Entrenchment in Second Language Learning

Brian MacWhinney

Abstract

This chapter examines the role of entrenchment in second language (L2) learning. It is generally recognized that language learning success declines with age. However, the exact nature and causes of this decline are not yet clear. The earliest accounts of age-related effects in L2 learning were based on the critical period concept derived from embryology and ethology. However, the gradual nature of declines in L2 learning outcomes, along with evidence of successful learning during adulthood, do not correspond well to this model. Neural network models succeed in simulating the gradual decline of language learning abilities by focusing on the role of entrenchment. However, these models also predict catastrophic interference during the learning of L2. To resolve this paradox, the Unified Competition Model reformulates the issue in terms of the interplay of four risk factors with four support or protective factors. Risk factors include entrenchment, transfer, overanalysis, and isolation. To counteract these risk factors, adult L2 learners can make overt use of the protective factors of resonance, decoupling, chunking, and participation. This model provides a more complete account of age-related declines in L2 learning and ways in which they can be mitigated.

1. Critical periods

In his landmark study of the biological foundations of language, Lenneberg (1967) postulated a critical period for the acquisition of language that terminates as a result of the onset of cerebral lateralization at puberty. Lenneberg thought of this critical period as applying to the learning of both the first language (L1) and a second language (L2). He attributed the loss of cerebral equipotentiality to the factors of myelination and neuronal commitment. Subsequent research (Feldman, MacWhinney, & Sacco, 2002) has supported Lenneberg's ideas about early cerebral equipotentiality and plasticity. Using fMRI scanning, Booth et al. (2000) showed that children who had left hemisphere brain lesions during gestation and infancy developed language by relying on right hemisphere regions such as the inferior frontal gyrus (IFG) that correspond to damaged regions on the left. Moreover, when these children reached ages 6-8, both the functional use of language and basic information processing skills underlying language use (Feldman et al., 2002) ended up being generally well within the normal range. Interestingly, this pattern of nearly optimal recovery of language despite often massive brain injury does not extend to other cognitive skills. In particular, damage to right hemisphere areas from early focal

lesions can result in major spatial deficits (Stiles, Trauner, Engel, & Nass, 1997), and there is a general tendency for language to be more well preserved than other cognitive skills (Aram & Eisele, 1994). This pattern has been called "cognitive crowding" because it indicates that the development of successful language skills can come at the expense of other skills such as math or reading that might be acquired later in development.

Lenneberg's proposal corresponds well with some facts about early equipotentiality. However, his reliance on lateralization as the mechanism underlying the termination of the critical period suffers from two basic problems. The first is that lateralization is present well before age 13. Anatomical and neuroimaging work has shown that lateralization is already present at birth (Molfese, Freeman, & Palermo, 1975), and that it increases during the first two years (Mills, Coffey-Corina, & Neville, 1997). A second problem is that Lenneberg's analysis relied on hemispherectomy data from Basser (1962) to support the idea that equipotentiality terminates at age 13. However, the cases reported by Basser involved confounds with ongoing epileptic status, size of epileptic focus, and language measurement that make it impossible to establish any close link between the onset of puberty and the loss of a right hemisphere ability to reacquire the first language.

Lenneberg also suggested that myelination could be a mechanism underlying a critical period transition at puberty. We know that there is progressive myelination throughout children's development (Gao et al., 2009) and that this myelination largely culminates by puberty. However, there are two problems with attempts to rely on myelination as an explanation for termination of a critical period for L2 acquisition. The first is that, even during adulthood, focused cognitive training can lead to measurable changes in white matter mass through myelination (Engvig et al., 2012; Posner, Tang, & Lynch, 2014; Sampaio-Baptista et al., 2013; Takeuchi et al., 2010), indicating that this form of neuronal plasticity extends well past the postulated critical period of childhood. Second, myelination is a process that impacts all major fibre tracts, not just those supporting language skills. However, no one would want to argue that all aspects of adult behavior are locked in by critical period effects.

More generally, it is questionable whether age-related decreases in L2 learning outcomes qualify as a critical period in the sense used in embryology (Browman, 1989) or ethology (Lorenz, 1958). In these areas, critical periods are defined by radical stimulus deprivation, such as blindness (Hubel & Weisel, 1963), deafness (Kral & Sharma, 2012), or social isolation (Marler, 1991). For second language learning, we are not talking about massive deprivation during infancy, but rather the ability to modify an already intact system across the lifespan. Moreover, the critical periods studied in animals are not shaped by some specific single neural process, such as lateralization, synaptic pruning, myelination, metabolic decline, increases in estrogen, etc. Instead, they arise from a complex set of interactions and modulations between cortical structures and basal ganglia, involving various neural transmitters, and these modularity developmental effects work differentially in particular areas of structures, such as segments of the auditory cortex (Kral & Sharma, 2012) or motor cortex (Yamamoto, Hoffman, & Strick, 2006).

Johnson and Newport (1989) conducted the first psycholinguistic test of Lenneberg's critical period analysis. They examined grammaticality judgments from 46 participants whose first languages were Chinese or Korean and who had arrived in the

United States at various ages from 3 to 39. They observed a significant correlation between age of arrival and success in grammaticality judgments, but only for early arrivers. However, in his review of additional studies replicating and extending this initial study, Birdsong (2005) concludes that the accuracy of grammaticality judgments shows a continuous linear decline for older ages of arrival, even in the data from Johnson and Newport. This suggests that the decline is not linked to a critical or even sensitive period, but simply to decreasing plasticity. Further evidence for this conclusion comes from a large census-based study (Hakuta, Bialystok, & Wiley, 2003; Wiley, Bialystok, & Hakuta, 2005) showing that the gap in L2-attainment resulting from later age of arrival is equal in size to the gap caused by lower levels of education.

