

Second Edition

research
use processing
linguistic hypothesis
social representation cognitive
interaction mind
skill

Second Language Acquisition Research Series

**THEORIES IN SECOND
LANGUAGE ACQUISITION**
AN INTRODUCTION

Edited by
Bill VanPatten and Jessica Williams

ROUTLEDGE



THEORIES IN SECOND LANGUAGE ACQUISITION

An Introduction

Second Edition

*Edited by Bill VanPatten and
Jessica Williams*

MICHIGAN STATE UNIVERSITY AND
UNIVERSITY OF ILLINOIS AT CHICAGO

CONTENTS

<i>Preface</i>	<i>ix</i>
<i>Acknowledgments</i>	<i>xi</i>
1 Introduction: The Nature of Theories <i>Bill VanPatten and Jessica Williams</i>	1
2 Early Theories in SLA <i>Bill VanPatten and Jessica Williams</i>	17
3 Linguistic Theory, Universal Grammar, and Second Language Acquisition <i>Lydia White</i>	34
4 One Functional Approach to SLA: The Concept-Oriented Approach <i>Kathleen Bardovi-Harlig</i>	54
5 Usage-Based Approaches to SLA <i>Nick C. Ellis and Stefanie Wulff</i>	75
6 Skill Acquisition Theory <i>Robert DeKeyser</i>	94
7 Input Processing in Adult SLA <i>Bill VanPatten</i>	113

viii Contents

8	The Declarative/Procedural Model: A Neurobiologically Motivated Theory of First and Second Language <i>Michael T. Ullman</i>	135
9	Processability Theory <i>Manfred Pienemann and Anke Lenzing</i>	159
10	Input, Interaction, and Output in Second Language Acquisition <i>Susan M. Gass and Alison Mackey</i>	180
11	Sociocultural Theory and Second Language Development <i>James P. Lantolf, Steven L. Thorne, and Matthew E. Poehner</i>	207
12	Complexity Theory <i>Diane Larsen-Freeman</i>	227
13	Second Language Learning Explained? SLA across 10 Contemporary Theories <i>Lourdes Ortega</i>	245
	<i>Contributors</i>	273
	<i>Glossary</i>	277
	<i>Index</i>	287

8

THE DECLARATIVE/ PROCEDURAL MODEL

A Neurobiologically Motivated Theory of First and Second Language¹

Michael T. Ullman

In evolution and biology, previously existing structures and mechanisms are constantly being reused for new purposes. For example, fins evolved into limbs, which in turn became wings, while scales were modified into feathers. Reusing structures to solve new problems occurs not only evolutionarily, but also developmentally, as we grow up. For example, reading seems to depend on previously existing brain circuitry that is coopted for this function as we learn to read.

It thus seems likely that language should depend at least partly, if not largely, on neurobiological systems that existed prior to language—whether or not those systems have subsequently become further specialized for this domain, either through evolution or development. In this chapter, I focus on long-term memory systems, since most of language must be learned, whether or not aspects of this capacity are innately specified. Specifically, we are interested in whether and how two memory systems, **declarative memory** and **procedural memory**, play roles in language. These are arguably the two most important long-term memory systems in the brain in terms of the range of functions and domains that they subservise. The declarative/procedural (DP) model simply posits that these two memory systems play key roles in language in ways that are analogous to the functioning of these systems in other domains.

Importantly, these memory systems have been well studied in both animals and humans, and thus are relatively well understood at many levels, including their behavioral, brain, and molecular correlates. This understanding leads to a wide range of predictions about language that one might have no reason to make based on the more limited study of language alone. For example, if a particular brain structure or gene is known to play a particular role in these memory systems, we might expect it to play an analogous role in language, even if one might have had no reason to make such a prediction based on our understanding of language alone. The DP model is thus a very powerful theoretical framework.

The Theory and Its Constructs

In this section, I first present an overview of the two memory systems and how they interact with each other. I then examine predictions for language that follow from this knowledge of the memory systems. Although I focus on predictions for second language (L2), I explicitly compare L2 and first language (L1), since the predictions for L2 are intimately bound up with those for L1. For more on the DP model, including background, predictions, and evidence, see Ullman (2001b, 2004, 2007, 2008, in press). For a focus of the model on second language, see Ullman (2001a, 2005, 2012) and Morgan-Short and Ullman (2011).

Before we delve into the two memory systems, however, I will first provide a quick tour of the brain, as an overview of the necessary neurobiological basics. The largest part of the brain, and the most important for cognition, including language, is the cerebrum (Figure 8.1). The cerebrum is composed of two hemispheres, each of which contains four lobes: the frontal, temporal, parietal, and occipital lobes. Each lobe contains many smaller structures known as gyri and sulci (singular: gyrus and sulcus). The gyri are the ridges on the surface of the brain, and the sulci the valleys that lie between them. These gyri and sulci form the outer part of the cerebrum, called cortex.

Although most studies of language focus on cortical regions such as **Broca's area** (which corresponds to the opercular and triangular part of the inferior

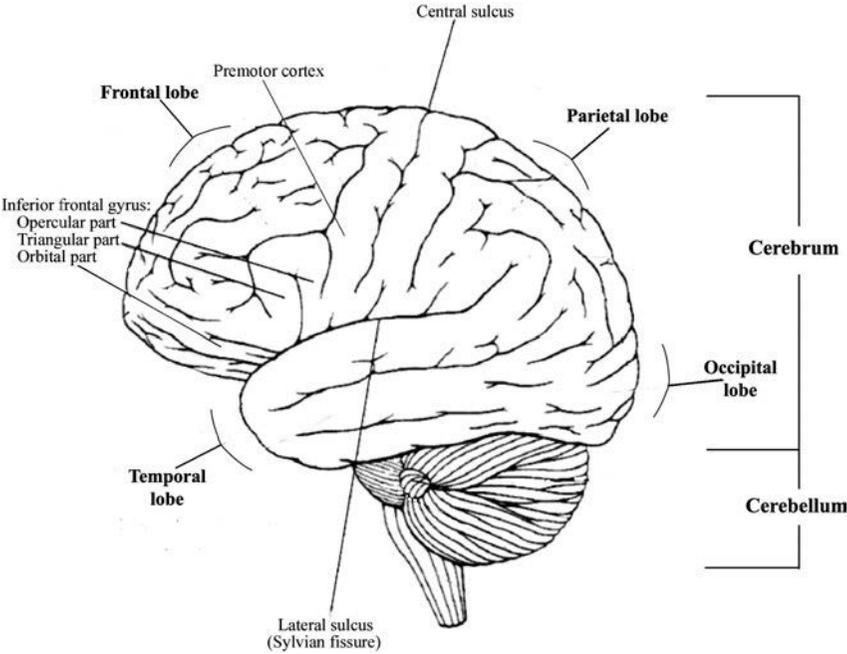


FIGURE 8.1 The left side of the brain, indicating structures referred to in the chapter.

frontal gyrus; see Figure 8.1), other brain structures are also important for language. The cerebellum, which lies below the cerebrum at the back of the brain, used to be thought of as being involved only in movement. However, we now know that it plays roles in cognition, including memory and language. There are also a number of structures deep inside the cerebrum itself (see Figure 8.2 in the color plate section). Of particular interest here are two sets of structures: first, the **basal ganglia** (including the caudate nucleus, and other portions not discussed here), which were previously thought to be mainly involved in movement, and second, the **hippocampus** and other medial (i.e., inner) temporal lobe regions, which were thought to underlie only memory. As we shall see, however, both of these sets of structures also play important roles in language.

