ENHANCING SECOND LANGUAGE ACQUISITION AND RECALL WITH MUSIC AND PROSODY

A Thesis

Presented

to the Faculty of

California State University, Chico

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Teaching International Languages

By

Fall 2013

ENHANCING SECOND LANGUAGE ACQUISITION AND RECALL

WITH MUSIC AND PROSODY

A Thesis

by

Suzanne S. Stephens

Fall 2013

APPROVED BY THE DEAN OF GRADUATE STUDIES

AND VICE PROVOST FOR RESEARCH:

____
Eun K. Park, Ph.D.

APPROVED BY THE GRADUATE ADVISORY COMMITTEE:

Hilda Hernández, Ph.D. Graduate Coordinator	Hilda Hernández, Ph.D., Chair
	Saundra K. Wright, Ph.D.
	Thomas A. Fox, Ph.D.

PUBLICATION RIGHTS

LIST OF FIGURES AND TABLES PAGE

TABLE OF CONTENTS

ABSTRACT

ENHANCING SECOND LANGUAGE ACQUISITION AND RECALL

WITH MUSIC AND PROSODY

by

© Suzanne S. Stephens 2013

Master of Arts in Teaching International Languages

California State University, Chico

Fall 2013

This paper examines evidence that music and language are so intertwined, so integral to one another that one cannot be learned without the other. Indeed, the study of music seems to enhance the acquisition of second languages because it provides phonological and metrical space cognition. Starting with the premise that language and music are ubiquitous, and that their respective grammars are implicitly known to nearly all people, researchers have parsed out the relevant topics into the major areas of cognition, memory and recall, and models of brain and functional modularity described by syntactic and semantic networks. How music and prosody influence language acquisition focuses on the following: The hierarchical grammars of music and linguistics; studies of the brain functions from neurolinguistics; language acquisition through psycholinguistics; and modern graph theory as applied to language networks. These models are examined for their predictions of growth and function. Recent interdisciplinary research in these fields has discovered correlative processing and shared resources between language and music that has implications for second language (L2) learners as a distinct class of language learners. Since the input for language, music and prosody is often under the L2 learner's control, strategic learning may improve memory and recall. First, this paper examines theories of infant language acquisition, including classic morpheme studies, prior to extending the theories to adult second language learners and use of music/language prosody. Thus, from the L2 learner's prospective, does music and language prosody improve or enhance the second language acquisition process, and what is the evidence? From neuroscience studies, is there a strong enough cognitive connection to include explicit music and prosody training? This needs evidence beyond the obvious sociolinguistic reasons, because the argument here is that music itself enhances cognitive functions. Or, is music and varieties of language prosody simply to be enjoyed as part of the cultural or social milieu, not necessary for language competence, just a frill. Using network theory to describe the syntactic bottleneck caused by incomplete or faulty knowledge of requisite morphemes, this paper extrapolates from the child language cascade to the adult L2 learner. For the adult second language learner, music and language prosody are strategic tools for acquiring these implicit rules ensuring language readiness.

CHAPTER I

INTRODUCTION

Overview of Language Learning Enhanced by Music

improvement on simple behavioral studies. The notion that people anticipate and predict based on an implicit syntax to produce semantic sense whatever the genre, music or language, is no longer speculative. The overwhelming social desire to synchronize motor movements and mimic behaviors is now seen as part of our animal nature, a natural force of biological selection.

Focus of the Research

This overview of the nexus between music and language cognition details the evidence of correlates in brain function between music and language. These functional correlates are used to predict the enhanced learning of second languages for adults through the inclusion of music and language prosody in the target language. A major assumption for this study is that there is a continuum of language and music in brain functions and that recent advances in neuroscience and network theory can provide that evidence.

Another assumption is that the adult language learner, in contrast to the infant or child, does initially remember the spatial-temporal circumstances of semantic learning (connections) in the lexicon between the L1 words and L2 words. This is because the words or chunks are so heavily invested in the situation. Later, as the syntactic network strengthens, the L1 to L2 syntactic connections become weaker; L2 to L2 connections take precedence in production. Words slip by with tags, and like breathing, they are amenable to willful control. Code switching, instead of being a necessity when circumlocution fails, becomes word play, much like punning and metaphorical embellishment. This paper looks at current evidence in the domains of cognition, music, language acquisition, adult learning of second languages, and network theory in support of the thesis that language without music is only half a language.

Theoretical Bases

From the L2 learner's prospective, how well does music, especially lyrical music, improve or enhance their second language acquisition process? Is there empirical evidence for using music in the classroom to improve learning? And, from neuroscience evidence, is there a strong enough cognitive connection between music and language such that second language instruction should naturally include music beyond sociolinguistic reasons, because music itself enhances cognitive functions? Most second language instruction is focused upon utility: How to pass the exit exam or score high enough on the entrance exam to continue on in higher education. Academic English, or how to write in standardized ways, or how to read academic texts are the norm for almost all school-based second language instruction in the western democracies. But learners, especially adults, are not required to spend their entire learning lives in classroom activities. They live and learn best outside that setting and can control, to some extent, the input to their 'living language'. Music, which can be accessed at will, and to any taste, is almost universally enjoyed by young and older learners; it just needs to be strategically employed.

One of the goals of this study is to provide L2 instructors with justification for music and language prosody analysis in their meta-language learning strategies for adults, so that learners can deliberately employ possibilities of enhanced cognition and memory through music, poetry and understanding of language prosody. It is hoped that upon learning the facts, language instructors will, despite testing imperatives, encourage learners to view music and language prosody as a normal part of language acquisition. This paper is an attempt to legitimatize music and language prosody through empirical evidence as to its efficacy. After establishing the empirical basis for music and language prosody in the curriculum, another goal is to synthesize basic research and hypotheses on brain networks and theories of modularity and domain-general sequence learning mechanisms to explain language spurts. Then, this explanation of language readiness can be applied to a new population--adult second language learners. Music and language prosody arms adult language learners with the technologies to strategically acquire and use the minimally necessary syntax and lexicon. Readiness for their own language cascade as an adult L2 learner follows the same path with the same steps as in their first language.

Purpose

Enhanced language networks utilizing lyrical music and prosody is at the heart of this paper. The proposition is based on four chapters, with topics as follows:

1. Language is both networked and modular in the brain,
2. Music is a technology shares many features of languages
3. Second language acquisition is implicit, but amenable to strategic alliances with music for learning, and
4. Learners must acquire implicit knowledge of certain requisite morphemes prior to real growth.