2. Entrenchment

A contrasting approach attributes age-related L2 learning effects to basic properties of learning in neural nets. Rather than referring to specific biological processes such as lateralization, myelination, synaptic pruning, or gene expression, this account attributes age-related effects to entrenchment, which is viewed as the emergent result of network computation (Zevin, 2012). Once a network has been trained to respond to a large variety of stimuli, it settles in to an attractor state that is very hard to modify. When such states are reached, the relevant patterns apply quickly and uniformly during processing (Segalowitz, Segalowitz, & Wood, 1998) and are difficult to modify developmentally (Brooks & Tomasello, 1999).

Entrenchment has been demonstrated both for Parallel Distributed Processing networks (Seidenberg & Zevin, 2006) and for self-organizing feature maps (Li, Farkas, & MacWhinney, 2004). The nature of this effect in the emergence of word classes can be demonstrated graphically in self-organizing feature maps for lexical development of the type created by the DevLex simulation (Li et al., 2004). After 50 cycles of learning, these maps have a rather loose and unstable organization of lexical items by part of speech. However, as learning progresses, more items are added to the map and the boundaries of the different parts of speech categories become sharper. Moreover, during the first stages of L1 learning, the categories migrate around the map. During later stages there is no more movement of category boundaries, because the shape of the parts-of-speech and the location of particular items on the feature map have become entrenched.

These neural network simulations of the entrenchment process represent existence proofs showing that critical periods can arise from the strengthening of organized connections between units. On the neuronal level, this effect would be based on the strengthening of synaptic connections according to the Hebbian (Pulvermüller, 2003) rule that "neurons that fire together, wire together". Entrenchment is further supported by synaptic pruning which can disallow certain patterns that are not needed for first language functioning. Myelination, hormonal changes, and lateralization can play further secondary roles in terms of locking in patterns of connection between major brain areas. Neural network simulations have not yet been configured to represent the interactions between these various physiological changes. Rather current simulations only capture the overall trajectory of entrenchment.

The computational account of entrenchment can be used to account for critical period effects in L2 learning of audition, articulation, syntax, and morphology. Such accounts do not need to postulate any sudden or sharp end to the critical period. In that way, they provide a satisfying account for the observed window of decreasing plasticity that slowly closes over a long period of time.

Entrenchment can also account for findings regarding articulatory development. A study of Italian immigrants to Canada (Munro, Flege, & Mackay, 1996) found that, unless these participants had arrived before age 6, it was possible to detect some features of an Italian accent in their English. However, it is possible that these effects arise not so much from entrenchment, as from the fact that immigrants who arrived at a later age tended to interact more frequently with the local Italian community, thereby maintaining an Italian accent in English. An interesting feature of this study is the suggestion that entrenchment could close off a critical period for articulation well before the onset of puberty. However, it is possible to modify L1 influences on articulation in adults through concentrated training (Flege, 1995).

Even earlier entrenchment effects have been postulated for the development of auditory perception. Studies by Werker and Tees (2005) and Kuhl (2004) have demonstrated a loss of certain perceptual contrasts during the first year of life. These effects led Kuhl to formulate the Perceptual Magnet hypothesis, according to which a child loses the ability to make discriminations between sounds that are not contrastive in the native language. According to this hypothesis, the loss arises because sounds that vary in perceptual detail are assimilated to the core of a given phonemic category. This account attributes age-related changes in perception to the entrenchment of the perceptual magnet. Although this model works well for many empirical results, there are several phenomena that it cannot handle. One problem is that certain sounds, such as clicks and nasals, are reliably detected at all ages (Werker, 1995). Another problem involves the fact that, after age 3, children begin to improve in their ability to detect certain contrasts (Werker, 1995). Moreover, deficits in perceptual accuracy in adult bilinguals do not link up directly to the age at which they began learning a second language during childhood, instead showing wide individual variation (Flege & MacKay, 2004).

The most intensely studied auditory entrenchment effect involves the inability of many Japanese speakers to discriminate English /r/ and /l/. This inability may stem from the fact that discrimination of phonemic contrasts in Japanese does not require attention to the third formant, although the distinction between English /r/ and /l/ is made clearest through processing of the third formant (F3). Several studies (Bradlow, Akahni-Yamada, Pisoni, & Tokhura, 1999; Ingvalson, Holt, & McClelland, 2012; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002) have succeeded in training perception of this contrast using artificial stimuli. However, generalization to speech stimuli has been limited, suggesting that more work is needed here using real speech as training materials.

Entrenchment can also be used to account for the problems that L2 learners have with the learning and usage of grammatical morphology. It is often observed that learners of languages such as German or Spanish fail to acquire full control over grammatical markings for noun declension or verb inflection (Sagarra & Herschensohn, 2010, 2011; VanPatten, Keating, & Leeser, 2012). VanPatten has argued that this deficit is due to the termination of a critical period for the use of Universal Grammar. Approaching this issue from a rather different perspective, Newport (1990) hypothesized that children are better

able to learn grammar because they start with small chunks. However, initial evidence supporting this idea from Elman (1993) failed later replication (Rohde & Plaut, 1997). In fact, it seems that, if anything, children store larger units than do adults (McCauley, Monaghan, & Christiansen, 2014).

