studies – imaging, brain stimulation, and patient (agrammatic) studies all suggest LIFG involvement in AGL, and imaging and patient (PD) studies both suggest BG involvement in the standard SRT task. Whatever functions are carried out by LIFG and BG seem to be essential to normal levels of performance in these tasks. But lack of alignment is more common. PD patients are unimpaired on AGL and perceptual sequence learning despite there being good imaging evidence for BG involvement in these tasks. Amnesics are unimpaired in AGL and SRT tasks despite there being at least some imaging evidence for hippocampal involvement. However, amnesic deficits become apparent when the difficulty of the learning problem is increased above that used in imaging studies – as for a stronger chunk strength manipulation in AGL, and when probabilistic sequences are used in the standard SRT. There is no PD impairment when perceptual and motor components of the SRT are separated, only when they are combined in the standard SRT. Hence, for simpler versions of the task, patient studies may suggest that certain brain areas are not necessary, even though those areas appear to be active in imaging of the healthy brain. This may suggest that in the simpler case other brain regions can perform the required functions, implying duplication of function across different areas (see below).

It is also evident that comparisons between imaging and patient studies are sometimes confounded by procedural differences that affect the nature of the learning problem. Imaging studies of AGL have generally used sequential presentation, whereas patient studies have used whole string presentation, which may explain why BG involvement is only evident in imaging. The hippocampal involvement that is evident in imaging of the SRT task may be due to the demands placed by mixing of sequenced and random material during training, introducing variability which is not a feature of patient versions. At best such observations may provide clues to underlying functions, and at least they serve to emphasise that comparisons between patient and imaging studies require that exactly the same procedures are employed wherever possible.

Given this complex picture, what can be said about the functions of these different brain areas in implicit learning, and how these functions relate to natural language processing and learning?

*LIFG/FOP*: On the view that LIFG (or some part of it) is specialised for processing hierarchical structure and non-adjacent dependencies, the relatively strong evidence for LIFG involvement in implicit AGL is surprising. Fitch and Friederici (2012) point to the relatively low level of performance attained in such experiments and suggest that broad frontal (including LIFG) activations might reflect increased cognitive control demands of the kind characteristic of second language processing, and that “A clear separation in the brain activation for adjacent versus long-distance hierarchical dependencies may thus only be observable in a fully established, mature system” (p. 1949). The fact that all evidence for LIFG involvement in AGL comes from GJTs makes control a pertinent issue. It could be argued that identifying a string as ungrammatical imposes particularly high control demands, especially if chunk familiarity has to be suppressed – hence the increased LIFG, and possibly RIFG, activity for ungrammatical items obtained in AGL imaging studies, and the improvement in rejection rates of ungrammatical items under brain stimulation. Appealing to control would merely reflect one of the functions of LIFG in language processing more generally (see Introduction). Udden et al. (2018) also appeal to a control function as the explanation for the relatively consistent FOP activation on NG > G contrasts found across implicit and explicit AGL studies. It is also relevant to note that in the assessment of second language grammatical knowledge, there is evidence that the rejection of ungrammatical items draws on explicit knowledge (Ellis & Roever, 2018; Vafaei, Suzuki, & Kachisnke, 2017) the expression of which would be expected to impose control demands. In this view LIFG/FOP would not be expected to be involved in sequence and statistical learning simply because the focus is on brain activity during learning rather than post-exposure judgment tasks. Any LIFG involvement in auditory statistical learning would need to be seen as a reflection of different processes. This is self-evident anyway since in this case activation is greater to structured than random material, opposite to the NG > G contrast used in AGL. One possibility is that phonological short-term memory processes are engaged more by structured material, especially considering the incentive to work out the words intentionally.

On the other hand, there is some evidence that LIFG is interested in structure and not just control. LIFG activation has been obtained even for liking judgments (Folia & Petersson, 2014) where competition between grammaticality and chunk familiarity is less likely to be an issue. The studies of AGL learning in agrammatics

showed impaired acceptance of grammatical items (though the fact that not even the controls could reliably reject ungrammatical items raises questions over whether this reflects abstract grammatical knowledge). Looking beyond humans, in a study of primate AGL Wilson et al (2015) found (bilateral) activation in a region homologous to human FOP using a NG > G contrast for images obtained during passive listening. TMS to LIFG has been shown to impair perceptual-motor sequence learning when chunking effects are examined (Clerget et al., 2012; Alamia et al, 2016), and in a statistical learning paradigm representational similarity analyses suggest that both simple and complex associative structure is encoded in LIFG (Schapiro et al., 2012, 2013).

An alternative view of LIFG function is that it is a “generic on-line structured sequence processor that unifies information from various sources in an incremental and recursive manner” (Petersson et al., 2012, p. 85). This view predicts LIFG involvement in AGL using FSGs, with greater activation for NG items possibly being a reflection of increased processing demands caused by a breakdown of the parsing process (Forkstam et al., 2006). But an emphasis on generic sequential structure building raises the question of why there is so little evidence for LIFG involvement in sequence and statistical learning. It may be speculated that complexity of the learning problem is a factor. The systems used in sequence learning, statistical learning, and AGL could be seen as lying on a continuum of increasing abstraction. Sequence learning requires no abstraction since it involves a single repeated chain of associations whose strength is built up through repetition. In contrast in AGL a complex grammar involving branching and recursion has to be induced from diverse examples. Statistical learning is at an intermediate level of complexity since the relevant units (usually triplets) are simple, but they need to be abstracted from a continuously varying input stream (the order of the triplets being pseudo-random). Community structure is an example of a more complex abstraction problem in statistical learning as would be chunking in sequence learning, and these have both yielded evidence for LIFG involvement (Schapiro et al., 2013; Clereget et al., 2012).

