13
The logic of the unified model

Brian MacWhinney

Historical discussion

Many people believe that learning a second language (L2) is fundamentally different from learning a first language (L1). Evidence of this fundamental difference comes from the fact that first language acquisition almost invariably produces full native speaker competence, whereas many second language learners achieve only partial success in learning their new language. Some researchers believe that this difference in levels of ultimate attainment result arises because, after the expiration of a certain critical period, the learning mechanisms that subserve first language learning atrophy or expire.

The Unified Competition Model (UCM) (MacWhinney, 2008b) takes a different approach to this issue. Instead of attributing differences between first and second language learning to the effects of a critical period, these differences are attributed to the differential interplay between risk-generating processes and protective, support processes. For L1 learning, the five risk factors are entrenchment, parasitism, misconnection, negative transfer, and isolation. To overcome these five risk factors, adults can rely on the support processes of resonance, internalization, chunking, positive transfer, and participation. All of these risk factors and support processes are available to children, as well as adults. What differs between L1 and L2 learning is the way in which these processes are configured.

There are three obvious differences between L1 and L2 learners. First, while infants are learning language, they are also engaged in learning about how the world works. In contrast, L2 learners already have a full understanding of the world and human society. Second, infants are able to rely on a highly malleable brain that has not yet been committed to other tasks (MacWhinney *et al.*, 2000). In contrast, second language learners have to deal with a brain that has already been committed to the task of processing the first language. Third, infants can rely on an intense system of social support from their caregivers (Snow, 1999). In contrast, L2 learners are often heavily involved in L1 social and business commitments that distract them from L2 interactions.

Together, these three differences might suggest that it would make little sense to try to develop a unified model of first and second language acquisition. In fact, many researchers have decided that the two processes are so different that they account for them with totally separate theories with separate processes. For example, Krashen (1994) sees L1 learning as involving "acquisition" and L2 learning as based instead on "learning." Clahsen and Muysken (1986) hold that Universal Grammar (UG) is available to children up to some critical age, but not to older learners of L2. Paradis (2004) and Ullman (2004) argue that children learn language implicitly as a proceduralized,

automated skill, whereas adults learn language through explicit declarative controlled processes. Using analyses and arguments such as these, Bley-Vroman (1989, 2009) articulated the fundamental difference hypothesis (FDH), which holds that first and second language acquisition are so fundamentally different that trying to explain them through a single unified theory would make no sense. In this vein, Chomsky (Searchinger, 1995) remarks that learning to speak a second language is as unnatural as learning to ride a unicycle.

Despite these analyses, there are good reasons to question the FDH. The many parallels between L1 and L2 learning are more striking than the differences. Both groups of learners need to segment speech into words. Both groups need to learn the meanings of these words. Both groups need to figure out the patterns that govern word combination in syntactic constructions. Both groups have to interleave their growing lexical and syntactic systems to achieve fluency. Both groups are trying to learn the same target language. Thus, both the overall goal and the specific subgoals involved in reaching that goal are the same for both L1 and L2 learners. Both groups are enmeshed in social situations that require a continual back and forth of communication and learning. Furthermore, both groups of learners rely on the same underlying neuronal hardware.

One could recognize these parallels, but still emphasize the idea that the remaining differences are fundamental. The question is whether those remaining differences are great enough to motivate two separate theories for learning and processing. The thesis of the UCM is that the inclusion of L1 and L2 learning in a single unified model produces a more coherent and insightful analysis. The fact that L2 learning is so heavily influenced by transfer from L1 means that it would be impossible to construct a model of L2 learning that did not take into account the structure of the first language. Unless the two types of learning and processing share virtually no important commonalities, it is conceptually simpler to formulate a unified model within which the specific areas of divergence can be clearly distinguished from the numerous commonalities.

Core issues

The UCM is an extension of the Competition Model of the 1980s (Bates and MacWhinney, 1982; MacWhinney, 1987). The original model was designed to account for the end state of first and second language learning, but not the details of the learning process. The classic Competition Model dealt effectively with processes of transfer and the growth of cue strength. However, it had three major gaps. First, it was unable to provide insights into the ways in which proceduralization of language processes can lead to increases in fluency and the avoidance of fossilization. Second, it failed to incorporate information from our continually growing understanding of the neuroscience of language. Third, it provided no central role to social processes in L2 acquisition. In order to explain the workings of the newer version of the model, it is best to begin with a review of the core concepts of competition, cues, cue strength, and cue validity from the original Competition Model. These concepts remain unchanged in the new model, and the empirical basis of these concepts remains as it has been described in the literature. However, the processes of the classic version of the model are now supplemented by formulations of additional cognitive, neural, and social mechanisms. These additional mechanisms provide additional determination of cue strength, which is the major explanatory principle in the model.

Competition

Competition is a fundamental construct in many psychological theories. Freud viewed the ego as mediating a competition between the impulses of the Id and the restrictions of the superego.

Modern cognitive theories view competition as arising whenever two cues for a given decision point in opposite directions. When this occurs, the strength of the resultant decision is a function of the competing strengths of the input cues. A classic example is the Stroop effect (Stroop, 1935) in which the name for a word competes with the color in which that word is written. Theories based on competition have been articulated in areas as diverse as visual perception (Brunswik, 1956), reading (Seidenberg and McClelland, 1989), social psychology (Kelley, 1967), cognitive development (Anderson, 1981), infant attachment (van Geert, 1991), motor control (Carlson *et al.*, 1989), and auditory perception (Massaro, 1987).

Cues

Traditionally, the Competition Model has focused on the ways in which cues compete for thematic role assignment in sentences with transitive verbs. For example, in the sentence *the boys chase the ball*, the two nouns (*boys* and *ball*) are possible candidates for the role of the agent or subject of the verb. However, the candidacy of *the boys* for this role is favored by three strong cues—preverbal positioning, subject–verb agreement, and animacy. None of these cues favors the candidacy of *ball*. Therefore, native speakers uniformly conclude that *the boys* are the agents. However, in certain ungrammatical sentences, the competition between the noun phrases can become tighter. The ungrammatical sentence *the ball are chasing the boys* illustrates this effect. In this sentence, the strong cue of preverbal positioning favors *the ball* as agent. However, the cues of subject–verb agreement and animacy favor *the boys* as the agents. Given a competition sentence of this type, listeners are often quite unsure which of the two noun phrases to choose as agent, since neither choice is perfect. As a result, listeners, as a group, are slower to make this choice, and their choices are nearly evenly split between the two possibilities.

