The neural basis of lexicon and grammar in first and second language: the declarative/procedural model*

MICHAEL T. ULLMAN
Georgetown University

Theoretical and empirical aspects of the neural bases of the mental lexicon and the mental grammar in first and second language (L1 and L2) are discussed. It is argued that in L1, the learning, representation, and processing of lexicon and grammar depend on two well-studied brain memory systems. According to the declarative/procedural model, lexical memory depends upon declarative memory, which is rooted in temporal lobe structures, and has been implicated in the learning and use of fact and event knowledge. Aspects of grammar are subserved by procedural memory, which is rooted in left frontal/basal-ganglia structures, and has been implicated in the acquisition and expression of motor and cognitive skills and habits. This view is supported by psycholinguistic and neurolinguistic evidence. In contrast, linguistic forms whose grammatical computation depends upon procedural memory in L1 are posited to be largely dependent upon declarative/lexical memory in L2. They may be either memorized or constructed by explicit rules learned in declarative memory. Thus in L2, such linguistic forms should be less dependent on procedural memory, and more dependent on declarative memory, than in L1. Moreover, this shift to declarative memory is expected to increase with increasing age of exposure to L2, and with less experience (practice) with the language, which is predicted to improve the learning of grammatical rules by procedural memory. A retrospective examination of lesion, neuroimaging, and electrophysiological studies investigating the neural bases of L2 is presented. It is argued that the data from these studies support the predictions of the declarative/procedural model.

In the study of language, a fundamental distinction is drawn between the memorized "mental lexicon" and the computational "mental grammar". The lexicon contains memorized words – that is, pairings of sound and meaning. It must contain at least those words whose phonological forms and meanings cannot be derived from each other (i.e., the sound–meaning pairings are arbitrary), such as the non-compositional word *cat*. It may also contain other non-compositional forms, smaller or larger than words: bound morphemes (e.g., the -*ed* past tense suffix) and idiomatic phrases (e.g., *kick the bucket*). The grammar contains rules, including operations and constraints, which underlie the productive sequential and hierarchical combination of lexical forms and abstract representations into complex structures, including complex abstract representations, words, phrases, and sentences. That is, the grammar subserves the computation of compositional linguistic forms whose meanings are transparently derivable from their structures. For example, a mental rule which specifies that English past tense forms are derived from the concatenation of a verb stem and an -*ed* suffix would allow us to productively compute past tenses from new words (e.g., *fax* + -*ed* → *faxed*) and from novel forms (e.g., *blick* + -*ed* → *blicked*) (Chomsky, 1995; Pinker, 1994). Rule-derived forms can thus be computed in real-time, and so do not need to be memorized – although even compositional forms (*walked*) could in principle be memorized.

Here I present a mental model of lexicon and grammar. The model addresses representational, computational, and neural aspects of the two language capacities. Although it was developed to explain native language (L1), this paper focuses on extending the model's predictions to non-native language (L2). The model's claims are evaluated in the context of existing empirical evidence. This paper focuses on the neural correlates of lexicon and grammar; therefore only neurolinguistic (no psycholinguistic) evidence will be discussed.

Address for correspondence
Michael Ullman, Department of Neuroscience, Research Building, Georgetown University, 3900 Reservoir Road, NW, Washington, DC 20007
Email: *michael@georgetown.edu*

The neural bases of lexicon and grammar in L1

The *declarative/procedural model* of lexicon and grammar posits that, for native speakers of a language, aspects of the lexicon/grammar distinction are tied to the distinction between two well-studied brain memory systems (Ullman, 2000a, 2001; Ullman et al., 1997), each of which has been implicated in a particular set of non-language functions (Mishkin, Malamut, and Bachevalier, 1984; Schacter and Tulving, 1994; Squire, Knowlton, and Musen, 1993).

One system is often referred to as the "declarative" memory system. It has been implicated in the learning, representation, and use of knowledge about facts ("semantic knowledge") and events ("episodic knowledge"). It may be particularly important for learning arbitrarily related information – that is, for the associative/contextual binding of information. It has been argued that the information learned by this system is not informationally encapsulated, being accessible to multiple response systems (Squire and Zola, 1996). Moreover, this information may be consciously ("explicitly") recollected. The memory system is subserved by medial temporal lobe regions (the hippocampus and related structures), which are connected largely with temporal and parietal neocortical regions (Suzuki and Amaral, 1994). Learning new information involves all parts of the system. However, the medial temporal components appear to be required to consolidate new memories. Memories eventually become independent of the medial temporal lobe structures, and dependent upon neocortical regions, particularly in the temporal lobes, possibly primarily (but not only) in the left hemisphere (see Hodges and Patterson, 1997; Schacter and Tulving, 1994; Squire et al., 1993; Squire and Zola, 1996).

The other system is usually referred to as the "procedural memory", "skill", or "habit" system. It has been implicated in the learning of new, and the control of long-established, motor and cognitive "skills" or "habits" (e.g., from simple motor acts to riding a bicycle to skilled game playing). Neither the learning nor the remembering of these procedures appears to be accessible to conscious memory – thus the system is often referred to as an "implicit memory" system (see De Renzi, 1989; Mishkin et al., 1984; Squire et al., 1993). It has been argued that the procedural system is largely "informationally encapsulated" (see Fodor, 1983), having relatively little access to other response systems (Squire and Zola, 1996). The system may be particularly important for learning and performing skills involving sequences (Graybiel, 1995; Willingham, 1998). The procedural system is rooted in frontal/basal-ganglia structures, with a possible role for inferior parietal regions (De Renzi, 1989; Squire et al., 1993). Evidence from impairments in expressing established motor skills (in ideomotor apraxia) suggests that the system may be especially dependent upon left hemisphere structures (De Renzi, 1989; Heilman, Watson, and Rothi, 1997). Inferior parietal structures may serve as a repository of stored knowledge of skills, including information on stored sequences (Heilman et al., 1997). The basal ganglia and the Supplementary Motor Area (SMA) may play a particularly important role in the processing of sequences (Kandel, Schwartz, and Jessell, 2000; Willingham, 1998). The basal ganglia also appear to play an important role in the learning of new skills and habits (Schacter and Tulving, 1994; Squire and Zola, 1996). The basal ganglia circuits are connected largely with frontal cortex (Alexander, Crutcher, and DeLong, 1990; De Renzi, 1989; Squire et al., 1993). These circuits appear to be parallel and functionally segregated, each receiving projections from a particular set of cortical and subcortical structures, and projecting via the thalamus to a particular cortical region, apparently largely in frontal cortex. Thus a basal ganglia "motor circuit" projects to frontal motor areas, while other circuits project to other frontal areas. The different basal ganglia circuits may have similar synaptic organizations, suggesting that similar neuronal operations might be performed at comparable stages of each circuit (Alexander et al., 1990). It has been independently hypothesized that the frontal lobes, to which the basal ganglia project, may also be organized in a similar manner, with distinct topographically organized regions playing the same or similar computational roles in different domains (Shimamura, 1995).

The declarative/procedural model posits that the declarative and procedural memory systems underlie the learning, representation, and use of aspects of lexical and grammatical knowledge, respectively. The model is motivated by a number of commonalities between the two memory systems and the two respective language capacities (see Ullman, 2000a; Ullman, 2001).

According to the declarative/procedural model, the declarative memory system subserves an associative memory that underlies stored knowledge about words, including their sounds, their meanings, and other memorized information. The consolidation of new words relies on medial temporal lobe structures. Eventually the knowledge of words becomes independent of these structures, and dependent on neocortex, particularly in temporal and temporo-parietal regions. Temporal lobe regions may be particularly important in the storage of word meanings, whereas

temporo-parietal regions may be more important in the storage of word sounds, including their phonological sequence information. Extrapolating from evidence from the study of declarative memory leads to the suggestion that lexical memory is not informationally encapsulated, being accessible to multiple response systems.

It is posited that the procedural memory system subserves the non-conscious (implicit) learning and use of aspects of a symbol-manipulating grammar, across grammatical sub-domains, including syntax, non-lexical semantics, morphology, and phonology. This system may be particularly important in the learning and computation of sequential and hierarchical structures (i.e., in grammatical structure building). Once learned, knowledge of sequences may depend upon left inferior parietal (that is, temporo-parietal) regions, which thus may serve as a locus of convergence between the declarative and procedural systems. The learning of rules is expected to be at least partially dependent upon basal ganglia structures. One or more particular basal ganglia circuits or sub-circuits, projecting to particular frontal region(s), may subserve grammatical processing, and perhaps even finer-grained distinctions, such as morphological (morpho-phonological) versus syntactic structure building. On this view, the frontal/basal-ganglia structures are domain-general in that they subserve non-linguistic as well as linguistic processes, but contain parallel domain-specific circuits.

The declarative/procedural model contrasts with two previously proposed theoretical frameworks, both of which have addressed the computational and neural bases of language.

Previously proposed (here termed "traditional") "dual-mechanism" theories posit distinct cognitive and neural components for the two capacities (Chomsky, 1995; Pinker, 1994). On this view, the learning, representation, and/or processing of words and associated information in a rote or associative memory is subserved by one or more components, which may be specialized and dedicated ("domain-specific") to these functions (Chomsky, 1995; Fodor, 1983; Forster, 1979; Levelt, 1989). It has been claimed that the use of stored words may be especially dependent upon left posterior regions, particularly temporal and temporo-parietal structures (Damasio and Damasio, 1992). In contrast, the learning, knowledge, and/or processing of grammar are posited to be subserved by one or more components that are specialized and dedicated to their linguistic functions, and whose computations depend upon innately specified constructs (Chomsky, 1995; Fodor, 1983; Pinker, 1994). The use of grammar has been claimed to be dependent on left frontal cortex,

particularly Broca's area (the inferior left frontal gyrus, which contains the cytoarchitectonic Brodmann's areas 44 and 45 (Damasio, 1992)) and adjacent anterior regions (Caramazza, Berndt, Basili, and Koller, 1981; Grodzinsky, 2000).