Another explanation for adult problems with morphology comes from Ullman (2004) who argues that increased estrogen levels at puberty lead to weakening of the procedural learning system involved in combinatorial morphology, and compensatory strengthening of declarative memory. This hypothesis predicts stronger learning of irregular morphology in adulthood and weaker learning of regular morphology. However, studies of the training of grammatical morphology in adult L2 learners (Presson, Sagarra, MacWhinney, & Kowalski, 2013) and experiments with artificial grammars (Kürten, De Vries, Kowal, Zwitserlood, & Flöel, 2012) have demonstrated exactly the opposite effect, indicating that there is no obvious procedural learning deficit in adults. Together, these results indicate that a more likely account is that adults seek to learn lexical stems in a highly analytic manner that reduces concomitant acquisition of attached grammatical markers. Thus, this is a problem with learning style, rather than a basic maturational limitation.

We have seen that the timing and nature of the age-related effects attributed to the operation of entrenchment vary markedly across the domains of audition, articulation, syntax, and morphology. Moreover, when we look in greater detail at each of these areas, we find additional variation across structures and processes. If entrenchment is to work as an account for all of their effects, we need to understand how it can be that some neural networks entrench much more rapidly than others. In large part, these variations in entrenchment speed and rigidity can be explained by the fact that some skills are learned before others. For example, the child works on the distinction between /p/ and /b/ long before learning to produce *wh*-questions in English or mark the subjunctive in Spanish. However, we may also need to refine the theory of entrenchment to account for differential entrenchment in various cortical areas. Clear evidence of such differences comes from work on plasticity in motor cortex. Studies using infection with rabies virus (Yamamoto et al., 2006) have revealed one network of neurons capable of ongoing developmental plasticity along with another network that is no longer plastic during adult development. It is likely that local variations in plasticity during adulthood can be triggered by interactions between cortical areas, particularly in the highly interactive areas subserving higher language functions.

Apart from the need to consider complex patterns of neuronal variation, the entrenchment hypothesis suffers from two other problems. The first involves the phenomenon of catastrophic interference (McCloskey & Cohen, 1989). When a neural net that has been trained on one set of inputs is suddenly shifted to a new set of inputs, the weights derived from the earlier training set are overwritten. For second language learning, this would mean that learning of a second language erases memory of the first language. Such effects may occur when a child is adopted into a new language community and not given any opportunity to continue use of the first language (Pallier et al., 2003). However, for older children and adults, the termination of use of a first language leads not to total forgetting, but rather gradual attrition (Steinkrauss & Schmid, this volume). One could argue that maintenance of the first language in older learners reflects a higher level of entrenchment. However, explaining how entrenchment or

stability can exist alongside plasticity has been a problem for neural network modeling. Although a variety of technical solutions to this problem have been explored (Hamker, 2001; McClelland, McNaughton, & O'Reilly, 1995; Richardson & Thomas, 2008), it appears that successful learning must treat the learning of L2 as a separate task encoded in a partially separate network (Carpenter, Grossberg, & Reynolds, 1991). Achieving this separation requires control from attentional areas, and it is reasonable to assume that this separation between languages occurs in real-life second language learning.

The obverse of catastrophic interference is irreversible entrenchment. If networks are configured in ways that maximize the increase of stability or entrenchment over time, they become incapable of new learning. In the extreme, the entrenchment account predicts that adults should be nearly incapable of learning second languages. In fact, some adults are able to learn second languages quite well, often to the level of native speaker competence (Birdsong, 2005; Bongaerts, 1999). For audition, we know that adults can acquire an ability to respond quickly to second language auditory contrasts as reflected in the early mismatch negativity (MMN) component of the EEG (Winkler et al., 1999) after even just 6 hours of training (Kraus et al., 1995). Moreover, in adult rats, noise exposure can reinstate critical period plasticity, demonstrating that at least some of the effects of entrenchment are reversible (Zhou, Panizutti, de Villers-Sidani, Madeira, & Merzenich, 2011). Thus, although entrenchment provides a reasonable account for a decline in learning abilities, it overestimates the extent to which the brain loses plasticity.

We can summarize this brief review in terms of three major points.
1. The linkage of age-related effects in language learning directly to the onset of lateralization, increases in estrogen, or the effects of myelination has not been supported.
2. Age-related effects in L2 learning can be partially attributed to entrenchment in neural nets. However, learning of a second language must involve control processes that avoid triggering of catastrophic interference effects through construction of partially separate systems.
3. Patterns of plasticity, early canalization, and later learning vary markedly across individuals, language domains, and specific target structures. This indicates that entrenchment is not a uniform process across cortical areas, but rather a feature that varies markedly in onset, duration, and tenacity, depending on additional neurodevelopmental interactions.

Together, these results point to the need for a model that treats entrenchment as only one component of a more complex account of age-related effects in second language learning. In other words, entrenchment alone cannot explain age-related effects in L2 acquisition.

3. The Unified Competition Model

The Unified Competition Model or UCM (MacWhinney, 2012) seeks to address the issue of age-related variation in L2 learning by highlighting the interplay between risk factors and support processes. For L2 learning past early childhood, the model postulates the four risk factors of entrenchment, transfer, overanalysis, and isolation. To overcome these four risk factors, adults can rely on the support processes of resonance, decoupling,

chunking, and participation. These processes will be described in detail in Sections 4 to 7 below. This analysis is similar to that presented in MacWhinney (2012) with three important modifications. First, The risk factor of parasitism is treated here as a logical consequence of transfer, rather than as a separate process. Second, the earlier account discussed misconnection between cortical areas as a risk factor. Because the exact neurological basis of misconnection between cortical areas is still unclear, reference to that risk factor is now removed from the model and left as a topic for future research. Third, the current version places a stronger emphasis on the risk factor of overanalysis. Other aspects of the previous analysis remain the same.