Therefore, L2 learners should strategically use the language-music nexus to develop implicit knowledge of requisite morphemes, phonemes, and intonation. This language readiness allows learners to utilize other technologies such as reading to strengthen those implicit semantic and syntactic networks to quickly increase deposits in the lexicons. Thus, an adult language spurt is predicted on the hypothesis that music, especially lyrical music, is easier to learn because of its built-in redundancy, expectations and affective behavior mechanisms. Children easily acquire the requisite morphemes and move from one or two word utterances to complete command of most language syntax before attending school. It is predicted that adults should be able to do the same thing, with the same ease and proficiency as infants. The following sections of this Introduction provide a brief overview of the research that supports the main arguments for using music and language prosody and how that research points to this potentially useful tool.

Overview of Supporting Studies

Major players in this literature review of music and language research at the neurocognitive level look at the connections between music training and enhanced language skills, while others tease out the neural networks of language and music through syntactic activities. Researchers also attempt to differentiate explicit music knowledge from implicit music memory comparing it to the implicit knowledge of language. Equally important are studies examining the neural substrates of music and semantic memory, and one major player (Patel, 2008) has built an intricate model of music perception based on neural correlates. Long debated, the putative innate language facility is connected to the innate capacity for music. Interestingly, one research study argues for a specific area of brain activation from liked versus disliked music (Montag, Reuter & Axmacher, 2011). This paper will reveal through several studies that linguistics and music are intertwined and that songs aid in language acquisition. Arguing for the predictability of spoken language success, researchers examining musical training and brain plasticity not surprisingly link these skills to increased ability to discriminate sounds. There are many others, as this branch of neuroscience has morphed from just identifying regions of the brain activated under language or music stimuli to the discovery of many overlaps in cognitive processing. What had been a sterile dead end, a single language acquisition device (LAD) has become a multifaceted structure(s) that is probably uniquely human as a continuum of song *and* language.

Shared Syntax Processing

Aniruddh Patel (2003, 2008) proposed his Shared Syntactic Integration Resource Hypothesis (SSIRH), in several papers and his seminal book *Music, Language and the Brain* (2008). Patel's hypothesis describes the repurposing of brain resources for language and music. The theory was predicated on experimental findings that neural resources were consumed during activation of irregular or unexpected words or notes. This assumption of shared processing resources is behind most current research in music/language studies and argues for the irreducibility of the oral/aural communicative streams into separate domains. But these theories do not argue for co-mingled storage, which supports the brain lesion studies wherein music survives with impaired cognitive and language output as seen in Alzheimer's syndrome. Survival depends on the ability to discern change, in fields of view, range of hearing, etc. Our brains automatically tune-in or tune-out incoming perceptual data. Any unanticipated change predicts an existential threat causing override, forcing attention. It makes sense from an adaptation bias to have correlative processing, and that evidence for music/language syntactical correlates is presented in this paper. Language and music, for the purposes of this paper, are defined as projections outward to our social group our internal reactions and predictions, thus enhancing cohesion and cooperation.

Less abundant are empirical studies supporting the inclusion of music, especially, songs with lyrics in the L2 curriculum. There is a gap between the brain science indicating a commonality and applied science in the L2 classroom. This is a result of the difficulty of performing basic research in the chaos of real world language learning, where most language acquisition is done outside the walls of the classroom. Experimental disambiguation of all of the neural processes involved in language perception and production has also proven elusive. Therefore, this paper looks to the neurocognitive evidence of a common syntactic network between language and music, and relies on that positive neurological evidence to advocate for music's inclusion in the L2 curriculum. Additional motor learning areas involving synchrony and heightened limbic expression through mimicry during live performances are also discussed.

Underlying all theories of cognition and language are theories of networked modules for language memory, such as implicit knowledge of morphemes, a specially tuned phonemic knowledge, and at least several lexicons. The bases for networked theories of the mind are physical, as described in the topological mapping done over the last 200 years ago. Lesion studies, based on deficits observed after a specific injury, started research in pinpointing brain areas, like Broca's or Wernicke's, (networked neighborhoods). In the modern age, these have been tested by intracranial stimulation during brain surgery, and imaged using fMRI (functional Magnetic Resonance Imaging) and other types of recordings[1]. These studies are designed to reveal synaptic firings across the presumed networks. Most published images are actually compilations of averages over many test subjects, not snapshots like x-rays. Nevertheless, these pixilated maps associated with specific kinds of brain activity are rigorously compared with data

[1] The most common are electroencephalograms (EEGs) and differences in event-related electrical potentials (ERPs), but are only indirect measures of activity provoked by single stimuli.

from deficit case studies. These case studies of curious accounts of injuries or strokes continue to challenge theories and reflect the increasing complexity of neuronal mapping. The current issue with mapping of the brain is that it is difficult to tease out the causes for the differences being recorded. As research becomes more accurate, the number of neighborhoods in the brain showing activity during audition or production is only increasing, reflecting the complexity of language, cognitive and motor processes, and the disparate places for specific kinds of memory.

Semantic word maps, currently deployed in language arts classrooms, may be misleading when construed as reflecting brain activity when 'searching' for words. Equally misleading are constructs such as the world-wide web as a models for brain function, memory and learning. Though interesting as a model of an electronic network, the web does not begin to model language acquisition or functional attributes of language such as polysemy and generativity. What we do know is that the human brain is continually re-inventing itself. Accidental or purposeful physical insults, normal maturation, even learning and life experiences forces constant reorganization of the networks of the brain.

Requisite Morphemes and the Language Spurt

Researchers describe the cascade effect in youngsters at about 18 months of age and its subsequent language spurt. This is then followed by a pre-teen, a late teenage/young adult reorganization and finally the senescence of old age associated with Alzheimer's syndrome. Other major reorganizations have been identified. At every one of these stages, music and its language analog plays an important cultural, learning, cohesion, and identity role. For the child, there is a transition from me-centered to we. Why not add another reorganization, this time to adult L2 learning? This paper provides an extended description of one of these phases, the infant language cascade. It then applies the same cascade to the adult L2 learner, using network theory to describe the bottleneck caused by not having implicit knowledge of requisite morphemes. It is argued that in order to support both lexical and syntactic growth, music and language prosody provide the L2's crucial tool in acquiring these rules, and making them implicit. Along with these early rules, a certain level of first language competency is assumed in order to be ready to truly acquire an adult lexicon. This minimal proficiency is necessary to successfully navigate whole language in immersion settings.

What Is Special About Music?