Declarative Memory

The declarative memory brain system has been intensively studied in both humans and animals. The hippocampus and other medial temporal lobe structures (i.e., structures located toward the inner part of the temporal lobe) are critical for learning and consolidating new knowledge in this system (consolidation refers to the strengthening of memories after learning, for example, during sleep). These medial temporal lobe (MTL) structures may be not just involved, but actually *required* for learning idiosyncratic (unpredictable) bits of information and linking them together; evidence for this comes from studies of patients with extensive MTL damage. These individuals, such as the famous patient H.M., cannot learn new, idiosyncratic information. Although MTL structures are critical for acquiring new knowledge in this system, eventually this knowledge seems to depend less on MTL structures and more on neocortex, especially in the temporal lobes (neocortex refers to all cerebral cortex outside the MTL; for example, all of the cortex you see in Figure 8.1 is neocortex). Additionally, a region corresponding largely to the triangular and orbital parts of the inferior frontal cortex—often simplified to Brodmann’s Areas (BAs) 45 and 47, respectively—may underlie the encoding of new memories as well as their later recall. The molecular bases of declarative memory are also beginning to be understood. For example, various genes (e.g., for the proteins BDNF or APOE) play important roles in declarative memory and hippocampal function, as does the hormone estrogen (higher levels are associated with better declarative memory). (For more information about the declarative memory brain system and its functions, see Eichenbaum, 2012; Henke, 2010; Squire & Wixted, 2011; Ullman, 2004, in press.)

The functions of this network of brain structures are reasonably well characterized. Their role in learning idiosyncratic knowledge across a wide range of modalities and domains may explain why they are important for learning information about facts (semantic knowledge) and events (episodic knowledge), such as the fact that French, Swiss-German, Italian, and Romansch are all spoken in Switzerland, or that yesterday evening I roasted chestnuts in the fireplace with my

daughter Clemi while it was snowing outside. Knowledge can be learned very rapidly in declarative memory, with as little as a single exposure of the stimulus (the reader now knows what I did yesterday evening), although additional exposures of course strengthen these memories. This knowledge is at least partly, although *not* completely, explicit (available to conscious awareness): The system also underlies implicit (nonconscious) knowledge (Henke, 2010); also see “The Explicit/Implicit Debate” later in this chapter. Nevertheless, declarative memory appears to be the only long-term memory system that underlies explicit knowledge; thus, any explicit knowledge must have been learned in declarative memory.

Finally, a number of subject-level factors appear to modulate learning and retention in this system. Of particular interest for second language acquisition, learning in declarative memory seems to improve during childhood and plateaus in adolescence and early adulthood, after which it declines. Thus an older child or young adult tends to be better at learning in this system than a young child. Sex is also a factor, with evidence suggesting that females have an advantage at declarative memory over males, possibly due to their higher estrogen levels. Other factors that seem to affect declarative memory include handedness (left-handedness may be associated with an advantage at declarative memory), sleep (memory consolidation seems to improve during sleep), and exercise (which may enhance declarative memory).

Procedural Memory

Although procedural memory is not as well understood as declarative memory, the workings of this memory system are becoming clearer. The system is composed of a network of interconnected brain structures rooted in frontal/basal ganglia circuits. The basal ganglia, and especially the caudate nucleus, play a critical role in the learning and consolidation of new motor and cognitive skills. In contrast, frontal regions may be more important for processing skills that have been automatized—in particular premotor cortex and the opercular part of the inferior frontal cortex (Figure 8.1), often simply referred to as BA 6 and BA 44, respectively. The cerebellum also seems to play a role, though exactly how and in what ways remain unclear. Some aspects of the molecular bases of procedural memory are also beginning to emerge. For example, certain genes playing roles in procedural memory have been identified, including for the proteins FOXP2, DARPP-32, and DRD2. (For more on this memory system, see Ashby, Turner, & Horvitz, 2010; Doyon et al., 2009; Eichenbaum, 2012; Packard, 2008; Ullman, 2004, in press.)

This brain circuitry underlies the implicit (nonconscious) learning and processing of a wide range of activities and functions, including sequences, rules, categories and routes (for navigation). The system may be specialized for learning to predict, perhaps especially probabilistic outcomes—for example, the next item in a sequence or the output of a rule. Learning in the system requires extended practice, which seems to eventually result in more rapid and automatic processing of skills and knowledge than does learning in declarative memory. Various factors

seem to affect procedural memory, including, of particular interest to L2 learning, one's age: Unlike declarative memory, learning and consolidation in procedural memory seem to be already robust early in childhood, though they appear to decline during childhood/adolescence, resulting in poorer learning/consolidation abilities in adulthood.

Interactions between the Memory Systems

Declarative and procedural memory also interact with each other. (For more information, see Packard, 2008; Poldrack & Packard, 2003; Ullman, 2004, in press.) First, with important consequences for L2 learning, the two systems can complement each other in acquiring the same or analogous knowledge, including knowledge of sequences and rules. Thus, they play at least partly *redundant* roles, in that they can at least partly learn and process the same knowledge, though generally in different ways from each other. Various factors appear to modulate which of the two systems is relied on more. Declarative memory may acquire knowledge initially, thanks to its rapid acquisition abilities, while the procedural system gradually learns analogous knowledge, which is eventually processed rapidly and automatically. The learning context can also affect which system is relied on more. Explicit instruction (e.g., of sequences), or conscious attention to input stimuli and an attempt to understand underlying rules or patterns, can increase learning in declarative memory. Conversely, a lack of explicit instruction, as well as manipulations that reduce attention to the stimuli, or a high level of complexity of rules or patterns (thus decreasing the learner's ability to explicitly detect patterns), can all shift learning toward procedural memory.

Second, animal and human studies suggest that the two systems also interact competitively, resulting in a "seesaw effect." For example, the dysfunction or attenuation of one system may lead not only to an increased dependence on the other system for those tasks and functions that can depend on either one, but *also* to the enhanced functioning of this other system.

Finally, some evidence suggests that the learning and/or retrieval of knowledge in declarative memory may block (inhibit) the learning and/or retrieval of analogous knowledge in procedural memory. The converse may hold as well. For example, neuroimaging evidence in humans suggests that learning in declarative memory may inhibit, or at least override, learning of analogous knowledge in procedural memory. And in rats, even when a task is learned initially in declarative memory, it can be overridden by procedural memory when the task is subsequently learned in that system.

Predictions for Language

The DP model posits that the declarative and procedural memory systems should play roles in language that are largely analogous to the roles they play in other

domains. Thus, our independent knowledge of the two memory systems, as laid out above, leads to quite specific predictions for language. Here I lay out a number of these predictions, specifying where they are common to L1 and L2, and where they hold particularly for L2.

Predictions: Declarative Memory

Since declarative memory is important, and perhaps necessary, for learning arbitrary pieces of information and associating them, this memory system should be crucial for all learned idiosyncratic knowledge in language. This should hold for both L1 and L2. Thus, simple content words (e.g., *cat*, *devour*), including their phonological forms, meanings, subcategorization frames (e.g., *devour* requires a complement), and mappings between them (e.g., the sound–meaning mappings), should be learned in this system. Some sort of knowledge about irregular morphological forms, both inflectional and derivational (e.g., *dig-dug*, *solemn-solemnity*) should also be stored in declarative memory, as should idiosyncratic knowledge about idioms, proverbs, and so on (e.g., *jump the gun*, *a bird in the hand is worth two in the bush*).

Since declarative memory is so flexible in what it can learn, it should also be available for learning nonidiosyncratic, rule-governed aspects of language. Thus, just like simple and irregular words, one should be able to store, in some manner (e.g., as “chunks”), at least some rule-governed complex forms (e.g., “*walked*,” “*the cat*”), together with their meanings. Grammatical rules and constraints should also be learnable by this system (implicitly or explicitly), even though these are generally acquired by procedural memory. As we shall see, complex forms and grammatical knowledge should generally depend more on declarative memory (and less on procedural memory) in L2 than L1, due to factors such as age of acquisition and learning context.