"Single-mechanism" theories posit that the learning and use of the words and rules of language depend upon a single computational system with broad anatomic distribution (Bates and MacWhinney, 1989; MacDonald, Pearlmutter, and Seidenberg, 1994). On this view, there is no categorical distinction between non-compositional and compositional forms. Rather, rules are only descriptive entities, and the language mechanism gradually learns the entire statistical structure of language, from the arbitrary mappings in non-compositional forms to the rule-like mappings of compositional forms. Modern connectionism has offered a computational framework for the single mechanism view. It has been argued that the learning, representation, and processing of grammatical rules as well as lexical items take place over a large number of interconnected simple processing units. Learning occurs by adjusting weights on connections on the basis of statistical contingencies in the environment (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, and Plunkett, 1996; Rumelhart and McClelland, 1986; Seidenberg, 1997).

Thus the declarative/procedural model of language differs from both traditional dual-mechanism theories and single-mechanism theories. Although the model shares the perspective of traditional dual-mechanism theories in positing that lexicon and grammar are subserved by distinct (separable) computational systems, with posterior and anterior neural correlates, respectively, it diverges from these theories where they assume components dedicated (domain-specific) to each of the two capacities. Conversely, while the model shares with single-mechanism theories the view that the two capacities are subserved by domain-general circuitry, it diverges from them where they link both capacities to a single mechanism with broad anatomic distribution.

The three perspectives make different theoretical claims with respect to four issues: separability, computation, domain-generality, and anatomical localization. The differing theoretical claims in turn lead to distinct predictions, allowing the theories to be distinguished empirically.

Separability. Both traditional dual-mechanism models and the declarative/procedural model posit separability – that lexicon and grammar are subserved by separable cognitive systems, with at least partially distinct neural correlates. Thus these two models predict double dissociations between the two

language capacities. In contrast, single mechanism models do not posit separate underpinnings for lexicon and grammar, and therefore do not predict double dissociations between the two capacities.[1]

Computation. The declarative/procedural model's assumptions that one component is an associative memory, and that the other underlies symbol-manipulation, is consistent with traditional dual-mechanism models – although these often adopt the distinct perspective that lexical memory is a rote list of words. Thus psychological or neural markers of memory (e.g., frequency effects), and in particular of associative memory (e.g., phonological neighborhood effects) (Prasada and Pinker, 1993; Prasada, Pinker, and Snyder, 1990; Ullman, 1993; Ullman, 1999), should be found with memorized lexical items, but not with linguistic forms posited to be computed by grammatical rules. In contrast, single-mechanism models predict associative memory effects for all linguistic forms.

Domain-Generality. According to the declarative/ procedural model, but to neither competing theoretical framework, lexicon and grammar are subserved by distinct systems, each of which underlies a specific set of non-language functions. Only the declarative/ procedural model predicts associations – in learning, representation, and processing – among lexicon, facts, and events, and among grammar, skills, and habits.

Localization. The declarative/procedural model makes specific claims about links between lexicon and particular temporal/temporo-parietal structures, and between grammar and left frontal/basal-ganglia structures, as a function of the roles of those neural structures in the two memory systems (see above). Traditional dual-mechanism models expect similar links, although the particular neuroanatomical claims made by the declarative/procedural model are not predicted by traditional dual-mechanism models. Single-mechanism models do not predict the same function–structure associations.

Thus the declarative/procedural model, but not traditional dual-mechanism or single-mechanism models, predict double dissociations, with (1) associations among associative memory markers, lexical items, facts and events, and temporal/temporo-parietal regions (including medial temporal lobe structures in the learning of new information); and (2) a distinct (dissociated) set of associations among grammar, motor and cognitive skills and habits (especially sequences), and frontal/basal-ganglia structures.

These predicted independent sets of associations (what might be termed "dissociations of associations") are supported by evidence from a number of different methodological approaches in both children and adults: psycholinguistic studies using various types of methodologies; behavioral studies of cognitively impaired populations with adult-onset or developmental disorders; and neuroimaging studies investigating the hemodynamic, electrophysiological, and magnetophysiological basis of lexicon and grammar. These have been discussed in some detail elsewhere (for a review, see Ullman, 2001), and therefore will not be presented here.

The neural bases of lexicon and grammar in L2

A large literature suggests that linguistic abilities are sensitive to the age of exposure to language. People who learn at later ages, particularly past late childhood or puberty, do not generally learn a language as well as younger learners (Birdsong, 1999; Johnson and Newport, 1989). This finding holds for exposure not only to a first language but also to second and subsequent languages. However, age of exposure does not equally affect all language capacities. Of interest here, in both first and second language acquisition, the use of grammar is affected much more adversely by later ages of exposure than is the use of lexical items that do not play an important grammatical role (Birdsong, 1999; Johnson and Newport, 1989).

It is posited here that this greater age-of-exposure sensitivity of grammar than of lexicon leads to a shift of reliance from procedural memory in L1 to declarative memory in L2. The declarative/procedural model claims that in L1, the learning and use of grammar depend largely upon procedural memory, whereas the memorization and use of words depends upon declarative memory. In L2, by contrast, age-of-exposure sensitivity that affects grammatical computation is posited to involve the procedural memory system – either the procedural memory system itself, such as the basal ganglia circuits or the frontal cortical regions they project to, or perhaps closely linked components. Intriguingly, evidence suggests that procedural memory, and basal ganglia-dependent learning in particular, may be subject to critical

---

[1] Although some single-mechanism models predict certain dissociations, these dissociations crucially do *not* separate grammatical and lexical forms or processes, but rather make other cuts, based on other factors (e.g., Joanisse and Seidenberg, 1999). Indeed, this is not surprising, given that they deny the existence of a distinct grammatical system. Moreover, there appear to be no existing connectionist simulations that have successfully modeled the double dissociations between grammatical and lexical processes that have been empirically demonstrated in humans (for a discussion, see Ullman et al., in press).

period effects in rodents (Fredriksson, 2000; Walton, Lieberman, Llinas, Begin, and Llinas, 1992; Wolansky, Cabrera, Ibarra, Mongiat, and Azcurra, 1999). In contrast, declarative memory function (that is, the learning of new memories) may actually *improve* with age during childhood (see Di Giulio, Seidenberg, O'Leary, and Raz, 1994) – although it begins to decline in early adulthood (Bjork and Bjork, 1996; Kirasic, Allen, Dobson, and Binder, 1996). It is proposed that, because grammatical computations relying on procedural memory become relatively difficult to learn, whereas the learning ability of declarative memory function remains relatively strong, later learners of language, particularly those exposed after late childhood or puberty, may differ in crucial ways from earlier learners. Whereas earlier learners rely largely on procedural memory for grammatical computations, later learners tend to shift to declarative memory for the same "grammatical" functions, which are moreover learned and processed differently than in the earlier learners. Thus the processing of linguistic forms that are computed grammatically by procedural memory in L1 is expected to be dependent to a greater extent upon declarative memory in L2.

This dependence is posited to take at least two forms. First, in the absence of grammatical rules, linguistic forms that are compositionally computed by grammar in L1 may simply be memorized in the lexicon, like words or idioms. Thus regular morphologically complex forms like *walked*, or even phrases and sentences – particularly those of high frequency, because higher levels of exposure should increase the likelihood of memorization – should be stored in and retrieved from declarative/lexical memory. To the extent that associative lexical memory can generalize patterns from memorized forms to new ones (Pinker, 1999; Prasada and Pinker, 1993), such lexicalization would be expected to lead to productivity in the language. The memorization of these forms does not imply that they are unanalyzed and unstructured strings. Indeed, the type of structured representations that are found in the lexicon in L1 speakers might be expected in L2 speakers as well. This would presumably encompass any structured lexical representations of morphologically complex forms (Pinker, 1999; Ullman, 2000a, 2001) as well as phrases and sentences. Moreover, to the extent that such structured representations are more abstract, containing (sub-) category specifications (rather than specifications for particular lexical items), one might expect an even greater degree of productivity to ensue.

Second, at least some rules may be learned in declarative memory. Crucially, these are not "grammatical" rules in that they do not depend at all upon procedural/grammatical computations. Indeed, what they specify could in principle differ radically from the implicitly learned grammatical rules of L1. They may be consciously (explicitly) learned (e.g., in a pedagogic environment), although they need not be, just as words or facts can be explicitly learned but need not be. Similarly, they may be consciously applied, but need not be, just as word or facts may be consciously accessed, but might also be used implicitly.

Both of these predicted declarative/lexical memory effects have been observed in L1 in Specific Language Impairment (SLI). SLI refers to a developmental disorder of language in the absence of other cognitive impairments (Leonard, 1998). It has been claimed that at least some groups of SLI subjects have grammatical deficits (Leonard, 1998). Two groups of SLI subjects with syntactic impairments were investigated for their use of regular and irregular past tense production. According to the declarative/procedural model, regulars are real-time procedural/grammatical (affixation) rule products, whereas irregulars are learned in and retrieved from associative declarative/ lexical memory (Ullman, 2001; Ullman et al., 1997).