One of the best explanations of the nature of music is its rhythm, meter, tonality, melody and harmony. Its ability to invoke emotion in listeners is without challenge, as are the maps to major cognitive structures invoked by music. However, there are still skeptics of parallel structures between language and music syntax. As one of the affective adjuncts to communication that expresses affective messages (along with gesture, facial expressions and other postures), music is explored through musical structure and cognitive structures, musical grammar (musical idiom), and acquisition of musical grammar. Inevitably, the ghost of the innateness hypothesis for the inherited capacity for the grammar of music is addressed. Not all theorists support a shared syntax between music and language, but there is unusual agreement on the innate ability to acquire proficiency in music and language. Poetry, naturally, is seen as very close to song since it has meter and stress, impelling completion of chunked phrases. Because it is often aligned with its particular language, poetry can be utilized to encourage implicit acquisition of patterns, phonemes and morphemes. Larger structures, such as phrase and sentence endophora become entwined in the prosody and thereby, easily recalled.

Patel, (2008) a neuroscientist and musician, suggests that a cross-cultural translation of music is futile. We should enjoy the music syntax for what it is. Unlike poetry, in music syntax there are no nouns and verbs (only in accompanying lyrics) but there are event hierarchies built into music's structure, marking it as having syntax. Hierarchical structure is one of the hallmarks of language, and one of the early cognitive challenges for child language acquisition. Patel argues for a cognitive and neural overlap in the processing of music and language syntax based on the definition of syntax as a learned recurring pattern.

For multiple language speakers, all lexicons are probably called up in language production and reading. Recall of these other words is not amenable to conscious control (Koda, 2004, p. 67). Even if there were a single lexicon for multilingual verbs with language tags, for example, and separate storage for each of the closed sets of irregulars, this extra load entails significant additional processing effort. Getting past explicit calculation to implicit production is crucial to language acquisition.

Most poetry and music share meter as a structural element. Jackendoff and Lerdahl (2006, p. 43) remark that "poetic meter can be viewed as a metrical grid to which the stress grid in the text is optimally aligned," as in nursery rhymes, limericks, and metrical poetry. Patel (2010, p. 1) suggests that because music can have measurable effects on brain structure and non-musical functions, music should be considered a "transformative technology of the mind." Patel, Jackendoff and Lerdahl equate music to reading as a technology. Each generation learns anew how to read or play a musical instrument. Both sets of researchers believe music might have evolved to promote social cohesion, and is completely absent in other primates.

Let's compare the exclusion of music from the curriculum, to the lessons of Hall's *Silent Language*. After nearly six decades his precepts are now firmly part of sociolinguistic curricula. Posture and demeanor might not always be 'readable' by others in any overt sense, but we instinctively 'mirror' crossed arms, or lean in to convey interest, etc. Persons from other cultures almost always miss subtle clues, either visual or aural. 'Stealth' perception must be explicitly taught to outsiders. No language curriculum would sensibly exclude language pragmatics. Music needs to be seen by the learner as an affective adjunct to language because it invites temporal mimicry and synchronicity. Perhaps Hall's 'silent language' needs to be reconnected to melody, harmony and synchrony in the second language classroom through music.

Like the range of proficiency in L2 learners, there is also a perceived wide range of competency in music production. Contrast this with universal L1 acquisition and even effortless child bilingualism. This competency phenomenon is a key connection to the study of the intersection of language and music in brain function. Warning of overreach, Jackendoff and Lerdahl (2006, p. 35) say that "our current knowledge of relevant brain function, while growing rapidly, is still limited in its ability to address matters of sequential and hierarchical structure," which is by definition, musical syntax. They detail the regularity of meter and its deliberate violation, its structural-based variation, and its uniqueness to humans. These structures cross boundaries between music and language, as discussed above, but not with other cognitive systems such as vision. The metrical 'alternating strong and weak beats' does not appear to be shared widely with beasts, (except maybe cockatoos) but is evident in small children. Even though many people claim they can't sing, it is generally embarrassment, not the lack of vocal ability. People seem to be able to follow a melodic contour, even when it is transposed to another key. This puts singing into the realm of universal competency, albeit slightly out of tune for some. Most people, not just musicians, have an ability to recall both melody and meter for many musical pieces, and can easily replicate complicated metrical forms by finger or foot tapping.

Another recent study by key researchers in this field Carras, Koelsch, and Bhattacharya (2011) starts from the premise that subtle changes in brain activity during higher order cognitive functioning may have been overlooked (p. 51). Neuronal oscillations that are produced during cognitive tasks, starting from the baseline theory of separate storage regions for music and language, require a "dynamical integration of information" between the two regions. Measurements of that neural activity identify local networks and clues to the intricacies of the wiring assembly. In this study, the ability to do time-frequency analysis, resulted in experimental evidence of communication between 'neuronal populations,' also known as neighborhoods in network theory. This supports the shared resource hypothesis. By producing events through the mechanism of eliciting incongruities, i.e., ambiguous cloze, and discordant sounds, this study also has the unfortunate potential for masking the true effects, which the researchers acknowledge. Because of the phenomenon of making 'sense' as we go along--often not under conscious control--their study reveals a basic flaw in the current methodology (Carrus, Koelsch & Bhattacharya (2011) p. 51).

In a review of music perception contemporaneous to the work above, Koelsch (2011) gathers brain studies into a model of the entire music perceptual process and its consequent memory formation. There appears to be an overlap in the language and music syntactic processing, what he calls the hierarchical structural organization of grammars from phrases. Koelsch (2011) reiterates the now accepted scientific evidence that like language, non-musicians exhibit "highly sophisticated (implicit) knowledge about music syntax" (Koelsch, 2011 p. 8). This implicit knowledge is known as temporal pattern recognition.

The Leading Shared Resources Hypothesis

Patel's SSIRH (2003, 2008) hypothesis, using "online structural integration" of music and word elements is a frequently referenced hypothesis that predicts the "shared, limited processing resources" for music and word-category processes. Word category, or lexical category, is a structural concept in grammar usually associated with parts of speech. Behind this idea of shared word category processes with music is the philosophical dispute between cognitive linguistics and generative linguistics, beyond the scope for this paper. But it does indicate future trends in the cognitive sciences research as desktop predictions of a single language acquisition device (LAD) meet headlong with results of experimental imaging, patterns of electrical impedances and brain chemistry. Language functions seem to have invaded and parasitized all sensorimotor modalities, along with its handmaiden, music.

Lexical semantics, or the properties of word meanings, and conceptual semantics will be shown to be in alignment with studies of syntactic processing of language and music. This domain of co-located words and phrases has cornered research using modern graph theory, and has contributed to our understanding of shades of meaning through association, priming, and habituation. It has even triggered hypotheses on the variability and predictability of particular usage, deterministic in nature.

However, there is a kind of tension caused by generative grammar being viewed as too dependent on syntax, leaving phonology and semantics out of the generative equation. Other explanations for the conjoined processing of language and music have resulted in researchers looking more closely at the function of the auditory system and pitch-related abilities which can be learned. Thus, the current thesis for music enhancing linguistic abilities is based on the proposition of shared resources, multiple domain recall, and prediction.