However, natural languages involve levels of abstraction and complexity that go beyond even the relatively ‘complex’ FSGs (i.e., hierarchically organised constituents, long distance dependencies, and word classes). The problem is that the imaging experiments that have used grammars with these properties have employed explicit training regimes, and so we cannot be sure that the broad similarity with

LIFG activations in FSGs reflect the same kinds of processes. Clearly, what is needed is neuroscientific work looking at implicit learning of these more complex and abstract systems. But the prospects for obtaining implicit learning of such grammars are uncertain. Opitz and Hofmann (2015) trained participants on Brocanto under implicit conditions, and though GJTs were above chance even for phrase structure violations, further analyses suggested that decisions were based on similarity, not rule knowledge. Systems with word classes and central embedding (A1 A2 A3 B3 B2 B1, where letters stand for item classes and numbers dependency relations) have also been shown to not be amenable to implicit learning when performance depends on recovering the underlying non-adjacent dependencies (Perruchet & Rey, 2005). On the other hand, de Vries, Petersson, Geukes, Zwitserlood, and Christiansen (2012) found that a crossed nested dependency grammar (A1 A2 A3 B1 B2 B3) was amenable to implicit learning, as was a hierarchical grammar with less than 3 embeddings (A1 A2 B2 B1, a limitation that they argue is reflected in natural language processing). In this context it is highly significant that, using tDCS, Udden et al (2017) found evidence for LIFG involvement for crossed nested dependencies. This suggests that imaging studies of various kinds of hierarchical grammars under implicit conditions would be worthwhile. Without such studies it is unclear whether the LIFG activity that is evident under (assumed) implicit learning conditions is indicative of processes that are powerful enough to deal with the complexity and abstractness of natural language. One may wonder whether to achieve this in what is effectively a second language learning situation (since all of the experiments reviewed here were performed on adults) additional explicit processes might indeed be required. But in that case, the activations revealed could not be said to be indicative of first language processing and learning.

LIFG involvement in implicit learning can therefore be variously attributed to structure building, control, or working memory, all of which are functions that have been ascribed to this region in relation to natural language processing (though as noted above, the activations to AGL are less left lateralized). This ambiguity appears to stem from the use of judgement tasks, and insufficient control over whether the learning process, and the resulting knowledge, is implicit. More research is needed that obviates the need for judgment tasks and utilizes truly incidental learning paradigms whilst at the same time systematically manipulating the complexity of the structures to be learned.

*Basal ganglia:* The general learning literature emphasises the role of the BG, and in particular the caudate, in feedback-based learning, or reinforcement learning, e.g. in maze navigation, motor sequence learning, and categorisation. Learning is driven by computation of the difference between the expectation of reward for some action and the actual reward received (“reward prediction error”). This error is signalled by dopamine, which strengthens coding of stimulus-response associations when appropriate (see Lim, Fiez, & Holt, 2014 for an overview). In the context of category learning, where stimuli that vary on a range of dimensions have to be associated with responses, posterior striatum (caudate body and tail) appears to be involved, acting through cortico-striatal loops with, say, visual, or auditory cortex depending on the nature of the stimuli (see Patterson & Knowlton, 2018 for a meta-analysis). To explain the involvement of BG in the SRTT it must be assumed that implicit predictions are rewarded by the appearance of the next stimulus at the expected position. To account for BG involvement in purely perceptual sequence learning and statistical learning it would have to be assumed that the computation of reward prediction error extends to situations in which no overt responses are made or rewarded and stimuli follow in very rapid succession (although the idea that prediction error drives learning is familiar from connectionist models of sequence learning, (although the idea that prediction error drives learning is familiar from connectionist models of sequence learning, Elman, 1990). Applying this logic to AGL is more problematic unless sequential letter presentation is used, as happens to be the case in most imaging studies (unlike in PD studies where no learning deficit is observed). Of course, just because BG is responsive to prediction error does not mean that it is responsible for learning the basis for those predictions, which could occur in cortical areas to which BG is connected through cortico-striatal loops. However, some evidence that BG is indeed learning structure comes from a perceptual category learning experiment in which generalisation ability was related to BG activity, revealing its potential for information compression (Lim et al., 2019).

Note that the computation of prediction error in learning is likely to be distinct from its role in language processing. In situations of conflicting interpretation or ungrammaticality the BG increases in activation (Mestres-Misse et al., 2014; Mestres-Misse et al., 2012) whereas in the learning experiments reviewed here the BG are

more active in response to structured than unstructured (or ungrammatical) material.<sup>9</sup> This would be more consistent with the use of low prediction errors to effect larger learning-related neural changes as opposed to the use of high prediction error to trigger controlled repair processes, as has been suggested for language processing, and would appear to put the BG at the heart of the implicit learning process rather than performing an ancillary function. However, any such view runs up against the observations of unimpaired implicit learning in PD patients. In the case of AGL, as suggested above, this might be because non-sequential stimulus presentation makes learning less BG-dependent, which is to imply that the predictive function of the BG is only triggered by sequential material. In the case of sequence learning it may be because residual function within the BG, or compensation from other systems, is sufficient to deal with situations that place low demands on information integration. But these remain important unresolved exceptions in establishing the role of the BG in implicit learning.