Competition Model experiments use sentences in which cues have been randomly combined to measure the strength of the underlying cues. The same method has been used in 45 empirical studies involving 18 different languages. Across these various experiments and languages, the cues involved come from a very small set of linguistic devices. Languages mark case roles using basically five possible cue types: word order, case marking, agreement, intonation, and verb-based expectations. For simple transitive sentences with two nouns and a verb, the possible word orders are NNV, NVN, and VNN. In addition, the marking of the cases or thematic roles of nouns can rely on affixes (as in Hungarian or Turkish), postpositions (as in Japanese), prepositions (as in Spanish), or articles (as in German). Agreement marking displays correspondences between the subject and the verb (as in English) or the object and the verb (as in Hungarian and Arabic). Some of the features that can be marked through agreement include number (as in English), definiteness (as in Hungarian), gender (as in Arabic), honorific status (as in Japanese), and other grammatical features. Intonation is seldom a powerful cue in thematic role identification, although we have found that it plays a role in some non-canonical word order patterns in Italian and in the topic marking construction in Hungarian. Verb-based expectations vary markedly across verb types. High activity transitive verbs like *push* and *hit* tend to serve as cues for animate agents and inanimate patients. Stimulus-experiencer verbs like *amaze* and *surprise* cue animate patients and either animate or inanimate agents.

Competition Model experiments put these various cues into systematic conflict with one another using orthogonalized analysis of variance designs. The extent to which cues dominate or control the choices of agent nouns in these experiments is the measure of their *cue strength*. The core claim of both the classic and revised versions of the Competition Model is that cue strength is determined by cue validity. Cue strength is defined through experimental results; cue validity is defined through corpus counts. Using conversational input data such as those available from the

CHILDES (http://childes.psy.cmu.edu) or TalkBank (http://talkbank.org) corpora, we can define *cue reliability* as the proportion of times the cue is correct over the total number of occurrences of the cue (Ellis, Chapter 12, this volume). Cue *availability* is the proportion of times the cue is available over the times it is needed. The product of cue reliability and cue availability is overall *cue validity*.

Early in both L1 and L2 learning, cue strength is heavily determined by availability, because beginning learners are only familiar with cues that are moderately frequent in the language input (Matessa and Anderson, 2000; Taraban and Palacios, 1993). As learning progresses, cue reliability becomes more important than cue availability. In adult native speakers, cue strength depends entirely on cue reliability. In some cases, we can further distinguish the effects of *conflict reliability*. When two highly reliable cues conflict, we say that the one that wins is higher in conflict reliability. For example, in the case of Dutch pronouns, only after age 8 do L1 learners begin to realize that the more reliable cue of pronoun case should dominate over the more frequent, but usually reliable, cue of word order (McDonald, 1986).

When adult native speakers have sufficient time to make a careful decision, cue strength is correlated at levels above 0.90 with cue reliability. However, when cue strength is measured online during the actual process of comprehension, before the sentence is complete, other factors come into play. During online processing, listeners tend to rely initially on a single cue with good reliability and high availability without integrating the effects of that core cue with other possible cues. This happens, for example, during online processing of sentences in Russian (Kempe and MacWhinney, 1999). Cue strength is also heavily influenced during the early phases of learning by the factors of *cue cost* and *cue detectability*. Cue cost factors arise primarily during the processing of agreement markers, because these markers cannot be used to assign thematic roles directly. For example, in an Italian sentence such as *il gatto spingono i cani* (the cat push the dogs), the listener may begin by thinking that *il gatto* is the agent because it occurs in preverbal position. However, because the verb *spingono* requires a plural subject, it triggers a search for a plural noun. The first noun cannot satisfy this requirement and the processor must then hope that a plural noun will eventually follow. In this example, the plural noun comes right away, but in many cases it may come much later in the sentence. This additional waiting and matching requires far more processing than that involved with simple word order or case marking cues. As a result of this additional cost for the agreement cue, Italian children are slow to pick it up, despite its high reliability in the language (Bates *et al.*, 1982).

Cue detectability factors (VanPatten, Chapter 16, this volume) play a major role only during the earliest stages of learning of declensional and conjugational patterns. For example, although the marking of the accusative case by a suffix on the noun is a fully reliable cue in both Hungarian and Turkish, Three-year-old Hungarian children show a delay of about ten months in acquiring this cue, when compared to young Turkish children. The source of this delay seems to be the greater complexity of the Hungarian declensional pattern and the weaker detectability of the Hungarian suffix. However, once Hungarian children have "cracked the code" of accusative marking, they rely nearly exclusively on this cue. Because of its greater reliability, the strength of the Hungarian case-marking cue eventually comes to surpass the strength of the Turkish cue.

Although Competition Model experiments have focused on the issue of thematic role assignment in simple transitive sentences, the principle of competition is a very general one that can be elaborated into a full model of language processing (MacDonald *et al.*, 1994; MacWhinney, 1987). For example, in a sentence, such as *the women discussed the dogs on the beach*, there is a competition between the attachment of the prepositional phrase *on the beach* to the verb or the noun *the dogs*. In this case, the competition can be resolved either way. However, in a sentence, such as *the communist farmers hated died*, the competition between the adjectival and nominal reading of

communist is initially resolved in favor of the adjectival reading, because of the presence of the following noun *farmers* and then the verb *hated*. However, once the second verb is encountered, the listener realizes that the adjectival reading has taken them down a garden path. At that point, the weaker nominal reading of *communist* is given additional strength and the alternative reading is eventually obtained.

The view of language processing as fundamentally competitive has three important consequences for a variety of issues in second language learning. One consequence, which has already been discussed, is that second language learning is viewed as a data-driven process in which the forces of cue validity, detectability, and reliability play a major role. A second consequence is the learner's ability to recover from errors and overgeneralizations can be directly related to variations in the strengths of competing constructions, thereby resolving the core issue in the Logical Problem of Acquisition (MacWhinney, 2004). Finally, the fact that cue strength can be influenced by additional inputs from other factors allows us to extend the model, while still maintaining a focus on detailed studies of cue usage during language processing.

