Carnegie Mellon University
Research Showcase @ CMU

Department of Psychology | Dietrich College of Humanities and Social Sciences

2009

Emergent Fossilization

Brian MacWhinney
Carnegie Mellon University, macw@cmu.edu

Follow this and additional works at: http://repository.cmu.edu/psychology

This Book Chapter is brought to you for free and open access by the Dietrich College of Humanities and Social Sciences at Research Showcase @ CMU. It has been accepted for inclusion in Department of Psychology by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Emergent Fossilization

Brian MacWhinney

Psychology, CMU

In: Perspectives on Fossilization, Han, Z. and Odlin, T. Clevedon, UK: Multilingual Matters. Mail to: Editorial Manager, Multilingual Matters Ltd, Frankfurt Lodge, Clevedon Hall, Victoria Road, Clevedon, BS21 7HH, England. 8000 words

Aging is an inescapable fact of human life. In most areas of our lives, aging leads to an unremarkable, gradual decline in physical ability. For example, no one questions why a 40-year-old runner can no longer compete in the Olympics. As the body ages, metabolism slows, joints wear out, and energy is diminished. Aging also has a uniform natural effect on the learning of new skills. As a result, we are not surprised to find that someone who tries to learn soccer at age 30 makes less progress than someone who begins learning at 12. We approach these gradual age-related physical declines and losses in learning ability with equanimity, since few of these skills are crucial for everyday functioning. No one would suggest that these declines represent the sudden expiration of a some innate ability linked to a specific biological time fuse.

However, when we look at the decline in language learning abilities that comes with age, we assume a somewhat different position. We are distressed to find that a 35-year-old Romanian immigrant to the United States is unable to lose her Romanian accent,

saying that this may limit her ability to adjust to the new society. We may wonder whether the observed fossilization represents the final expiration of some special gift for language learning. Or we may worry that a 28-year-old graduate student from Japan has trouble learning to use English articles. If these error patterns continue year after year, we say that the language spoken by these immigrants has "fossilized" (Selinker, 1972).

Fossilization can also affect the young. For example, American university students in Japan in their twenties often make good progress in learning for two or three years and then level off before attaining native speaker competence. This level of partial fossilization for young immigrants is seldom reported for students visiting other countries, leading one to wonder what features of Japanese language and culture may be inducing these problems. We should distinguish this type of true immigrant fossilization from cases of incomplete learning or forgetting when students go abroad to study a language and then later show fossilized or diminished abilities when they return home and cease regular practice of the language.

In truth, fossilization is not an across-the-board phenomenon (Birdsong, this volume; Han, this volume). Rather, we find continual growth in some areas and relative stability of error in others. For example, older "fossilized" Hungarian learners of English may continue to pick up new verbs, constructions, and phrases, while continuing to pronounce English *water* as *vater*. Somehow, we tend to focus our attention more on these ongoing errors than the continuing new acquisitions. However, for those particular areas that show little change, it is accurate enough to think about localized fossilization.

Critical Periods

One traditional approach treats fossilization as a consequence of the expiration of a critical period for language acquisition. Customarily, this idea has been referred to as the Critical Period Hypothesis (CPH). For biologists, the concept of a critical period in grounded on studies of the maturation of tissues in the embryo. For example Sperry (1956) showed how eye cells in the frog embryo could be transplanted at an early period and still form the correct pattern of connectivity with the brain. This happens because the cells are induced into appropriate connectivity by the surrounding tissue. However, if the eyes are transplanted after the critical period, then they will have committed to their previous position and will wire up incorrectly to the brain.

These embryological characterizations of critical periods depend on an understanding of the unfolding of the epigenetic landscape(Waddington, 1957) during the embryogenesis. By extension, similar processes are thought to occur in infant animals during the first days of life. For example, greylag geese will imprint on the first face they see after hatching, whether it be that of their mother or Konrad Lorenz (1958). Salmon hatchlings will imprint on the location of their home pond for later breeding.

A fundamental difference between prenatal and postnatal critical periods is that the latter require specific external stimuli as input. Because of this, we have to speak about experience-dependent or experience-expectant processes during the postnatal period. Eventually, as the critical period widens and the shape of the triggers broadens, we begin to talk about sensitive periods rather than critical periods. As we move away from the traditional embryological critical period to the postnatal sensitive period, the biochemical basis of the period often becomes increasingly complex. This is not to say that epigenesis

is necessarily limited to the embryo and the infant. Stages such as puberty and menopause could well form the backdrop for critical period events later in life. For example, the rapid increase in members of the opposite sex that we find in teenagers could be analogized to the critical period for the greylag geese. However, hopefully young boys or girls will not immediately and irrevocably imprint on their true love on the first day of puberty.

A further problem with the notion of a critical period is that, without further definition and analysis, it would apply equally across all linguistic levels and systems. However, all of us – linguists, psychologists, and educators alike – would agree that language involves control of a diverse set of systems for articulation, audition, lexicon, grammar, and meaning. It is difficult to imagine how a single biological mechanism could have a uniform impact across all of these systems.

We can avoid all of these conceptual and empirical problems by speaking about the effects of the Age of Arrival (AoA) variable, rather than the expiration of critical period. The AoA variable measures the age at which an immigrant arrives a new country and begins serious exposure to L2. It is AoA, rather than length of residence, that most strongly predicts the extent of achievement of nativelike proficiency in L2 (Birdsong, in press). By focusing our attention on accounting for AoA effects, rather than critical period effects, we create a level playing field for the equal consideration of neurological, psychological, physiological, and sociological determinants of localized fossilization.

Possible Accounts

We would like to be able to construct and test models that account for the observed AoA and fossilization patterns using clearly stated mechanisms for processing and

learning. Fortunately, the last decade has seen a proliferation of possible accounts. We can now distinguish at least these ten concrete proposals.