Background sounds in our lives, the sound pollution modern humans endure, is easily ignored. However, human voices, especially music can stop us in our tracks, demanding attention. Why are music and voices able to do this? According to Koelsch (2011, p. 11), it is because “listeners automatically engage social cognition.” That is, they are forced to decode intent when listening to music. Intent, structure, and motivation are in play. That which cannot be ignored is the heightened symbolic meaning attached to some music, like national anthems, where both words and music shape the culture and control culture, as well. This is undoubtedly true of other cultural/musical icons, like the Beatles, or Bach mania, hence the term cultural imperialism, whether the affect is intended or not. It is not just the music rattling around in our heads; it is the ideas and the singing them aloud in a social context that fixes the phrases. According to Koelsch (2011, p. 11), what is interesting about music is its influence on the semantic processing of words. Music conveys more than emotional content, and its decoding engages the same cognitive processes. For example, the social experience of music bonding is legend with Japanese karaoke, a cross-cultural phenomenon, during which bosses and minions indulge in group humiliation. Communal excess leads to social cohesion.

Koelsch, (2011) suggests that in social situations this emotional power of music accounts for its evolution as a direct method for socializing without necessarily having physical contact. He concludes by saying that “communication of meaning is not exclusively a linguistic domain, but that music can also convey meaningful information” (Koelsch, 2011, p. 14). He supports the thesis that language and music fall along a continuum, that they are “different aspects of the same domain” and warns that separating them will only result in what he calls a rather “artificial construct” (Koelsch, 2011, p. 14).

Future Research

Music is enjoyable principally because it activates reward centers in the brain, (limbic system). That same neuronal reward network is also instrumental in learning. Instead of a natural instinct for language, we have a natural curiosity for finding out what other people are thinking. One avenue for future research is based on the observed resistance to memorization when being explicitly taught. Lyrical music and language prosody provide such an avenue through direct performance. Another interesting avenue for further research is whether personality affects this neuronal reward network. Specifically, researchers have suggested that learners, who classify themselves as 'self-forgetful', i.e., easily transported or absorbed by music or art, should register a positive correlation to activity in the ventral striatum when listening to favorite songs. As this has not yet been established, it might again point to a systemic weakness in brain behavioral experiments. These experiments are deliberately stripped of context as researchers try to control for extraneous inputs. Groussard, et al., (2009, p. 2764, p. 2771) found that musical cues of familiar songs and semantic memory for well-known proverbs activated a common left temporal area of the neo-cortex. This caused a 'topographical preference' showing more activation in the anterior for music and the reverse, with posterior activation for the phrases. Additional right brain activity was attributed to "retrieval of melodic traces" in perceptual memory. To some researchers this suggests that indeed, there are two neural networks, and both are activated by music. Such experiments support the significance of musical memory for learning.

Extrapolating from studies such as these, and based on network theory and neighborhoods, these bilateral connections can be deliberately exploited by L2 learners. Learners can extend implicit knowledge of language through music and language prosody by exploiting brain plasticity and hemispheric connectivity through greater bilateral activations. L2 learners can do this by engaging in activities that lay down more perceptual (from the senses) memory traces rather than memorizing reified rules. Future research might test the perceptual traces with recall of full phrases from songs, targeting requisite morphemes for correct production in either music phrases or conversation. Concurring with the notion that "linguistic and melodic components of songs are processed in interaction," Schön, Gordon and Besson (2005, p. 72) call for testing using more natural materials when using fMRI. Results from odd test stimuli that do not reflect real musical or semantic processing may be misleading. Musicians and non-musicians are likely to have different foci of attention and need to be tested separately. Three years later, Schön, et al., (2008) called for consideration of songs as an "aid for language acquisition," having by then confirmed their hypothesis in the laboratory. They highlighted the word segmentation hypothesis of the statistical properties of language learning, comparing speech sequence learning to song sequence learning. Their conclusion: There was a benefit. People learn by exposure to language in action, whether as children or as adults. Language learning entails guessing at probable next-syllable sounds, and eventually learning cues for word boundaries. Research suggests that the innate statistical learning mechanism may be the same for tonal segmentation in music as for speech (Schön, et al., 2008, p. 976). This common learning device was seen in experiments that compared learning based on songs versus speech.[2] Schoen, et al., (2008) reiterate the accepted theory of speech repetition: "...redundant information in general is easier to process, not only across the linguistic and musical domains, but more generally throughout cognitive domains," (Schoen, et al., 2008 p. 982) affirming our thesis that for the new L2 learner, lyrical music can by its very structure and emotive affect improve learning outcomes for speech segmentation.

A further line of inquiry might be to determine how the production of lyrical music is able to strip accent from language and how this might be utilized by the new language learner. Unlike phoneme drills, lyrical music is inherently pleasing in its repetitions. Forgotten phrases can be hummed or tapped, and rejoined at their logical place, in synchrony, body and soul.

Limitations of this Review

As a study of current models of neuroscience and its implications for L2 practice, this paper is mainly focused on research and language/music models from the last decade. The emphasis will be on the evidence reported in the science literature, not education case studies, because of difficulties controlling for untoward affects in doing empirical research in the classroom. This paper provides a synthesis of some of the relevant scientific research. It also provides some practical suggestions by extrapolating from empirical evidence that which might be applied to practice in adult L2 learning in the areas of music as culture, language pragmatics, and of course language acquisition

---

[2] This also describes another hypothesis, inter-sensory redundancy, which also underpins current pedagogical theories of multimodal learning.

CHAPTER II

LANGUAGE AS NETWORKED AND MODULAR

Overview

Networks are not just descriptions of how one thing is connected to another, a map of what exists. Network explanations of language must be predictive, i.e., the theory must be able to predict language acquisition and reflect observed stages in children as well as adult L2 acquisition as their vocabulary and application of syntax expand to meet semantic demands. There are two major sticking points between the language acquisition of native infants and second or multi-lingual adults. These are the limitations imposed on the immature brain and biases imposed by the already settled language(s) of the adult learner. That they might be different phenomena demands a short reference to other implicit structural acquisitions such tonal music. In the West, we are biased towards the symmetry of octave scales, but this is not universal. In infants, but not adults, exposure to Western music seems to engender a preference for asymmetric scales. Infants seem to prefer the new and unusual. But for adults, regularity is the norm. Adult preference for symmetry is probably caused by a cognitive bias towards the tonal center upon which one bases relative pitch, "making it easier to learn and remember complex melodic sequences (Patel, 2008, p. 34). Adult biases in music and language depend upon experienced reference points. For the infant, any point of reference can be utilized, because experiential growth has not yet settled upon a schema, but rather upon what slot happens to be available for anchoring the sensorimotor experience. Rather than a blank slate, the newborn has immature neuronal mapping in place, but these maps are tentative and subject to experience. There are many facile explanations of the human mind as learning machines, computational machines, etc.; but not just any machine makes a good model. Rather, growth and snipping of unused alternative routes is a better fit for the various biological phenomena, as experience modifies the growth of networks.