Inputs to competition

The unified version of the Competition Model (MacWhinney, 2008b) extends the basic model by providing characterizations of additional neurocognitive, developmental, and social (Bayley and Tarone, Chapter 3, this volume) forces that control the core competition. As MacWhinney (2005a) notes, these forces operate on very different time scales, varying from seconds to years. However, in the end, these forces have their effect at the moment of speaking by imparting strength to particular cues and by affecting the timing of the interaction between cues. Some of these forces operate to restrict the smooth acquisition of second languages. We can refer to these as "risk factors." Other forces serve to promote both first and second language learning. We can refer to these as support factors. Table 13.1 presents these factors in terms of these two dimensions.

Entrenchment and resonance

Entrenchment is a basic neurodevelopmental process. At birth, the cerebral cortex of the human infant is uncommitted to specific linguistic patterns. However, across the first years, neural territory becomes increasingly committed to the patterns of the first language. These processes of commitment and entrenchment can be modeled using self-organizing maps (SOMs) (Kohonen, 2001), a computational formalism that reflects many of the basic facts of neural structure. Simulations of lexical learning from real input to children (Hernandez *et al.*, 2005) have shown how the organization of lexical fields into parts of speech becomes increasingly inflexible across learning. If the structure of the second language is extremely close to that of the

Table 13.1 Risk factors and support factors for second language learning

Risk Factors	Support Factors
Entrenchment	Resonance
Misconnection	Proceduralization
Parasitism	Internalization
Negative transfer	Positive Transfer
Isolation	Participation

Table 13.2 Levels of linguistic processing

Map	Area	Processes	Theory
Audition	Auditory cortex	Extracting units	Statistical learning
Articulation	IFG, motor cortex	Targets, timing	Resonance, gating
Lexicon	Wernicke's area	Phonology to meaning	DevLex
Syntax	Inferior Frontal Gyrus	Slots, sequences	Item-based patterns
Gram. Roles	DLPFC	Role binding, lists	Attachment, roles
Mental Models	Dorsal cortex	Deixis, perspective	Perspective

first language, then this commitment to lexical structure will cause few errors. However, if the languages are very different, the entrenchment of grammatical categories in the lexicon can lead to problems.

The detailed operation of entrenchment has been modeled most explicitly for lexical (Li *et al.*, 2007) and auditory (Guenther and Gjaja, 1996) structure. However, the UCM holds that cortical maps exist for each of the structural levels recognized by traditional linguistics, including syntax (Pulvermüller, 2003), grammatical roles (Jackendoff, 1983), and mental models (MacWhinney, 2008a), as given in Table 13.2. Research has shown that entrenchment in cortical maps presents the greatest risk factor for second language learners in the areas of auditory phonology (Kuhl *et al.*, 2005), articulatory phonology (Major, 1987), and the interactions of syntax with the lexicon (DeKeyser, 2000).

Resonance

The risk factor of entrenchment can be counteracted by the support factor of resonance. Resonance provides new encoding dimensions to reconfigure old neuronal territory, permitting the successful encoding of L2 patterns. Because this encoding operates against the underlying forces of entrenchment, special configurations are needed to support resonance. Resonance can be illustrated most easily in the domain of lexical learning. Since the days of Ebbinghaus (1885), we have understood that the learning of the associations between words requires repeated practice. However, a single repetition of a new vocabulary pair such as *mesa–table* is not enough to guarantee robust learning. Instead, it is important that initial exposure be followed by additional test repetitions timed to provide correct retrieval before forgetting prevents efficient resonance from occurring (Pavlik and Anderson, 2005). Because robustness accumulates with practice, later retrieval trials can be spaced farther and farther apart. This is the principle of "graduated interval recall" that was formulated for second language learning by Pimsleur (1967).

The success of graduated interval recall can be attributed, in part, to its use of resonant neural connections between cortical areas. While two cortical areas are coactive, the hippocampus can store their relation long enough to create an initial memory consolidation. Repeated access to this trace (Wittenberg *et al.*, 2002) can further consolidate the memory. Once initial consolidation has been achieved, maintenance only requires occasional reactivation of the relevant retrieval path-way. This type of resonance can be used to consolidate new forms on the phonological, lexical (Gupta and MacWhinney, 1997), and construction levels.

The success of graduated interval recall also depends on correctly diagnosing the point at which a new memory trace is still available, albeit slightly weakened. At this point, when a learner attempts to remember a new word, sound, or phrase, some additional work will be needed to

generate a retrieval cue. This retrieval cue then establishes a resonance with the form being retrieved. This resonant cue may involve lexical analysis, onomatopoeia, imagery, physical responses, or some other relational pattern. Because there is no fixed set of resonant connections (Ellis and Beaton, 1995), we cannot use group data to demonstrate the use of specific connections in lexical learning. However, we do know that felicitous mnemonics provided by the experimenter (Atkinson, 1975) can greatly facilitate learning.

Orthography provides a major support for resonance in L2 learning. When a learner of German encounters the word *Wasser*, it is easy to map the sounds of the word directly to the image of the letters. Because German has highly regular mappings from orthography to pronun-ciation, calling up the image of the spelling of *Wasser* is an extremely good way of activating its sound. When the L2 learner is illiterate, or when the L2 orthography is unlike the L1 orthography, this backup orthographic system is not available to support resonance. L2 learning of Chinese by speakers of languages with Roman scripts illustrates this problem. In some signs and books in mainland China, Chinese characters are accompanied by romanized pinyin spellings. This allows the L2 learner a method for establishing resonant connections between new words, their pronunciation, and their representations in Chinese orthography. However, in Taiwan and Hong Kong, characters are seldom written out in pinyin in either books or public notices. As a result, learners cannot develop resonant connections from these materials. In order to make use of resonant connections from orthography, learners must focus on the learning of Chinese script. This learning itself requires constructing other resonant associations, because the Chinese writing system is based heavily on radical elements that have multiple potential resonant associations with the sounds and meanings of words.