Both groups of subjects had trouble generating novel regular past tense forms (e.g., *plam–plammed*), and produced few if any over-regularizations (e.g., *dig–digged*). These results suggest that they had difficulty computing -*ed*-affixation (Ullman and Gopnik, 1994, 1999; van der Lely and Ullman, 2001). Importantly, whereas healthy control subjects show frequency effects for irregular but not regular past tense forms (Prasada et al., 1990; Ullman, 1999), suggesting that the former but not the latter are retrieved from memory, both groups of SLI subjects showed frequency effects for regulars as well as irregulars (Ullman and Gopnik, 1994; Ullman and Gopnik, 1999; van der Lely and Ullman, 2001). These data have been taken to suggest that, due to a dysfunction of procedural memory, they have difficulty learning grammatical rules such as those underlying -*ed*-suffixation, but, because their lexical memory is relatively intact, they memorize regular as well as irregular forms (Ullman and Gopnik, 1999). This view is strengthened by the fact that members of one of the groups have been shown to have motor deficits, particularly of motor sequencing (Vargha-Khadem, Watkins, Alcock, Fletcher, and Passingham, 1995), suggesting a deficit of procedural memory. Moreover, this group is associated with frontal and basal ganglia abnormalities, especially on the left side, including Broca's area (Vargha-Khadem et al., 1998).

The group of SLI subjects with sequencing deficits and frontal/basal-ganglia abnormalities also showed

evidence of consciously learned and consciously used "explicit rules" (e.g., adding a /d/ sound to verbs when the action or event occurred "yesterday", whether or not the form was already past-marked: *swimmed, swammed*) (Goad and Rebellati, 1994; Ullman and Gopnik, 1994, 1999). It has been posited that these "explicit rules" are learned in declarative memory (Ullman and Gopnik, 1999). These findings from SLI suggest that a procedural memory dysfunction can lead to an increased dependence on declarative memory, which can occur in two ways. First, linguistic forms that are normally compositionally computed can simply be memorized. Second, SLI subjects can engage in the explicit learning and use of rule-like behavior.

A recent study of recovery in L1 aphasia further underscores the claim that a dysfunction of procedural memory can lead to the memorization of linguistic forms that are normally compositionally computed in L1 (Ullman, 2000b). Nine agrammatic non-fluent aphasics with frontal lesions and five anomic fluent aphasics with posterior lesions were asked to read aloud regular and irregular past-tense forms (see Ullman, Izvorski, Love, Yee, Swinney, and Hikock, in press). All subjects were native speakers of North American English. The posterior aphasics, like healthy subjects, showed past-tense frequency effects for irregulars, but not for regulars. In contrast, the anterior aphasics (like the SLI subjects described above) showed frequency effects for regulars as well as irregulars, suggesting both were memorized. Moreover, a positive correlation between the number of years since the occurrence of the lesion and frequency correlation r-values (that is, the measure of the strength of the frequency effect) was obtained only for the anterior aphasics and only with regulars, suggesting that the anterior aphasics memorized regular past-tense forms following lesion-onset. These results support the hypothesis that anterior aphasics can memorize regular past-tenses following lesion onset, whereas posterior aphasics are like unimpaired subjects in depending upon frontal structures to compute regulars compositionally. The data underscore the relative plasticity of the temporal lobe system and its role in the recovery of grammatical function.

In L2, this shift of dependence from procedural to declarative memory is expected to be a function of age of exposure. Later-exposed L2 learners should rely more on declarative memory, and less on procedural memory, than earlier exposed learners. However, age of exposure is not expected to be the only factor explaining the degree of relative dependence on the two memory systems. Because learning in procedural memory improves with practice

(Schacter and Tulving, 1994; Squire et al., 1993), the declarative/procedural model predicts that practice with L2 should increase the language's relative dependence on procedural memory for grammatical computations. Thus even older learners may show a degree of dependence on procedural memory if they have had a relatively large amount of practice – that is, a fairly substantial amount of use of the language. This is not to deny age-of-exposure effects. Rather, the claim is that practice as well as age of exposure should affect both grammatical proficiency and the degree of dependence on procedural memory for grammatical computations.

A number of empirical predictions fall out of the declarative/procedural model of L1 and L2.[2] L2 speakers should show a different pattern than L1 speakers with respect to the computational and neural bases of lexicon and grammar. The lexicon/grammar dissociations of associations posited for L1 (see above) should be weaker or perhaps even absent in L2. Moreover, the later the age of exposure to L2, and the less practice with it, the weaker the dissociations should be. Following are specific empirical predictions:

Separability. L2 speakers should show less separability than L1 speakers between the two types of linguistic forms – that is, those forms posited to rely on either lexicon or grammar in L1. Thus within L2 speakers, one would expect weaker lexicon/grammar double dissociations of the sort found in L1, or no

[2] The declarative/procedural model is quite similar to, but also differs from, the model proposed by Paradis (e.g., Paradis, 1994; Paradis, 1995a). Both models emphasize a shift from procedural to declarative memory between L1 and L2. Paradis discusses this shift largely in terms of greater automatization and implicitness of L1 than L2, across various domains of language, including the lexicon. In contrast, the declarative/procedural model proposed here posits a specific shift from procedural to declarative memory of grammatical but not lexical processes (which are posited to depend on declarative memory even in L1). Note however, that Paradis also acknowledges that there may be some shift, in particular for function words, from implicit grammar to a declarative-memory-based vocabulary (Paradis, 1998). However, unlike the declarative/procedural model proposed here, this vocabulary does not include grammatical aspects of the lexicon, such as sub-categorization information (Paradis, personal communication). The two views also diverge somewhat with respect to neuroanatomy. Whereas Paradis focuses on medial and sub-cortical structures, in particular the hippocampus (and related structures) and the basal ganglia, the present model includes neocortical structures in the implicated brain systems – namely frontal as well as basal-ganglia structures in the procedural/grammatical system, and neocortical temporal/temporo-parietal as well as medial temporal in declarative/lexical memory. In summary, the two models differ in certain respects, but also share many assumptions, and are perhaps best thought of as complementary rather than competing models (Paradis, personal communication).

such double dissociations at all. L1 can also be compared to L2 within subjects: separability of grammatical but not lexical processing is predicted between L1 and L2.

Computation. In L2 speakers, the computation of linguistic forms that are grammatically computed in L1 should be similar to that of forms that are expected to depend upon declarative/lexical memory in both L1 and L2. Both types of forms may be learned in and computed over associative memory, and thus both may show associative memory effects (e.g., frequency and phonological neighborhood effects). L1 grammatical computation should differ from that of not only L1 and L2 lexical processing, but also of L2 grammatical processing.

Domain-Generality. Within L2, both types of forms should tend to pattern with facts and events, and neither with skills and habits.

Localization. Both types of forms should be linked to temporal/temporo-parietal structures, primarily in the left hemisphere, but also in the right. Medial temporal lobe structures (the hippocampus and related structures) are expected to underlie the learning of both types of forms. Neocortical temporal/temporo-parietal regions are expected to underlie the use of already-learned forms of both types. Because the procedural/grammatical system is more left-lateralized than declarative/lexical memory, a general decrease in left-lateralization (and concomitant increase in right-lateralization) should be observed in L2 as compared to L1.

Evidence from aphasia

According to the declarative/procedural model, brain damage to left (and to a lesser extent, right) neocortical temporal/temporo-parietal regions should be associated with the following pattern of impairments in the processing of linguistic forms that are posited to depend upon grammatical computation in L1: greater difficulty with a later-learned and less-used L2 than with an L1 or an earlier-learned or well-practiced L2. In contrast, damage to left-hemisphere (but not right-hemisphere, in right-handers) frontal or basal ganglia structures should lead to greater grammatical impairments in L1 or an earlier-learned or well-practiced L2 than to a less-used or later-learned L2, whereas lexical performance should be similarly affected in all cases.

There is a very large literature examining aphasia in bilinguals (see Albert and Obler, 1978; Paradis, 1995b). Unfortunately, many studies of bilingual aphasia have not carefully identified damaged and undamaged brain structures, have not separately probed grammatical and lexical performance, or

have omitted details of language use, proficiency, or acquisition order. Thus it is difficult to test the predictions of the declarative/procedural model by examining this literature. Nevertheless, some general findings, as well as several specific ones, provide evidence which appears to support the model.

The bilingual aphasia literature indicates that, just as in L1, in L2 the left hemisphere is dominant for language (Paradis, 1995b). However, some evidence suggests that there may be greater right-hemisphere participation in the bilingual than the monolingual, consistent with the declarative/procedural model (see Albert and Obler, 1978; but for possible problems with this conclusion, see Zatorre, 1989).

Evidence also suggests that the temporal lobes may play a more important role in L2 than L1, as expected by the declarative/procedural model. One recent study examined the language impairment of a right-handed 16-year-old native Chinese speaker who had been living in the US since age 10 (Ku, Lachmann, and Nagler, 1996). Following the onset of herpes simplex encephalitis involving the left temporal lobe, he lost the ability to speak or comprehend English, but retained these abilities in Chinese – including syntactic abilities, although he used relatively short sentences. This study is appealing because herpes simplex encephalitis is associated with relatively circumscribed damage to temporal lobe structures, whereas most studies of aphasia examine patients who have suffered a stroke, which tends not to respect such anatomical boundaries. Moreover, both CT and MRI (as in the study being discussed here) produce extremely accurate images of the brain regions involved in herpes simplex encephalitis, whereas this is not the case with most other neuro-pathological processes (Brandt, Caplan, Dichgans, Diener, and Kennard, 1996; Damasio and Damasio, 1997). Thus this study suggests that left temporal lobe structures may be more important in L2 than L1. However, the pattern of better recovery of the native or more proficient language than the non-native or less proficient language is very common in bilingual aphasia (Paradis, 1995b). Therefore it would be important to investigate bilingual patients whose lesions were circumscribed to left frontal or basal ganglia structures, to demonstrate the opposite pattern, of worse grammatical performance at L1 or a proficient second language than a less proficient L2.