Chapter II reviews the concept of language as networked to prepare the ground for music and language prosody as inseparable from second language acquisition in adults. Even though there are major differences in infant versus adult language acquisition caused by maturational differences and biases introduced from the first and subsequent languages, both infants and adults go through comparable stages that can be tracked. Fluency in L1 and in L2 (and multi-linguals) is functionally undifferentiated in current fMRIs, and likely to remain so. Lexicons for multiple languages are congruent and possibly undifferentiated, even for disparate language families. Networks inside of networks interlaced with short-cutting strands seems to describe the functioning of language production, with children providing clues to how adults might be learning. Finally, the stranding model of weighted shortcuts between webs will be introduced in Chapter II as an original model to explain the flow of conceptual knowledge through the networks of language production.

General Theory of Networks

Barabási's (2003) first scale-free network model had four nodes evenly distanced and connected in a diamond shape. Any one of the four corners could be a hub, and in three-dimensional space these four nodes connect in six edges. (Think of a pyramid with planes.) For modeling purposes, this is the smallest unit of a network, expanding exponentially, from 4 nodes to 16, 32, 64 …etc., clustering into hierarchical modules. In his model, the hubs (any pointy end on the pyramid) maintain communication between modules. Barabási takes the 4-sided cluster and builds a self-similar structure by duplicating, exponentially, and connecting the multiples with symmetry on each central 'hub'.

There is insufficient evidence to determine if Barabási's (2003) four-node modular structure is either optimal or even commonly seen in nature. For example, might not a pentagon work as well, and why do the distances between nodes have to be the same in the model? This makes mechanical drawing easier, but obscures the disparate functional areas in the brain whereby long distances are traveled and ignores hemispheric redundancy. The (three) degrees of separation are clearly not the norm when the entirety of the brain or even discrete functional areas of brain specialization are considered. Any Euclidian ratios could be just as rational as Barabási's four-sided cluster. There are other competing descriptions of organic growth. These are seen throughout gross anatomical shapes seen in nature such as leaf-shapes as described by the Mandelbrot equation. Barabási also utilizes another simple shape, the pentagon. In a pentagon, the ratio of the whole to the shortcut (skipping across nodes) is the same ratio as any bisected line between nodes. This shortcut allows hubs to downgrade into nodes, essential for growth and re-indexing in simple taxonomies as more 'data' are added. Using this equality of ratios in a pentagonal shape creates a starting point for modeling calls and retrievals in a network, and might also provide an explanation of how Barabási's synonyms (p. 237) can be stored clustered in a network or a neural relational database of the mechanical type. Barabási says the 4-node model, like 5-node model, still results in the famous three degrees of separation. (Barabási, & Ravasz, 2008). Described is a scale-free semantic network of synonyms, with a degree exponent of $\Upsilon = 3.25$, it indicates "...that language has a hierarchical organization." (p. 3).

Straight lines between nodes (edges) provide a reasonable schematic for copper, silicon chip or fiber optic connectivity but they are not a model from nature. Since nature tends to repeat itself, a neuronal network needs to be modeled on biological models, where new growth and expanded robustness takes place, not human directed manufacturing. Optimal growth determines roots and branches of many organisms; we call this growth dendritic or arborization. Many organic networks exhibit enlargement of pathways that are successful and atrophy to those that are not. Further, strength of signal and interference from the environment or converging webs is not accounted for in Barabási's early models, irrespective of four or five nodes. Five nodes is preferable because of the irregularity or asymmetry creating interstices or gaps along the edges. One of this paper's themes is that asymmetry is the normal, with symmetry only fleeting and brief: As knowledge increases with experience, discrete 'facts' have less relevance. This is a cautionary reminder that network models sometimes obscure, not simplify thinking about hugely complicated dynamical systems.

Granovetter’s (1978) ‘strength of weak ties’ in a social model, now part of zeitgeist of threshold theories, also predicts the microbiological model of language networks. Cognitive networks are momentarily connected and mediated by chemical expression creating a difference in electrical potential, measured by FMRi or EEG. For the latter, experimenters place flexible 'caps' on human subjects to record electrical impulses detected during experimental conditions. These differences in electrical potentials (from a baseline, since the brain is always active) are used to create the cortical maps that are gradually revealing both the topology and the anatomy of functional areas of the brain. Interference or mediation by multiple chemicals through networks might have a cumulative effect even if there are many connections of varying length, coming in from different modules. A multiplicity of weak ties might triumph over a single stronger tie. In a fight or flight decision, the preponderance of evidence from the senses demands action, a likely biological model for multiple weak ties. Specialists in brain modules based on functions such as hearing or sight tend to build models that reflect the complexity of animal responses to this single module, ignoring processing of stimuli from other modules. Therefore, models of integrated cognitive functions in humans tend to be more speculative, and rely on measurements such as event-related potentials (ERPs) that measure the difference and timing of electrical activity. These studies are conducted on specific, repeatable, and thereby isolated language or music events, attempting to measure the onset of electrical activity between established Brodmann[3] areas for language.

It is not surprising that many researchers have developed their own cognitive models based upon their specialized interests. Alas, researchers are themselves agents in a vast network of research, and subject to decisions and explanations based on their own perceptions of where the science is going. Examples can be found in the literature from philosophy of language to psycholinguistics of contagious fads such as nativism, chaos theory, connectionism and ecophysics. There are many parallels between the evolution of ideas and the production of science to support them. But the inescapable fact is that science currently has neither a central theory of mind nor of language. Working around the edges is network theory, boosted by the visible networks of the brain (chemicals producing electrical charges), man-made networks (mechanics producing electrical charges) and the networked connective activity of human agents, (software creating social networks).