1. the lateralization hypothesis (Lenneberg, 1967),
2. the neural commitment hypothesis (Lenneberg, 1967),
3. the parameter-setting hypothesis (Flynn, 1996),
4. the metabolic hypothesis (Pinker, 1994),
5. the reproductive fitness hypothesis (Hurford & Kirby, 1999),
6. the aging hypothesis (Barkow, Cosmides, & Tooby, 1992),
7. the fragile rule hypothesis (Birdsong, in press),
8. the starting small hypothesis (Elman, 1993),
9. the entrenchment hypothesis (Marchman, 1992),
10. the entrenchment and balance hypothesis (MacWhinney, in press),

My goal in the current paper is to evaluate each of these ten proposals against the whole range of age-related effects in language learning, including not only fossilization, but also earlier changes throughout the lifespan.

We could classify the first seven of these accounts as supporting a nativist view of language learning and the last three (or perhaps four) as tending toward empiricism or emergentism. However, little is to be gained from this characterization, since each of the views makes reference both to biological and psychological processes (MacWhinney, 2002). Thus, rather than trying to extract a set of binary features for evaluating these models (Newell, 1973), let us take a detailed look at each proposal in turn.

The lateralization hypothesis

The idea that AoA effects arise from maturational changes in the brain related to lateralization of function was first developed systematically by Lenneberg (1967). Lenneberg viewed that the two cerebral hemispheres as equipotential at birth for the acquisition of language. However, over time, the left hemisphere assumed dominance for language functions. Lenneberg viewed this process of gradual lateralization as providing biological limits on first and second language acquisition. He placed great emphasis on evidence from children undergoing hemispherectomy to correct epileptic seizures which showed that, up until age 13, language could be relearned after removal of the let hemisphere (Basser, 1962). The idea here is that the massive changes that occur at puberty serve to terminate the language learning abilities of the child by finalizing the process of lateralization. Johnson and Newport (1989) and others have pointed to the onset of puberty as the defining moment in terms of the loss of an ability to acquire nativelike L2 skills. However, it is not clear why a growth in lateralization should, by itself, lead to a decline in the ability to acquire a first or second language. One model, suggested in the 1970s (Hardyck, Tzeng, & Wang, 1978) was that the first language would be organized to the left hemisphere and that the second language would be represented in the right hemisphere. According to this account, continued plasticity of the right hemisphere was required for nativelike learning of L2. Once this plasticity disappeared, nativelike L2 learning would become impossible. However, this account has not received consistent empirical support. Moreover, recent evidence (Muter, Taylor, & Vargha-Khadem, 1997; Vargha Khadem, Isaacs, & Muter, 1994) has called into question the extent to which equipotentiality vanishes, even by puberty.

Research on language development during the first three years has demonstrated a variety of lateralization-related effects. There is abundant evidence suggesting that the left hemisphere is the preferred locus for speech from birth (Dennis & Whitaker, 1976; Kinsbourne & Hiscock, 1983; D. Molfese, Freeman, & Palermo, 1975; D. Molfese & Hess, 1978; V. Molfese & Betz, 1987; Wada, Clarke, & Hamm, 1975; Witelson, 1977; Witelson & Pallie, 1973; Woods & Teuber, 1977). However, dominance may shift to the right, if there are insufficient resources in the left hemisphere to allow it to perform its usual function (Zangwill, 1960). The timing of this damage will interact with the ongoing and progressive (Luria, 1973) process of language organization during the first years of life (Satz, Strauss, & Whitaker, 1990).

Together these studies point to important changes in lateralization for language during the first years of life. However, there is little in this literature suggesting that changes in lateralization are major determinants of age-related effects after the first three years. One exception is an fMRI study from our group (Booth et al., 2001) showing somewhat stronger lateralization for adults compared to 10-year-olds for the processing of difficult embedded relative clauses. However, this effect may be due to late learning of these structures, rather than overall changes in lateralization.

The commitment hypothesis

A second proposed mechanism for AoA effects involves a progressive maturational commitment of language areas to linguistic functioning. Recent advances in the study of neural plasticity and commitment make accounts of this type increasingly attractive. This hypothesis would claim that some specific brain region or set of regions is undergoing commitment in parallel with the period during which AoA effects are demonstrated for

learners. Data reviewed by Flege (1999) show a decline in the ability to acquire nativelike articulatory competency, beginning at age 5. The commitment hypothesis could account for these results if we could show that some area of the brain involved in the acquisition of phonology undergoes progressive commitment during this period.

To demonstrate this effect, we could conceivably use functional magnetic resonance imaging (fMRI) to track activation of brain areas across the age range of 4 to 20. Ideally, these fMRI studies would focus on the control of articulation. However, because articulation produces movement artifacts in the scanner, this study is not currently possible. Alternatively, the stimuli could be auditory sounds or words. We would expect that regions activated for these stimuli would become progressively narrower across this age range (Booth et al., 2001; Haier, 2001).

It is likely that we will eventually be able to demonstrate some commitment effects during the relevant period. We already know that there are increases in myelinization (Lecours, 1975) and white matter commitment (McArdle et al., 1987) up to age 7. These changes might well coincide with data on AoA effects for articulation (Flege et al., 1999), but they may not match up as well for earlier effects in audition (Werker, 1995) and later effects in lexical learning (Snow & Hoefnagel-Hohle, 1978).

The parameter setting hypothesis

Generative linguists (Flynn, 1996) have proposed a very different approach to AoA effects. This account suggests that first language learning depends on a process of parameter setting (Chomsky, 1981) specified by universal grammar (UG). For example, languages like English require that each verb have a subject, whereas languages like

Italian allow dropping of pronominal subjects. Generative grammar holds that children use a set of syntactic triggers to determine the correct set of parameters for their language. Once all of these settings are made for some small set of parameters, the language is fully identified. Years later, when the learner tries to learn a second language, these same parameters could perhaps be reset and be used to identify that second language.