The data reported by Fabbro and Paradis (1995) reveals such a pattern. This study discusses the language impairments of four bilingual or polyglot aphasics whose lesions were entirely or largely circumscribed to the left basal ganglia. Careful examination of the data reported by the authors reveals that all four patients showed a pattern of worse

grammatical performance in their native or more proficient language than their non-native or less proficient language. In contrast, these patients did not show this pattern in their lexical abilities. The two patients whose lesion was circumscribed to basal ganglia gray matter (E.M. and C.B.) showed the most striking pattern. Their linguistic deficits are described in detail here.

Prior to lesion onset, patient E.M. was highly proficient in her native language (Venetan) and far less proficient in her second language (Italian). She spoke Venetan almost exclusively, both with her family and at work. Following onset, analyses of her spontaneous speech revealed that she made almost twice as many omission errors of grammatical function words (auxiliaries, obligatory pronouns, articles, prepositions, and conjunctions) in obligatory contexts in L1 (10%) than L2 (6%). In contrast, she made fewer omission errors in L1 than in L2 of full verbs in obligatory contexts (0% in L1 vs. 7% in L2), she produced a similar number of words per minute in the two languages (74 in L1 vs. 78 in L2), and displayed a comparable number of instances of word-finding difficulties (3.7 vs. 3.0 as a percentage of the total number words spoken).

Patient C.B. had learned Friulian as her first language, but also had five years of childhood schooling in Italian. At age 22 she moved to England and married an Englishman, where presumably English became her most frequently used language. After 39 years she moved back to Italy with her son. Ten years later she suffered an ischemic stroke. Following onset, analyses of her spontaneous speech showed that she made many more omission errors of grammatical function words in obligatory contexts in her native language (Friulian; 24%) and her long-practiced L2 (English; 18%) than in her non-proficient L2 (Italian; 8%). She also substituted inflectional morphemes at a higher rate in Friulian (5%) and English (9%) than Italian (2%). In contrast, she omitted verbs in obligatory contexts at about the same rate in Friulian and Italian (3% vs. 2%), although at a much higher rate in English (21%).[3] She produced *more* words per minute in Friulian (96) than in English (70) or Italian (79), and likewise displayed the fewest instances of word-finding difficulties in her native language (as a percentage of words spoken: Friulian 2.9, English 5.1, and Italian 4.3).

[3] Patient C.B.'s high rate of verb omission in English is consistent with her results on both other reported measures of lexical abilities. In all cases, she performed worst in English, which she was first exposed to at age 22. This performance pattern may be explained by a decrement in lexical/declarative learning ability in early adulthood (see text).

The two patients whose lesions extended from the left basal ganglia to the thalamus (patient El.M.) or to insular cortex (patient O.R) showed similar speech patterns to those of the two patients with more circumscribed basal ganglia lesions. Thus all four patients with left basal ganglia lesions omitted more grammatical function words in obligatory contexts in spontaneous speech in L1 or in a long-practiced second language than in a later-learned or little-used L2. Grammatical function words play an important role in syntactic computations (Chomsky, 1965). Moreover, their omission in agrammatic aphasia has been taken as an indication of a grammatical deficit (Goodglass, 1976; Grodzinsky, 2000). Thus it is reasonable to interpret the observed pattern of function word omissions as evidence of grammatical difficulties. It remains to be seen whether the particular pattern of function word differences between L1 and L2 is consistent with specific proposals that explain function word deficits in agrammatic aphasia in L1 in terms of grammatical impairments (e.g., Grodzinsky, 2000; Hagiwara, 1995). In particular, further investigation may reveal whether these L1/L2 function word distinctions can be explained by a recent proposal suggesting that aspects of agrammatic speech in L1 are the result of sequencing deficits which lead to grammatical structure building difficulties in both syntax and morphology (Izvorski and Ullman, 1999).

Although the four subjects discussed above are particularly interesting because of the detailed information available for them, other reports also support the predictions of the declarative/procedural model. Fabbro (1999) presents a historical review of six cases of bilingual aphasia, all of whom showed a greater impairment in their native language than non-native languages(s) following brain damage (see also Paradis, 1983). Whereas all were non-fluent (agrammatic) in their first language, they were able to speak or read in one or more non-native languages – ranging from simple words and stock phrases to apparently fluent speech. At least three, and likely all six, had left frontal damage. This evidence further supports the prediction that left frontal damage leads to greater grammatical impairments in L1 than L2.

Evidence from functional neuroimaging: PET and fMRI

The declarative/procedural model predicts greater temporal/temporo-parietal activation, primarily in the left hemisphere, but also in the right, in L2 than L1 for tasks involving grammatical processing, but no differences for tasks involving only lexical processing.

Dehaene et al. (1997) used fMRI to examine the neural correlates of receptive sentence processing in first and second language. They tested right-handed male subjects who were native speakers of French and non-native speakers of English. The subjects had not been exposed to English before the age of seven, and had not lived in an English speaking country for more than one year. In the study, they listened to French (L1) and English (L2) stories, and, as a control condition, to Japanese speech played backwards. In both L1 and L2, most activation was found in temporal lobe regions, although frontal regions were also activated. Compared to L2, L1 yielded areas of temporal lobe activation that were more consistent across subjects (largely restricted to the superior temporal sulcus and nearby areas), and, within subjects, were more left-lateralized and smaller. For example, some subjects showed temporal lobe activation only in the left hemisphere in L1, and only in the right in L2. Others showed left temporal lobe activation in both conditions, but greater anatomical dispersion in L2. Thus compared to L1, L2 is associated with greater anatomical variability and less left-lateralization, in temporal lobe regions. Given that the task required processing that is expected to be dependent upon grammar in L1, the decrease in left-lateralization in L2 is predicted by the declarative/procedural model. Similarly, as will be argued below, the greater temporal lobe dispersion may also be taken as evidence for greater reliance on temporal lobe structures in L2 than L1.

Intriguingly, in frontal regions L2 yielded not only greater dispersion than L1, but also more *left*-lateralization, with most of the activation in posterior dorsolateral and inferior frontal regions. At first blush, this does not seem to be consistent with the declarative/procedural model, as one might expect that the model would predict *less left* frontal activation in L2 than L1 in sentence processing tasks. After all, the model predicts less frontal-based grammatical computation for L2 than L1. However, those neuroimaging studies which have yielded left frontal activation in syntactic processing (Bookheimer, Zeffiro, Gaillaird, and Theodore, 1993; Embick, Marantz, Miyashita, O'Neil, and Sakai, 1999; Stromswold, Caplan, Alpert, and Rauch, 1996) carefully held lexical processing constant, whereas this was not done in Dehaene et al. (1997). A simple explanation for the observed pattern of greater left-lateralization in frontal regions for L2 than L1 is consistent with the declarative/procedural model. Although Dehaene et al. (1997) attribute the inferior frontal activation in L2 to working memory rehearsal, this region, as well as more dorsal areas, has been strongly implicated in the retrieval, selection, search, or encoding of lexical/

semantic information (e.g., Buckner and Peterson, 1996; Desmond, Gabrieli, and Glover, 1997; Wagner et al., 1998; Wise, Chollet, Hadar, Friston, and Hoffner, 1991). Such an interpretation follows from the declarative/procedural model, which expects a greater degree of lexical processing in L2 than in L1, due to the increased L2 reliance on lexical/declarative memory for the processing of forms that are largely computed by the grammatical/procedural system in L1.

Perani et al. (1998) carried out two PET studies in which they investigated the receptive sentence processing of highly proficient second language speakers. In the first experiment they tested right-handed Italian native speakers who learned English after the age ten, who spent one to six years in an English speaking country, and who currently used English in their daily activities. These "late bilingual" subjects were scanned during five conditions: listening to Italian, English, or Japanese stories, listening to backwards Japanese, or attentive silence. In the second experiment, they examined right-handed Spanish and Catalan bilinguals who were able to speak both languages equally well, but who were exposed only to one or the other language before the age of two, and to both languages thereafter. These "early bilingual" subjects were scanned while listening to Spanish and Catalan stories played forwards and backwards. In both experiments, L1 and L2 yielded highly similar cortical activation patterns. This activation pattern similarity between L1 and L2, in comparison to the considerable L1–L2 activation differences found in low-proficiency L2 speakers in the same (Perani et al., 1996) or similar (Dehaene et al., 1997) task paradigms, is consistent with the declarative/procedural model's prediction that when L2 is well practiced, even in bilinguals with later L2 exposure, it may be subserved by largely the same brain systems as L1.