There is some common agreement on basic definitions of parts of networks, biological and man-made. As we have seen, nodes are individual units, be that a person, phoneme, word or even a static phrase like 'isa'. They can be smaller units, like IP addresses, phones or lexical entries like radicals or verb stems. In theory these small units are connected to other nodes or are themselves hubs. Processing is done in functional modules, passing extracted or enhanced results on to other modules, which may further

[3] Brodmann areas are smaller divisions of the larger brain lobe cortical areas. The temporal lobe language areas are also known as Broca's area and Wernicke's area, and roughly correspond to BA44 and BA45.

synthesize the extract. Neighborhoods tend to have like-minded individuals with many connections, whereas islands are created when opportunities to connect are either ignored or dead-ended. Lexical networks, like dictionary entries or thesaurus entries, are appealing to educators and learners alike because these network 'maps' add back visual components such as contiguity and clustering. These connections are lost in the reified codification of reality that is school-taught language, whether written or spoken. Semantic webs connect word meanings, and the average number of links (called degree of separation) is only three. Syntactic webs have hubs that are functional words such as particles. Linguistic ‘rules’ determine the acceptability of any matching colocation, much like the shapes in a tiling mosaic, a virus invasion, or the locking shapes of a jigsaw puzzle. For the purposes of this paper, they can be labeled matrices, or set modular chunks of language held together by syntactic rules. They function during the early periphrastic phase of acquisition, before inflection becomes duty-free. Stranding, described here as filamentous strings of connections, is adopted in this context for the purpose of modeling language. Strands in a network are like bundles of cables which link over long distances, but in contrast to the myelinated bundles of motor neurons, they have charged momentary effects on neighborhoods by local stimulation of association areas. Unlike models of strong and weak ties, it is the strength of the bundle that prevails.

Turning Up the Register

According to Barabási, (2003) synonyms hold the "various lexical modules together" (p. 237). But, it is clear that syntax governs their use, and probably limits the choice of words through the mechanism called stranding for the purposes of this paper. More research must be done on Barabási's synonyms theory to refine his definition of 'lexical modules' and extend how he addresses the massive numbers of synonyms for common words, which is the puzzle of polysemy. Perhaps dueling strands in competition with each other can be used to describe how and when a particular synonym is actually used.

Alternate spellings, even alternate words for British versus American English do not seem to bother us when encountered in reading, and slip through our fingers when typing or wiggle out in speech, like our use of mobile instead of cell phone. They must be connected to one another in our mental lexicon. But what switches them on and off? Registers of language, and the bilingual's effortless slipping in and out of L1 or L2 for example, indicate that the switch must be relative to the context, or socially mediated. An intriguing extension of Barabási's synonyms theory would make L1 and L2 words merely synonyms of one another.

Talking about a graspable object excites motor neurons for grasping as well as lexical entries and areas required for language production. This phenomenon was first discovered based upon a mirror neuronal response in monkeys and is now extended to language. Combining motor activity with sight and language to match the dyadic communicative environment is now accepted cognitive activity. Semantic and syntactic webs are not the same, meanings are fluid, while structure is rather rigid. Lexicons, and the kinds of networks that connect words, might be fundamentally different from the networks that provide for language acquisition, which here are predicated on a syntactic dependency model. Lexical models, such as Barabasi's synonyms model, in which the hundreds of meanings of common words are connected to each other, are well established. Syntactic webs, requiring rules, have to be inferred from examples. People, both small and tall, gain mastery, often without explicit knowledge of the rules. How this is done can be explained by the reorganization of webs when too many connections make predictions difficult. Reorganization is required to insert new hubs in the hierarchy just as biologists do by creating sub-species in taxonomic classifications.

Reorganization at age 18 months to two years, that period during which infants move from two or three word utterances to simple sentences, can be characterized as moving from a tree-like structure with almost random connections to the interconnected small world structure of hubs and hierarchies, according to Solé, et al., (2010). This provides evidence of acquisition of the obligatory syntax rules and functional particles prior to real sentences, defined as complete thoughts. The evidence suggests that the word association hubs are very different before and after cascade. Not only is language re-organize but utilizes a new taxonomy (morphemic/syntactic). All of the connections to the heterogeneity of the reference points in association areas of sensorimotor functional areas are also reorganized. What was a hub becomes just another node, expanding the hierarchy, or a new classification (hub) is added.

This reorganization also happens to adults immersed in a second language, such as working immigrants not in a school setting. Not knowing the rules, they listen for what they know, what is familiar. The dictum to students editing their own writing is to "Read it aloud". This has real potency when considering how much of language is acquired by how it sounds. With enough oral practice in a contextual setting in which behavior can be modeled, the learner gradually ‘hears’ and detects false notes and incongruities. The best setting for learning appears to be goal-oriented and interactive, but not necessarily other-corrected. This also adds complexity to the network model, since there is most certainly an outgoing sound signal with some self-monitoring, hence the introduction of matrices or chunks to the learning tools. When learning lyrical music, the sound stream is immediately and continuously correcting both expected tone and anticipated words, and is naturally chunked into phrases. What might be interesting to test is whether new learners ‘hear’ their own voice when subvocalizing, and how closely it matches their output during the learning session.

Power Law and Random Attachments

The connection between Zipf’s (1949) power law (ranking words by frequency distribution) to modern theories of language acquisition has also led to the sociological principles of network growth and preferential attachment. This narrative of applying graph theory to language theory is told from two perspectives in the network books of Watts (2003) and Barabási (2003). The back story is that random attachments to nodes eventually leads to an egalitarian state, even if a snapshot during growth nicely graphs onto a Poisson curve. This egalitarian state would mean that all words are equally distributed throughout the language network. But we know, because of Zipf's pioneering work, that all words are not equal, either in usage, based on corpus studies, nor in associations. Attachments in the language acquisition network model are never random, which Watts neatly describes through modeling the social situation in which the rich get richer effect, and Barabási and team solved by suggesting additional nodes, not just additional links. This results in older nodes having an advantage over newer nodes in the accumulation of riches, the so-called power law distribution. Applied to language, the power law distribution model describes the exponential growth of words from a tiny core of essential expressions to full-blown competence.

Zipf was the first to describe the power law notion by ranking the frequency of words in English from corpora. It is not surprising that the most frequent words are not content words, but function words, at least in English, even as other studies report there is a clear noun bias in L1 early acquisition. Depending on the corpora used, the words are something like the following: *I, and, the, that, a, you, to, of, it,* and *in* and as any caregiver would add, *me* and *mine*. The list varies slightly between British and American English. Plotting frequency rank to number of occurrences results in a power law distribution, (i.e., word frequency is inversely proportional to word rank.) Applied to cities or businesses, regardless of all of their incentives and constraints to growth, the largest keeps its initial superior status, building upon that larger base at a rate that is twice its nearest competitor, and three times the third ranked. If language learning is largely implicit, the success of early leaners, according to this theory, seems to rely on exposure to variety, not intensity. But there is a catch, those function words, the requisite morphemes, have to be mastered as discussed later in this paper. This constraint cannot be overcome until requisite morphemes become implicit and native-like, without recourse to rules.