Connection and misconnection

The negative effects of entrenchment on L2 learning are amplified by the fact that the major language processing areas are connected across white matter tracts. Unlike the digital computer, the brain has no system for assigning absolute and constant addresses to individual neurons. Consider a sentence, such as *the black dog chased the white dog*. When processing the two instances of *dog* in this sentence, the same lexical item is responding. But the brain needs to keep the first mention of the noun distinct from the second, so it can know that the one who did the chasing was the black dog and not the white dog. This means that the head noun and all of its associated arguments have to be bound together to operate as a unit. Solving this "binding problem" is a fundamental challenge for neuronal computation. Within individual cortical maps, local regional self-organization can solve a part of the binding problem, because items that behave similarly tend to cluster near each other. Differences between these items can be resolved by short, local connections that are relatively inexpensive in metabolic terms (Buzsaki, 2006). However, the connections between separate cortical maps must rely on long-distance neuronal projections that are more expensive metabolically and more difficult to repair, if they are broken. As a result, these connecting pathways are relatively less plastic and more committed to L1 functions than the areas within individual maps. For the language areas, important connecting pathways include the *arcuate fasciculus*, the *superior longitudinalis*, and the several subsegments of these major pathways (Friederici, 2009).

When activation is passed along these long-distance connections, there must be some method for the receiving units to process the identity of the sending units. The brain can address this problem by applying a method of parallel *topological organization* that has been widely documented for sensory and motor systems. Within the lexical map, words that share a similar part of speech and a similar meaning, such as *cut*, *chop*, *slice*, and *hack*, are located topologically near each other.

When items in this area connect to other processing regions, topological organization can help the receiving area identify the general shape of the sending units.

In aphasia, lesions to white matter tracts can result in various forms of language loss, and disorganization in the formation of these tracts during early development can lead to Specific Language Impairment. The task of reorganizing communication across these connections presents a major challenge to L2 learners. If L1 and L2 use roughly similar systems for part-of-speech assignment, communication between L2 lexical items and L2 syntactic processes will be relatively smooth. However, if L2 is typologically quite different from L1, it will be difficult for a learner to acquire the new mapping and there will be a persistent tendency for L2 learners to rely on L1 pathways for composing sentences.

Proceduralization and chunking

Second language learners can address these connectivity problems by relying on the support factors of proceduralization and chunking. Proceduralization (Anderson, 1993) is a cognitive process that transfers newly learned material into a smoothly operating procedure which then requires minimal attentional control. Proceduralization is closely related to the process of chunking (Rosenbloom and Newell, 1987) that takes a series of separate elements and welds them into a single processing unit or chunk.

Chunks function as single, unanalyzed wholes, whereas procedures may have some room for flexible variation. For example, in Spanish, L2 learners can learn *muy buenos días* "very good morning" as a chunk. This chunk is based on a series of connections between preexisting lexical items, stored within the lexical map in the posterior cortical areas in the temporal lobe. However, this pattern could also be learned as a flexible procedure triggered by the word *muy* "very" that would allow other completions such as *muy buenas tardes* "good afternoon" or *muy buenas noticias* "very good news."

Second language learners often fail to pick up sufficiently large chunks, seeking instead to analyze the input into small easily managed segments. For example, learners of German often learn the word *Mann* "man" in isolation. If, instead, they would learn phrases such as *der alte Mann, meines Mannnes, den jungen Männern*, and *ein guter Mann*, they would have a good basis for acquiring the declensional paradigm for both the noun and its modifiers. If learners were to store larger chunks of this type, then the rules of grammar could emerge from analogic processing of the chunks stored in feature maps (Bybee and Hopper, 2001; Ellis, 2002; MacWhinney, 1982; Tomasello, 2003). However, if learners analyze a phrase like *der alte Mann* into the literal string "the + old + man" and throw away all of the details of the inflections on "der" and "alte," then they will lose an opportunity to induce the grammar from implicit generalization across stored chunks.

Chunking focuses on storage in posterior lexical areas, whereas proceduralization relies on storage in frontal areas for sequence control (Broca's) that then point to lexical items in posterior areas. Proceduralization is initially less robust then chunking, but it is capable of greater extensibility and flexibility (Gobet, 2005) across constructions beyond the level of the item-based construction. For example, a Spanish phrase such as *quisiera comprar* … (I would like to buy …) can be used with any manner of noun to talk about things you would like to buy. In each of these cases, producing one initial combination, such as *quisiera comprar una cerveza* (I would like to buy a beer) may be halting at first. However, soon the result of the creation process itself can be stored as a chunk. In this case, it is not the actual phrase that is chunked, but rather the process of activating the predicate combination (*quisiera comprar*) and then going ahead and filling the argument. In other words, we develop fluency by repeated practice in making combinations.

Once learners have developed fluency in the combination of well-learned words, they can still experience disfluency when trying to integrate newly learned words into established constructions. For example, even if we have learned to use the frame *quisiera comprar* fluently with words such as *una cerveza* (a beer) or *un reloj* (a clock), we may still experience difficulties when we need to talk about buying "a round trip ticket to Salamanca" (*un billete de ida y vuelta para Salamanca*). In this selection, we might have particular problems when we hit the word "para" since the English concept of "for, to" can be expressed in Spanish using either por or para and our uncertainty regarding the choice between these two forms can slow us down and cause disfluency or error. In general, for both L1 and L2 learners, disfluencies arise from delays in lexical access, misordering of constituents, and selection of agreement markings. Fluency arises through the practicing of argument filling and improvements in the speed of lexical access and the selections between competitors.

Researchers such as Paradis (2004) or Ullman (2004) believe that L2 learners cannot effectively proceduralize their second language; and, as a result, L2 productions must remain forever slow and non-fluent. We can refer to this position as the Proceduralization Deficit Hypothesis (PDH). This hypothesis is a specific articulation of the general Critical Period Hypothesis (CPH) about which so much has been written in recent years (DeKeyser and Larson-Hall, 2005). Surveying this vast literature is beyond the scope of this brief review. However, we can point to a couple of recent findings that bear specifically on the PDH. Initial work by Hahne and Friederici (2001) indicated that, even after five or more years learning German, native Russian and Japanese speakers failed to show rapid early left anterior negativity (ELAN) responses to grammaticality violations in German sentences. These results suggested that, after the end of the critical period, comprehension could not be automated or proceduralized. However, further studies using artificial language systems (Friederici *et al.*, 2002; Müller *et al*, 2005) have shown that, if the rules of the target language are simple and consistent, L2 learners can develop proceduralization, as measured by ELAN, with a couple of months of training. Thus, it appears that proceduralization can be successful in adult learners, as long as cues are consistent, simple, and reliable (MacWhinney, 1997; Tokowicz and MacWhinney, 2005). This finding is in accord with the UCM analysis, rather than the PDH analysis, since it shows that the crucial factor here is not the age of the learner, but the shape of the input.