*Hippocampus:* Traditionally the hippocampus has been regarded as playing a role in explicit memory, and so evidence for its involvement in AGL and SRT is often attributed to explicit, rather than implicit, processes (e.g., see arguments in Batterink et al., 2019). However, we have seen that there are numerous examples of hippocampal contributions to implicit learning – most notable in sequence learning, but also in statistical learning, and in amnesic AGL. At the same time, there seems to be nothing in the implicit learning studies to contradict the idea, emerging from the explicit language learning studies reviewed in the Introduction, that the H is particularly involved in similarity-based processing, possibly during relatively early stages of learning.

When hippocampal involvement in implicit learning is revealed researchers invariably appeal to the notion of “relational encoding” (Eichenbaum, 2000). According to this theory, cortical areas form representations of individual items, but the hippocampus is involved in forming temporal and spatial associations between them. The crucial role of the hippocampus in conscious recollection of events derives

---

<sup>9</sup> BG activity in natural language processing also shows a tendency towards left lateralisation (Mestres-Misse et al., 2012; Moro et al., 2001), although bilateral activation was found in (Moreno et al., 2018) using sign language and second language). But in implicit learning there is a tendency towards right or bilateral activation (the notable exceptions are left lateralisation in Yang & Li, 2012, implicit > explicit, and Ling et al.’s (2015) auditory perceptual sequence learning).

merely from the fact that episodic recall requires associations between the to-be-recalled item and its context of occurrence – information about time and place – which act as retrieval cues and render the experience distinctive. But, over recent years, the emerging evidence for hippocampal involvement in implicit memory and learning suggests that the relational encoding function of the H is engaged even in the absence of conscious recall (see (Hannula & Greene, 2012) for a review). Moreover, the hippocampus has been implicated in the encoding of sequential information, as distinct from temporal information (DeVito & Eichenbaum, 2011). It is also well established that the hippocampus plays a vital role in the abstraction of regularities across experiences; that is, in forming generalisations (O'Reilly & Norman, 2002). Although this is usually thought to occur through interactions between the hippocampus and cortex during sleep, it has been suggested that the hippocampus might have an internal architecture that makes it possible to rapidly perform certain kinds of abstraction over the very short time scales that are typical of implicit learning experiments (Schapiro et al., 2016).

From the above it is striking that, at least in broad terms, very similar computational functions have been ascribed to LIFG, BG and H: learning of sequential information, and even extraction of generalisations has been mooted for each area. Recall also that structural sequence learning effects have even been found in STG for auditory, and even visual, statistical learning (McNealy et al., 2006; Plante et al., 2015; Turk-Browne et al., 2009; Schapiro et al., 2013), and in primary visual cortex and occipital cortex for visual statistical learning (Rosenthal et al., 2016; Turk-Browne et al., 2009). Hence it is possible that sequence information is registered, computed, or stored in a number of distinct, but interconnected, brain regions (Batterink et al., 2019). This broad similarity in function may explain why patients with damage to one specific area do not always display the expected learning deficit. Perhaps other regions can compensate for reduced function in one region by increasing their contribution. Or there may be sufficient redundancy and duplication of function that the learning system is resistant to local damage.

At the same time, patient studies do show that impairments can be revealed under particular circumstances, say for BG when the demands of information integration are high or feedback is provided (Smith & McDowall, 2006a, 2006b), and for H when the system can only be learned through chunking (Channon & Shanks,

2002), or when there is variability in the input (Vandenberg et al., 2006). Thus, whilst there may be basic functions that are duplicated, there may also be regional specialisations that can only be revealed under specific circumstances. But these specialisations are likely to be relatively subtle.

### Towards functional differentiation

A promising way forward is to link learning-related changes in distinct brain regions to mathematical models of what is learned. For example, Bornstein and Daw (2012) employed a perceptual sequence learning task with an underlying probabilistic structure and found that the basal ganglia response to each stimulus was correlated with its ‘forward entropy’ (the probability distribution of the possible next stimuli), whereas they suggested that hippocampal activity was related to “preparatory ‘prefetching’ of the anticipated next elements in the sequence” (specific predictions of sequence continuations). Using a similar sequence learning paradigm Konovalov and Krajbich (2018) found that the caudate response was related to prediction error; that is, violations of expectancy (which they align with Bornstein & Daw’s forward entropy), whereas they suggest that the hippocampal response was related to pattern encoding (but not retrieval, which seems to contradict Bornstein & Daw). Wang, Shen, Tino, Welchman, and Kourtzi (2017) used a probabilistic system and the participants’ task was to make explicit predictions about sequence continuations with no feedback. They found that dorsal caudate and putamen responses reflected a strategy of basing predictions on sequence knowledge, whereas hippocampal activations reflected a strategy of basing predictions on simple item frequency. These studies converge on the general idea that the BG are involved in processes related to prediction, but they all have a slightly different view of the role of the hippocampus. It is also not clear to what extent the imputed functions are connected to conscious processing, and so there is a clear need to adopt this approach in combination with rigorous awareness tests.