Consolidation a Self-Referencing Structure

The classical image of a fern 'leaf' can be described mathematically by a formula developed by Benoît Mandelbrot. This formula quantifies the notion of boundaries with self-referencing edges, infinitely replicative to provide for growth. The Mandelbrot Set, as each iteration is known, could be imagined as a root concept of a word and all its manifestations, including synonyms in L2. There are numerous researchers (and speculators) engaged in this branch of theoretical linguistics. An important direction for researchers and educators is to look at how concepts and words hold together (their stickiness), i.e., the process of creating edges or connections. Each construct could conceivably have a singular unique self-referencing, dynamical shape. Grammatical particles have an affinity for certain words, and the construction of phrases could be construed as 'charged' strands clinging to one another, and as a set slotted into an English matrix of S-V-O, for example. These co-occurrences or sets become implicit knowledge with even one exposure, just as single words do. After all, what is a 'word' in the sound stream of a new learner? Once the pattern is recognized, it allows for infinite recursion, as in the children's rhyme: This is the house that Jack built. This is the mouse that lives in the house that Jack built.

Other metaphors for language as networked are streams converging on rivers and multi-sensory groups, such as schools of fish or flocks of birds, or even predator-prey pursuit. These models indicate our preference for thinking as movement. Because current neuroscience cannot untangle thought from the many locations that show changes in electrical potentials on fMRIs, we may be confusing the metaphor with reality, which is likely to be much more diffuse. This is not just the modularity of the system imposed by our senses. Mammalian brains process input in remarkably similar ways. From an L2 language acquisition perspective, the more depositions from the senses, the better, since recall is not entirely under conscious control--think Madeleines and Proust's involuntary remembrance of things past. Smell, unlike other senses, connects directly to the emotional center in the limbic system. (Carter, 2010, p. 114) Alternate paths, that is, temporal experience connected through multiple modalities, may provide the most secure memories, because networked connections are multifarious. Temporal existence, yesterday, today and tomorrow is one of the earliest syntactic distinctions for infant grammatical morpheme acquisition. For adults, narratives are easier to remember because of their temporal nature, especially when connected to a pleasant experience.

When proficient speakers and writers use the process of how something sounds, rather than taking time to reference rules-based editing, is it the same phonological stream reference to what sounds ‘right’ that proficient L2s utilize? This is not an idle question, but has intrigued linguists and psycholinguists for decades. Grammar rules are descriptive of what is spoken. At the same time, these rules become prescriptive by utilizing another searchable network. To complicate matters, experimental evidence indicates that L1 lexical units are attached during recall of L2 lexical units. (Van Heuven, et al., 1998, as cited in Koda, 2004). This calls for additional mechanisms for switching and maintaining a particular language or register. This creates a traffic management issue as current L2 mapping suggests. Early, less proficient L2 repositories are not in same brain area as L1 or bilinguals. The evidence suggests the possibility that the early logjam phenomenon in less proficient speakers is the result of some L1 and L2 lexical items being clustered together, requiring a declarative rule to ferret out the right response. These roots and shoots should be accounted for in any model of second language acquisition. For example, etymology, or explicit memories for linguistic oddities like spelling of ‘separate,’ ‘license,’ ‘neighbor,’ and ‘neither’ have been mapped to separate areas of the brain from ‘normal’ processing, hence the theory of a separate computational network. For the L2 learner, the transition from pulling the words out of memory and lining them up in a sentence, to simply saying what one 'means' is at the heart of language proficiency. The adult L2 has to go backwards to a more child-like appreciation of the sounds of the language bereft of grammatical and syntactical rules. Music, especially lyrical music, but also metrical verse, provide this implicit knowledge by the imperative of the tempo, pitch and the necessity of synchrony. Language production sans melody seems to involve self-monitoring of what it sounds like, so educators should not expect early learners to produce sound strings or phrases that have not become familiar through multiple sense modalities.

CHAPTER III

MUSIC AND LANGUAGE AS TECHNOLOGY WITH SHARED RESOURCES

Overview

Adaptation theorists believe that when language evolved, it provided humans with the advantages that allowed them to dominate their landscape, both inner and outer. Language was the tool (technology) that provided the means for humans to distance themselves from the immediate. Evolutionary theories do not require a specialized language acquisition device, but rather a utilization of extant sensorimotor systems. We learn to use the technology for its existential imperative in a social environment. If we assume that a developed sense of self does not occur until well after the mobile infant encounters its image in a reflection, the question arises: To what should we ascribe the need to learn language? Could not the infant simply whine and then cry, escalating its neediness until fulfilled by its mother? After all, we see a powerful maternal instinct in many vertebrate species, such as birds, not just mammals. Each child must go through the repetition of its ontogeny, with language and song (motherese and crooning) comforting and quieting it. With the evidence of spontaneous signing among deaf children, it is apparent that no human children can thrive without language, given companions or caregivers. (The case is entirely different for ascetics or recluses who have taken a vow of silence. Post-language thinking, even when interfered with by monotones or chanting always involves 'thinking for speaking' even when one does not speak.) Some theorists say that this language imperative is based on the need to communicate; others say it is the need to play, especially role-play. Some researchers are looking for those brain correlates between what is clearly our specialized technology as compared to a baseline of our 'animal' selves. We share song with songbirds, and perhaps marine mammals, words and dance with parrots. Despite our best efforts with bonobos, who share much of our DNA, we have had more success in looking for the recruitment of established systems for language in birds, than in monkeys or apes.

What distinguishes language from mere vocalization is its internal order, or syntax. Music also has internal order, so applying syntactic rules to music can be seen as either provocative or logical. Patel (2008) has proposed a shared learning system for language and music, based on the similarities of the acoustical signal. Humans also have a special ability to pick out accents that differ from theirs. (Patel, 2008, 77-79) Thus, the imperative to distinguish friend from foe may have enhanced the repeated use of special sounds. (Passwords are a very primitive device in myth and our lives today.) If language is simply a technology, how are some technologies passed down through the generations for animals, but not others? For example, apes who learn to use tools seem to loose the special tool use between bands and generations, and have to learn it all again by trial and error. Other than song and language, there does not seem to be a technology that is passed down between the generations, unless one considers nest building, migratory paths, etc. When one considers these animal technologies, they were all considered innate. This accounts for the early emphasis on the innateness of language learning, and the theory's appeal, even if the language acquisition device is now recognized as spread across all sensorimotor brain functions. Few advanced behaviors in animals are now considered innate, having been shown to result from social learning and patterning.