In both L1 and L2, linguistic knowledge in declarative memory should be learned relatively rapidly, perhaps in some cases even from a single presentation of the information (e.g., if the information is simple enough), though repeated exposures should improve learning and retention. Linguistic knowledge learned in this system can be explicit or implicit, since both types can be learned by declarative memory. However, any explicit long-term knowledge of language must have been learned by declarative memory, since this is the only long-term memory system that seems to underlie explicit knowledge. Most importantly for L2, language learning that depends on declarative memory should ameliorate during childhood, plateau in adolescence/early adulthood, and then decline.

We can also make neurobiological predictions about language knowledge learned in declarative memory. The functional neuroanatomy of this knowledge, whether in L1 or L2, should reflect the functional neuroanatomy of declarative memory. Thus it should crucially depend on the hippocampus and other MTL structures, at least during learning. Eventually MTL structures should become less

important, with a corresponding increasing role for neocortical regions, especially in the temporal lobes. The area encompassing BA 45 and 47 should underlie the encoding of new linguistic information being learned in declarative memory, as well as the recall of that knowledge once it has been learned. Finally, the genes and molecules involved in declarative memory should play analogous roles in aspects of language learned by this system.

Predictions: Procedural Memory

Given what we know about procedural memory, this memory system may be expected to underlie the learning and processing of sequences and rules in both first and second language. Those that are probabilistic may depend especially on this system. The system may be particularly important in learning to predict in language, such as predicting the next item in a sequence or the output of a linguistic rule. Only rules or sequences that are implicit should be learned by procedural memory. Given that the mental grammar heavily involves implicit rules, in particular ones that involving sequencing (especially hierarchical sequencing), we would expect procedural memory to play a major role in grammar. This should hold across linguistic subdomains, including syntax, morphology, and phonology. Other aspects of L1 and L2 may also be learned in procedural memory. Given the role of this system in category learning, it might also underlie the acquisition of linguistic categories (e.g., syntactic categories). Other nongrammatical implicit learning in language may also depend on procedural memory, especially if it involves probabilistic patterns, sequences, and learning to predict. For example, the implicit learning of word boundaries in a speech stream may be expected to depend on procedural memory.

Since procedural memory learns with repeated exposure, linguistic knowledge learned in this system should be acquired gradually. Although only implicit (not explicit) knowledge in L1 and L2 should rely on procedural memory, not all implicit knowledge should involve this system, since there are other implicit memory systems, and, as we have seen, declarative memory also subserves implicit knowledge (also see “The Explicit/Implicit Debate”). Of particular importance for L2, language learning in procedural memory should be better in childhood than adulthood. Thus grammar should be easier to acquire in procedural memory in childhood (whether in L1 or L2) than in adulthood (generally as an L2).

As with declarative memory, neurobiological predictions for L1 and L2 follow largely from what we know about procedural memory from animal and human studies. Linguistic skills and knowledge learned in this system should involve frontal and basal ganglia structures, and perhaps the cerebellum. Learning and consolidation should engage the basal ganglia, especially the caudate nucleus. Once automatized, knowledge and skills should rely especially on neocortical regions, in particular premotor cortex and BA 44. And genes such as FOXP2 may modulate learning and processing in this system.

Predictions: Interactions between the Memory Systems

Our understanding of interactions between the two memory systems also leads to predictions for language and is of particular interest to L2. First, to some extent, we expect the two memory systems to acquire the same or analogous knowledge or skills, that is, to play at least partly redundant roles. In both L1 and L2, as in nonlinguistic domains, such redundancy should be found for any tasks or functions that can be subserved by both systems. Given the learning flexibility of declarative memory, and the fact that this system can underlie implicit as well as explicit knowledge, we might expect it to be able to at least partly support most, if not all, linguistic functions that are learned by procedural memory, including grammar. Declarative memory could support such knowledge in a variety of ways, including storing rule-governed complex forms as chunks (which could be structured or unstructured) or learning rules (explicitly or implicitly) (Ullman, 2005, 2006a).

Various factors should modulate which of the two memory systems is relied on more for linguistic knowledge that can be learned by either system. Such knowledge, in particular for grammar, should be learned first by declarative memory, but more slowly and in parallel by procedural memory, which should eventually lead to highly automatized knowledge. Since learning in declarative memory improves during childhood up to adolescence and young adulthood, while learning in procedural memory seems to become less effective during this period (perhaps due to the seesaw effect), young adult L2 learners should on average rely more on declarative and less on procedural memory for grammar than (L1 or L2) child learners, holding constant their exposure to the language. And even though the grammar should eventually become at least somewhat proceduralized in both child and adult learners, this process should occur faster and more completely in children. Thus, even after years of exposure, adult L2 learners might not attain the degree of proceduralization of their grammar as L1 or early L2 learners. It is worthwhile pointing out that adult L2 learners and adult L1 learners are often compared in empirical studies examining L1 and L2 processing, even though this comparison probes the two groups at different points in the learning trajectory; that is, L2 learners at earlier stages than L1 learners.

Explicit language instruction, or attention to language stimuli or patterns in the input, may increase language acquisition in declarative memory, while a lack of such instruction or attention, and greater complexity of rules or patterns (e.g., more complex grammatical rules or constraints), may lead to a greater relative dependence on procedural memory. Thus, explicit instruction of grammar, as is often given in classrooms to L2 learners, should encourage learning in declarative memory (which may then inhibit learning or processing in procedural memory). Conversely, exposure to the L2 without explicit instruction, as often occurs in immersion contexts, might enhance grammar acquisition in procedural memory, and thus lead to more L1-like processing of grammar.

Summary of Predictions

Here I summarize some of the main nonneurobiological predictions, focusing on similarities and differences between L2 and L1. First of all, in some ways, the predictions are similar in first and second language. In both L1 and L2, declarative memory should underlie the learning, storage, and use of all idiosyncratic knowledge in language. Thus idiosyncratic lexical knowledge should always be stored in this system, across linguistic subdomains (e.g., simple words and their meanings, irregular morphology, syntactic complements, idioms). In both L1 and L2, aspects of grammar should initially be learned in declarative memory. In parallel, procedural memory should also gradually learn grammatical knowledge. After sufficient experience with the language, procedural memory-based grammatical processing should tend to take precedence over analogous declarative knowledge, resulting in increasing automatization of the grammar.

However, L2 acquisition is also expected to differ in important ways from L1 acquisition. Perhaps most importantly, grammar should tend to depend more on declarative memory and less on procedural memory in L2 than L1, for several reasons. First, L2 learners will always have had less language exposure than L1 learners at the same age, simply because they began learning the L2 later. The later the L2 age of acquisition, the more pronounced this difference. Since declarative memory learns quite rapidly, while procedural memory learns only gradually, at any given age a learner's L2 grammar should be less proceduralized and should thus depend more on declarative memory than their L1 grammar. Thus, just for this reason alone, L2 grammar should tend to rely more on declarative memory than L1 grammar.

Second, because learning in procedural memory seems to be established early and then declines, while declarative memory shows the opposite pattern, L1 learners (and early L2 learners) should tend to rely particularly on procedural memory for learning grammar (especially after a reasonable amount of language exposure), while later (L2) learners should rely more on declarative memory, and indeed may never proceduralize their grammar to the same extent as L1 learners. Importantly, this pattern should hold even after the same amount of language exposure in L1 and L2. However, most neurocognitive studies do not compare L1 children with L2 adults (e.g., both after 10 years of language exposure). Rather, as pointed out earlier, most studies comparing the neurocognition of L2 with L1 examine both groups at the same age (e.g., a given subject's L1 and L2), and thus at different points in the learning trajectory. This is not problematic per se, but it must be kept in mind when interpreting the data.