It is important not to confuse proceduralization with implicit learning. Although first language learning relies primarily on implicit learning, second language learning involves a complex interaction of both explicit and implicit learning (VanPatten, Chapter 16, this volume). In formal contexts such as classrooms, a second language may be learned through explicit methods. However, this knowledge can then become proceduralized and automatized, producing good fluency. A simple example of this process comes from a study by Presson and MacWhinney (under review) based on use of a computerized tutorial system for teaching the gender of French nouns. In this experiment, if naïve learners who have never studied French are given simple cues to gender, they are able to achieve 90 percent accurate gender assignment after only 90 minutes of computerized practice. Moreover, this ability is retained across three months without any further training.

In a review of the role of explicit rule presentation, MacWhinney (1997) argued that L2 learners can benefit from explicit cue instruction, as long as the cues are presented simply and clearly. Once a simple pattern has been established in explicit declarative form, repeated exposures to a cue can use the scaffolding of the explicit pattern to establish proceduralization. As in the case of lexical learning, the method of graduated interval recall can further support proceduralization. In addition, error correction can help to tune cue weights (McDonald and Heilenman, 1991). Of course, proceduralization can be achieved without scaffolding from explicit instruction. However, if explicit scaffolding is available, learning will be faster.

Positive and negative transfer

Entrenchment and connectivity have important consequences for L2 learning, because new forms must be entered into maps that are already heavily committed to L1 patterns. One way of solving this problem is by aligning L2 forms with analogous L1 forms. When the forms align well, mapping an L1 form to L2 will result in *positive transfer*. However, when there are mismatches, then the alignment produces at least some *negative transfer*. In the terms of the overall analysis of risk and support factors, negative transfer functions as a risk factor and positive transfer as a support factor.

The UCM holds that L2 learners will attempt to transfer any pattern for which there is some perceptual or functional match between L1 and L2. The match need not be exact or complete, as long as it is close enough. It is often easy to transfer the basic pragmatic functions that help structure conversations and the construction of mental models (Bardovi-Harlig, Chapter 9, this volume). The transfer of lexical meaning from L1 to L2 is also largely positive, although there will be some mismatches in meaning (Dong *et al.*, 2005) and translation ambiguities (Prior *et al.*, 2007). We also expect a great deal of transfer of L1 patterns stored on the auditory and articulatory maps. It is reasonable enough to map a Chinese /p/ to an English /p/, even though the Chinese sound has a different time of voicing onset and no aspiration. The result of this type of imperfect transfer is what leads to the establishment of a foreign accent in L2 learners. As Eckman (Chapter 6, this volume) explains, patterns in learners' phonologies are determined both by transfer and universal principles of markedness. Moreover, the unit of phonological transfer reaches beyond the segment, including syllable structure and other prosodic patterns.

Transfer is also easy enough for the semantics of lexical items (Kroll and Tokowicz, 2005). In this area, transfer is often largely positive, particularly between languages with similar linguistic and cultural patterns. In the initial stages of L2 word learning, this type of transfer requires very little reorganization, because L2 forms are initially parasitic upon L1 forms.

However, transfer is difficult or impossible for item-based syntactic patterns (MacWhinney, 2005b), because these patterns cannot be readily matched across languages. For the same reason, transfer is unlikely for the formal aspects of conjugational or declensional patterns and classes. The fact that transfer is difficult for these systems does not mean that they are easy for L2 learners, but rather that they must be learned from the bottom up without any support from the L1.

When learners have several possible L1 forms that can transfer to L2, they tend to prefer to transfer the least marked forms (Eckman, 1977; Major and Faudree, 1996). For example, as Pienemann *et al.* (2005) have noted, Swedish learners of German prefer to transfer to German the unmarked Swedish word order that places the subject before the tense marker in the German equivalent of sentences such as *Peter likes milk today*. Although Swedish has a pattern that allows the order *Today likes Peter milk*, learners tend not to transfer this pattern initially, because it is the more marked alternative.

Parasitism and internalization

In her Revised Hierarchical Model, Kroll has emphasized the extent to which beginning second language learners depend on preexisting L1 pathways for mediating the activation of L2 lexical items (Kroll and Sholl, 1992). For example, when hearing the word *perro* "dog" in Spanish, the learner may first translate the word into English and then use the English word to access the meaning. At this point, the use of the Spanish word is *parasitic* on English-based knowledge. Later on, the word *perro* comes to activate the correct meaning directly.

In order to move from this parasitic use of L2 to direct access of meaning, the learner needs to strengthen the direct pathways between the new forms and the preexisting functions. The process of internalization can serve to counteract the forces of parasitism. Internalization (Pavlenko and Lantolf, 2000) involves the use of L2 by learners in their inner speech (Vygotsky, 1934). When we activate inner speech, we are using language to build up mental models to control our thinking and plans. Vygotsky (1934) observed that young children would often give themselves instructions overtly. For example, a two-year-old might say, "pick it up" while picking up a block. At this age, the verbalization tends to guide and control the action. By producing a verbalization that describes an action, the child sets up a resonant connection between vocalization and action (Asher, 1969). Later, as Vygotsky argues, these overt instructions become inner speech and continue to guide our cognition. L2 learners go through a process much like that of the child. At first, they use the language only with others. Then, they begin to talk to themselves in the new language and start to "think in the second language." At this point, the second language begins to assume the same resonant status that the child attains for the first language.

Once a process of internalization is set into motion, it can also be used to process new input and relate new forms to other forms paradigmatically. For example, if I hear the phrase *ins Mittelalter* (in the Middle Ages) in German, I can think to myself that this means that the stem *Alter* must be *das Alter*. This means that the dative must take the form *in welchem Alter* (in which age) or *in meinem Alter* (in my age). These form-related exercises can be conducted in parallel with more expressive exercises in which I simply try to talk to myself about things around me in German, or whatever language I happen to be learning. Even young children engage in practice of this type (Berk, 1994; Nelson, 1998). Internalization also helps us understand the growth of the ability to engage in code switching. If a language is being repeatedly accessed, it will be in a highly resonant state. Although another language will be passively accessible, it may take a second or two before the resonant activation of that language can be triggered by a task (Grosjean, 1997). Thus, a speaker may not immediately recognize a sentence in a language that has not been spoken in the recent context. On the other hand, a simultaneous interpreter will maintain both languages in continual receptive activation, while trying to minimize resonant activations in the output system of the source language.