The problem with this first version of L2 parameter setting is that it fails to predict any AoA effects, since resetting for L2 would not be influenced by transfer from L1 or any loss in ability to access UG. To deal with this, some UG researchers have favored a second model that holds that the parameters of UG that were available to the child are no longer available to the older second language learner. This account does not propose any specific biological mechanism for this maturational phenomenon, although one could invoke either commitment or lateralization as possible mechanisms.

The major problem with this second account, as currently articulated, is that there are no independent theoretical grounds for understanding why a given parameter would become either available or unavailable for the L2 learner. Lacking any independent biological or psychological grounding, this theory must rely on linguistic constraints to determine these predictions. If the constraints are set on the basis of evidence derived from L2 learning and fossilization, then this second version becomes circular and vacuous.

Both versions of the UG parameter-setting analysis suffer from another problem. This is the fact that UG only determines certain core features of a grammar, such as branching direction or binding, leaving other features such as agreement or case to the periphery.

However, both core UG features and peripheral non-UG features are subject to AoA effects (Johnson & Newport, 1991; McDonald, 2000). Thus, we would have to supplement any UG theory of AoA and fossilization with a parallel external theory, perhaps following different principles.

The metabolic hypothesis

Pinker (1994) has proposed an innovative account of AoA effects based on notions from cognitive neuroscience and evolution. He suggests (p. 293) that "a decline in metabolic rate and the number of neurons during early school-age years" is a probable cause of the loss of language-learning ability. However, Pinker's analysis places the neurological cart before the neurological horse. It is certainly true that both subjects show greater metabolic activity for problems that are being learned (Haier, 2001; Merzenich, 2001) and less for ones that are fully controlled. However, this difference is not one that is shut down after early childhood. Rather, whenever learners – even older adults – are confronted with new tasks they show widespread metabolic activity during learning and narrower regions of activation for automated tasks (Raichle et al., 1994). Thus, there is no evidence for any overall loss in metabolic activity of the type Pinker is suggesting.

The reproductive fitness hypothesis

Hurford and Kirby (1999) present an account of AoA effects in both first and second language learning based on an analysis of evolutionary considerations. They reason that, over the course of human evolution, the attainment of complete fluency in a first language was a major determinant of reproductive fitness. If a child had not successfully acquired language by the age of sexual maturity, they would not be as attractive to a

sexual partner and would therefore be less likely to produce offspring. Conversely, those who had acquired a high level of language ability would be highly attractive and would reproduce. Hurford and Kirby simulate this effect by imagining that a given target L1 has a fixed size expressed as a number of units. Children with a given amount of language learning ability acquire a fraction of these total units on each learning cycle. Because there is no advantage to completing learning before puberty, the simulation settles in with a pattern in which most learners acquire the full language just before puberty.

Initially, it would seem that assumptions and claims of this model are strikingly at odds with established wisdom from the field of child language. Child language researchers often consider that the core of a language is learned by age 4 (Brown, 1973). In fact some generativists (Poeppel & Wexler, 1993; Wexler, 1998) believe that full competence is acquired even by the time of the first productive syntactic combinations. However, Hurford and Kirby's analysis can also be given an interpretation that escapes from these problems. This is not an interpretation that they themselves present, but it does seem consistent with their account. According to this extended co-evolutionary account, full control of a language involves acquisition of a variety of non-core language-related skills such as oratory, poetry, song, reading, and verbal memory. To the degree that progressive advances in social structure over the last 10,000 years have led to an increased evolutionary pressure for attainment of these extended linguistic abilities, we should indeed expect to see an continuing movement toward consolidation of these skills just before puberty.

Unlike Pinker, Hurford and Kirby are interested in determining the age by which language must be learned, rather than the age after which a second language cannot be

learned. There is little in their analysis that would suggest that the ability to learn language should decline after puberty, only that first language learning must be completed by puberty. In fact, evolutionary considerations suggest that there should be strong pressures in some groups for ongoing maintenance of language learning abilities. For example, many tribal groups in Australia, Southeast Asia, and South America practice a form of alternating bride exchange between villages. In such cases, husband and wife will often speak different dialects or even different languages. Cases of this type show that there is continuing co-evolutionary pressure toward maintenance of language learning abilities into adulthood. Moreover, the fact that dominant males continue to procreate up to age 60 and beyond suggests that there should be evolutionary pressure away from a decline in language learning loss during adulthood.

These pressures for reproductive diversity and exogamy are in direct competition with another set of pressures forcing group loyalty and cohesion (MacWhinney, 2004). These pressures tend to favor mates from a related lineage, preferring endogamy to exogamy. The clearest way of establishing in-group membership is to lock in on a phonological accent, perhaps during early childhood. In this sense, a phonological accent functions in human communities much like dialect variation in song birds (Marler, 1991). However, if reproductive fitness were only conditioned on conservative maintenance of a local dialect, it would be difficult to account for the fact that reproductively attractive adolescent males often drive processes of phonological change (Labov, 1994)

Because of the diversity of these various pressures and the complexity of written languages with rich literatures, it is difficult to imagine that any single evolutionary mechanism would determine all aspects of adult linguistic ability. However, the wide

range of individual differences in successful mastery of a second language after early children does indeed suggest that a variety of fairly recent evolutionary pressures have been operating to produce the observed population diversity.

The aging hypothesis

The most uncontroversial account of AoA and fossilization effects is one that emphasizes the physiological and neurological changes that occur with aging. As they age, most people begin to experience a marked slowdown in metabolic activity, energy, and flexibility. Hormonal processes slow down; arteries become blocked by platelets leading to strokes; Parkinsonism produces a loss of motor control; hearing acuity diminishes; and rheumatism and osteoporosis can lead to physical collapse. Core cognitive functions such as the storage of new memories and the retrieval of old memories can be disrupted by degeneration in the hippocampus and temporal lobe (Scheibel, 1996).