Nevertheless, important activation pattern differences between L1 and L2 were reported by Perani et al. (1998). In both studies, L2 was associated with different temporal lobe patterns than L1. The late bilinguals showed a greater number of activated regions in the left temporal lobe, and greater right temporal activation in L2 than L1. The early bilinguals did not show this pattern of an increased number of activated left temporal lobe regions, but did show greater right temporal lobe activity in L2 than L1. Thus even highly practiced L2 learners display the greater temporal lobe dispersion found in low-proficiency L2 learners with fMRI (Dehaene et al., 1997), both within the left hemisphere, and extending to the right hemisphere. Moreover, the dispersion was less striking in the early than late

proficient bilinguals. These data, together with the overall similar activation patterns between L1 and L2 in both of the highly practiced subject groups, is consistent with the view that practice and age of exposure both affect the relative dependence on declarative and procedural memory. Indeed, like Dehaene et al. (1997), Perani et al. (1998) found somewhat greater activation for L2 than L1 in left frontal cortex, as would be predicted by a somewhat increased reliance on lexical processing in L2 (see above).

Thus Perani et al. (1998) reported greater dispersion, within and between the temporal lobes in later-than in earlier-learning proficient bilinguals. It is also striking that these later-learning proficient bilinguals' temporal lobe dispersion was in turn less apparent than in that of the low-proficiency bilinguals in the fMRI study (Dehaene et al., 1997), even though the age of L2 exposure of these low-proficiency bilinguals may have been no later than that of the proficient late bilinguals. Together, these results suggest that both age of acquisition and practice are independently predictive of the same type of neural pattern – that is, greater temporal lobe dispersion. However, the comparison of the PET and fMRI studies must be treated with caution, because in the former but not the latter data is averaged over subjects, which can result in reducing the apparent dispersion if it varies over subjects. Indeed, this may explain why the amount of temporal lobe dispersion was actually lower in the non-proficient L2 learners reported in Perani et al. (1996) than in the two proficient L2 learners reported in Perani et al. (1998): the dispersion in the proficient learners may have been low enough to yield activated areas common to several subjects, whereas the non-proficient learners may have had too much dispersion to be reflected in averaged data.

Other neuroimaging studies of L1 and L2 have focused on tasks expected to involve only single-word lexical processing, with no grammatical processing (Chee, Tan, and Thiel, 1999; Illes et al., 1999; Klein, Milner, Zatorre, Meyer, and Evans 1995; Klein, Milner, Zatorre, Zhao, and Nikelski, 1999; Klein, Zatorre, Milner, Meyer, and Evans, 1994; Klein, Zatorre et al., 1995). In all of these studies, no consistent differences in activation patterns were reported between L1 and L2. These data strengthen the declarative/procedural model: the lack of L2/L1 activation pattern differences in these single word processing tasks, in contrast to the striking L2/L1 differences in the sentence processing tasks that are expected to involve grammar (see above), suggest that the L2/L1 differences found in those tasks can be attributed to differences in how "grammar" is processed. In particular, the contrast suggests that the increased L2 temporal lobe dispersion and left frontal activation in the sentence processing tasks is unlikely to be explained by any L2/L1 differences involving single word processing, and therefore is more likely to be due to a greater reliance on declarative/lexical memory for linguistic forms that are grammatically computed in L1.

Evidence from electrophysiology: event-related potentials

The declarative/procedural model predicts that ERP components associated with grammatical processing and left frontal structures in L1 should be absent or modified in L2 speakers. In contrast, ERP effects associated with lexical–conceptual processing and temporal lobe structures should differ minimally if at all between L1 and L2. Finally, L2 speakers may show ERP components associated with lexical–conceptual processing for the processing of linguistic forms which depend upon grammatical computations in L1.

Weber-Fox and Neville (1996) reported an ERP study of syntactic and lexical–conceptual violations in bilingual subjects. The subjects, who were adult right-handed native speakers of Chinese, were categorized by their age of first exposure to English – at 1–3 years of age, 4–6 years, 7–10 years, 11–13 years, or greater than 16 years. It is important to note, however, that age of exposure was highly correlated with number of years of experience with English – although the 11–13 and >16 age-of-exposure groups did not differ in their years of experience of English. In comparisons among these groups and with monolingual native English speakers, later exposure to English was associated with worse behavioral performance on judging syntactically anomalous sentences and their correct control sentences, and on independent grammatical tests. For some measures this pattern even held for the 1–3 age-of-exposure group, as compared to monolinguals. In contrast, judgments of lexical–conceptual anomalies were relatively impervious to age of exposure: only the subject group with the latest exposure (>16 years) showed worse performance than monolinguals, and even these late-exposed subjects were not impaired at judging the lexical–conceptual control sentences.

A similar lexical/grammatical contrast was found in the ERP data. In all subject groups the lexical–conceptual violations yielded a typical "N400" pattern, an enhanced negativity peaking around 400 milliseconds after the onset of the anomalous word, with a posterior maximum and largely bilateral symmetry, but with a slightly larger effect on the right

hemisphere. The distribution and amplitude of the N400 were not affected by age of exposure, although there was a slight increase in the N400's latency (about 20 milliseconds) for subjects in the 11–13 and >16 age-of-exposure groups (also see Ardal, Donald, Meuter, Muldrew, and Luce, 1990; Kutas and Kluender, 1994). The N400 is a well-studied central/posterior negative ERP component which is associated with manipulations of word sounds and meanings (Hagoort and Kutas, 1995; Kutas and Hillyard, 1980), and has been linked to temporal lobe structures, primarily in the left hemisphere (Newman, Pancheva, Ozawa, Neville, and Ullman, 2001; Nobre, Allison, and McCarthy, 1994; Papanicolaou, Simos, and Basile, 1998). It has also been tied to non-lexical conceptual processing (Barrett and Rugg, 1990; Olivares, Bobes, Aubert, and Valdes-Sosa, 1994).

In contrast, syntactic violation ERP effects were highly sensitive to age of exposure. Later ages of L2 exposure were associated with changes in the actual presence as well as the distribution, amplitude, and latency of the ERP components that were elicited by syntactic violations in monolinguals. In monolinguals (see Neville, Nicol, Barss, Forster, and Garrett, 1991), syntactic phrase structure violations (e.g., The scientist criticized Max's *of* proof the theorem), in comparison to control sentences (e.g., The scientist criticized Max's proof *of* the theorem) lead to an early left anterior negativity, peaking around 125 ms post-stimulus. After a brief convergence of the two conditions (at the P200 component), a second left negativity is found. This negativity ranges from anterior to parietal sites, lasts from approximately 300 to 500 ms post-stimulus, and is similar to the left anterior negativities (LANs) found in many other studies of syntactic violations (for a brief review, see Friederici, Hahne, and Mecklinger, 1996). Finally, a late central-posterior bilaterally symmetric positivity is found, ranging from 500 to 700 ms, similar to the late positivity found in numerous studies of syntactic violations, often termed the P600 (for reviews, see Hagoort and Kutas, 1995; Osterhout, McLaughlin, and Bersick, 1997).

Subjects learning English as a second language – even those exposed to English as young as 1–3 – showed a different pattern to syntactic phrase structure violations. The very early left negativity was eliminated in early learners of L2 (1–3, 4–6, 7–10), and was larger in the right than left hemisphere in later learners (11–13, >16). In the 300–500 ms time window, the negativity showed increasing bilateral symmetry and augmented amplitude with increasing age of exposure. In fact, in the latest-exposed group (>16) this negativity was somewhat right-lateralized.

Finally, the late positivity was normal for early learners (1–3, 4–6, 7–10), showed increased latency for the 11–13 group, and was not observed for the >16 subjects.

These findings support the declarative/procedural model of first and second language. The behavioral and ERP measures of lexical–conceptual processing were relatively unaffected by age of exposure and years of experience (practice). In contrast, at least four lines of evidence suggest that grammatical measures were highly sensitive at least to age-of-exposure, and possibly also to experience. Moreover, in later learners, these measures reflected a dependence upon declarative/lexical memory.

First, later exposure led to substantially worse behavioral performance at grammatical tasks, suggesting abnormal grammatical processing in later and less-practiced learners. Second, the early anterior negativity, which may be the same early anterior negativity that has been taken to reflect automatic syntactic processing (Friederici et al., 1996; Hahne and Friederici, 1999), and which has been linked to left frontal structures (Friederici, Hahne, and von Cramon, 1998; Friederici, von Cramon, and Kotz, 1999), was absent or right-hemisphere shifted, indicating an abnormality of automatic syntactic processes that normally depend on left frontal regions.

Third, there was a striking similarity between the distribution and latency of the N400 and of the increasingly symmetrical (and even somewhat right-lateralized, in the >16 age group) 300 to 500 ms negativities found in subjects with later ages of exposure.[4] This suggests that these later learners may depend more on lexical–conceptual processes than native speakers do for the computation of syntactic structures, supporting the declarative/procedural model of L2. The model is further strengthened because the N400 has been tied both to non-linguistic conceptual/semantic processing and to temporal lobe structures (see above).

Fourth, the P600 was relatively preserved in the non-native speakers: it was unaffected for the three early-exposed age groups, and, although it showed an increased latency for the 11–13 age group (that is, in the 700–900 ms rather than 500–700 ms time window), its centro-parietal maximum here was typical of the P600 (Hagoort and Kutas, 1995). Finally, because the 11–13 age group showed an increased latency over the earlier-exposed age groups, and because there was no reported examination of

---

[4] This similarity must be treated with caution because the small number of electrodes used in the study leaves open the possibility that there were two bilateral negativities rather than one central negativity (Weber-Fox, personal communication).

possible waveforms for the time window past 900 ms, the >16 group may in fact have elicited an even later positivity which simply went unreported. Thus for all but the latest learners, the late positivity was observed, and even in the latest learners it may have been present.