There may be other clues that help distinguish functions. For example, BG and H are classically thought to be associated with slow and fast learning rates respectively. However, this contrast is based more on the distinction between the development of automaticity (or “habit learning” in the learning literature) and ‘one shot’ learning in the formation of episodic memories, leading to the classic distinction

between procedural and declarative learning. Given the short time scales of most implicit learning experiments it should be clear that any BG involvement is not a reflection of habit formation in this sense. Acquisition of automaticity in situations of ‘overtraining’ is generally associated with putamen, whereas, as we have seen here, implicit learning over shorter time scales is associated with the caudate (Patterson & Knowlton, 2018). Automaticity is also reflected in decreased activation to structured material (e.g., Poldrack et al., 2005) rather than the increases shown in the present studies. Nevertheless, differences in time course may be evident over shorter time scales. In SRT learning Gheysen et al. (2010) found learning-related activity in the hippocampus after 3 minutes of exposure, whereas caudate activity was not evident until after 51 minutes. Schendan et al. (2003) found that caudate activity increased over runs but hippocampal activity decreased (being maximal in the first run, i.e. in the first 90 seconds). Albouy et al. (2008) found a reduction in H activation but no reported increase in learning-related caudate activation over a 7 minute training period. As noted by Albouy et al., a decreasing H response is incompatible with the idea that it is associated with explicit knowledge (which would be predicted to strengthen over training on this time scale), but is compatible with the idea that it reflects rapid learning of novel (and often, second-order) associations, there being more to learn early in training, and less to learn as more of the structure is discovered. On the other hand, if BG activation reflects prediction errors then any effect would be expected to increase the better the underlying structure is learned. Appealing as this picture is, as ever, there is contradictory evidence. In a study that expressly set out to distinguish learning rates in implicit perceptual learning, (Bornstein & Daw, 2012) identify the BG with a fast learning rate and the hippocampus with a slow rate, precisely contrary to SRT learning. Ling et al. (2015) did not find any difference in time course of BG and H activity (both emerging after about 8 minutes), and Turk-Browne et al. (2002) found a slightly delayed hippocampal response (after 48 secs compared to 32 secs for caudate). And of course, extreme caution is necessary when relating these patterns to the reduction in hippocampal activity found in language learning experiments over much larger time scales (e.g., Opitz & Friederici, 2003, 2004). Nor is this to say that hippocampal activation cannot also reflect explicit memory processes in some cases, or that deactivations could be indicative of a strategy of suppressing explicit item-based memories when making GJTs. But given

the evidence that the H can also be involved in implicit learning it is necessary that such claims are backed up by actual awareness measures.

Conclusion: From modules to networks

In all of the above the function of each region has been treated in isolation. Though the contribution of other regions is clearly recognised, the general approach has been to try and differentiate the specific function of one region from that of the others. Perhaps, though, it is more appropriate to take a holistic approach in which discrete areas are seen as parts of interconnected networks. After all, we have already seen that the interconnectedness of brain regions is related to learning – as reflected in the integrity of fibre tracts emanating from LIFG (in AGL), LIFG to STG connectivity through the arcuate fasciculus (in auditory statistical learning), and between the hippocampus and LIFG (in visual statistical learning), and possibly negative correlations between the H and BG in AGL and SRT. Unsurprisingly, communication between regions is integral to learning. There is also emerging evidence of connectivity, and a certain sharing of function between BG and another subcortical structure, the cerebellum, two structures that were previously thought to be independent (Bostan & Strick, 2018). No region can be considered in isolation, damage in one region can have knock-on effects on function in another, or one region may compensate for malfunction in another. One consequence of this is that the load on different systems will change in response to task demands, as the system dynamically reconfigures (Ferbinteanu, 2019) (see (Poldrack et al., 2001; Wang et al., 2017) for examples of how the balance between BG and H can shift according to task or strategy). In the extreme, computational functions emerge from activity in multiple regions, making it difficult to identify specific functions with specific regions. For example, the ‘evolutionary accretion’ model (Murray, Wise, & Graham, 2016) relates the evolution of the brain to a progression towards ever more complex forms of memory and learning – from simple conditioning in the oldest brain structures (BG), to navigation (hippocampus), through to semantic and episodic memory (neocortex). As the system evolves, each new region builds on and extends the computational power of the more primitive systems to solve new problems. But what this means is that the accomplishment of a function cannot be attributed to a specific region, but depends on cooperation between regions. The whole system has evolved over time to

achieve more and more complex forms of learning. Natural language is at the pinnacle of this evolutionary development, but the capacities that make it possible cannot be separated from more 'primitive' computations in the neural bedrock. Implicit learning experiments may seem to represent test tube learning problems that on the face of it have little to do with natural language. But they provide an environment in which we can investigate how the basic ingredients for achieving language learning in the brain combine and interact.