Although we typically think of these effects are impacting people over age 65, many of these effects begin even by age 45. However, even if we allow for an early onset of some of this decline, it seems difficult to attribute all AoA effects to aging. On the one hand, there is a clear and progressive decrement in language attainment that occurs well before 45. On the other hand, there are clear cases of partially successful language learning after age 65. I can relate a case I observed my Romanian relative from the city of Arad who began learning English when he retired at age 65. Married to a Hungarian woman in this bilingual community, he had spoken fluent Romanian and Hungarian all his life, and had some knowledge of German. With no input from native speakers, he began learning English from textbooks and dictionaries at 65. At first his speaking and

writing was extremely difficult to follow. However, after 10 years of practice, particularly through letters mailed to relatives in the United States, his speaking and writing became increasingly comprehensible. In fact, he continued to progress in his language learning until his death at age 7 with no evidence of fossilization. Cases of this show that, although biological wear and tear undoubtedly leads to loss of language learning ability in some speakers, not all elderly people are equally impacted.

This illustrates important principle. Many researchers in SLA would like to identify a single hypothesis that could account single-handedly for all of the observed age-related effects in language learning. Aging is certainly not going to account for all age-related effects, since it tells us nothing about changes before 45. However, it is clearly a contributory factor to some fossilization. Therefore a plausible account will have to show how aging combines with other factors. In general, our goal must be to provide independent support for each factor that we believe contributes to fossilization and to see how fossilization patterns emerge from these combinations.

The fragile rote hypothesis

Birdsong (in press) suggests that, with increasing age, learners may have more problems acquiring irregular forms as opposed to regular forms. He suggests that irregular forms may include not only words with irregular inflections, but also irregular use of particles and prepositions in phrasal verbs. Birdsong seeks to ground the decline in learning of irregulars on neruoanatomical changes in the parts of the brain subserving the declarative memory system.

Currently, there is little data available to evaluate this hypothesis. Even in first language learners, irregulars can pose problems. However, it does seem reasonable to believe that these problems may increase in adult learning. On the other hand, the declarative memory loss that Birdsong invokes only plays a major role after age 50. Therefore, this account leaves unexplained the declines in language learning with AoA that we see before that time as well as changes in attainment of nativelike fluency that occur during childhood. If this hypothesis could be grounded on a mechanism that may change earlier in life, such as auditory memory, it might be able to do a better job accounting for observed patterns in age-related effects on language learning.

The starting small hypothesis

Having now considered seven hypotheses that emphasize biological determinants of AoA effects, we can now turn to a consideration of three hypotheses that go beyond biology to consider the role of psychological processing mechanisms. The first hypothesis we will examine in this area is the "starting small" hypothesis of Newport (1990). Newport argued that "language learning declines over maturation precisely because cognitive abilities increase." The idea is that children have a smaller short-term memory span and that this shorter span makes it more difficult for them to store large chunks of utterances as formulaic items. As a result, children are forced to analyze language into its pieces.

Before turning to an analysis of the predictions of this model, it is worth noting that the idea that there is a growth in working memory during childhood is well accepted and documented by developmental psychologists (Halford et al., 1995; MacWhinney, Feldman, Sacco, & Valdes-Perez, 2000). Although a basic growth in working memory

capacity is not disputed, it is not at all clear how this growth in memory capacity translates into changes in processing of sentences. In particular, children appear to have an auditory memory roughly equivalent to that of adults (Aslin, Saffran, & Newport, 1999). On the other hand, their ability to pick up new lexical items is more limited (Gupta & MacWhinney, 1997; Snow & Hoefnagel-Hohle, 1978). It is unclear whether adults differ from children in terms of ability to piece together syntactic strings. How these discrepant abilities on these different levels interact with the growth of short-term memory capacity is unclear.

Whatever its claims about these interactions, the starting small hypothesis makes at least three clear predictions. The first is that young children will not acquire complex multimorphemic words, preferring instead to pick up monosyllables. This claim is difficult to evaluate in languages like English or Chinese where the majority of words are only one or two syllables long. However, in languages like Hungarian (MacWhinney, 1974), Inuktitut (Crago & Allen, 1998), or Navajo, children seldom reduce four or five syllable words to single syllable components. Moreover, we often find the opposite effects in which young children learning English pick up formulaic utterances and clusters of words (Clark, 1974; Peters, 1983). Thus, it seems clear that this first basic prediction is at least partially incorrect.

A second prediction of this model is be that the size of formulaic chunks will increase with age. This prediction seems to be supported in studies of school-age immigrant children learning English as a second language (Wong-Fillmore, 1976). These children do indeed pick up a variety of phrasal formulas such as “why don’t you __” or “I wish I could just ___.” However, the 8 and 10-year-old children who show these patterns are

likely to end up learning their new language quite well, with non-native features being more evident in phonology than in sentence construction. Thus, although this prediction is generally supported, it fails to match up well to the overall course of age-related effects in second language learning.

A third prediction of the starting small hypothesis is that adults should be particularly good at acquiring larger phrasal chunks. For example, learners of German should be successful in picking up phrases such as *ins Mittelalter* or *des meine Mannes*, rather than learning these phrases as combinations of *in*, *das*, *Mittelalter*, *meine*, and *Mannes*. However, if adult learners were really using their memory to analyze these phrases as chunks, they be able to use this database of adjectives and prepositions with nouns to acquire accurate use of German gender-number-case marking. In fact, adult learners have terrible trouble learning German gender, because they pick up each new noun as a separate analyzed unit, rather than as a part of a richer phrase. It is children who pick up the longer phrases and succeed thereby in acquiring correct use of gender. Thus, this third prediction of the model also seems to be wrong.

A fourth prediction of the model is that children with a weaker working memory should actually learn language better than those with a larger memory. It has often been shown that children with Specific Language Impairment have a smaller working memory capacity. This limitation should actually be a strength, but there is no evidence that learning works this way.