This relative preservation of the late positivity, together with several lines of evidence which link the P600 to lexical memory, declarative memory, and temporal/temporo-parietal structures, converge to support the declarative/procedural model of L2. First, evidence suggests that the P600 may be sensitive to lexical knowledge (i.e., verb subcategorization information) (Osterhout, Holcomb, and Swinney, 1994). Second, evidence from brain-damaged patients suggests that the P600 may depend upon left temporal/temporo-parietal structures (Friederici et al., 1998). Third, evidence suggests that the P600 may reflect (consciously) controlled aspects of processing (Hahne and Friederici, 1997), which would be consistent with its dependence upon declarative memory – even in a native language. Fourth, it has been argued that the P600 is a member of the well-studied P300 family, whose components range from about 300 to 700 ms post-stimulus. In particular, the P600 may reflect – perhaps only partially; see Osterhout and Hagoort (1999) – neural events that also underlie the P3b component of the P300 family (see Friederici, Mecklinger, Spencer, Steinhauer, and Donchin, 2001). This view is supported by the fact that the P600 and P3b are similar in having a large positive amplitude and a centro-parietal distribution, and show similar task-related patterns (Coulson, King, and Kutas, 1998; Gunter, Stowe, and Mulder, 1997). A large amount of evidence implicates temporal and temporo-parietal structures, including hippocampal/medial-temporal-lobe regions, as neural generators of the P3b. Just as the declarative memory system has been identified in both humans and animals, the P3b is also found in animals as well as humans. Evidence suggests that the P3b may be involved in tasks which appear to probe declarative memory: P3b-like positivities have been observed to be greater for items that are subsequently remembered than for items that are not remembered. This "subsequent memory effect" is found with both semantic and non-semantic knowledge, but is greater with semantic knowledge. Finally, the positivity is larger when the remembered information can be explicitly recalled, in comparison to being produced in an implicit memory task (i.e., stem completion). (For a review of the P300, P3b, and memory-related ERPs, see Knight, 1997; Rugg, 1995.) Thus several lines of evidence converge to suggest that the P600 elicited by syntactic violations depends upon declara-

tive memory. Therefore the relative preservation of the P600 in second language learners supports the declarative/procedural model of second language.

Similar results have been obtained in two recent ERP studies of lexical–conceptual and syntactic phrase structure violations in second language learners of German. In the first study, Hahne (2000; submitted) tested adult native speakers of German as well as adult Russian native speakers. The Russian subjects learned German as a second language, with a mean of six years of instruction, and a mean of five years since they had been living in Germany. German (L1) yielded the expected ERP components: an N400 for lexical–conceptual violations, and an early LAN and a later P600 for syntactic violations. The second language was also associated with an N400 for lexical–conceptual violations and a P600 for syntactic violations, although, compared to L1, both components were delayed, and the N400 had a reduced amplitude. In contrast, no LAN was observed in L2. Unlike the study reported by Weber-Fox and Neville, no bilateral negativity was observed in the syntactic condition.

In the second study, Hahne and Friederici (in this issue) reported a similar ERP study, with similar stimuli. Subjects were native German speakers and native Japanese speakers who learned German as a second language. The Japanese speakers were all right handed, began to learn German at a mean age of 21, and had a mean of about two and a half years since they had started living in Germany and since they had begun formal instruction in German. Lexical–conceptual violations yielded an N400 for both L1 and L2. The N400 effect was similar in the two groups, although it lasted longer in the L2 group. Syntactic violations yielded an early LAN and a P600 for the L1 speakers, but neither component for the L2 speakers.[5] Visual inspection of the

---

[5] The lack of a P600 in this study of native Japanese speakers as compared to the study of native Russian speakers may be explained by either of two factors. First, it may be a function of differences in proficiency between the two groups, given that the Japanese had lower proficiency than the Russians. The Japanese had learned German for a mean of two and a half years, as compared to six years for the Russians. Moreover, the Japanese had much higher error rates than the Russians at judging sentences in the ERP study. With such lower proficiency, the Japanese may have simply given up, failing to (re-)analyze the incorrect German sentences, which could lead to the disappearance of the P600 (see Hahne and Friederici, 1999). Second, the lack of the P600 for the Japanese may be explained by the fact that the target words in their *correct* sentences yielded a P600–like positivity, with similar amplitude to the positivities generated by both the L1 and L2 speakers' syntactic violations. This L2 positivity for correct sentences is not surprising, given that evidence (from L1) suggests that the P600 may reflect syntactic integration difficulty (Kaan, Harris, Gibson, and

reported figures indicates that there was a trend for a central-right negativity (beginning after 500 ms) in the syntactic violation condition for the Japanese speakers. Although it did not reach significance, its distribution was similar to that of the negativity in the semantic violation condition. Its lack of significance may be attributed to a problem of small sample size: many fewer sentences were analyzed for the syntactic than the semantic condition, because only correct sentences were considered, and the error rates were almost twice as high for the syntactic condition (33%) than the semantic condition (18%).

In a brief abstract, Osterhout and McLaughlin (2000) report an ERP study of semantic and syntactic anomalies presented to adult native French speakers and to adult second language learners of French. In native speakers, the semantic anomalies yielded an N400 effect, whereas the syntactic anomalies produced a P600 effect (no anterior negativity was reported). In the French L2 learners, after only four weeks of instruction, semantic anomalies already elicited an N400 effect. Strikingly, the syntactic anomalies yielded either an N400 or no effect, across subjects. Thus L2 learners, but not L1 speakers, may show an N400 for syntactic anomalies, supporting the declarative/procedural model. After four months of instruction, semantic anomalies still yielded an N400, whereas the syntactic violations elicited P600 effects rather than an N400. This further supports the assertion that, unlike the LAN, which evaporates in later learners, the P600 remains robust.

Neville, Mills, and Lawson (1992) examined ERPs of both open-class (content) words and closed-class (function) words, which play important grammatical roles (see above). Native English speakers displayed a left anterior negativity approximately 280 ms post-stimulus for closed-class words (N280), and an N400 for open-class words. Congenitally deaf late learners of English did not differ from the native English speakers in the distribution, amplitude, or latency of the N400 for open-class words. In contrast, the deaf subjects failed to show the early left anterior negativity for closed-class words. Weber-Fox and Neville (submitted) also examined the processing of open- and closed-class English words, in native English speakers and in Chinese native speakers who learned English as a second language. The L1 speakers showed the same closed-class N280 and open-class N400 pattern as that reported by Neville, Mills, and Lawson (1992). For open-class words, the L2 learners

Holcomb, 2000), and presumably the L2 speakers had greater difficulty with syntactic integration than the L1 speakers. Thus the enhanced centro-parietal positivity for correct sentences, compared to an equal sized positivity for incorrect sentences, would fail to yield a difference between the two.

produced an N400 that did not differ in latency or distribution from the L1 speakers' N400. However, for closed-class words, the L2 learners' negativity had a longer peak latency than that of the native speakers. Thus in both studies of the processing of (meaning-laden) open-class and (grammatical) closed-class words (Neville et al., 1992; Weber-Fox and Neville, submitted), the N400 component yielded by open-class words did not differ between L1 and L2 speakers, whereas the N280 yielded by closed-class words in L1 speakers was either absent or delayed among L2 speakers. Moreover, Weber-Fox and Neville (submitted) reported that the latency of the closed-class word negativity was significantly correlated with an independent standard test of grammatical ability – the higher the score on the test, the earlier the negativity, linking both closed-class words and the anterior negativity to grammatical processing.

Summary and conclusion

It has been argued that two well-studied brain memory systems underlie the learning and use of language. In the native language (L1), the mental lexicon and the mental grammar are posited to each rely on one of the two memory systems. The memorization, storage, and processing of the stored sound–meaning pairings of lexical memory are subserved by declarative memory, a brain system rooted in temporal lobe structures, and implicated in the learning and use of knowledge about facts and events. In contrast, the learning, representation, and processing of aspects of grammar depend largely upon procedural memory, a distinct brain system rooted in left frontal/basal-ganglia structures, and implicated in the learning and expression of motor and cognitive skills and habits. This theoretical perspective, referred to as the declarative/procedural model, is supported by a number of lines of psycholinguistic and neurolinguistic evidence.

In later-acquired second and subsequent languages (L2), especially those learned after late childhood or puberty, it is argued that the two brain memory systems tend to play a somewhat different role. Later exposure to language may impair the ability of the procedural memory system to learn or compute aspects of grammar. Instead, linguistic forms that are computed grammatically in procedural memory in L1 may depend largely on declarative/lexical memory in L2. This dependence is posited to occur in at least two ways. First, linguistic forms that are compositionally computed in L1 (e.g., *walk* + *-ed*) may be memorized in their entirety in L2. Productivity may emerge from the ability of the associative lexical

memory to generalize patterns to new forms, and from the learning of abstract structured representations whose frames specify word (sub-)categories. Second, L2 speakers may learn explicit rules in declarative memory, and use those rules to construct linguistic forms. Evidence for such a shift from a dysfunctional procedural/grammatical system to declarative/lexical memory has independently been shown in L1 in Specific Language Impairment and in agrammatic anterior (non-fluent) aphasia. Importantly, the shift is not absolute. Younger L2 learners are expected to depend more on procedural memory and less on declarative memory than older learners. Moreover, the strong practice effects of procedural memory learning lead to the prediction that, in addition to age of exposure, an increasing amount of experience (i.e., practice) with a language should lead to better learning of grammatical rules in procedural memory, which in turn should result in higher proficiency in the language.