Third, even the type of language experience may influence learning and the learner's relative dependence on the two memory systems. As we have seen, explicit, classroom-like instruction of the grammar may encourage learning in declarative memory, perhaps at the expense of learning in procedural memory. Conversely, the lack of explicit instruction, as often occurs in immersion contexts, may encourage learning in procedural memory. These predictions should hold

for both L1 and L2 learners. However, L1 learning generally occurs primarily in an immersion (naturalistic) context, further encouraging proceduralization of the grammar in L1 speakers and eventual automatization. In contrast, since L2 learners vary considerably with respect to the type of exposure, this factor should in general have a larger impact on the neurocognition of L2 and should tend to increase the reliance of grammar on declarative memory in L2.

Note that much of the literature on the neurocognition of L2 grammar, and whether L1-like neurocognitive grammatical processing can be attained, has focused on two factors: age of acquisition and proficiency. However, proficiency is somewhat of a problematic variable. First, it is operationalized and measured quite differently across studies. Perhaps more importantly, in the vast majority of studies proficiency is highly confounded with other variables, in particular the amount of exposure and even the type of exposure, since higher proficiency is associated with higher exposure and even in many studies with more immersion experience. As we have seen, the DP model makes separate predictions for both the amount and the type of exposure. In contrast, the model does not take a strong position on proficiency itself to the extent that it may vary independently from these other variables. For example, it might (or might not) be that higher proficiency is associated with greater grammatical proceduralization, even holding constant the amount and type of experience. Future research will hopefully elucidate this issue.

What Counts as Evidence?

Multiple types of evidence can help test the predictions laid out earlier. This includes evidence from different methodologies, different language paradigms (e.g., natural languages, artificial languages, artificial grammars), different tasks, and different experimental designs. Importantly, every methodology, paradigm, task, and design has both strengths and weaknesses. Thus, it is crucial to obtain evidence from multiple approaches to test for *converging evidence*. Only with converging evidence should we begin to have confidence in a theory. Here I discuss several lines of relevant evidence.

Behavioral Evidence: Correlational Studies

Various types of behavioral evidence can be used to test the predictions of the DP model. One of the most straightforward and intuitive behavioral approaches is examining correlations across subjects, between how well they learn in the memory systems and how well they learn or process language. For example, if people who are better at learning in declarative memory are also better at word learning, this may be taken to suggest that word learning depends on declarative memory. However, we have to be careful, because correlation does not imply causation. A correlation between word learning and declarative memory could be explained

not just by words being learned in declarative memory, but instead by some general cognitive process that underlies both word learning and declarative memory.

One way to address this problem is to hold such other factors constant (e.g., in statistical analyses), if one can identify them. Another way is to show the specificity of the correlation. If word learning or processing correlates with declarative memory but *not* with procedural memory, this suggests that lexical memory has a particular link to declarative memory that is not found with all learning systems. Moreover, if the converse holds, that is, performance at grammar learning or processing correlates with learning in procedural but not declarative memory, this would further strengthen the specificity both of the relation of lexicon to declarative memory and of grammar to procedural memory. Indeed, such **double dissociations** have been found (Kidd, 2012; Lum, Conti-Ramsden, Page, & Ullman, 2012).

Similarly, one can use correlations to test the DP model's prediction that grammar depends particularly on declarative memory at low L2 exposure, but on procedural memory at high exposure: across subjects, at low L2 exposure grammar measures (e.g., syntactic judgment) should correlate with one's ability to learn in declarative (but not procedural) memory, whereas at high exposure grammar should correlate with learning in procedural (but not declarative) memory. This pattern of correlations was in fact found in a recent study, supporting the predictions of the DP model (Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter, & Wong, 2014).

Finally, other types of behavioral evidence can also be used to test the DP model in L1 and L2, for example, the examination of priming and frequency effects; for more information on these, see Babcock, Stowe, Maloof, Brovotto, and Ullman (2012).

Neurological Evidence: The Lesion Method

If a person suffers from damage (lesions) to particular brain structures, and then loses the ability carry out certain cognitive functions, one might infer that the lost functions previously depended on the damaged structures. Using this approach to understand which brain structures normally underlie which functions is referred to as the lesion method. For example, the fact that lesions to the occipital lobes consistently lead to visual deficits supports the conclusion that the occipital lobes are important for vision. By analogy, if you damage your lungs, you will have trouble breathing, whereas if you damage your stomach, you will probably have trouble with your digestion. This shows that one cannot perform these functions without these particular organs, and so these organs are necessary for these functions.

The lesion method can be used to test the DP model. Patients with lesions limited to the medial temporal lobes, including the hippocampus, should have trouble learning not only facts and events in declarative memory, but also words. This is indeed the case, for example for the famous amnesic patient H.M. (Postle & Corkin, 1998). Additionally, patients with lesions that extend to temporal neocortex (e.g.,

from Alzheimer's disease) should have more trouble with an L2 grammar than with an L1 grammar. Conversely, patients with lesions to frontal/basal ganglia circuits (e.g., from strokes or from Parkinson's disease) should have greater grammatical impairments in their L1 than their L2. And in any such patients who have more than one L2, we should see evidence of greater grammatical impairment in the L2 to which they had more exposure. Note that these predictions are striking given that the L1 (and the higher exposure L2) had presumably been better learned than the L2 (and the lower exposure L2), and yet are predicted to be more impaired. Indeed such double dissociations have been found, supporting the DP model (Hyltenstam & Stroud, 1989; Johari et al., 2013; Zanini, Tavano, & Fabbro, 2010).

Like all other methods, however, the lesion method has its weaknesses. Clearly, we cannot go around causing brain injury in people willy-nilly. Rather, we must test patients who have already had a brain injury. But such 'accidental experiments' are not ideal. One cannot choose the location of the lesion, which is moreover often large and involves multiple brain structures, complicating structure-function inferences: How do you know which brain structure does what when many structures are damaged? Timing is also an issue. If one waits too long after the onset of an acute brain lesion, other structures may take over some of the functions that the damaged structure used to perform. Such compensation confuses one's inferences, since a lesioned structure may indeed have been critical for a function, but compensation by a different part of the brain leads to normal functioning, and thus could lead to the false conclusion that the lesioned structure is not important for the function. On the other hand, if one tests a patient too quickly after a stroke or head injury, the loss of function can be much greater than is attributable to the damaged regions, because nearby regions are often temporarily affected by a number of factors, such as tissue swelling. In practice, researchers tend to err on the side of longer periods of time, usually waiting months or even a year or more after acute brain damage before testing a patient.

Electrophysiological Evidence: Event-Related Potentials

Event-Related Potentials (ERPs) are measures of brain activity, specifically the electrical activity that constitutes the basis of brain function. In an ERP study, just as in a sleep study measuring EEGs (electroencephalograms), electrical potentials from brain activity are recorded from electrodes placed on the scalp. ERPs are simply the EEG activity that occurs right after a person hears or sees a word, sees a picture, and so on. The presentation of such a stimulus is called an "event," hence the name Event-Related Potentials.

ERPs offer several advantages over some other methodologies. First, unlike functional neuroimaging methods like fMRI or PET (see later), ERPs provide excellent **temporal resolution**, with millisecond measurements that allow one to examine the actual time course of language processing. On the down side however, localizing the neuroanatomical source of ERPs in the brain is quite difficult.

Second, ERP research has revealed a set of widely-studied language-related ERP patterns (“ERP components”) in L1, whose characteristics and underlying functions are reasonably well understood: primarily the N400, LAN and P600 (see later). Moreover, lexical and grammatical processing in the L1 are each associated with largely distinct ERP components. These components thus provide a relatively clear way of comparing the neurocognition of language processing between L2 and L1, in particular for these language domains. Finally, since ERPs can be sensitive to effects that are not actually observed with behavioral measures, including in language learning studies, they can potentially reveal L2-L1 differences and similarities that might not be found with behavioral approaches. (For reviews of ERP research in L2, see Bowden, Steinhauer, Sanz, & Ullman, 2013; Kotz, 2009; Morgan-Short & Ullman, 2011; Steinhauer, White, & Drury, 2009.)