Given the empirical problems facing this model, one might wonder why it has received so much attention. In part, this may be a result of the fact that there have been several successful attempts to model this process (Goldowsky & Newport, 1993; Kareev,

1995) (Elman, 1993) and also some failures to replicate the successes (Rohde & Plaut, 1999). However, these models all present a rather limited view of the acquisitional process, focusing on the learning of abstract syntactic patterns from 'predigested" input. For such abstract patterns, it may well be the case that a memory filter furthers attention to certain covariation patterns. Certainly, adults will do better at the extraction of lexical forms and rules if they are not overloaded with too much input (Cochran, McDonald, & Parault, 1999; Kersten & Earles, 2001). However, the assistance that this filtering provides may be counterproductive in the long run, blocking the acquisition of larger chunks as in the case of the German example given above.

Overall, we can say that filters could be useful in terms of forcing attention to covariation patterns. However, a more powerful approach to language learning combines filtering with an opposed ability to pick up large unanalyzed chunks. More work is needed to see how these abilities change across the lifespan.

The entrenchment hypothesis

The most intuitive account for fossilization focuses on the notion of entrenchment. When we practice a given skill thousands of times, we soon find that it has become automated or entrenched. The more we continue to practice that skill, the deeper the entrenchment and the more difficult it becomes to vary or block the use of the skill. Entrenchment occurs in neural networks when a high frequency pattern is presented continuously in the input training data. For example, Marchman (1992) shows how irregular morphological forms such as *went* can become entrenched in a network learning English morphology. The entrenchment of a form such as *went* can serve to block overregularizations such as *goed*.

Entrenchment can be observed in many areas of our lives. Consider the case of a Hungarian peasant who has learned the dance forms of the Hungarian plain or Alföld with its csárdás steps and emphasis on straight posture in couple dances. For that dancer, the more flowery, style of Transylvania with its leaps and twists will be a bigger challenge. If the dancer only begins to learn the Transylvanian style after age 35, having danced the Alföld style for 25 years, there will be an unmistakable Alföld dance accent (Sándor Timár, folk ethnographer, personal communication) and we may even see a fossilization of dance style. If, however, the dancer learns both styles early on, then there will be no clear fossilization and the dancer will be a "balanced bilingual" in dance styles. However, if that same bilingual now turns to learning Thai dance at age 40, we can expect that the entrenchment of Hungarian dance styles and postures will produce strong fossilization during the learning of Thai patterns.

Within a single system, entrenchment can work smoothly to block overregularizations and speed responses. However, when there is a radical shift in the input to an entrenchment system, neural network systems can suffer from "catastrophic interference." For example, Janice Johnson and I taught a neural network to assign grammatical roles to 178 different English sentence patterns. Once the network had learned these patterns, we shifted training to Dutch. However, the shift to a second language led to a catastrophic decline in the network's ability to process English, even though there were units in the training corpora that clearly identified the language of each sentence.

Within a single system, entrenchment can work smoothly to block overregularizations and speed responses. However, when there is a radical shift in the input to an

entrenchment system, neural network systems can suffer from "catastrophic interference." For example, Janice Johnson and I taught a neural network to assign grammatical roles to 178 different English sentence patterns. Once the network had learned these patterns, we shifted training to Dutch. However, the shift to a second language led to a catastrophic decline in the network's ability to process English, even though there were units in the training corpora that clearly identified the language of each sentence.

Not all neural networks are subject to catastrophic interference. Architectures that use local organization, such as self-organizing maps (SOM), can pick up new words as variants of old forms. Consider a feature map that has already encoded the word *table* on both a phonological map and a separate semantic map (Li, Farkas, & MacWhinney, in press) along with an association between the two maps. When this system begins to learn the Spanish word *mesa*, it will enter first into the phonological map as just a new word (although possibly with English rather than Spanish phonology). This form will then be associated to the preexisting pattern for *table* in the semantic map. In this form of learning, *mesa* becomes parasitic on the meaning of *table*, because it is acquired simply as another way of saying *table*.

However, if we turn to syntactic learning, the problem is more serious. In that area, most models rely on back propagation. For example, in order to avoid catastrophic interference, Johnson and MacWhinney had to interleave training of Dutch and English from the beginning. This will work for some purposes, but it is clearly not a general solution to the problem, since many cases of real second language learning involve major shifts in language environment. To deal with these problems, neural network models of

syntactic learning will need to shift to a lexicalist focus, as discussed in MacWhinney (2000). This focus emphasizes the extent to which syntax can be controlled through item-based patterns. For example, when learning the Spanish adjective grande, the system will not only encode the new word as a variant of English *big*, but will also encode its positioning as following the noun as in *una mesa grande*. This item-based positional pattern is encoded directly as a property of grande that does not interfere with related patterns for English. Eventually, a set of item-based patterns of this type will yield a general construction that places Spanish adjectives after their nouns, but this construction will be dependent on the collection of new Spanish word forms and will not interfere with the existing English system for placing the adjective before the noun. I should note that we have not yet implemented this system, but the nature of activation patterns in self-organizing maps indicate that it should operate in this way.

The entrenchment and balance hypothesis

We see then that catastrophic interference can be solved by systems that emphasize the lexical and item-based nature of second language learning. Moreover, these lexically-grounded systems can also illustrate another important aspect of second language learning. This is the parasitic nature of L2 learning when L1 is already well consolidated. Parasitism occurs because the L1 form is already well consolidated and entrenched by the time the learner tries to add the L2 form to the map. But what happens when both L1 and L2 are acquired simultaneously during childhood (Cenoz & Genesee, 2001). In this case, the LX and the LY forms should compete for nearby territory in the semantic space. In some cases, the LX form might be parasitic on the LY form. In other cases, the LY forms may be parasitic on the LX form. In still other cases, the learner may enter the two

forms in related areas of the map along with additional features that distinguish particular properties of the words. For example, the Spanish noun *vaso* will be encoded as referring only to a container from which we drink, whereas the English noun *glass* will refer both to the container and the material used in bottles and windows. In this way, simultaneous bilingual acquisition tends to minimize the misleading effects of transfer and parasitism.