In a retrospective examination of studies pertaining to the neural basis of L2, it was argued that this view was largely borne out. Studies of aphasics, PET and fMRI studies of brain activation patterns, and electrophysiological investigations examining event-related potentials (ERPs), all converged to support the declarative/procedural model. The findings can be summarized as follows.

Aphasia. Temporal lobe damage can lead to worse "grammatical" performance at L2 than L1 – that is, worse performance in the processing of linguistic forms which are posited to depend upon grammatical computation in L1. In contrast, left basal ganglia or left frontal lobe damage can lead to worse performance at such "grammatical" processing in L1 than L2, as well as in a more practiced than less practiced L2, while these L1/L2 differences are not found in lexical processing.

Neuroimaging (PET and fMRI). Tasks involving "grammatical" processing show greater temporal lobe involvement, in both the left and right hemispheres, in L2 than in L1, as evidenced by a larger number of areas of activation in these regions. The extent of temporal lobe involvement appears to be greater for later than earlier L2 learners, and possibly in L2 speakers with less practice than those with more practice. In addition, these tasks yield greater left frontal activation in L2 than L1 in brain regions which have independently been implicated in lexical/conceptual retrieval and encoding. In contrast, tasks involving only lexical processing do not yield different activation patterns in L1 and L2.

Electrophysiology (ERPs). In L1, tasks hypothesized to involve grammatical computation yield left anterior negativities (LANs), which have been linked to automatic processing and left frontal structures. LANs are modified or even absent in L2 speakers. In contrast, the P600 component, which is linked to controlled late syntactic processing in L1, as well as temporal lobe regions and declarative memory, is generally displayed by L2 speakers. Similarly, the N400 component, which is associated with lexical–conceptual processing and temporal lobe structures, is present in both L1 and L2. Finally, in grammatical tasks which yield a LAN in L1 speakers, L2 speakers may instead show N400-like negativities.

The evidence suggests the following. Aspects of grammatical processing are less dependent upon left frontal and basal ganglia structures in L2 than in L1. Linguistic forms that are largely grammatically/procedurally computed in L1 are more dependent in L2 than in L1 on declarative/lexical memory – although in L1 as well as in L2, some processing of complex representations may depend on declarative memory, as evidenced by the P600. The shift from procedural to declarative memory is reflected by an increased involvement of left (and, to a lesser extent, right) temporal/temporo-parietal structures. This shift of dependence from the procedural/grammatical system to declarative/lexical memory appears to be a function both of age of exposure and of practice.

The declarative/procedural model may have important implications. In particular, the numerous studies of the two memory systems in animals as well as humans are expected to elucidate the computational and neural bases of the learning, representation, and processing of words and rules in both first and subsequent languages.

References

Albert, M. L. & Obler, L. K. (1978). *The bilingual brain: neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.

Alexander, G. E., Crutcher, M. D. & DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor oculomotor "prefrontal" and "limbic" functions. In H. B. M. Uylings, C. G. Van Eden, J. P. C. DeBruin, M. A. Corner, & M. G. P. Feenstra (eds.), *Progress on brain research*, vol. 85, pp. 119–146. New York: Elsevier Science.

Ardal, S., Donald, M. W., Meuter, R., Muldrew, S. & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, 39 (2), 187–205.

Barrett, S. E. & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, 14 (2), 201–212.

Bates, E. & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (eds.), *The crosslinguistic study of sentence processing*, pp. 3–73. Cambridge: Cambridge University Press.

Birdsong, D. (1999) (ed.). *Second language acquisition and the critical period hypothesis*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bjork, E. L. & Bjork, R. A. (1996) (eds.). *Memory* (2nd ed.). New York: Academic Press.

Bookheimer, S. Y., Zeffiro, T. A., Gaillaird, W. & Theodore, W. (1993). Regional cerebral blood flow changes during the comprehension of syntactically varying sentences. *Society for Neuroscience Abstracts*, 19 (1), 843.

Brandt, T., Caplan, L. R., Dichgans, J., Diener, H. C. & Kennard, C. (1996) (eds.). *Neurological disorders: course and treatment*. New York: Academic Press.

Buckner, R. L. & Peterson, S. E. (1996). What does neuroimaging tell us about the role of prefrontal cortex in memory retrieval? *Seminars in the Neurosciences*, 8, 47–55.

Caramazza, A., Berndt, R. S., Basili, A. G. & Koller, J. J. (1981). Syntactic processing deficits in aphasia. *Cortex*, 17, 333–348.

Chee, M. W., Tan, E. W. & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience*, 19 (8), 3050–3056.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Coulson, S., King, J. W. & Kutas, M. (1998). Expect the unexpected: event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13 (1), 21–58.

Damasio, A. R. (1992). Aphasia. *New England Journal of Medicine*, 326 (8), 531–539.

Damasio, A. R. & Damasio, H. (1992). Brain and language. *Scientific American*, 267 (3), 88–95.

Damasio, H. & Damasio, A. R. (1997). The lesion method in behavioral neurology and neuropsychology. In T. E. Feinberg & M. J. Farah (eds.), *Behavioral neurology and neuropsychology*, pp. 69–82. New York: McGraw-Hill.

De Renzi, E. (1989). Apraxia. In F. Boller & J. Grafman (eds.), *Handbook of neuropsychology*, vol. II, pp. 245–263. New York: Elsevier Science.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P. F., Lehericy, S. & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, 8 (17), 3809–3815.

Desmond, J. E., Gabrieli, J. D. E. & Glover, G. H. (1997). Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and search. *Neuroimage*, 7 (4), 368–376.

Di Giulio, D. V., Seidenberg, M., O'Leary, D. S. & Raz, N. (1994). Procedural and declarative memory: a developmental study. *Brain and Cognition*, 25 (1), 79–91.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996). *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W. &

Sakai, K. L. (1999). An fMRI study of ungrammaticality in an error-detection task. *Society for Neuroscience Abstracts*, 25 (2), 1813.

Fabbro, F. (1999). *The neurolinguistics of bilingualism*. Hove, East Sussex: Psychology.

Fabbro, F. & Paradis, M. (1995). Differential impairments in four multilingual patients with subcortical lesions. In M. Paradis (ed.), *Aspects of bilingual aphasia*, vol. III, pp. 139–176. Oxford: Pergamon.

Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: MIT Press.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (eds.), *Sentence processing: psycholinguistic studies presented to Merrill Garrett*, pp. 27–85. Hillsdale, NJ: Lawrence Erlbaum.

Fredriksson, A. (2000). Maze learning and motor activity deficits in adult mice induced by iron exposure during a critical postnatal period. *Developmental Brain Research*, 119 (1), 65–74.

Friederici, A. D., Hahne, A. & Mecklinger, A. (1996). The temporal structure of syntactic parsing: early and late effects elicited by syntactic anomalies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (5), 1219–1248.

Friederici, A. D., Hahne, A. & von Cramon, D. Y. (1998). First-pass versus second-pass parsing processes in a Wernicke's and a Broca's aphasic: electrophysiological evidence for a double dissociation. *Brain and Language*, 62 (3), 311–341.

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K. & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: a spatiotemporal analysis of event-related brain potentials. *Cognitive Brain Research*, 11 (2), 305–323.

Friederici, A. D., von Cramon, D. Y. & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain*, 122 (6), 1033–1047.

Goad, H. & Rebellati, C. (1994). Pluralization in familial language impairment: affixation or compounding?, *The McGill working papers in linguistics: linguistic aspects of familial language impairment*, vol. X, pp. 24–40. Montreal: McGill.

Goodglass, H. (1976). Agrammatism. In H. Whitaker & H. A. Whitaker (eds.), *Studies in neurolinguistics*, vol. I, pp. 237–260. New York: Academic Press.

Graybiel, A. M. (1995). Building action repertoires: memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology*, 5, 733–741.

Grodzinsky, Y. (2000). The neurology of syntax: language use without Broca's area. *Behavioral and Brain Sciences*, 23 (1), 1–71.

Gunter, T. C., Stowe, L. A. & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, 34 (6), 660–676.

Hagiwara, H. (1995). The breakdown of functional categories and the economy of derivation. *Brain and Language*, 50 (1), 92–116.

Hagoort, P. & Kutas, M. (1995). Electrophysiological in-

sights into language deficits. In F. Boller & J. Grafman (eds.), *Handbook of neuropsychology*, vol. X, pp. 105–134. Amsterdam: Elsevier.

Hahne, A. (2000). Auditory sentence processing in foreign language learners as revealed by ERPs. Paper presented at the 13th Annual CUNY Conference on Human Sentence Processing, San Diego.

Hahne, A. (submitted). What's different in second language processing? Evidence from event-related brain potentials.

Hahne, A. & Friederici, A. D. (1997). Two stages in parsing: early automatic and late controlled processes. *Experimental Brian Research*, 117, 47.

Hahne, A. & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11 (2), 194–205.

Heilman, K. M., Watson, R. T. & Rothi, L. G. (1997). Disorders of skilled movements: limb apraxia. In T. E. Feinberg & M. J. Farah (eds.), *Behavioral neurology and neuropsychology*, pp. 227–235. New York: McGraw-Hill.

Hodges, J. R. & Patterson, K. (1997). Semantic memory disorders. *Trends in Cognitive Sciences*, 1 (2), 68–72.

Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D., Glover, G. H., Poldrack, R., Lee, C. J. & Wagner, A. D. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and Language*, 70 (3), 347–363.

Izvorski, R. & Ullman, M. T. (1999). Verb inflection and the hierarchy of functional categories in agrammatic anterior aphasia. *Brain and Language*, 69 (3), 288–291.