Isolation and participation

The fifth risk factor for older L2 learners is social isolation. As we get older, full integration into a second language community can become increasingly difficult. There are at least three reasons for this. First, as we age, it can become increasingly difficult to set aside L1 allegiances and responsibilities. Second, L2 communities tend to be more immediately supportive of younger L2 learners. As children get older, peer groups become increasingly critical of participants who fail to communicate in accepted ways. Third, as we age, we may develop images regarding our social status that make it difficult to accept corrective feedback, teasing, or verbal challenges, even though these are excellent sources of language input. The cumulative effect of these social factors is that positive support for language learning can decrease markedly across the lifespan. Unless older learners focus directly on making friends in the new community and developing a full L2 persona (Pavlenko and Lantolf, 2000), they can become isolated and cut off from learning.

The fifth support factor for older L2 learners is the obverse of the risk factor of social isolation. Older learners can increase their participation (Pavlenko and Lantolf, 2000) in the L2 community in a variety of ways. They can join religious groups, athletic teams, or work groups. Often these

groups are highly motivated to improve the language abilities of new members, so that they can function smoothly within the group. Any method that can promote interaction with native speakers can facilitate learning (Mackey, Abbuhl, and Gass, Chapter 1, this volume).

Older learners can also engage in formal study and expose themselves to L2 input through books, films, and music. When these methods for increasing participation operate in concert with the processes of chunking, resonance, and internalization, L2 learning will lead to increasingly high levels of proceduralization and correctness. Instruction can also incorporate insights from activity theory (Engeström, 1999; Ratner, 2002) to guide a contextualized curriculum. Many syllabi already make use of a simple form of activity theory when they compose units based on specific activities such as ordering food at a restaurant, asking for directions, dealing with car problems, or transferring money across bank accounts. Multimodal video materials linked to transcripts can be used to further support this type of activity-based learning of vocabulary, pragmatics, and syntax.

Data and common elicitation measures

As we noted earlier, Competition Model studies place cues into competition using orthogonalized analysis of variance designs. Each factor in the design represents a particular cue. Consider the example of a study that examines the competition between word order, case marking, and agreement in Hungarian. The three levels of the word order factor will be NNV, NVN, and VNN with the verb in different positions vis-a-vis the two nouns. For example, the VNN sentence could be the Hungarian equivalent of *chases the dog the cat*. For the case-marking cue, the three levels would involve marking on the first noun, the second noun, or neither noun. For the agreement cue, the three levels would involve agreement of the verb with the first noun, the second noun, or neither noun. This combination of factors would then yield a $3 \times 3 \times 3$ design with 27 cells. In order to achieve greater reliability of measurement, the study might have three replications in each cell for a total of 81 trials. The dependent variables would be percentage choice of the first noun as agent and reaction time for this decision. In the 1980s, these studies were conducted using pictures; work with very young children still uses small toy objects. However, with adults we now use computers to present both the picture stimuli and the sentences. In this mode, subjects select one of two pictures on the screen by pressing a key corresponding to the selected picture. The results of these studies are analyzed using either analysis of variance or maximum likelihood analysis.

Competition Model work also involves computation of cue validity from texts. This is done by examining selected corpora and looking for each case of a relevant competition. To compute the validity of cues for agent choice, we examine sentences with a transitive verb. We then list for each cue whether it is (a) available in the sentence, (b) contrastive, (c) reliable, and (d) reliable in direct conflicts with other cues. In this way, we compute the proportions for simple availability, contrast availability, reliability, and conflict reliability. We then use these values as predictors of the relative strengths of the cues.

Apart from the data provided by the classic version of the Competition Model, the UCM relies on new data gathered from online studies of L2 learning. These studies are designed to examine the effects of resonance, chunking, and internalization on L2 learning. These studies have shown that methods that promote these three processes achieve higher levels of language learning for basic skills such as French nominal declension (Presson and MacWhinney, under review), Japanese sentence patterns (Yoshimura and MacWhinney (2007), vocabulary (Pavlik *et al.*, 2007), and pinyin dictation (Zhang 2009).

By looking at how children, adult monolinguals, and adult bilinguals speaking 18 different languages process various types of sentences, we have been able to reach these conclusions, regarding competition during sentence comprehension:

(1) When given enough time during sentence comprehension to make a careful choice, adults assign the role of agency to the nominal with the highest cue strength.
(2) When there is a competition between cues, the levels of choice in a group of adult subjects will closely reflect the relative strengths of the competing cues.
(3) When adult subjects are asked to respond immediately, even before the end of the sentence is reached, they will tend to base their decisions primarily on the strongest cue in the language.
(4) When the strongest cue is neutralized, the next strongest cue will dominate.
(5) The fastest decisions occur when all cues agree and there is no competition. The slowest decisions occur when strong cues compete.
(6) Children begin learning to comprehend sentences by first focusing on the strongest cue in their language.
(7) As children get older, the strength of all cues increases to match the adult pattern with the most valid cue growing most in strength.
(8) As children get older, their reaction times gradually get faster in accord with the adult pattern.
(9) Compared to adults, children are relatively more influenced by cue availability, as opposed to cue reliability.
(10) Cue strength in adults and older children (8–10 years) is not related to cue availability (since all cues have been heavily encountered by this time), but rather to cue reliability. In particular, it is a function of conflict reliability, which measures the reliability of a cue when it conflicts directly with other cues.
(11) Older learners tend to transfer cue strengths from L1 to L2.

A bibliography of studies supporting these conclusions can be found on the Web at http://psyling. psy.cmu.edu/papers.