A preliminary assessment

We have now examined ten proposed accounts of fossilization and AoA effects in second language learning. The majority of these accounts generate predictions that are inconsistent with the observed patterns. For example, the lateralization hypothesis targets effects occurring at puberty despite the fact that the major changes in lateralization occur during the first three years. The parameter setting hypothesis, when coupled with the idea that UG becomes inaccessible at a discrete moment, would predict a rapid drop in L2 learning that is not observed in the data. The metabolic decline hypothesis does not align well with evidence from neuroscience that decreased metabolic activity actually represents increased learning and increased automaticity. The reproductive fitness hypothesis is in sharp disagreement with basic facts regarding early child language learning. A further problem with each of these first four hypotheses is that they treat language ability as a single undifferentiated whole, failing to distinguish different age-related patterns for phonology, lexicon, syntax, morphology, and pragmatics.

The fragile rote hypothesis is linked to the decline in memory abilities that occurs after age 50. As such, it may account for some fossilization effects, but not for AoA and other age-related effects at younger ages. The same can be said with the aging hypothesis

with which it is closely related. The starting small hypothesis succeeds in accounting for some adult experimental data, but makes a series of incorrect predictions regarding child language learning.

Of the first eight models we have examined, the one that seems most nearly in accord with observed AoA and fossilization effects is the neuronal commitment hypothesis. However, there are two reasons why we should be careful in evaluating this apparent success. First, within specific evidence about how specific neuronal areas lose their plasticity at particular ages, we can simply invoke this hypothesis to explain any observed pattern. Second, it is different to draw a sharp conceptual difference between neuronal commitment and entrenchment. Rather, entrenchment seems to be the psychological result of processes of neuronal commitment operating on the cellular level. Given this, it is perhaps best to focus on an evaluation of entrenchment. We have already seen that an entrenchment account, by itself, will not provide a full account of basic AoA and fossilization effects. Instead, we clearly need to view entrenchment as working in the context of the additional mechanism of parasitism.

Thus it seems that the best currently available account of AoA and fossilization effects is one that combines the concepts of entrenchment and parasitic transfer. This is also the account most in accord with the Unified Competition Model (MacWhinney, in press) – a model that emphasizes the role of transfer in second language learning, the use of item-based patterns to avoid catastrophic interference, and chunking and item learning as methods for automatization and entrenchment.

The entrenchment-and-transfer account predicts a gradual decline of L2 attainment beginning as early as age 5 and extending through adulthood. It predicts no sharp drop,

but rather a slow, gradual decline. These predictions are in good accord with the basic shape of observed AoA patterns. Moreover, the specific predictions of the model are further differentiated across linguistic areas or arenas.

The Unified Competition Model holds that transfer should be particularly strong on the level of phonology, since new phonological words are initially learned as combinations (Gupta & MacWhinney, 1997) of old L1 segments and syllables. For lexical learning, there is also massive initial transfer of old meanings. Syntax shows some item-based transfer effects, but less than the other areas, since these patterns are dependent on new lexical items. Moreover, even when there is a match between languages (Pienemann, Di Biase, Kawaguchi, & Håkansson, in press), transfer does not operate in terms of whole sentence patterns, but only individual word combinations such as adjective + noun (MacWhinney, in press). Finally, in the area of morphological marking, we only expect transfer of grammatical function, if these functions have a close match, not transfer of specific grammatical forms or patterns. Entrenchment also operates differentially across linguistic areas, with the strongest entrenchment occurring in output phonology and the least entrenchment in the area of lexicon, where new learning continues to occur in L1 in any case.

The social stratification hypothesis

Although an account based on transfer and entrenchment captures the overall shape of fossilization phenomena, it fails to predict the observed diversity of outcomes among adult learners. It also fails to predict the extent to which some adult learners achieve nearly complete mastery, whereas others seldom advance beyond the lowest levels. To account for these additional patterns in adult learners, I believe that we need to invoke

two further processes. The first involves the social positioning of the older learner and the shape of the input that this positioning will provide. The second process involves the strategic mechanisms that adults need to employ to overcome the effects of entrenchment and parasitism. I will refer to these two additional processes as the social stratification hypothesis and the compensatory strategies hypothesis.

It is difficult to overestimate the potential impact of social structures on language acquisition by both older children and adults. In most modern cultures, children acquire a first language within the context of a nuclear family group. Although there are strong variations between and within cultural and socioeconomic groups, virtually all children benefit from rich and consistent input from their parents, the extended family, co-wives, or older children. Whatever the exact configuration of the input social group, children are always treated as apprentices who need to be guided through the language learning process. Their silly mistakes and inarticulate productions are considered cute and lovable and they are never made to feel embarrassed or inadequate. Age-matched peers are also engaged in language learning and are in no position to make fun of errors from their playmates.

By the time children enter school at age 5, the situation has begun to change. An immigrant child arriving in a new country at this age will immediately seem strange and out of place. At school, they cannot express themselves and at home their parents may provide little in terms of L2 input. During the early school years, the child's best approach is to be withdrawing and silent at first, only entering into social groups after picking up a few basic phrases. By the time the child reaches adolescence, the pressure of the L1 community on the immigrant child can be truly massive and sometimes even

vicious. In order to become exposed to adequate peer group input, the adolescent must be willing to suffer a major degradation in status and rights. However, even this will not guarantee full acceptance. Although the rewards of improved L2 competence are great, the barriers to attainment of that competence also increase greatly.

Acculturation during adulthood involves a far greater diversity of situations. In some cases, immigrants may marry into the L1 community, thereby guaranteeing basic acceptance and access to input. However, even in these cases, they will not be treated in as supportive a manner as a parent treats a child. When they are overtly corrected, they will feel a certain loss of prestige that can strain social relations. Outside of marriage, L2 acculturation may succeed through work groups and casual social groups. However, in many other contexts the immigrant withdraws at least partially from the dominant language and remains tightly within the L1-speaking immigrant community. Because of the growing importance of English as an international language, L1 speakers may place pressure on English-speaking immigrants to converse in English, thereby further blocking potential L2 acquisition. This effect can be further exacerbated by modern communication systems such as email, phone, television, and the Internet that allow travelers to remain within an English language capsule from the beginning to the end of even month-long stays abroad.