The UCM holds that all of these risk factors and support processes are available to both children and adults. In that sense, there is no fundamental difference (Bley-Vroman, 2009) between children and adults as language learners. What differs between language learning in childhood and adulthood is the way in which risk and support processes are configured and the need for explicit invocation of support processes by adult learners.

There are four obvious differences between child and adult language learners. First, during the process of first language learning, infants are also engaged in learning about how the world works. In contrast, adult second language learners already have a basic understanding of the world and human society. Second, infants are able to rely on a brain that has not yet undergone entrenchment. In contrast, adult second language learners have to deal with a brain that has already been dynamically configured for 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, adult second language learners are often heavily involved in L1 social and business commitments that distract them from L2 interactions. Fourth, children have not yet developed adult-like methods for executive control of attention. Although the executive control areas of the brain are active at birth (Doria, Beckman, Arichi, & Merchant, 2010), they continue to develop through childhood and adolescence (Asato, Terwilliger, Woo, & Luna, 2010; Casey, Giedd, & Thomas, 2000). Regularity and inhibitory control over behavior increases in complexity and refinement across the whole period of childhood and adolescence (Munakata, McClelland, Johnson, & Siegler, 1997). To the degree that language and language learning depends on executive control, we can expect differences between adults and children from these sources, although there is no sharp transition point.

Along with these four areas of difference, there are many shared features between L1 and L2 learners. Both groups are trying to learn the same target language; both need to segment speech into words; both need to learn the meanings of these words; both need to figure out the patterns that govern word combination in syntactic constructions; and both have to interleave their growing lexical and syntactic systems to achieve fluency. Thus, both the overall goal and the specific subgoals involved in reaching that goal are the same for both L1 and L2 learners. In addition, the neurocognitive mechanisms available to solve these problems are the same for the two groups. Both rely on episodic memory to encode new forms and chunks; both have access to embodied encodings of actions and objects; both use statistical learning (see Jost & Mortensen, this volume) and generalization (see Cordes, this volume) to extract linguistic patterns; and both solidify knowledge and procedures through routine and practice. Both groups are enmeshed in social situations that require a continued back and forth of communication, imitation, and learning, as well as understandings regarding shared intentions and common ground. One

could recognize the shared nature of all those mechanisms and processes, but still claim that the remaining differences are the ones that are fundamental (Bley-Vroman, 2009). 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.

4. Entrenchment and Resonance

As long as a neural network continues to receive input in the first language, entrenched first language patterns will block learning of conflicting second language patterns. However, if input from the first language tapers off or stops, the network will undergo catastrophic interference leading to loss of the first language. In real L2 learning, this is not what we observe. Thus, the computational mechanism of entrenchment can only be used as an account of age-related effects in the case of ongoing use of L1 during learning of L2.

As noted earlier, the effects of entrenchment in neural network models vary markedly across language processing levels. The detailed operation of entrenchment in self-organizing feature maps has been modeled most explicitly for lexical and phonological structure using the DevLex model (Li, Zhao, & MacWhinney, 2007) and auditory structure using the DIVA model (Guenther & Gjaja, 1996). However, the Unified Competition Model holds that cortical maps exist for each of the structural levels recognized by traditional linguistics, including syntax (Pulvermüller, 2003), mental models (MacWhinney, 2008), and discourse patterning (Koechlin & Summerfield, 2007) as given in Table 1.

Table 1: Levels of linguistic processing

Map	Area	Processes	Theory
1. Audition	Auditory cortex	Extracting units	Statistical learning
2. Articulation	IFG, motor cortex	Targets, timing	Gating
3. Lexicon	STG, ATL	Sound to meaning	DevLex
4. Syntax	IFG	Slots, sequences	Item-based patterns
5. Mental Models	DLPFC	Deixis, perspective	Perspective, roles
6. Discourse	Frontal cortex	Conversation	CA, pragmatics

This table presents the major brain areas involved in each level of language processing, some of the processes involved, and the theories that have been developed to account for each type of processing. For audition, auditory cortex is the focus of initial processing. However, according to Hickok and Poeppel (2004), additional areas are involved in linking auditory patterns to lexical items. Studies of statistical learning focus on the ways

in which these patterns become encoded in auditory and lexical cortex. Articulation involves communication between gestural patterns in the pars opercularis of the inferior frontal gyrus (IFG) and detailed activation of patterns in motor cortex. The actual firing of these patterns is gated by a monitor (Roelofs, 2011) that checks to see if output patterns align with lexical targets. In the DevLex framework, lexical patterns are organized phonologically in the superior temporal gyrus (STG) and conceptually in the anterior temporal lobe (ATL). Syntactic processing involves communication between general constructions involving the pars triangularis of the inferior frontal gyrus (IFG) and the ATL in accord with the theory of item-based patterns (MacWhinney, 2014). Mental model construction relies on frontal areas such as the dorsal lateral prefrontal cortex (DLPFC) to assign case roles and track perspectives. Additional frontal areas are involved in tracking the overall pragmatic flow and sequencing of conversation.