Applications

The findings and formulations of the UCM have important implications for the teaching of second languages. For learners in the pre-school and early school years, the risk factors of entrenchment, misconnection, transfer, and isolation are not yet serious concerns. Young learners can acquire additional languages using the same methods they used to pick up their first language. At this age, instruction should focus on providing rich input to implicit learning processes. The principle danger is that, once instruction or exposure to a language ceases, children will soon lose their ability to use that language (Burling, 1959). For immigrant children, the major challenge during this period is to provide social situations that allow them to integrate fully into peer group contexts (McLaughlin, 1985).

During the later school years, second language instruction should become increasingly explicit. For ten year olds, instruction can still rely principally on songs, phrases, and games. However, adolescents should begin to learn in adult mode by relying on chunking, resonance, internalization, and participation. For adolescent and adult learners, instruction should include both contextualized and decontextualized components. Decontextualized components should focus on the resonant practice of basic skills in auditory phonology, articulatory phonology, lexicon, and syntactic constructions. This type of basic skills practice can be controlled through computerized presentation with the results tailored to the individual student level (Pavlik et al., 2007) and relying

on the method of graduated interval recall to maximize efficiency. We have implemented systems of this type (http://talkbank.org/pslc) for learning Chinese sound patterns through Pinyin dictation, Chinese vocabulary, French dictation, and French gender (Presson and MacWhinney, under review). These online systems automatically provide the instructor with students' scores to allow them to monitor students' progress through each phase of each module.

Basic skills training can focus first on chunking (Yoshimura and MacWhinney, 2007) and resonance. As these basic skills become consolidated, learners can begin to focus increasingly on internalization and participation for consolidating L2 fluency. For these levels, instruction should be increasingly contextualized. Methods that rely on computerized presentation of contextually realistic videotaped interactions linked to transcripts can be particularly effective, as in the DOVE transcript browser illustrated at http://talkbank.org/pslc.

Future directions

Further elaboration and testing of the UCM's approach to the issues of fluency, competition, transfer, entrenchment, and internalization will need to address the following high priority research questions:

(1) Competition Model experiments typically treat all transitive verbs as a single group. However, the model emphasizes the item-based nature of syntactic learning (MacWhinney, 2005b; McDonald and MacWhinney, 1995). This means that we need to devise experimental methods that can measure cue competition more accurately across smaller lexical groups for both nouns and verbs.
(2) To deepen the grounding of the model on neuroscience, we need to extend the DevLex model in three major ways. First, we need to show how morphological markers can emerge through the processing of lexical forms. Second, we need to develop a SOM model of the acquisition of syntactic patterns from item-based frames. Third, we need to construct DevLex simulations of early bilingual learning that show how the two languages are merged on the level of deep semantics, but separate on the level of lexical semantics.
(3) We need to devise methods for evaluating the mechanistic effects of the support factors of proceduralization, resonance, transfer, internalization, and participation. Much of this work is now in progress. We have some precise characterizations of some of these mechanisms, but a great deal of careful, empirical work will be needed to complete this picture.
(4) In terms of pedagogical applications, we need to parcel out the effects of transfer, markedness, and cue strength on early skill learning. For example, we know that beginning learners find some French gender cues easier than others (Carroll, 2005; Presson and MacWhinney, under review). Similarly, we know that learners with different L1 backgrounds face very different problems in the learning of Chinese phonology (Zhang, 2009). In order to maximize the efficiency of computerized instruction, we need to develop models that base training on information about these differences.
(5) We need to study the retention of items and basic skills across longer time spans, using standard pre-test/post-test designs. Specifically, we need to know whether emphases on chunking and resonance produce robust learning of L2 patterns.

References

Anderson, J. (Ed.) (1981). *Cognitive skills and their acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Asher, J. (1969). The total physical response approach to second language learning. *The Modern Language Journal*, *53*, 3–17.

Atkinson, R. (1975). Mnemotechnics in second-language learning. *American Psychologist*, *30*, 821–828.

Bates, E. and MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner and L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). New York: Cambridge University Press.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., and Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, *11*, 245–299.

Berk, L. E. (1994). Why children talk to themselves. *Scientific American, November*, *273*, 78–83.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. Gass, and J. Schachter (Eds.), *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press.

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition*, *31*, 175–198.

Brunswik, E. (1956). *Perception and the representative design of psychology experiments*. Berkeley, CA: University of California Press.

Burling, R. (1959). Language development of a Garo and English speaking child. *Word*, *15*, 45–68.

Buzsaki, G. (2006). *Rhythms of the brain*. Oxford: Oxford University Press.

Bybee, J. and Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

Carlson, R., Sulllivan, M., and Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 517–526.

Carroll, S. (2005). Input and SLA: Adults' sensitivity to different sorts of cues to French gender. *Language Learning*, *55*, 79–138.

Clahsen, H. and Muysken, P. (1986). The availability of UG to adult and child learners: A study of the acquisition of German word order. *Second Language Research*, *2*, 93–119.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition studies. *Studies in Second Language Acquisition*, *22*, 499–533.

DeKeyser, R. and Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll and A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press.

Dong, Y. -P., Gui, S. -C., and MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition*, *8*, 221–238.

Ebbinghaus, H. (1885). *Über das Gedächtnis*. Leipzig: Duncker.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*, 315–330.

Ellis, N. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*, 143–188.

Ellis, N. and Beaton, A. (1995). Psycholinguistic determinants of foreign language vocabulary learning. In B. Harley (Ed.), *Lexical issues in language learning* (pp. 107–165). Philadelphia: John Benjamins.

Engeström, Y. (1999). Activity theory and individual social transformation. In Y. Engeström, R. Miettinen, and R. L. Punamiki (Eds.), *Perspectives on activity theory* (pp. 1–15). New York: Cambridge University Press.

Friederici, A. (2009). *Brain circuits of syntax: From neurotheoretical considerations to empirical tests. Biological foundations and origin of syntax*. Cambridge, MA: MIT Press.

Friederici, A., Steinhauer, K., and Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, *99*, 529–534.

Gobet, F. (2005). Chunking models of expertise: Implications for education. *Applied Cognitive Psychology*, *19*, 183–204.

Grosjean, F. (1997). Processing mixed languages: Issues, findings and models. In A. M. B. de Groot and J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–254). Mahwah, NJ: Lawrence Erlbaum.

Guenther, F. and Gjaja, M. (1996). The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America*, *100*, 1111–1121.

Gupta, P. and MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language*, *59*, 267–333.