The compensatory strategies hypothesis

The variance we find in adult L2 attainment must certainly be explained in large part by variation in the social contexts facing immigrants and the ways in which they deal with these contexts socially. However, even in unfavorable social situations, learners can make use of compensatory learning strategies. These strategies are designed to directly

combat the effects of increased L1 entrenchment, as well as the effects of biological aging. There are at least three major strategies: input maximization, recoding, and resonance.

Input maximization involves a whole series of strategies designed to obtain good learning input. The learner may try to use dictionaries, work through grammars, take classes, watch movies, listen to lectures, talk to friends, or just go shopping. The goal here is to simply practice use of the language.

Recoding involves the construction of alternative images of new words and phrases. Because of the likely decline in raw auditory memory with age and the possible further decline in declarative storage, older learners need to construct methods of keeping new words and phrases in memory while they are being learned. The easiest way to do this is to represent the new word orthographically. Orthographic learning has two important roles for older learners. First, it provides a solid recoding of transient auditory input. Second, it opens access to input from books, signs, and product labels. This second effect is relatively weaker in the learning of languages such as Navajo or Inuktitut. Although these languages have an ample written literature, there is relatively less use of signage and written instructions in these rural cultures then in more urban communities.

When both L1 and L2 use the Roman alphabet, it is relatively easy to recode L2 into L1, although there may be some glitches in this recoding, as in the recoding of French, using English grapheme-phoneme correspondences. However, these glitches can be readily repaired. Even the mapping from Roman onto phonetic alphabets such as Cyrillic for Russian, Hangul for Korean, or Kana for Japanese is within reach of older learners. However, acquisition of a non-phonemic script such as Chinese Hanzi characters is a

major challenge for older learners. Because younger learners rely relatively less on recoding, they will be able to pick up Chinese more directly. Thus older learners attempt to overcome entrenchment are at a particular disadvantage until they have mastered the new writing system. As a result, older learners must either acquire the new orthography or rely relatively more on alternative compensatory strategies.

Because learning through resonant connections is highly strategic, L2 learners will vary markedly in the constructions they can control or which are missing or incorrectly transferred (Birdsong, this volume). In addition to the basic forces of entrenchment, transfer, and strategic resonant learning, older learners will be affected by problems with restricted social contacts, commitments to ongoing L1 interactions, and declining cognitive abilities. None of these changes predict a sharp drop at a certain age in L2 learning abilities. Instead, they predict a gradual decline across the life span.

A final important strategy available to adult learners is resonance. Resonance involves the establishment of a series of associative relations between words and meanings that can allow the learner to maintain a vivid image of the word until the relations are consolidated. In this sense, resonance is really the general case of recoding. Examples of resonance include the use of the keyword mnemonic or imagery method for learning new words (Atkinson, 1975). For example, Italian *pomidoro* "tomato" can be encoded as "Dora sitting under a palm tree eating a tomato." Or the learner might recognize *pomidoro* as a golden (doro) apple (pomme). Or the learner may create some idiosyncratic relation between the sound of *pomidoro* and the action of slicing a tomato with its sounds and smells. It doesn't matter if these connections are real or fanciful. All that matters is that they help to maintain a resonant trace of the vividness of the word and

its meaning until the new item can be consolidated. Similar methods can be used to acquire longer phrases that encode for properties such as gender, tone, or particle usage.

Conclusions

In this paper, we have examined a set of twelve hypotheses regarding the etiology of age of acquisition (AoA) and fossilization effects. Of the various candidate hypotheses, the one that matches most clearly with the basic data on a gradual decline in learning ability is the hypothesis that combines the effects of ongoing L1 entrenchment with the notion that L2 develops at first as parasitic or dependent on L1. Although this account correctly predicts the overall gradual decline in L21earning, it fails to predict the diversity of fossilization patterns we see among older learners. To account for these additional effects, we need to look at both the effects of social stratification on older immigrants and the extent to which they can use compensatory strategies to combat the effects of entrenchment.

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1999). Statistical learning in linguistic and nonlinguistic domains. In B. MacWhinney (Ed.), *The emergence of language* (pp. 359-380). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Barkow, J., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

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.

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

Booth, J. R., MacWhinney, B., Thulborn, K. R., Sacco, K., Voyvodic, J. T., & Feldman, H. M. (2001). Developmental and lesion effects during brain activation for sentence comprehension and mental rotation. *Developmental Neuropsychology, 18*, 139-169.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard.

Cenoz, J., & Genesee, F. (2001). *Trends in bilingual acquisition*. Amsterdam: John Benjamins.

Chomsky, N. (1981). *Lectures on government and binding*. Cinnaminson, NJ: Foris.

Clark, R. (1974). Performing without competence. *Journal of Child Language, 1*, 1-10.

Cochran, B. P., McDonald, J. L., & Parault, S. J. (1999). Too smart for their own good: The disadvantage of superior processing capacity for adult language learners. *Journal of Memory and Language, 41*, 30-58.

Crago, M. B., & Allen, S. E. M. (1998). Acquiring Inukitut. In O. Taylor & L. Leonard (Eds.), *Language acquisition across North America: Cross-cultural and cross-linguistic perspectives* (pp. 245-279). San Diego, CA: Singular Publishing Group.

Dennis, M., & Whitaker, H. (1976). Linguistic superiority of the left over the right hemisphere. *Brain and Language, 3*, 404-433.

Elman, J. (1993). Incremental learning, or the importance of starting small. *Cognition, 48*, 71-99.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78-104.

Flynn, S. (1996). A parameter-setting approach to second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 121-158). San Diego: Academic Press.

Goldowsky, B. N., & Newport, E. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In E. Clark (Ed.), *Proceedings of the 24th Annual Child Language Research Forum* (pp. 124-138). Stanford: CSLI.