Entrenchment in cortical maps is the strongest risk factor for auditory phonology (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005), articulatory phonology (Major, 1987), and syntax (DeKeyser, 2000). For audition and articulation, the power of entrenchment arises from the fact that myriads of detailed coordinations are involved in attaining nativelike proficiency. For articulation, there are many muscles to be controlled and the ways in which these must be sequenced in words to achieve fluency are very numerous and require frequent practice. Moreover, many of the connections between motor cortex and the spinal cord are fixed early in development. For audition, learning of L1 involves adjustment to a wide range of speaker variations, as well as the ability to use top-down information to reconstruct sounds in noise. On the lexical level, interference between L1 and L2 leads to relatively fewer direct clashes. As a result, L2 learning of new words suffers minimal age-related declines. In fact, adults can acquire new L2 vocabulary more efficiently than children (Snow & Hoefnagel-Hohle, 1978).

In a thorough review of studies comparing L1 and L2 learning of syntax, Clahsen and Felser (2006) found no significant differences for the majority of morphological and syntactic processes. This finding is in line with the claim of the Unified Competition Model regarding the fundamental similarity of L1 and L2 learning. However, Clahsen and Felser also noted that L2 learners often encounter problems in their processing of complex syntactic patterns involving discontinuities and gaps. Based on these data, they formulated the Shallow Structure Hypothesis which holds that L2 learners are not able to extract deeper, more abstract levels of syntactic structure, whereas L1 learners eventually acquire this ability. However, it is possible that the movement and gapping structures in question may present challenges to learners because of interference from entrenched L1 processes, particularly in Asian languages.

The risk factor of entrenchment can be counteracted during L2 learning by explicit reliance on the support factor of resonance. Resonance can be defined as the consolidation of new linguistic patterns through support from existing relations. It provides new encoding dimensions to reconfigure old neuronal territory, permitting the successful encoding of L2 patterns. Because this encoding must operate 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 & Anderson, 2005). Because robustness accumulates with practice, later retrieval trials can be spaced farther and farther apart. This is the principle of "graduated internal 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 (see Takkashima & Bakker, this volume). Repeated hippocampal access to this trace (Wittenberg, Sullivan, & Tsien, 2002) can further consolidate the memory. Once initial consolidation has been achieved, maintenance only requires occasional reactivation of the relevant retrieval pathway (Nadel, Samsonovich, Ryan, & Moscovitch, 2000). This type of resonance can be used to consolidate new forms on the phonological, lexical (Gupta & 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 & 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. During reading, one can activate the sounds of words (Perfetti, Bell, & Delaney, 1988) and the sounds of words can activate their orthographies (Share, 1995). When an L2 learner of German encounters the word *Wasser,* the regularity of German spelling makes it easy to map the sounds of the word directly to the image of the letters. However, 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. Delays in L2 learning of Chinese by speakers of languages with Roman scripts illustrate this problem.

5. Transfer and Decoupling

The UCM holds that L2 learners will attempt transfer whenever they can perceive a match between an item in L1 and a corresponding item in L2. If the match between L1 and L2 is close enough, the result is positive transfer, which is not a risk factor. For example, it is often easy to transfer the basic pragmatic functions (Bardovi-Harlig, 2013) that help structure conversations and the construction of mental models. Transfer is also easy enough for the semantics of lexical items (Kroll & Tokowicz, 2005), for which 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

(Kroll, 1992). However, apart from the issue of lexical parasitism, transfer can encounter some mismatches in meaning (Dong, Gui, & MacWhinney, 2005) and translation ambiguities (Prior, MacWhinney, & Kroll, 2007).

There is also a great deal of transfer from L1 to L2 in terms of patterns on both 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.

However, transfer is difficult or impossible for item-based syntactic patterns (MacWhinney, 2005), 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. Apart from these formal aspects, the semantic distinctions involved in grammatical structures can readily transfer from L1 to L2. For example, the conceptual notion of an indirect object can transfer from English to German, although the actual use of the dative case in German involves many additional complexities not found in English.

When learners have several possible L1 forms that they can transfer to L2, they tend to prefer to transfer the least marked forms (Eckman, 1977; Major & Faudree, 1996). For example, as Pienemann, Di Biase, Kawaguchi, and Håkansson (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.

To correct the errors produced by transfer, learners must construct an independent L2 lexicon, phonology, and syntax. We can refer to this process of creating this independent system as *decoupling*. Decoupling involves creating new L2 forms and patterns and linking them not to L1 items, but to each other in the emerging separate grammar. To strengthen links between L2 items, learners must use the language either in real conversations or for inner speech (Vygotsky, 1934). Because items that fire together wire together, this will create new connections within the emerging L2 system. This type of co-activation can help 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.

Decoupling can be further promoted by invoking resonance to link up forms and meanings. For example, if I hear the phrase *ins Mittelalter* ('into the Middle Ages') in German, I can think to myself that this means that the stem *Alter* must be neuter gender, i.e. *das Alter*. This means that the dative must take the form *in welchem Alter* ('in which (Dat.) age') or *in meinem Alter* ('in my (Dat.) 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).

6. Overanalysis, proceduralization, and chunking

Adult L2 learners excel in the learning of new vocabulary items (Folse, 2004; Nation, 2001). Language textbooks encourage this process by introducing new lexical items in terms of their basic stems, sometimes along with annotations regarding additional morphological forms, such as plurals or case markers. However, learners often just ignore these additional markings – an example of what we can call overanalysis. Children who are learning L1, on the other hand, have to discover the relations between alternative forms of words on their own. Although this would seem to impose an additional burden on the child learner, it ends up conferring them with an advantage, because they pick up words in a fuller context that helps them link into patterns of declension, conjugation, and collocation. In effect, children learn chunks from which they then extract items, whereas adults learn items from which they must then reconstruct larger patterns, often without having the contextual context needed for the reconstruction. For example, adult L2 learners of German often learn the word *Mann* in isolation as the translation for English *man*. If, instead, they would learn phrases such as *der alte Mann* ('the old man')*, meines Mannes* ('my (Gen.) man (Gen.), of my man')*, den jungen Männern* ('the (Dat., Pl.) young (Dat. Pl.) men (Dat.)')*,* and *ein guter Mann* ('a good man'), 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 & 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.