## References

- Abla, D., & Okanoya, K. (2008). Statistical segmentation of tone sequences activates the left inferior frontal cortex: A near-infrared spectroscopy study. *Neuropsychologia, 46*, 2787-2795.
- Alamia, A., Solopchuk, O., D'Ausilio, A., Van Bever, V., Fadiga, L., Olivier, E., & Zenon, A. (2016). Disruption of Broca's Area Alters Higher-order Chunking Processing during Perceptual Sequence Learning. *Journal of Cognitive Neuroscience, 28*, 402-417.
- Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., Darsaud, A., Ruby, P., Luppi, P.-H., Degueldre, C., Peigneux, P., Luxen, A., & Maquet, P. (2008). Both the Hippocampus and Striatum Are Involved in Consolidation of Motor Sequence Memory. *Neuron, 58*, 261-272.
- 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*, 1667-1674.
- Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *Neuroimage, 42*, 525-534.
- Batterink, L. J., Paller, K. A., & Reber, P. J. (2019). Understanding the Neural Bases of Implicit and Statistical Learning. *Topics in Cognitive Science, 0*.
- Bornstein, A. M., & Daw, N. D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience, 35*, 1011-1023.
- Bostan, A. C., & Strick, P. L. (2018). The basal ganglia and the cerebellum: nodes in an integrated network. *Nature Reviews Neuroscience, 19*, 338-350.
- Bruck, A., Aalto, S., Nurmi, E., Vahlberg, T., Bergman, J., & Rinne, J. O. (2006). Striatal subregional 6- F-18 fluoro-L-dopa uptake in early Parkinson's disease: A two-year follow-up study. *Movement Disorders, 21*, 958-963.
- Channon, S., Shanks, D., Johnstone, T., Vakili, K., Chin, J., & Sinclair, E. (2002). Is implicit learning spared in amnesia?: Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia, 40*, 2185-2197.
- Christiansen, M. H., Kelly, M. L., Shillcock, R. C., & Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition, 116*, 382-393.
- Clark, G. M., Lum, J. A. G., & Ullman, M. T. (2014). A meta-analysis and meta-regression of serial reaction time task performance in Parkinson's disease. *Neuropsychology, 28*, 945-958.
- Clerget, E., Poncin, W., Fadiga, L., & Olivier, E. (2012). Role of Broca's Area in Implicit Motor Skill Learning: Evidence from Continuous Theta-burst Magnetic Stimulation. *Journal of Cognitive Neuroscience, 24*, 80-92.
- Cope, T. E., Wilson, B., Robson, H., Drinkall, R., Dean, L., Grube, M., Jones, P. S., Patterson, K., Griffiths, T. D., Rowe, J. B., & Petkov, C. I. (2017). Artificial grammar learning in vascular and progressive non-fluent aphasia. *Neuropsychologia, 104*, 201-213.
- Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The Necessity of the Hippocampus for Statistical Learning. *Journal of Cognitive Neuroscience, 30*, 680-697.

- Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., Pujol, J., & Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, *48*, 541-553.
- Curran, T. (1997). Higher-Order Associative Learning in Amnesia: Evidence from the Serial Reaction Time Task. *Journal of Cognitive Neuroscience*, *9*, 522-533.
- Daselaar, S. M., Rombouts, S. A. R. B., Veltman, D. J., Raaijmakers, J. G. W., & Jonker, C. (2003). Similar network activated by young and old adults during the acquisition of a motor sequence. *Neurobiology of Aging*, *24*, 1013-1019.
- de Vries, M. H., Barth, A. C. R., Maiworm, S., Knecht, S., Zwitserlood, P., & Flöel, A. (2010). Electrical Stimulation of Broca's Area Enhances Implicit Learning of an Artificial Grammar. *Journal of Cognitive Neuroscience*, *22*, 2427-2436.
- de Vries, M. H., Petersson, K. M., Geukes, S., Zwitserlood, P., & Christiansen, M. H. (2012). Processing multiple non-adjacent dependencies: evidence from sequence learning. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *367*, 2065-2076.
- DeVito, L. M., & Eichenbaum, H. (2011). Memory for the Order of Events in Specific Sequences: Contributions of the Hippocampus and Medial Prefrontal Cortex. *The Journal of Neuroscience*, *31*, 3169.
- Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: distinguishing structural knowledge and judgment knowledge. *Psychological Research*, *69*, 338-351.
- Dominey, P. F., Inui, T., & Hoen, M. (2009). Neural network processing of natural language: II. Towards a unified model of corticostriatal function in learning sentence comprehension and non-linguistic sequencing. *Brain and Language*, *109*, 80-92.
- Eichenbaum, H. (2000). A cortical–hippocampal system for declarative memory. *Nature Reviews Neuroscience*, *1*, 41-50.
- Ellis, R., & Roeper, C. (2018). The measurement of implicit and explicit knowledge. *The Language Learning Journal*, 1-16.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179-211.
- Ergorul, C., & Eichenbaum, H. (2006). Essential Role of the Hippocampal Formation in Rapid Learning of Higher-Order Sequential Associations. *The Journal of Neuroscience*, *26*, 4111.
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. *Current biology : CB*, *22*, 2059-2062.
- Ferbinteanu, J. (2019). Memory systems 2018 - Towards a new paradigm. *Neurobiology of learning and memory*, *157*, 61-78.
- Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *367*, 1933-1955.
- Flinker, A., Korzeniewska, A., Shestyuk, A. Y., Franaszczuk, P. J., Dronkers, N. F., Knight, R. T., & Crone, N. E. (2015). Redefining the role of Broca's area in speech. *Proceedings of the National Academy of Sciences*.
- 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*, 1974-1981.
- Folia, V., Udden, J., Forkstam, C., Ingvar, M., Hagoort, P., & Petersson, K. M. (2008). Implicit Learning and Dyslexia. *Learning, Skill Acquisition, Reading, and Dyslexia*, *1145*, 132-150.

- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78-84.
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences*, 103, 2458.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences*, 19, 117-125.
- Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., & Fias, W. (2010). Hippocampal contribution to early and later stages of implicit motor sequence learning. *Experimental Brain Research*, 202, 795-807.
- Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., & Fias, W. (2011). The Neural Basis of Implicit Perceptual Sequence Learning. *Frontiers in Human Neuroscience*, 5.
- Goschke, T., Friederici, A. D., Kotz, S. A., & van Kampen, A. (2001). Procedural Learning in Broca's Aphasia: Dissociation between the Implicit Acquisition of Spatio-Motor and Phoneme Sequences. *Journal of Cognitive Neuroscience*, 13, 370-388.
- Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, 4.
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, 28, 136-141.
- Hannula, D., & Greene, A. (2012). The hippocampus reevaluated in unconscious learning and memory: at a tipping point? *Frontiers in Human Neuroscience*, 6.
- Hartwigsen, G., Neef, N. E., Camilleri, J. A., Margulies, D. S., & Eickhoff, S. B. (2018). Functional Segregation of the Right Inferior Frontal Gyrus: Evidence From Coactivation-Based Parcellation. *Cerebral Cortex*, 29, 1532-1546.
- Hauser, M. F. A., Hofmann, J., & Opitz, B. (2012). Rule and similarity in grammar: Their interplay and individual differences in the brain. *NeuroImage*, 60, 2019-2026.
- Hendricks, M. A., Conway, C. M., & Kellogg, R. T. (2013). Using Dual-Task Methodology to Dissociate Automatic From Nonautomatic Processes Involved in Artificial Grammar Learning. *Journal of Experimental Psychology-Learning Memory and Cognition*, 39, 1491-1500.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393.
- Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language*, 97, 243-257.
- Hsu, N. S., Jaeggi, S. M., & Novick, J. M. (2017). A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language*, 166, 63-77.
- Jablonowski, J., Taesler, P., Fu, Q., & Rose, M. (2018). Implicit acoustic sequence learning recruits the hippocampus. *PLOS ONE*, 13, e0209590.
- January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and Syntactic Ambiguity Resolution in Broca's Area: Implications for the Neural Basis of Sentence Processing. *Journal of Cognitive Neuroscience*, 21, 2434-2444.

- Julia, U., & Claudia, M. (2018). Artificial Grammar Learning and Its Neurobiology in Relation to Language Processing and Development. In R. Shirley-Ann & M. G. Gaskell (Eds.), *Oxford Handbook of Psycholinguistics*: Oxford University Press.
- Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*, 46-54.
- Kawashima, S., Ueki, Y., Kato, T., Ito, K., & Matsukawa, N. (2018). Reduced striatal dopamine release during motor skill acquisition in Parkinson's disease. *Plos One*, *13*.
- Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2018). Connectivity of the hippocampus and Broca's area during acquisition of a novel grammar. *Neuroimage*, *165*, 1-10.
- Kinder, A., & Assmann, A. (2000). Learning artificial grammars: no evidence for the acquisition of rules. *Memory and Cognition*, *28*, 1321-1332.
- Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact Artificial Grammar Learning in Amnesia: Dissociation of Classification Learning and Explicit Memory for Specific Instances. *Psychological Science*, *3*, 172-179.
- Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 169-181.
- Knowlton, B. J., & Squire, L. S. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 79-91.
- Konovalov, A., & Krajbich, I. (2018). Neurocomputational Dynamics of Sequence Learning. *Neuron*, *98*, 1282-1293.
- Kürten, J., De Vries, M. H., Kowal, K., Zwitserlood, P., & Flöel, A. (2012). Age affects chunk-based, but not rule-based learning in artificial grammar acquisition. *Neurobiology of Aging*, *33*, 1311-1317.
- 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. *Journal of Cognitive Neuroscience*, *16*, 427-438.
- Lieberman, P. (2007). The evolution of human speech: its anatomical and neural bases. *Current Anthropology*, *48*, 39-66.
- Lim, S.-J., Fiez, J. A., & Holt, L. L. (2014). How may the basal ganglia contribute to auditory categorization and speech perception? *Frontiers in Neuroscience*, *8*.
- Ling, X., Guo, X., Zheng, L., Li, L., Chen, M., Wang, Q., Huang, Q., & Dienes, Z. (2015). The neural basis of implicit learning of task-irrelevant Chinese tonal sequence. *Experimental Brain Research*, *233*, 1125-1136.
- Longworth, C. E., Keenan, S. E., Barker, R. A., Marslen-Wilson, W. D., & Tyler, L. K. (2005). The basal ganglia and rule-governed language use: evidence from vascular and degenerative conditions. *Brain*, *128*, 584-596.
- 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*, *110*, 13168.
- McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing. *The Journal of Neuroscience*, *26*, 7629-7639.