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

Haier, R. (2001). PET studies of learning and individual differences. In J. McClelland & R. Siegler (Eds.), *Mechanisms of cognitive development* (pp. 123-148). Mahwah, NJ: Lawrence Erlbaum Associates.

Halford, G. S., Smith, S. B., Dickson, J. C., Mayberry, M. T., Kelly, M. E., Bain, J. D., et al. (1995). Modeling the development of reasoning strategies: The roles of analogy, knowledge, and capacity. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hardyck, C., Tzeng, O., & Wang, W. (1978). Cerebral lateralization of function and bilingual decision processes: Is thinking lateralized? *Brain and Language, 5*, 56-71.

Hurford, J., & Kirby, S. (1999). Co-evolution of language size and the criticial period. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 39-63). Mahwah, NJ: Lawerence Erlbaum Associates.

Johnson, J., & Newport, E. (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.

Johnson, J., & Newport, E. L. (1991). Critical period effects on universal properties of language: the status of subjacency in the acquisition of a second language. *Cognition, 39*, 215-258.

Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition, 56*, 263-269.

Kersten, A. W., & Earles, J. L. (2001). Less really is more for adults learning a miniature artificial language. *Journal of Memory and Language, 44*, 250-273.

Kinsbourne, M., & Hiscock, M. (1983). The normal and deviant development of functional lateralization of the brain. In P. Mussen, M. Haith & J. Campos (Eds.), *Handbook of Child Psychology*. New York: Wiley.

Labov, W. (1994). *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

Lecours, A. R. (1975). Myelogenetic correlates of the development of speech and language. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development: A multidisciplinary approach* (Vol. 1, pp. 121-136). New York: Academic Press.

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

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

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

Luria, A. R. (1973). *The working brain*. New York: Basic Books.

MacWhinney, B. (1974). *How Hungarian children learn to speak*. University of California, Berkeley.

MacWhinney, B. (2000). Lexicalist connectionism. In P. Broeder & J. Murre (Eds.), *Models of language acquisition: Inductive and deductive approaches* (pp. 9-32). Cambridge, MA: MIT Press.

MacWhinney, B. (2002). Language emergence. In P. Burmeister, T. Piske & A. Rohde (Eds.), *An integrated view of language development - Papers in honor of Henning Wode* (pp. 17-42). Trier: Wissenschaftlicher Verlag Trier.

MacWhinney, B. (2004). Language evolution and human development. In D. Bjorklund & A. Pellegrini (Eds.), *Child development and evolutionary psychology*. New York: Cambridge University Press.

MacWhinney, B. (in press). A Unified Model of Language Acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

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

Marchman, V. (1992). Constraint on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience, 4*, 215-234.

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

McArdle, C. B., Richardson, C. J., Nicholas, D. A., Mirfakhraee, M., Hayden, C. K., & Amparo, E. G. (1987). Developmental features of the neonatal brain: MR imaging -- Part 1. Gray white matter: Differentiation and myelination. *Radiology, 162*, 162-223.

McDonald, J. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*, 395-423.

Merzenich, M. (2001). Cortical plasticity contributing to child development. In J. McClelland & R. Siegler (Eds.), *Mechanisms of cognitive development* (pp. 67-96). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Molfese, D., & Hess, T. (1978). Hemispheric specialization for VOT perception in the preschool child. *Journal of Experimental Child Psychology, 26*, 71-84.

Molfese, V., & Betz, J. (1987). Language and motor development in infancy: three views with neuropsychological implications. *Developmental Neuropsychology, 3*, 255-274.

Muter, V., Taylor, S., & Vargha-Khadem, F. (1997). A longitudinal study of early intellectual development in hemiplegic children. *Neuropsychologia, 35*, 289-298.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. Chase (Ed.), *Visual information processing*. New York: Academic Press.

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

Peters, A. (1983). *The units of language acquisition*. New York: Cambridge University Press.

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

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

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language, 69*, 1-33.

Raichle, M. E., Fiez, J. A., Videen, T. O., Macleod, A. K., Pardo, J. V., T., F. P., et al. (1994). Practice-related changes in human brain functional anatomy during non-motor learning. *Cerebral Cortex, 4*, 8-26.

Rohde, D., & Plaut, D. (1999). Language acquisition in the absence of explicit negative evidence: how important is starting small? *Cognition, 72*, 67-109.

Satz, P., Strauss, E., & Whitaker, H. (1990). The ontogeny of hemispheric specialization: Some old hypotheses revisited. *Brain and Language, 38*, 596-614.

Scheibel, A. B. (1996). Structural and functional changes in the aging brain. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging, fourth edition* (pp. 105-128). San Diego: Academic Press.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*, 209-231.

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

Sperry, R. (1956). The eye and the brain. *Scientific American, 194*, 48-52.

Vargha Khadem, F., Isaacs, E., & Muter, V. (1994). A review of cognitive outcome after unilateral lesions sustained during childhood. *Journal of Child Neurology, 9*, 2s67-62s73.

Wada, J. A., Clarke, R., & Hamm, A. (1975). Cerebral hemispheric asymmetry in humans. *Archives of Neurology, 32*, 239-246.

Waddington, C. H. (1957). *The strategy of the genes*. New York: MacMillan.

Werker, J. F. (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.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*, 23-79.

Witelson, S. (1977). Early hemisphere specialization and interhemisphere plasticity: An empirical and theoretical review. In S. Segalowitz & F. Gruber (Eds.), *Language Development and Neurological Theory*. New York: Academic Press.

Witelson, S., & Pallie, W. (1973). Left hemisphere specialization for language in the newborn. *Brain, 96*, 641-646.

Wong-Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Stanford: Stanford University Press.

Woods, B., & Teuber, H. (1977). Changing patterns of childhood aphasia. *Annals of Neurology, 3*, 273-280.

Zangwill, O. L. (1960). *Cerebral dominance and its relation to psychological function.* Ediburgh: Oliver and Boyd.