There is an understandable confusion in the literature regarding the relations between entrenchment, chunking, and proceduralization. Proceduralization involves taking two or more items or patterns that are initially separate and joining them together into a unified whole. Often, the resultant procedure is referred to as a chunk (Newell, 1990); see Gobet, this volume). However, it would be less confusing to refer to the process as proceduralization and to the result as a procedure. Once pieces are unitized, they tend to dominate over analytic alternatives. In other words, the stronger the whole, the weaker the parts (Bybee & Brewer, 1980; Hadley & Healy, 1991; Kapatsinski & Radicke, 2009; Sosa & MacFarlane, 2002).

Chunks are created through a very different non-combinatorial process in which a series of items is pulled out of the speech stream as a whole. For example, children may learn *do you want any more* as a chunk, at least receptively. Later on, they may learn the sequence *any more* and use that to further analyze the larger chunk. Adult L2 learners are likely to avoid chunk acquisition in the first place, moving directly to analysis of the phrase into separate lexical items. To distinguish these two processes, we could use proceduralization to refer to the first and chunk extraction to refer to the second.

Proceduralization operates most obviously in terms of production, whereas chunk extraction is a perceptual process. Neither of these processes should be confused with entrenchment, as it is understood here, which can operate equally well on analyzed or unanalyzed items.

To compensate for the risk factor of overanalysis, L2 learners can rely on explicit invocation of the protective processes of chunking and proceduralization. To the degree that L2 learners can focus their attention on larger input strings, such as full preposition phrases or verb-particle combinations, they will be able to pick up chunks corresponding to the forms learned by the child. However, they can also rely on overt control of proceduralization by combining known pieces. 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 proceduralization process can be stored as a unit. In this case, it is not the actual phrase that is stored, 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.

Paradis (2004) argues that L2 learners cannot proceduralize their second language. Ullman (2004) agrees that L2 learners have problems with proceduralization, but he believes that L2 proceduralization is possible, albeit difficult and delayed. As a result, L2 productions may remain forever slow and non-fluent. We can refer to the Paradis-Ullman position as the *Proceduralization Deficit Hypothesis* (PDH). In support of this idea, a study 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, Steinhauer, & Pfeifer, 2002; Müller, Hahne, Fujii, & Friederici, 2005) have shown that, if the rules of the target language are simple and consistent, L2 learners can develop proceduralization, as measured by ELAN, within 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 &

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 reliability of the patterns in the input.

Crucially, proceduralization relies on the ability of the basal ganglia to form new connections between preexisting cortical patterns (Dominey & Boussaoud, 1997; Graybiel, 1995, 1998). The fact that this system is designed to create sequential recombinations of patterns encoded in cortical areas provides us with a fundamental understanding how decoupling can occur and how catastrophic interference can be avoided. For articulation, the basal ganglia, including striate cortex, work to recombine L1 patterns stored in motor cortex into new chains for use in L2. For audition, this can be done by acquiring new sequential regularities based on statistical learning and new phonotactic patterns. For grammar, this involves new constructions of sentential patterns (Hinaut & Dominey, 2013), again through the action of cortico-striatal projections and reinforcement learning in the basal ganglia.

7. Isolation and participation

The fourth risk factor for older L2 learners is social isolation. As we get older, full integration into a second language community becomes 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 if 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 & Lantolf, 2000), they can become isolated and cut off.

The support factor that can counter isolation is participation. Older learners can increase their participation (Pavlenko & 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. Older learners can also engage in formal study and expose themselves to L2 input through books, films, and music.

8. Conclusions

Explanations of age-related declines in L2 learning outcomes have often focused on the notion of a critical period shaped by a small set of biological mechanisms. These accounts postulate a rapid termination of language learning ability that is not observed in real learners. Accounts based on gradual entrenchment in neural nets fare a bit better, but run into problems accounting for the lack of catastrophic interference effects in real learners. Entrenchment alone cannot explain age-related effects in L2 acquisition.

The Unified Competition Model provides an alternative account of declines in L2 learning outcomes that focuses on the interplay between the risk factors (entrenchment, transfer, overanalysis, and isolation) and support factors (resonance, decoupling, chunking, and participation). The basal ganglia can recombine entrenched cortical patterns to form new L2 structures, thereby permitting decoupling of L2 from L1 and a gradual escape from L1 entrenchment.

The successful application of each of the support factors requires conscious manipulation of input, metalinguistic reflection, good linguistic and pedagogical analysis, study of L2 materials, and ongoing spoken interactions in the new language. To facilitate this process, L2 instruction can provide focused instruction on basic language skills using online learning tutors (see e.g. http://sla.talkbank.org) along with community support for "language learning in the wild" (Clark, Wagner, Lindemalm, & Bendt, 2011) based on language learning tours, QR codes, translator apps, and on-site recordings, all blended with and linked to support from classroom instruction (MacWhinney, 2015).

References

Aram, D. M., & Eisele, J. A. (1994). Intellectual stability in children with unilateral brain lesions. *Neuropsychologia, 32*, 85-95.