Here I briefly outline the main ERP language components—the N400, LAN, and P600—and explain how they can be used to test the DP model. In L1, lexical manipulations, such as seeing or hearing an unexpected word (e.g., “Adrian likes to eat planets”) reliably leads to N400 ERP components: that is, negative (hence the N) potentials that are generally found about 400 milliseconds after the presentation of the word, mainly at electrodes on the top of the head (see Figure 8.3 in the color plate section). It has been argued that the N400 reflects, at least in part, the processing of knowledge learned in declarative memory (Ullman, 2001a). Since the DP model predicts that lexical memory depends on declarative memory not only in L1, but also in L2, lexical manipulations should consistently elicit N400s in L2 as well as in L1. Indeed, this is the case.

Disruptions of rule-governed grammatical processing (of syntactic word order, or of morphosyntax such as agreement or tense, as in the sentence “Yesterday my father Fred walk all around Prague”) often produce two ERP components in L1. First, they can elicit early Anterior Negativities (“ANs”; Figure 8.3), which are often larger in the left hemisphere. These can begin as early as 100 milliseconds after the critical word (e.g., “walk”) and often continue for hundreds of milliseconds. It has been suggested that ANs may partially reflect the processing of knowledge learned in procedural memory. Second, grammatical disruptions often also produce P600s, positive potentials that often begin around 600 milliseconds (Figure 8.3). The P600 seems to reflect conscious processing of syntax and is not posited to depend on procedural memory; thus ANs are more relevant for testing the DP model. The DP model predicts that grammar depends more on declarative memory in L2 than L1, in particular in low exposure L2, but can be proceduralized at high exposure L2. Thus, grammatical disruptions should elicit N400s in L2, especially at lower levels of L2 experience, but ANs at higher levels. The evidence thus far is indeed consistent with this pattern.

Functional Neuroimaging Evidence

Functional neuroimaging methods such as PET (positron emission tomography) and the more common fMRI (functional magnetic resonance imaging) have also

been widely used to examine the neural bases of both L1 and L2. These techniques detect changes in blood flow (PET) or blood oxygen levels (fMRI) that are known to correlate with changes in neuronal activity. For example, if during grammar processing neurons fire particularly in the area of BA 44, this region should show particular changes in blood flow and oxygen levels (since firing neurons need more oxygen), suggesting that this region is especially important for grammar. The primary benefit of functional neuroimaging techniques is their excellent **spatial resolution**, allowing one to pinpoint activity to within a few millimeters in the brain. In contrast, such changes in the blood are too slow to allow the detection of real-time processing changes, so (unlike with ERPs) one cannot use functional neuroimaging to measure real-time language processing in the brain. For a summary of fMRI and other neuroimaging methods, including their pros and cons, see Ullman (2006b).

Functional neuroimaging can be used to test the DP model. The model predicts that word learning should show activation initially in the MTL, including the hippocampus, whereas once words are learned neocortical regions, especially in the temporal lobe, should be more active. Grammar should also initially yield MTL activation. However, as learning proceeds during L2 acquisition, activation should decrease in the MTL, and should increase both in neocortical regions subserving declarative memory and in procedural memory structures. In particular, activation should be found in the basal ganglia, especially the caudate nucleus (since it is involved in procedural learning), but increasingly in BA 44 (for the processing of already-learned procedures). At high L2 exposure, it is possible that declarative memory structures will drop out for grammar, and the basal ganglia will no longer be reliably engaged, leaving only BA 44 activation. The specificity of these predictions allows the DP Model to be clearly tested, and potentially falsified (i.e., shown to be incorrect).

Thus far, the neuroimaging evidence from fMRI and PET has been somewhat inconsistent. Nevertheless, some patterns seem to be emerging. Word learning does engage MTL structures, including the hippocampus (Breitenstein et al., 2005; Davis & Gaskell, 2009). In contrast, MTL regions are not reliably engaged in lexical/semantics in adults (Binder, Desai, Graves, & Conant, 2009; Ullman, 2004). Grammar learning may also engage MTL structures as well as the caudate nucleus at very early stages, with continuing activation of the caudate nucleus and later engagement of BA 44 (Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004; Ruschemeyer, Fiebach, Kempe, & Friederici, 2005; Ullman, in press). However, further studies are needed, ideally with better controls for factors such as age of acquisition and the amount and type of L2 experience.

Common Misunderstandings

Here I address two common misunderstandings about the DP model and also discuss how this model differs from other neurocognitive perspectives of L2. I will

discuss the various misunderstandings regarding the relation between the declarative and procedural memory systems on the one hand, and explicit and implicit knowledge on the other, in the section “The Explicit/Implicit Debate.”

First, there is a common misunderstanding regarding the domain generality of the two memory systems. On one hand, both systems are “domain general” in that they underlie multiple cognitive domains. However, this does not preclude subspecialization for language within either system, which could come about either evolutionarily or during development. Indeed, evidence from other domains suggests that subspecialization can occur in both systems. For example, different portions of the MTL and different regions of temporal neocortex underlie different types of information (Ullman, in press). Likewise, different frontal/basal ganglia circuits subserved different sorts of information (Middleton & Strick, 2000). Nevertheless, at this time there is no convincing evidence for *domain-specific* circuitry for language (i.e., dedicated to this domain), either within structures involved in the two memory systems or elsewhere in the brain (Ullman, Lum, & Conti-Ramsden, in press). Future research may further clarify this issue.

Second, another common misconception is that the changes in the reliance of grammar from declarative to procedural memory are due to some sort of “transformation” of knowledge from one to the other system. This is not the case. Rather, the two systems seem to independently acquire knowledge. Thus, proceduralization of grammar does not constitute the “transformation” of declarative into procedural representations but rather the gradual acquisition of grammatical knowledge in procedural memory, which is increasingly relied on, with an accompanying decrease in reliance on declarative memory.

Finally, to clarify any potential misconceptions regarding differences between the DP model and other neurocognitive models of L2, here I will compare the models. The DP model lies within one of three broad classes of neurocognitive models of L2. One class of models posits that the neurocognitive mechanisms underlying L2 are essentially the same as those subserving L1 (Abutalebi, 2008; Ellis, 2005; Green, 2003; Hernandez, Li, & MacWhinney, 2005; Indefrey, 2006; MacWhinney, 2011; Perani & Abutalebi, 2005). Second, it has been suggested that the mechanisms underlying L2 are fundamentally different from those of L1 (Bley-Vroman, 1989). A third group of models hypothesize that L2 learners initially depend heavily on different substrates than L1, but, with increasing experience or proficiency, gradually rely more on L1-like neurocognitive mechanisms. This group of theories includes the views espoused by Paradis and by Clahsen as well as the DP model. Although these views are similar in certain respects, they also differ. Paradis (2004, 2009) suggests that a shift between neurocognitive systems can take place both for rule-governed grammatical processes, and at least some lexical properties, specifically, grammatical properties of lexical items that are generally implicit in L1. More generally, Paradis takes a traditional view equating the explicit/implicit distinction with the declarative/procedural memory distinction, a view that is not tenable given what we know about the

memory systems (see details on the memory systems above, and the later section “The Explicit/Implicit Debate”). Clahsen proposes a model that is quite similar to the DP model in many respects, though with less of an expectation that the processing of grammar can become L1-like (Clahsen & Felser, 2006a, 2006b; Clahsen, Felser, Neubauer, Sato, & Silva, 2010). Additionally, Clahsen’s model focuses on psycholinguistic processing claims, rather than the neurocognitive bases of language.