- Mestres-Misse, A., Bazin, P. L., Trampel, R., Turner, R., & Kotz, S. A. (2014). Dorsomedial striatum involvement in regulating conflict between current and presumed outcomes. *Neuroimage*, *98*, 159-167.
- Mestres-Misse, A., Turner, R., & Friederici, A. D. (2012). An anterior-posterior gradient of cognitive control within the dorsomedial striatum. *Neuroimage*, *62*, 41-47.
- Meulemans, T., & Van der Linden, M. (2003). Implicit learning of complex information in amnesia. *Brain and Cognition*, *52*, 250-257.
- Moreno, A., Limousin, F., Dehaene, S., & Pallier, C. (2018). Brain correlates of constituent structure in sign language comprehension. *Neuroimage*, *167*, 151-161.
- Morgan-Short, K., Faretta-Stuttenberg, M., Brill-Scheutz, 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*, 56-72.
- Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., & Fazio, F. (2001). Syntax and the Brain: Disentangling Grammar by Selective Anomalies. *NeuroImage*, *13*, 110-118.
- Murray, E. A., Wise, S. P., & Graham, K. S. (2016). *The evolution of memory systems: Ancestors, anatomy, and adaptations*. : Oxford University Press.
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, *6*, 774-781.
- Naismith, S. L., Lagopoulos, J., Ward, P. B., Davey, C. G., Little, C., & Hickie, I. B. (2010). Fronto-striatal correlates of impaired implicit sequence learning in major depression: An fMRI study. *Journal of Affective Disorders*, *125*, 256-261.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive Affective & Behavioral Neuroscience*, *5*, 263-281.
- Nurmi, E., Ruottinen, H. M., Bergman, J., Haaparanta, M., Solin, O., Sonninen, P., & Rinne, J. O. (2001). Rate of progression in Parkinson's disease: A 6- F-18 fluoro-L-dopa PET study. *Movement Disorders*, *16*, 608-615.
- O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, *6*, 505-510.
- 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. *The Journal of Neuroscience*, *24*, 8436.
- Opitz, B., & Friederici, A. D. (2007). Neural basis of processing sequential and hierarchical syntactic structures. *Human Brain Mapping*, *28*, 585-592.
- Opitz, B., & Hofmann, J. (2015). Concurrence of rule- and similarity-based mechanisms in artificial grammar learning. *Cognitive Psychology*, *77*, 77-99.
- Papagno, C., Comi, A., Riva, M., Bizzi, A., Vernice, M., Casarotti, A., Fava, E., & Bello, L. (2017). Mapping the brain network of the phonological loop. *Human Brain Mapping*, *38*, 3011-3024.
- Pasquini, J., Durcan, R., Wiblin, L., Gersel Stokholm, M., Rochester, L., Brooks, D. J., Burn, D., & Pavese, N. (2019). Clinical implications of early caudate

- dysfunction in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 90, 1098.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6, 674-681.
- Patterson, T. K., & Knowlton, B. J. (2018). Subregional specificity in human striatal habit learning: a meta-analytic review of the fMRI literature. *Current Opinion in Behavioral Sciences*, 20, 75-82.
- Peigneux, P., Meulemans, T., Van der Linden, M., Salmon, E., & Petit, H. (1999). Exploration of implicit artificial grammar learning in Parkinson's disease. *Acta Neurol. Belg.*, 99, 107-117.
- Peñaloza, Claudia, Benetello, A., Tuomiranta, L., Heikius, I.-M., Järvinen, S., Majos, M. C., Cardona, P., Juncadella, M., Laine, M., Martin, N., & Rodríguez-Fornells, A. (2015). Speech segmentation in aphasia. *Aphasiology*, 29, 724-743.
- Perruchet, P., & Rey, A. (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review*, 12, 307-313.
- Plante, E., Patterson, D., Gomez, R., Almryde, K. R., White, M. G., & Asbjørnsen, A. E. (2015). The nature of the language input affects brain activation during learning from a natural language. *Journal of Neurolinguistics*, 36, 17-34.
- Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, 414, 546-550.
- Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The Neural Correlates of Motor Skill Automaticity. *The Journal of Neuroscience*, 25, 5356.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855-863.
- Reber, P. J., Martinez, L. A., & Weintraub, S. (2003). Artificial grammar learning in Alzheimer's disease. *Cognitive Affective & Behavioral Neuroscience*, 3, 145-153.
- Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory*, 1, 217-229.
- Reber, P. J., & Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *Journal of Cognitive Neuroscience*, 10, 248-263.
- Reber, P. J., & Squire, L. R. (1999). Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience*, 113, 235-242.
- Rogalsky, C., & Hickok, G. (2011). The Role of Broca's Area in Sentence Comprehension. *Journal of Cognitive Neuroscience*, 23, 1664-1680.
- Rose, M., Haider, H., Salari, N., & Büchel, C. (2011). Functional Dissociation of Hippocampal Mechanism during Implicit Learning Based on the Domain of Associations. *The Journal of Neuroscience*, 31, 13739.
- Rosenthal, C. R., Andrews, S. K., Antoniadis, C. A., Kennard, C., & Soto, D. (2016). Learning and Recognition of a Non-conscious Sequence of Events in Human Primary Visual Cortex. *Current Biology*, 26, 834-841.
- Ruitenbergh, M. F. L., Duthoo, W., Santens, P., Notebaert, W., & Abrahamse, E. L. (2015). Sequential movement skill in Parkinson's disease: A state-of-the-art. *Cortex*, 65, 102-112.

- Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., & Halgren, E. (2009). Sequential Processing of Lexical, Grammatical, and Phonological Information Within Broca's Area. *Science*, *326*, 445.
- Sambin, S., Teichmann, M., Balaguer, R. D., Giavazzi, M., Sportiche, D., Schlenker, P., & Bachoud-Levi, A. C. (2012). The role of the striatum in sentence processing: Disentangling syntax from working memory in Huntington's disease. *Neuropsychologia*, *50*, 2625-2635.
- Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The Necessity of the Medial Temporal Lobe for Statistical Learning. *Journal of Cognitive Neuroscience*, *26*, 1736-1747.
- Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Current Biology*, *22*, 1622-1627.
- Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*, 486-492.
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, *26*, 3-8.
- Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An fMRI Study of the Role of the Medial Temporal Lobe in Implicit and Explicit Sequence Learning. *Neuron*, *37*, 1013-1025.
- Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 322-327.
- Schuchard, J., & Thompson, C. K. (2014). Implicit and Explicit Learning in Individuals with Agrammatic Aphasia. *Journal of Psycholinguistic Research*, *43*, 209-224.
- Schuchard, J., & Thompson, C. K. (2017). Sequential learning in individuals with agrammatic aphasia: evidence from artificial grammar learning. *Journal of Cognitive Psychology*, *29*, 521-534.
- Seger, C. A., & Cincotta, C. M. (2005). The Roles of the Caudate Nucleus in Human Classification Learning. *The Journal of Neuroscience*, *25*, 2941.
- Seger, C. A., Prabhakaran, V., Poldrack, R. A., & Gabrieli, J. D. E. (2000). Neural activity differs between explicit and implicit learning of artificial grammar strings: An fMRI study. *Psychobiology*, *28*, 283-292.
- Shanks, D. R., Channon, S., Wilkinson, L., & Curran, H. V. (2006). Disruption of sequential priming in organic and pharmacological amnesia: A role for the medial temporal lobes in implicit contextual learning. *Neuropsychopharmacology*, *31*, 1768-1776.
- Shanks, D. R., & St. John, M. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367-447.
- Shin, J. C., & Ivry, R. B. (2003). Spatial and temporal sequence learning in patients with Parkinson's disease or cerebellar lesions. *Journal of Cognitive Neuroscience*, *15*, 1232-1243.
- Smith, J., Siegert, R. J., & McDowall, J. (2001). Preserved implicit learning on both the serial reaction time task and artificial grammar in patients with Parkinson's disease. *Brain and Cognition*, *45*, 378-391.

- Smith, J. G., & McDowall, J. (2006a). The implicit sequence learning deficit in patients with Parkinson's disease: A matter of impaired sequence integration? *Neuropsychologia*, *44*, 275-288.
- Smith, J. G., & McDowall, J. (2006b). When artificial grammar acquisition in Parkinson's disease is impaired: The case of learning via trial-by-trial feedback. *Brain Research*, *1067*, 216-228.
- Smits-Bandstra, S., & Gracco, V. (2015). Retention of Implicit Sequence Learning in Persons Who Stutter and Persons With Parkinson's Disease. *Journal of Motor Behavior*, *47*, 124-141.
- Squire, L. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195-231.
- Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism-Language and Cognition*, *17*, 277-293.
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism-Language and Cognition*, *16*, 367-382.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness. *Journal of Cognitive Neuroscience*, *21*, 1934-1945.
- Tyler, L. K., Cheung, T. P., Devereux, B. J., & Clarke, A. (2013). Syntactic computations in the language network: Characterising dynamic network properties using representational similarity analysis. *Frontiers in Psychology*, *4*.
- Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., van Elswijk, G., Hagoort, P., & Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain Research*, *1224*, 69-78.
- Udden, J., Ingvar, M., Hagoort, P., & Petersson, K. M. (2017). Broca's region: A causal role in implicit processing of grammars with crossed non-adjacent dependencies. *Cognition*, *164*, 188-198.
- Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews: Neuroscience*, *2*, 717-726.
- Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, *92*, 231-270.
- Vafae, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, *39*, 59-95.
- van de Meerendonk, N., Rueschemeyer, S. A., & Kolk, H. H. J. (2013). Language comprehension interrupted: Both language errors and word degradation activate Broca's area. *Brain and Language*, *126*, 291-301.
- Van Tilborg, I. A. D. A., Kessels, R. P. C., Kruijt, P., Wester, A. J., & Hulstijn, W. (2011). Spatial and nonspatial implicit motor learning in Korsakoff's amnesia: evidence for selective deficits. *Experimental Brain Research*, *214*, 427-435.
- Vandenberghe, M., Schmidt, N., Fery, P., & Cleeremans, A. (2006). Can amnesic patients learn without awareness?: New evidence comparing deterministic and probabilistic sequence learning. *Neuropsychologia*, *44*, 1629-1641.
- Wang, R., Shen, Y., Tino, P., Welchman, A. E., & Kourtzi, Z. (2017). Learning Predictive Statistics: Strategies and Brain Mechanisms. *Journal of Neuroscience*, *37*, 8412-8427.

- Werheid, K., Ziessler, M., Nattkemper, D., & von Cramon, D. Y. (2003). Sequence learning in Parkinson's disease: The effect of spatial stimulus-response compatibility. *Brain and Cognition, 52*, 239-249.
- Willingham, D. B., Salidis, J., & Gabrieli, J. D. E. (2002). Direct Comparison of Neural Systems Mediating Conscious and Unconscious Skill Learning. *Journal of Neurophysiology, 88*, 1451-1460.
- Wilson, B., Kikuchi, Y., Sun, L., Hunter, D., Dick, F., Smith, K., Thiele, A., Griffiths, T. D., Marslen-Wilson, W. D., & Petkov, C. I. (2015). Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. *Nature Communications, 6*, 8901.
- Wilson, B., Marslen-Wilson, W. D., & Petkov, C. I. (2017). Conserved Sequence Processing in Primate Frontal Cortex. *Trends in Neurosciences, 40*, 72-82.
- Witt, K., Nuhsman, A., & Deuschl, G. (2002a). Dissociation of habit-learning in Parkinson's and cerebellar disease. *Journal of Cognitive Neuroscience, 14*, 493-499.
- Witt, K., Nuhsman, A., & Deuschl, G. (2002b). Intact artificial grammar learning in patients with cerebellar degeneration and advanced Parkinson's disease. *Neuropsychologia, 40*, 1534-1540.
- Yang, J., & Li, P. (2012). Brain Networks of Explicit and Implicit Learning. *Plos One, 7*.