Asato, M. R., Terwilliger, R., Woo, J., & Luna, B. (2010). White matter development in adolescence: a DTI study. *Cerebral Cortex, 20*, 2122-2131.

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

Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning, 63*(s1), 68-86.

Basser, L. S. (1962). Hemiplegia of early onset and the faculty of speech with special reference to the effects of hemispherectomy. *Brain, 85*, 427-460.

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

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. F. Kroll & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York, NY: Oxford University Press.

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

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Booth, J., MacWhinney, B., Thulborn, K., Sacco, K., Voyvodic, J., & Feldman, H. (2000). Developmental and lesion effects during brain activation for sentence comprehension and mental rotation. *Developmental Neuropsychology, 18*, 139-169.

Bradlow, A., Akahni-Yamada, R., Pisoni, D., & Tokhura, Y. i. (1999). Training Japanese liseners to identify English /r/ and /l/: Long term retention of learning perception and production. *Perception and Psychophysics, 61*, 977-985.

Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language, 75*, 720-738.

Browman, H. I. (1989). Embryology, ethology and ecology of ontogenetic critical periods in fish. *Brain, behavior and evolution, 34*(1), 5-12.

Bybee, J., & Brewer, M. (1980). Explanation in morphophonemics: Changes in Provencal and Spanish preterite forms. *Lingua, 52*, 271-312.

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

Carpenter, G., Grossberg, S., & Reynolds, J. (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks, 4*, 565-588.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54*, 241-257.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3-42.

Clark, B., Wagner, J., Lindemalm, K., & Bendt, O. (2011). *Språkskap: Supporting second language learning "in the wild"*. Paper presented at the INCLUDE, London.

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

Dominey, P. F., & Boussaoud, D. (1997). Encoding behavioral context in recurrent networks of the fronto-striatal system: a simulation study. *Cognitive Brain Research, 6*(1), 53-65.

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

Doria, V., Beckman, C. F., Arichi, T., & Merchant, N. (2010). Emergence of resting state networks in the preterm human brain. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 20015-20020.

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

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71-99.

Engvig, A., Fjell, A. M., Westlye, L. T., Moberget, T., Sundseth, Ø., Larsen, V. A., & Walhovd, K. B. (2012). Memory training impacts short-term changes in aging white matter: A Longitudinal Diffusion Tensor Imaging Study. *Human brain mapping, 33*(10), 2390-2406.

Feldman, H., MacWhinney, B., & Sacco, K. (2002). Sentence processing in children with early unilateral brain injury. *Brain and Language, 83*, 335-352.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 229-273). Baltimore, MD: York Press.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*(1), 1-34.

Folse, K. S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. Ann Arbor, MI: University of Michigan Press.

Friederici, A., Steinhauer, K., & 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.

Gao, W., Lin, W., Chen, Y., Gerig, G., Smith, J., Jewells, V., & Gilmore, J. (2009). Temporal and spatial development of axonal maturation and myelination of white matter in the developing brain. *American journal of neuroradiology, 30*(2), 290-296.

Graybiel, A. M. (1995). Building action repertoires: memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology, 5*(6), 733-741.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of learning and memory, 70*(1), 119-136.

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

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

Hadley, J., & Healy, A. (1991). When are reading units larger than the letter? Refinement of the unitization reading model. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 1062-1073.

Hahne, A., & 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.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical period hypothesis for second language acquisition. *Psychological Science, 14*, 31-38.

Hamker, F. H. (2001). Life-long learning Cell Structures -- continuously learning without catastrophic interference. *Neural Networks, 14*, 551-573.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*, 67-99.

Hinaut, X., & Dominey, P. F. (2013). Real-time parallel processing of grammatical structure in the fronto-striatal system: a recurrent network simulation study using reservoir computing. *PLoS ONE, 8*(2), e52946.

Hubel, D., & Weisel, T. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology, 26*, 994-1002.

Ingvalson, E., Holt, L., & McClelland, J. (2012). Can native Japanese listeners learn to differentiate /r-l/ on the basis of F3 onset frequency? *Bilingualism: Language and Cognition, 15*, 255-274.

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*, 60-99.

Kapatsinski, V., & Radicke, J. (2009). Frequency and the emergence of prefabs: Evidence from monitoring. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatley (Eds.), *Formulaic language. Volume 2: Acquisition, loss, psychological reality, and functional explanations* (pp. 499-522). Amsterdam: John Benjamins.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences, 11*, 229-235.

Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences, 35*, 111-122.

Kraus, N., McGee, T., Carrell, T. D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience, 7*(1), 25-32.

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

Kuhl, P. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience, 5*(11), 831-843.

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

Kürten, J., De Vries, M. H., Kowal, K., Zwitserlood, P., & Flöel, A. (2012). Age affects chunk-based, but not rule-based learning in artificial grammar acquisition. *Neurobiology of aging, 33*(7), 1311-1317.

Lenneberg, E. (1967). *Biological foundations of language*. New York, NY: Wiley.

Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks, 17*, 1345-1362.

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

Lorenz, K. Z. (1958). The evolution of behavior. *Scientific American, 199*, 95-104.

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

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

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

MacWhinney, B. (2012). The logic of the Unified Model. In S. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 211-227). New York: Routledge.

MacWhinney, B. (2014). Item-based patterns in early syntactic development. In T. Herbst, H.-J. Schmid, & S. Faulhaber (Eds.), *Constructions collocations patterns* (pp. 33-70). Berlin: de Gruyter Mouton.