Exemplary Study: Morgan-Short et al. (2012)

A major limitation of L2 research that examines the L2 learning trajectory is the time it takes learners to reach high proficiency. This makes it impractical to examine subjects over the full course of language learning (i.e., in a longitudinal study). As a result, the trajectory of learning has almost always been examined between groups of learners at different proficiency or exposure levels. However, like any between-subjects design, this approach is not ideal. The difficulty in selecting and matching different subject groups on critical factors of L2 exposure and use, let alone on other factors that may influence language learning (e.g., genotype), can introduce noise and inconsistency, reducing confidence in the findings.

To address these weaknesses, some studies have turned to artificial grammars or artificial languages. An artificial grammar typically involves presenting subjects with letter or tone sequences that are generated by some grammar. Artificial grammars can be learned to high proficiency quickly, in minutes to hours. However, even though the rules of artificial grammars can be consistent with the rules of natural languages, they are not fully language-like, since they lack vocabulary and the sequences have no meanings. Additionally, unlike a natural language, one does not speak or comprehend artificial grammars. Artificial languages address some of these concerns. An artificial language contains a small, meaningful lexicon and a limited number of grammatical rules, generally consistent with those found in natural languages. The sentences have meanings, and the language can be spoken and understood. Crucially, their small size makes them learnable to high proficiency within hours, allowing one to longitudinally examine the L2 learning trajectory to high proficiency.

A recent study using an artificial language paradigm examined L2 learning longitudinally to high proficiency (Morgan-Short, Steinhauer et al., 2012). Monolingual native English-speaking adults were trained to speak and comprehend an artificial language, Brocanto2. The words in this language refer to the pieces and moves of a computer based game, and the rules follow those of natural languages. Half the subjects in this study were given “explicit”, instructed, classroom-like training, and half were given an equivalent amount of “implicit,” uninstructed, immersion-like training. ERPs on violations of syntactic word order (discussed here) and morpho-syntactic agreement (see Morgan-Short, Sanz, Steinhauer, & Ullman, 2010) were measured three times: at low proficiency, high proficiency, and

again about 5 months later, to test for the neurocognition of L2 retention (retention testing is reported in Morgan-Short, Finger, Grey, & Ullman, 2012).

Behavioral analyses showed that both the explicitly and implicitly trained groups learned the language to high proficiency and then retained it 5 months later, and did not differ from each other at any of these time points. In contrast, ERPs showed clear group differences (here I discuss word order violations; for agreement violations, see Morgan-Short et al., 2010). At low proficiency and exposure the implicitly trained group showed an N400, whereas the explicitly trained group showed no detectable ERP effects (Figure 8.3). At high proficiency and exposure, the implicitly trained group showed an AN/P600 biphasic pattern (although the AN was not significantly left-lateralized), with the AN continuing as a late anterior negativity. In contrast, the explicitly trained group showed only an anterior positivity (not typical of native language) followed by a P600. At retention testing 5 months later, the implicitly trained group showed a more robust and left-lateralized AN than at high proficiency, the explicitly trained group no longer showed the (non-L1-like) anterior positivity and developed a more robust P600, and both groups showed a stronger late anterior negativity.

In sum, L1-like processing of syntactic word order, including ANs, was more likely for implicit, uninstructed (immersion-like) training than for explicit, instructed (classroom-like) training, and more likely at retention testing than at high proficiency/exposure than at low proficiency/exposure. Specifically, N400s were only found at low proficiency, suggesting a reliance of syntax early in the learning trajectory on declarative memory. The fact that no N400 or any other ERP component was reliably found in the explicitly trained subjects at low proficiency may be due to greater temporal variability (i.e., in when the component occurs) for explicit, conscious, strategies, resulting in the lack of any consistent ERP components in any given time window (Morgan-Short, Steinhauer et al., 2012). At high proficiency more native-like grammatical ERP components were found, including an AN. This is consistent with proceduralization, and more generally with greater L1-like grammatical processing emerging with greater exposure and proficiency. The findings that both training groups showed more native-like syntactic processing at retention testing may have been due in part to continuing consolidation of the grammar in procedural memory (Morgan-Short et al., 2012). Finally, the greater native-like processing resulting from implicit than explicit training is consistent with immersion leading to more native-like processing and proceduralization than explicit instructed classroom training (Bowden et al., 2013).

This study is exemplary in several respects. First, the use of an artificial language allows one to control for the amount, type, and timing of L2 exposure. Second, the fact that an artificial language rather than an artificial grammar was examined, and moreover one that subjects learned to speak and comprehend, and that followed the rules of natural languages, suggests that the results are reasonably likely to generalize to natural languages, which is of course what we actually care about

understanding. Third, the measurement of ERPs as well as behavioral assessments provides a variety of advantages (see earlier), including revealing ERP differences that were not found in behavior. Fourth, examining and contrasting instructed, explicit, classroom-like training and uninstructed, implicit, immersion-like training, moreover in a tightly controlled design, elucidates neurocognitive effects of the type as well as the amount of input. Fifth, examining retention, moreover after quite an extended period, provides insights into longer-term outcomes of language learning. Since people usually learn an L2 to retain it (at least for a reasonable period), this is particularly important.

Explanation of Observed Findings in SLA

Observation 1: Exposure to input is necessary for SLA; Observation 2: A good deal of SLA happens incidentally. As discussed earlier, the evidence suggests that not only is exposure to input necessary for learning an L2, but the amount and even the type of input is important. Specifically, more exposure (correlating with higher proficiency), and immersion experience (which presumably is associated with incidental learning) may be critical for proceduralization of the grammar and the attainment of L1-like neurocognitive grammatical processing.

Observation 5: Second language learning is variable in its outcome; Observation 6: Second language learning is variable across linguistic subsystems. According to the DP model, both behavioral and neural correlates of L2 learning should vary on the basis of multiple factors, including biological variables (e.g., sex and genetic variability), input variables (e.g., amount and type of L2 exposure), and linguistic subsystems (e.g., lexicon vs. grammar). Moreover, a number of these variables likely interact. Some of these factors have already been reasonably well examined (in particular, lexicon vs. grammar, and input variables), and indeed the evidence suggests that they influence SLA. A host of other variables should be examined in future studies.

Observation 9: There are limits on the effects of instruction on SLA. As we have seen, implicit, uninstructed immersion-like L2 training appears to be more effective than instructed classroom-like training in the attainment of L1-neurocognition of grammar.

The Explicit/Implicit Debate

At first blush the distinction between the declarative and procedural memory brain systems seems to parallel that between explicit and implicit knowledge. Indeed, explicit knowledge is subserved only by declarative memory, while procedural memory underlies implicit knowledge. However, the parallel largely falls apart at this point.

First, the DP model is based on claims about brain systems, whereas the explicit/implicit distinction is premised on claims about awareness. This latter distinction is somewhat problematic in that awareness is difficult not only to define, but also

to test (DeKeyser, 2003; Schmidt, 1994). In contrast, the distinction between the declarative and procedural brain systems is relatively clear, and the dichotomy can be tested, as we have seen, with a variety of methodological approaches.

Second, the mapping between declarative/procedural memory, on one hand, and explicit/implicit knowledge on the other, is by no means isomorphic (one-to-one). On the one hand, information stored in declarative memory can be explicit (accessible to conscious awareness in some sense). Indeed, as we have seen, this brain system appears to be the only long-term memory system to underlie explicit knowledge—a finding that is useful since it allows us to identify declarative memory as the locus of any long-term explicit knowledge. However, this system also underlies implicit knowledge. Although declarative memory was historically associated only with explicit knowledge, this was always a highly problematic assumption (even though this problem was rarely discussed). It was never shown (how would one do so?) that this *brain system* does *not* underlie implicit knowledge. Indeed, work in nonhuman animals such as rats and monkeys on this brain system did not assume that learning involved explicit knowledge, since testing animals' conscious awareness of what they have learned would clearly be very difficult. And of course it is also highly unwarranted to simply define a biological entity such as a brain system as having particular behavioral characteristics, in this case that it only underlies explicit knowledge. Rather this is an empirical question. Thus, not only was it always the case that the assumption that declarative memory underlies only explicit knowledge was unwarranted, but evidence now indicates that this assumption was not correct, and that this declarative memory also underlies implicit knowledge (Henke, 2010; Ullman, in press). In sum, although declarative memory appears to be the only long-term memory system in the brain that underlies explicit knowledge, it also underlies implicit knowledge.