Joanisse, M. F. & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: a connectionist model. *Proceedings of the National Academy of Sciences, USA*, 96 (13), 7592–7597.

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21 (1), 60–99.

Kaan, E., Harris, A., Gibson, E. & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15 (2), 159–201.

Kandel, E. R., Schwartz, J. H. & Jessell, T. M. (2000) (eds.). *Principles of neural science*. New York: McGraw-Hill.

Kirasic, K. C., Allen, G. L., Dobson, S. H. & Binder, K. S. (1996). Aging, cognitive resources, and declarative learning. *Psychology and Aging*, 11 (4), 658–670.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E. & Evans, A. C. (1995). The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Sciences, USA*, 92 (7), 2899–2903.

Klein, D., Milner, B., Zatorre, R. J., Zhao, V. & Nikelski, J. (1999). Cerebral organization in bilinguals: a PET study of Chinese–English verb generation. *Neuroreport*, 10 (13), 2841–2846.

Klein, D., Zatorre, R. J., Milner, B., Meyer, E. & Evans,

A. C. (1994). Left putaminal activation when speaking a second language: evidence from PET. *Neuroreport*, 5 (17), 2295–2297.

Klein, D., Zatorre, R. J., Milner, B., Meyer, E. & Evans, A. C. (1995). The neural substrates of bilingual language processing: evidence from postitron emission tomography. In M. Paradis (ed.), *Aspects of Bilingual Aphasia*, pp. 23–36. Oxford: Pergamon.

Knight, R. T. (1997). Electrophysiological methods in behavioral neurology and neuropsychology. In T. E. Feinberg & M. J. Farah (eds.), *Behavioral neurology and neuropsychology*, pp. 101–120. New York: McGraw-Hill.

Ku, A., Lachmann, E. A. & Nagler, W. (1996). Selective language aphasia from herpes simplex encephalitis. *Pediatric Neurology*, 15 (2), 169–171.

Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, 207 (1), 203–205.

Kutas, M. & Kluender, R. (1994). What is who violating? A reconsideration of linguistic violations in light of event-related brain potentials. In H.-J. Heinze, T. F. Münte & G. R. Mangun (eds.), *Cognitive electrophysiology: basic and clinical applications*, pp. 183–210. Boston: Birkhäuser.

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.

MacDonald, M. C., Pearlmutter, N. J. & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101 (4), 676–703.

Mishkin, M., Malamut, B. & Bachevalier, J. (1984). Memories and habits: two neural systems. In G. Lynch, J. L. McGaugh, & N. W. Weinburger (eds.), *Neurobiology of learning and memory*, pp. 65–77. New York: Guilford.

Neville, H. J., Mills, D. L. & Lawson, D. S. (1992). Fractionating language: different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2 (3), 244–258.

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I. & Garrett, M. F. (1991). Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3 (2), 151–165.

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J. & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguistic Research*, 30 (2).

Nobre, A. C., Allison, T. & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, 372, 260–263.

Olivares, E., Bobes, M. A., Aubert, E. & Valdes-Sosa, M. (1994). Associative ERP effects with memories of artificial faces. *Brain Research Cognitive Brain Research*, 2 (1), 39–48.

Osterhout, L. & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: counterarguments to Coulson, King, and

Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*, 14 (1), 1–14.

Osterhout, L., Holcomb, P. J. & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (4), 786–803.

Osterhout, L. & McLaughlin, J. (2000). What brain activity can tell us about second-language learning. Paper presented at the 13th Annual CUNY Conference on Human Sentence Processing, San Diego.

Osterhout, L., McLaughlin, J. & Bersick, M. (1997). Event-related brain potentials and human language. *Trends in Cognitive Science*, 1 (6), 203–209.

Papanicolaou, A. C., Simos, P. G. & Basile, L. F. H. (1998). Applications of magnetoencephalography to neurolinguistic research. In B. Stemmer & H. A. Whitaker (eds.), *Handbook of neurolinguistics*, pp. 143–158. San Diego, CA: Academic Press.

Paradis, M. (1983) (ed.). *Readings on aphasia in bilinguals and polyglots*. Quebec: Marcel Didier.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: implications for bilingualism and SLA. In N. C. Ellis (ed.), *Implicit and explicit learning of languages*, pp. 393–419. London: Academic Press.

Paradis, M. (1995). Introduction: the need for distinctions. In M. Paradis (ed.), *Aspects of bilingual aphasia*, vol. III, pp. 1–9. Oxford: Pergamon.

Paradis, M. (1995b) (ed.). *Aspects of bilingual aphasia*. Oxford: Pergamon.

Paradis, M. (1998). Language and communication in multilinguals. In B. Stemmer & H. A. Whitaker (eds.), *Handbook of neurolinguistics*, pp. 417–430. San Diego, CA: Academic Press.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F. & Mehler, J. (1996). Brain processing of native and foreign languages. *Neuroreport*, 7 (15–17), 2439–2444.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F. & Mehler, J. (1998). The bilingual brain: proficiency and age of acquisition of the second language. *Brain*, 121 (10), 1841–1852.

Pinker, S. (1994). *The language instinct*. New York: William Morrow.

Pinker, S. (1999). *Words and rules: the ingredients of language*. New York: Basic.

Prasada, S. & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8 (1), 1–56.

Prasada, S., Pinker, S. & Snyder, W. (1990). *Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated*. Paper presented at the 31st Annual Meeting of the Psychonomics Society, New Orleans.

Rugg, M. D. (1995). ERP studies of memory. In M. D. Rugg & M. G. H. Coles (eds.), *Electrophysiology of mind*, pp. 132–170. Oxford: Oxford University Press.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & PDP Research Group (eds.), *Parallel distributed processing: explorations in the microstructures of cognition*, vol. II, pp. 216–271. Cambridge, MA: Bradford/MIT Press.

Schacter, D. L. & Tulving, E. (1994) (eds.). *Memory systems 1994*. Cambridge, MA: MIT Press.

Seidenberg, M. S. (1997). Language acquisition and use: learning and applying probabilistic constraints. *Science*, 275, 1599–1603.

Shimamura, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (ed.), *The cognitive neurosciences*, pp. 803–813. Cambridge, MA: MIT Press.

Squire, L. R., Knowlton, B. & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology*, 44, 453–495.

Squire, L. R. & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences, USA*, 93, 13515–13522.

Stromswold, K., Caplan, D., Alpert, N. & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language*, 52, 452–473.

Suzuki, W. A. & Amaral, D. G. (1994). Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferants. *Journal of Comparative Neurology*, 350 (4), 497–533.

Ullman, M. T. (1993). *The computation of inflectional morphology*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Ullman, M. T. (1999). Acceptability ratings of regular and irregular past tense forms: evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes*, 14 (1), 47–67.

Ullman, M. T. (2000a). A mental model of morphology: the psychological and neural bases of the representation and computation of complex words. In K. K. Grohmann & C. Struijke (eds.), *Special Issue: Proceedings of the Maryland Mayfest on Morphology 1999*, vol. X, pp. 127–156. College Park, MD: University of Maryland, Linguistics Department.

Ullman, M. T. (2000b). Neural plasticity and morphological processing in agrammatic aphasia: implications for recovery. *Journal of Cognitive Neuroscience* (Supplement), 53.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30 (1), 37–69.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J. & Pinker, S. (1997). A Neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9 (2), 266–276.

Ullman, M. T. & Gopnik, M. (1994). The production of inflectional morphology in hereditary specific language impairment. In J. Matthews (ed.), *The McGill working papers in linguistics: linguistic aspects of familial*

language impairment, vol. X, pp. 81–118. Montreal: McGill.

Ullman, M. T. & Gopnik, M. (1999). Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics*, 20 (1), 51–117.

Ullman, M. T., Izvorski, R., Love, T., Yee, E., Swinney, D. & Hickok, G. (in press). Neural correlates of lexicon and grammar: evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language*.

van der Lely, H. K. J. & Ullman, M. T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, 16 (2), 177–217.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences, USA*, 92 (3), 930–933.

Vargha-Khadem, F., Watkins, K. E., Price, C. J., Ashburner, J., Alcock, K. J., Connelly, A., Frackowiak, R. S., Friston, K. J., Pembrey, M. E., Mishkin, M., Gadian, D. G. & Passingham, R. E. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences, USA*, 95 (21), 12695–12700.

Wagner, A., Schacter, D., Rotte, M., Koutstaal, W., Maril, A., Dale, A., Rosen, B. & Buckner, R. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281 (5380), 1188–1191.

Walton, K. D., Lieberman, D., Llinas, A., Begin, M. & Llinas, R. R. (1992). Identification of a critical period for motor development in neonatal rats. *Neuroscience*, 51 (4), 763–767.

Weber-Fox, C. & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–256.

Weber-Fox, C. & Neville, H. J. (submitted). Sensitive periods differentiate processing subsystems for open and closed class words: an ERP study in bilinguals.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review*, 105 (3), 558–584.

Wise, R., Chollet, F., Hadar, U., Friston, K. & Hoffner, E. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain* (Pt 4), 114, 1803–1817.

Wolansky, M. J., Cabrera, R. J., Ibarra, G. R., Mongiat, L. & Azcurra, J. M. (1999). Exogenous NGF alters a critical motor period in rat striatum. *Neuroreport*, 10 (13), 2705–2709.

Zatorre, R. J. (1989). On the representation of multiple languages in the brain: old problems and new directions. *Brain and Language*, 36 (1), 127–147.