MacWhinney, B. (2015). Multidimensional SLA. In S. Eskilde & T. Cadierno (Eds.), *Usage-based perspectives on second language learning* (pp. 22-45). New York, NY: Oxford University Press.

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

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

Marler, P. (1991). Song-learning behavior: the interface with neuroethology. *Trends in Neuroscience, 14*, 199-206.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. (2002). Success and failure in teaching the /r/-/l/ contrast to Japanese adults: Test of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience, 2*, 89-108.

McCauley, S., Monaghan, P., & Christiansen, M. (2014). Language emergence in development: a computational perspective. In B. MacWhinney & W. O'Grady (Eds.), *Handbook of Language Emergence* (pp. 415-436). New York, NY: Wiley.

McClelland, J., McNaughton, B., & O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419-457.

McCloskey, M., & Cohen, N. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. Bower (Ed.), *The psychology of learning and motivation* (Volume 24, pp. 109-166). New York, NY: Academic Press.

Mills, D., Coffey-Corina, S., & Neville, H. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology, 13*, 397-445.

Molfese, D. L., Freeman, R. D., & Palermo, D. S. (1975). The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain and Language, 2*, 356-368.

Müller, J., Hahne, A., Fujii, Y., & 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.

Munakata, Y., McClelland, J., Johnson, M., & Siegler, R. (1997). Rethinking infant knowledge: An adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104*, 686-713.

Munro, M. J., Flege, J. E., & Mackay, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics, 17*, 313-334.

Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus, 10*(4), 352-368.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

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

Newell, A. (1990). *A unified theory of cognition*. Cambridge, MA.: Harvard University Press.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex, 13*, 155-161.

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

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

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

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27*, 59-70.

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

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

Posner, M. I., Tang, Y.-Y., & Lynch, G. (2014). Mechanisms of white matter change induced by meditation training. *Frontiers in Psychology, 5*.

Presson, N., Sagarra, N., MacWhinney, B., & Kowalski, J. (2013). Compositional production in Spanish second language conjugation. *Bilingualism: Language and Cognition, 16*, 808-828.

Prior, A., MacWhinney, B., & 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.

Richardson, F., & Thomas, M. (2008). Critical periods and catastrophic interference effects in the development of self-organizing maps. *Developmental Science, 11*, 371-389.

Roelofs, A. (2011). Modeling the attentional control of vocal utterances: from Wernicke to WEAVER+. In J. Guendozi, F. Loncke, & M. Williams (Eds.), *The handbook of psycholinguistic and cognitive processes: Perspectives in Communication Disorders* (pp. 189-208). New York, NY: Psychology Press.

Rohde, D. C., & Plaut, D. (1997). Simple recurrent networks and natural language: How important is starting small? *Proceedings of the 19th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*(8), 2022-2039.

Sagarra, N., & Herschensohn, J. (2011). Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning, 61*, 80-116.

Sampaio-Baptista, C., Khrapitchev, A. A., Foxley, S., Schlagheck, T., Scholz, J., Jbabdi, S., ..., Thomas, N. (2013). Motor skill learning induces changes in white matter microstructure and myelination. *The Journal of Neuroscience, 33*(50), 19499-19503.

Segalowitz, S., Segalowitz, N., & Wood, A. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics, 19*, 53-67.

Seidenberg, M. S., & Zevin, J. D. (2006). Connectionist models in developmental cognitive neuroscience: Critical periods and the paradox of success. *Attention & Performance XXI: Processes of change in brain and cognitive development*, 585-612.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition, 55*, 151-218.

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.

Snow, C. E., & Hoefnagel-Hohle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development, 49*, 1114-1128.

Sosa, A., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word *of. Brain and Language, 83*, 227-236.

Stiles, J., Trauner, D., Engel, M., & Nass, R. (1997). The development of drawing in children with congenital focal brain injury: Evidence for limited functional recovery. *Neuropsychologia, 35*, 299-312.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., ..., Kawashima, R. (2010). Training of working memory impacts structural connectivity. *The Journal of Neuroscience, 30*(9), 3297-3303.

Tokowicz, N., & 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, MA: Harvard University Press.

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

VanPatten, B., Keating, G. D., & Leeser, M. J. (2012). Missing verbal inflections as a representational problem: Evidence from self-paced reading. *Linguistic Approaches to Bilingualism, 2*(2), 109-140.

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

Werker, J. (1995). Exploring developmental changes in cross-language speech perception. In L. Gleitman & M. Liberman (Eds.), *An Invitation to Cognitive Science. Language* (Volume 1, pp. 87-106). Cambridge, MA: MIT Press.

Werker, J., & Tees, R. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology, 46*, 233-251.

Wiley, E., Bialystok, E., & Hakuta, K. (2005). New approaches to using census data to test the Critical-Period Hypothesis for second language acquisition. *Psychological Science, 16*, 341-343.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., ..., Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology, 36*(5), 638-642.

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

Yamamoto, K., Hoffman, D. S., & Strick, P. (2006). Rapid and long lasting plasticity of input output mapping *Journal of Neurophysiology, 96*, 2797-2798.

Zevin, J. D. (2012). A sensitive period for shibboleths: The long tail and changing goals of speech perception over the course of development. *Developmental Psychobiology, 54*(6), 632-642.

Zhou, X., Panizutti, R., de Villers-Sidani, É., Madeira, C., & Merzenich, M. (2011). Natural restoration of critical period plasticity in the juvenile and adult primary auditory cortex. *The Journal of Neuroscience, 31*, 5625-5634.