Hahne, A. and Friederici, A. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141.

Hernandez, A., Li, P., and MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*, 220–225.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska Symposium on Motivation*, *15*, 192–238.

Kempe, V. and MacWhinney, B. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes*, *14*, 129–171.

Kohonen, T. (2001). *Self-organizing maps* (Third Edition). Berlin: Springer.

Krashen, S. (1994). The input hypothesis and its rivals. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45–78). San Diego: Academic.

Kroll, J. and Sholl, A. (1992). Lexical and conceptual memory in fluent and nonfluent bilinguals. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 191–206). Amsterdam: North-Holland.

Kroll, J. and Tokowicz, N. (2005). Bilingual lexical processing. In J. F. Kroll and A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

Kuhl, P., Conboy, B., Padden, D., Nelson, T., and Pruitt, J. (2005). Early speech perception and later language development: Implications for the "Critical Period". *Language Learning and Development*, *1*, 237–264.

Li, P., Zhao, X., and MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science*, *31*, 581–612.

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

MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Language acquisition: Vol. 1. Syntax and semantics* (pp. 73–136). Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (1997). Implicit and explicit processes. *Studies in Second Language Acquisition*, *19*, 277–281.

MacWhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language*, *31*, 883–914.

MacWhinney, B. (2005a). The emergence of linguistic form in time. *Connection Science*, *17*, 191–211.

MacWhinney, B. (2005b). Item-based constructions and the logical problem. *ACL*, 46–54.

MacWhinney, B. (2008a). How mental models encode embodied linguistic perspectives. In R. Klatzky, B. MacWhinney, and M. Behrmann (Eds.), *Embodiment, Ego-Space, and Action* (pp. 369–410). Mahwah: Lawrence Erlbaum.

MacWhinney, B. (2008b). A unified model. In P. Robinson and N. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B., Feldman, H. M., Sacco, K., and Valdes-Perez, R. (2000). Online measures of basic language skills in children with early focal brain lesions. *Brain and Language*, *71*, 400–431.

Major, R. (1987). The natural phonology of second language acquisition. In A. James and J. Leather (Eds.), *Sound Patterns in Second Language Acquisition* (pp. 207–224). Dordrect: Foris.

Major, R. and Faudree, M. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition*, *18*, 69–90.

Massaro, D. (1987). *Speech perception by ear and eye*. Hillsdale, NJ: Lawrence Erlbaum.

Matessa, M. and Anderson, J. (2000). Modeling focused learning in role assignment. *Language and Cognitive Processes*, *15*, 263–292.

McDonald, J. L. (1986). The development of sentence comprehension strategies in English and Dutch. *Journal of Experimental Child Psychology*, *41*, 317–335.

McDonald, J. L. and Heilenman, K. (1991). Determinants of cue strength in adult first and second language speakers of French. *Applied Psycholinguistics*, *12*, 313–348.

McDonald, J. L. and MacWhinney, B. J. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, *34*, 543–566.

McLaughlin, B. (1985). *Second-language acquisition in childhood*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Müller, J., Hahne, A., Fujii, Y., and Friederici, A. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience*, *17*, 1229–1244.

Nelson, K. (1998). *Language in cognitive development: The emergence of the mediated mind*. New York: Cambridge University Press.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Philadelphia: John Benjamins.

Pavlenko, A. and Lantolf, J. (2000). Second language learning as participation and the (re)construction of selves. In A. Pavlenko and J. Lantolf (Eds.), *Sociocultural theory and second language learning* (pp. 155–178). Oxford: Oxford University Press.

Pavlik, P. and Anderson, J. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, *29*, 559–586.

Pavlik, P., Presson, N., Dozzi, G., Wu, S., MacWhinney, B., and Koedinger, K. (2007). The FaCT (Fact and Concept Training) System: A new tool linking Cognitive Science with educators. *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 1379–1384). Nashville, TN: Cognitive Science Society.

Pienemann, M., Di Biase, B., Kawaguchi, S., and Håkansson, G. (2005). Processing constraints on L1 transfer. In J. F. Kroll and A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 128–153). New York: Oxford University Press.

Pimsleur, P. (1967). A memory schedule. *The Modern Language Journal*, *51*, 73–75.

Presson, N. and MacWhinney, B. (under review). Learning grammatical gender: The effects of rules and prototypes. *Applied Psycholinguistics*.

Prior, A., MacWhinney, B., and Kroll, J. (2007). Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity. *Behavior Research Methods*, *37*, 134–140.

Pulvermüller, F. (2003). *The neuroscience of language*. Cambridge: Cambridge University Press.

Ratner, C. (2002). *Cultural psychology: Theory and method*. New York: Kluwer/Plenum.

Rosenbloom, P. S. and Newell, A. (1987). Learning by chunking: A production system model of practice. In D. Klahr, P. Langley, and R. Neches (Eds.), *Production system models of learning and development* (pp. 221–286). Cambridge, MA: MIT Press.

Searchinger, G. (1995). *The Human Language Series*. New York: Equinox Films.

Seidenberg, M. and McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Snow, C. E. (1999). Social perspectives on the emergence of language. In B. MacWhinney (Ed.), *The emergence of language* (pp. 257–276). Mahwah, NJ: Lawrence Erlbaum Associates.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Taraban, R. and Palacios, J. M. (1993). Exemplar models and weighted cue models in category learning. In G. Nakamura, R. Taraban, and D. Medin (Eds.), *Categorization by humans and machines*. San Diego: Acdemic Press.

Tokowicz, N. and MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event–related potential investigation. *Studies in Second Language Acquisition*, *27*, 173–204.

Tomasello, M. (2003). *Constructing a first language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.

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

van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, *98*, 3–53.

Vygotsky, L. (1934). *Thought and language*. Cambridge: MIT Press.

Wittenberg, G., Sullivan, M., and Tsien, J. (2002). Synaptic reentry reinforcement based network model for long–term memory consolidation. *Hippocampus*, *12*, 637–647.

Yoshimura, Y. and MacWhinney, B. (2007). The effect of oral repetition in L2 speech fluency: System for an experimental tool and a language tutor. *SLATE Conference*, 25–28.

Zhang, Y. (2009). A tutor for learning Chinese sounds through pinyin. *Applied Psycholinguistics*.