There are also often confusions with respect to procedural memory. Brain researchers generally define procedural memory as it is defined here, that is, as a brain system rooted in particular brain structures. Importantly, procedural memory is only one of *several* brain systems that underlie implicit knowledge, including not just declarative memory, but also other systems (e.g., those underlying priming and habituation) (Eichenbaum, 2012; Squire & Wixted, 2011). Nevertheless, the terms procedural memory and implicit memory are still often used interchangeably in some fields, which can result in substantial confusion. To clarify: although procedural memory appears to only underlie implicit knowledge, several other brain systems, *including* the declarative memory system, also underlie implicit knowledge.

I have been discussing problems pertaining to explicit/implicit knowledge. However, the explicit/implicit distinction in other respects is at least as problematic. First, one may hear of a distinction between explicit and implicit learning. This distinction usually refers to whether knowledge is explicit or implicit, but during the learning period rather than subsequent to learning (i.e., the product). The terms explicit and implicit are also used with respect to the input (e.g., see the

exemplary study in this chapter). However, this terminology is perhaps even more problematic, since it also causes confusion as to where the explicit knowledge is supposed to lie, with the instructor (experimenter) or the learner (subject). If the knowledge lies with the teacher/experimenter, then it is uninteresting with respect to learning; if the knowledge lies with the learner/subject, then again the same issues described above apply. Clearer terms, such as instructed and uninstructed learning, may be more useful. In the exemplary study, we used the terms explicit and implicit training to be consistent with the use of these terms in the existing literature, though we attempted to further clarify them by specifying that the distinction can also be described as instructed/uninstructed and classroom-like/immersion-like.

Conclusion

The DP model appears to be a useful and informative theoretical approach. First, it is motivated by basic principles of evolution and biology. Second, it generates a wide range of behavioral and neurobiological predictions, for both L1 and L2, many of which would be unwarranted by the more limited study of language alone. Thus it is a very powerful theory. Third, it is highly testable by multiple methods. Fourth, converging evidence from different methods and experimental paradigms supports the basic predictions of the DP model.

Finally, the DP model likely has important applied implications. For example, the model seems likely to make useful predictions and offer explanatory accounts for factors that may lead to improvements in L2 acquisition, the attainment of native-like processing, and the retention of what has been learned. One line of research seems particularly promising. Studies of memory have shown that a number of variables and interventions can lead to better learning and memory in brain memory systems: not only intrinsic biological factors such as sex or genotype, but also external manipulations, that is, interventions, such as spaced versus massed presentation, the testing effect, and exercise (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Roediger & Karpicke, 2006; Stern & Alberini, 2012). The DP model specifically predicts that if these techniques enhance learning and retention in declarative or procedural memory, they should also enhance language learning, including in second language acquisition.

Discussion Questions

1. One of the major underlying tenets of the Declarative/Procedural model is that the memory systems used in nonlanguage learning are coopted for learning language. Does this perspective suggest that language learning is like learning anything else? Is it possible for language to make use of human memory systems and yet be “special” in the way that, say, Lydia White suggests language is special in Chapter 3?

2. Both Skill Theory and the DP Model make distinctions between declarative and procedural memory. What differences do you see between these constructs in the two approaches?
3. One of the predictions of the DP Model is that learning grammar in procedural memory becomes more difficult as age increases, over the course of childhood and adolescence. Compare and contrast this perspective with what is known as the Critical Period Hypothesis, which basically states that adults cannot make use of the same devices for language acquisition as children learning a first language.
4. An interesting finding is that immersion-like L2 experience seems to result in more L1-like (i.e., native-like) neurocognition than instructional experience. What do you make of this finding in the context of the DP Model?
5. Explain, in your own words, why one cannot equate declarative memory with explicit knowledge/learning and procedural memory with implicit knowledge/learning.
6. Read the exemplary study presented in this chapter and prepare a discussion for class in which you describe how you would conduct a replication study. Be sure to explain any changes you would make and what motivates such changes.

Note

1. I would like to thank Jarrett Lovelett, Harriet Bowden, Kara Morgan-Short, Kaitlyn Tagarelli, and Sarah Grey for help on this chapter.

Suggested Further Reading

Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences*, 14(5), 208–215.

Doyon, J., Bellec, P., Amiel, R., Penhune, V. B., Monchi, O., Carrier, J., . . . Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioural Brain Research*, 199, 61–75.

Both Doyon et al. and Ashby et al. provide overviews of aspects of procedural memory.

Squire, L. R., & Zola-Morgan, J. T. (2011). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, 34, 259–288.

An overview of long-term memory systems in the brain, focusing on declarative memory.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1–2), 231–270.

This paper gives an in-depth overview of the DP model and relevant evidence.

Ullman, M. T. (in press). How does language depend on general-purpose long-term memory systems in the brain? In G. Hickok & S. A. Small (Eds.), *The neurobiology of language*. New York, NY: Elsevier.

A recent overview of the DP model, its predictions, and relevant evidence.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, 128, 466–478.

- Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences*, *14*, 208–215.
- Babcock, L., Stowe, J. C., Maloof, C. J., Brovotto, C., & Ullman, M. T. (2012). The storage and composition of inflected forms in adult-learned second language: A study of the influence of length of residence, age of arrival, sex, and other factors. *Bilingualism: Language and Cognition*, *15*, 820–840.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*, 2767–2796.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Schacter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41–68). Cambridge, England: Cambridge University Press.
- Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*, 2492–2511.
- Breitenstein, C., Jansen, A., Deppe, M., Foerster, A.-F., Sommer, J., Wolbers, T., . . . Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage*, *25*, 958–968.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380.
- Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42.
- Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*, 564–570.
- Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21–43.
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London*, *364*, 3773–3800.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Malden, MA: Blackwell.
- Doyon, J., Bellec, P., Amsel, R., Penhune, V. B., Monchi, O., Carrier, J., . . . Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioural Brain Research*, *199*, 61–75.
- Eichenbaum, H. (2012). *The cognitive neuroscience of memory: An introduction* (2nd ed.). Oxford, England: Oxford University Press.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305–352.
- Green, D. W. (2003). Neural basis of the lexicon and the grammar in L2 acquisition. In R. V. Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The interface between syntax and the lexicon in second language acquisition* (pp. 197–208). Amsterdam, Netherlands: John Benjamins.
- Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, *11*, 523–532.
- Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*, 220–225.
- Hyltenstam, K., & Stroud, C. (1989). Bilingualism in Alzheimer's dementia: Two case studies. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss* (pp. 202–226). Cambridge, England: Cambridge University Press.

- Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? *Language Learning*, 56(Suppl. 1), 279–304.
- Johari, K., Ashrafi, F., Zali, A., Ashayeri, H., Fabbro, F., & Zanini, S. (2013). Grammatical deficits in bilingual Azari–Farsi patients with Parkinson’s disease. *Journal of Neurolinguistics*, 26, 22–30.
- Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. *Applied Psycholinguistics*, 33, 393–418.
- Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain & Language*, 109, 68–74.
- 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.
- Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, 48, 1138–1154.
- MacWhinney, B. (2011). The logic of the unified model. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 211–227). New York, NY: Routledge.
- Middleton, F. A., & Strick, P. L. (2000). Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition*, 42, 183–200.
- Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, 17, 56–72.
- Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS ONE*, 7, e32974.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, 60, 154–193.
- Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24, 933–947.
- Morgan-Short, K., & Ullman, M. T. (2011). The neurocognition of second language. In S. M. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 282–299). New York, NY: Routledge.
- Packard, M. G. (2008). Neurobiology of procedural learning in animals. In J. H. Byrne (Ed.), *Concise learning and memory: The editor’s selection* (pp. 341–356). London, England: Elsevier Science and Technology.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam, Netherlands: John Benjamins.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages* (Vol. 40). Amsterdam, Netherlands: John Benjamins.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15, 202–206.
- Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, 41, 245–251.
- Postle, B. R., & Corkin, S. (1998). Impaired word-stem completion priming but intact perceptual identification priming with novel words: Evidence from the amnesic patient H.M. *Neuropsychologia*, 15, 421–440.

- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255.
- Ruschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human Brain Mapping*, *25*, 266–286.
- Schmidt, R. W. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–209). London, England: Academic Press.
- Squire, L. R., & Zola-Morgan, J. T. (1991). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, *34*, 259–288.
- Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, *25*, 13–41.
- Stern, S. A., & Alberini, C. M. (2012). Mechanisms of memory enhancement. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *5*, 37–53.
- Ullman, M. T. (2001a). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*, 105–122.
- Ullman, M. T. (2001b). 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.
- Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141–178). Washington, DC: Georgetown University Press.
- Ullman, M. T. (2006a). The declarative/procedural model and the shallow-structure hypothesis. *Journal of Applied Psycholinguistics*, *27*, 97–105.
- Ullman, M. T. (2006b). Language and the brain. In J. Connor-Linton & R. W. Fasold (Eds.), *An introduction to language and linguistics* (pp. 235–274). Cambridge, England: Cambridge University Press.
- Ullman, M. T. (2007). The biocognition of the mental lexicon. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 267–286). Oxford, England: Oxford University Press.
- Ullman, M. T. (2008). The role of memory systems in disorders of language. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of the neuroscience of language* (pp. 189–198). Oxford, England: Elsevier.
- Ullman, M. T. (2012). The declarative/procedural model. In P. Robinson (Ed.), *Routledge encyclopedia of second language acquisition* (pp. 160–164). New York, NY: Routledge.
- Ullman, M. T. (in press). How does language depend on general-purpose long-term memory systems in the brain? In G. Hickok & S. A. Small (Eds.), *The neurobiology of language*. New York, NY: Elsevier.
- Ullman, M. T., Lum, J. A. G., & Conti-Ramsden, G. (in press). Domain specificity in language development. In P. Brooks, V. Kempe, & J. G. Golson (Eds.), *Encyclopedia of language development* (pp. 163–166). Los Angeles, CA: Sage.
- Zanini, S., Tavano, A., & Fabbro, F. (2010). Spontaneous language production in bilingual Parkinson's disease patients: Evidence of greater phonological, morphological and syntactic impairments in native language. *Brain & Language*, *113*, 84–89.

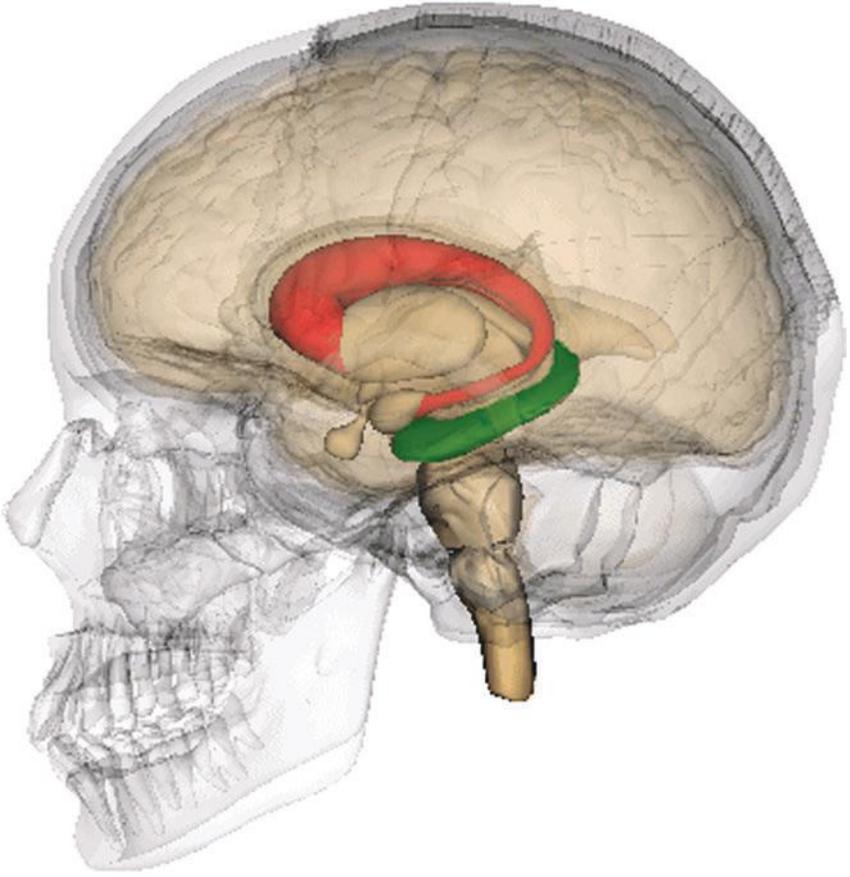


FIGURE 8.2 The caudate nucleus (red), part of the basal ganglia, and the hippocampus (green), in the medial temporal lobe.

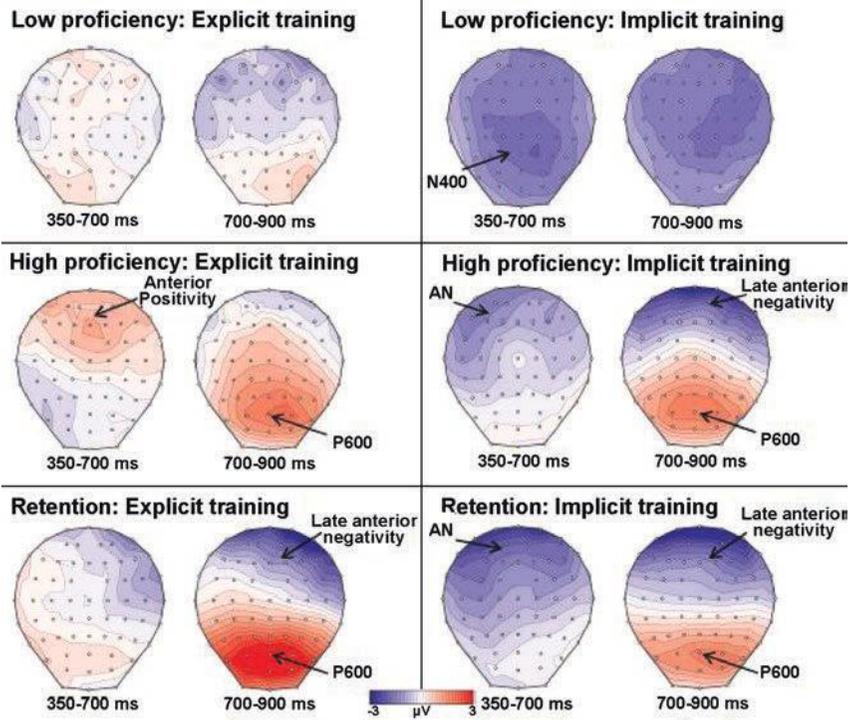


FIGURE 8.3 ERPs for word-order violations as compared to correct sentences, for subjects undergoing explicit (instructed) or implicit (uninstructed) training: at low proficiency, high proficiency, and 5 months later. Adapted from Morgan-Short, Finger, Grey, and Ullman (2012) and Morgan-Short, Steinhauer, Sanz, and Ullman (2012).