In the 20th century, learnability theories spawned the notion of statistical learning based on input (rather than tutelage) as well as a digression based on the theory of the paucity of input. For example, according to Patel (2008), a current descriptive hypothesis of language learning involves "tracking patterns in the environment and acquiring implicit knowledge of their statistical properties." (Patel, 2008, p. 84) Note the use of the term implicit knowledge. One does not seem to have to know anything about pitch, tone or chord progressions to play an instrument or sing. For Patel and Balaban (2001), "perceived pitch is not determined by the physical makeup of sound but is a perceptual construct" (p. 87). Knowledge of pitch requires exposure to music, and the same is true of language to accumulate what sounds right. Meaning exists without overt knowledge of the rules of syntax. What is being argued is that music, like language, is a perceptual construct, and is dependent upon the environment for determining what is a 'sour' note or wrong key, what rhythmic patterns are acceptable, and even the very definition of music (think Rap).

Memory and Music

Language learning and sensory-motor sequences share implicit memory locations, as does music learning. The proposition is that both music and language may share a common learning system that does not require formal training, unlike reading or playing the violin. We can all enjoy music and singing without instruction, but it is highly unlikely that we could either read or play the violin absent demonstrations of the technologies by proficient practitioners. There are seemingly exceptions to this. Fabulous descriptions of savants and prodigies abound in the literature: the blind child or the autistic who sits down at a piano and mimics complex musical scores. Parsing the language stream for segmentation into syllables and words is equated to implicit learning of musical structures. Phonemes are like chords, sentences like melodies. Exposure to natural language results in grammatically speaking adults with little knowledge of grammar; listening to music results in ability to discriminate sounds that are unexpected, discordant, un-melodic, without explicit instruction. So-called syntactic violations of music, detection of 'sour' notes or discordant chords (Ettlinger, et al., 2011, p.1) support the thesis that “implicit memory seems to play an important role in syntactic processing in both language and music.” This study also acknowledged the difficulties with conflicting evidence for the independence or dependence of the two functions. They reference Patel (2003) and his earlier model SSIR hypothesis that demands a separation of the syntactic representations (patterns) from the processing.

Emotional Notes

Since the limbic system is influenced by music, older learners might want to consider music therapy immediately following an intensive language immersion session. As hokey as it sounds, we seem to remember the last thing associated with an experience, be it good or bad. (See also Peretz, 2010 and Koelsch, 2010). Explicit long-term memory

(LTM) recall is enhanced by music in at least one kind of behavioral experiment. In this type of experiment, remembering words from a list followed by music after a 20-minute interval, increases list memory. Researchers surmised this was caused by physiological arousal. Judde and Rickard (2010) found that the music seems to have a priming effect, which aids in memory consolidation, whether the music is liked or not. Students who were more affective in personality type, that is empathetic, enthusiastic or positive in nature, seemed to retain more. From a pedagogical point of view, music that is the choice of a student might be a best.

Examining neuroanatomical differences, Wong and Ettlinger (2011) found that among successful learners, nearly two thirds of the variance in learning was based on brain structure, caused by early musical training. They observed thickening of the connection between the frontal and superior temporal lobes as evidence of "streamlined neural networks" (Wong & Ettlinger 2011, p. 565). This thickening of the connective networks was observed in both amateur and professional musicians. Learning differences were observed in the ability to process pitch patterns, a necessary skill for foreign language phonetic learning, especially in tonal languages like Chinese, but applicable to the acquisition of all languages. This area of research might provide an avenue for study of language fossilization, especially for accent improvement. Studies of dyslexia also show reduced phonological awareness, not just visual miscues. Applied researchers speculate that these deficits could be somewhat remediated with the addition of musical training to augment phonics drills, but the literature is almost always based on young learners. Prodigious learners of the past used rhyme and song to pass the great epic stories down through the generations. Today, we have Sesame Street to teach preschoolers. Both operate under the same rules of repetition, rhythm and rhyme to fix words and syntax in memory. In the past, it was dangerous to extrapolate from child acquisition studies to adults, but recent studies have extended brain plasticity well into adulthood. In fact, plasticity or neuronal reorganization has become the central dogma of cognitive function today.

Who's Got Rhythm?

When defining rhythm, Patel (2008) notes that periodicity is only one element along with accent, which is grouped systematically into temporal units. The resulting patterns can be compared and contrasted between language, only rarely having long stretches of systematic pattern, and music and poetry that do have these characteristics. Rhythm itself becomes the major organizing principle for much music and poetry. It is interesting that dyslexic kids often cannot tap out the rhythm to a song. This inability, besides being a quick diagnostic, should encourage testers and educators to always think in terms of spectrums or continuums for both deficits and exceptionalism. Consider the repeated head banging of autistic kids, which apparently brings some temporal anticipation and perhaps some solace. Finally, consider the use of Music Therapy to temporarily animate patients with degenerative brain disease. All of these are symptomatic of a suspected deficit in self-induced mental rhythmic animation.

Marching to a Different Drummer

Beat perception is actually anticipatory in humans. Comparison of music beats to recordings of finger or toe tapping indicates a slight prior brain and motor response. This could be explained by the necessity to synchronize by starting slightly early. It also is likely a survival technique as any student of Taekwondo Do or Karate can attest. The ability to avoid blows from an opponent would certainly convey an advantage. Patel (2008) points out that a sense of time is necessary for synchrony to succeed. This is likely to be unique to humans, according to Patel (2008, p. 102). Synchrony appears in many species, such as crickets, but it is not a voluntary act of synchrony, but rather the effect of trying to interject into an existing sound stream and missing the mark. 'Accommodation' to the speech rhythms of others is not just a matter of pragmatics or sociolinguistics. It is also evident in music, where melodies played at different tempos can be quite recognizable, even to infants. (Patel, 2008, p. 117) Contrast this with speech, which does not have periodicity, but does have "subjective intuitions". This leads us to believe there is something about timing in language prosody such as the speech of a master rhetorician Martin Luther King, Jr., something about the stress, timing, alternation of phrasal intonation that make certain orators sound almost musical and therefore more meaningful. However, the perception of speech isochrony can be explained by the human desire to find these musical periodicities in speech, despite empirical evidence against it, according to Patel (2008, p. 121). The rhythm we hear is actually the alternating long and short vowels (reduced vowels) plus a second separation in timing of syllables based on pitch (Patel, 2008, p. 124-126). This results in a description of English as having a more scattered cadence with a variable duration between consonants. This leads to "stress-shift" in order to avoid "accent-clash", which is a confounding non-trivial problem for L2 adult learners. Dutch and English are both characterized as having more consonants and fewer vowels to accommodate this spread (Patel, 2008, p. 128). This has relevance for varieties of Global English, since the ability to switch between stress-timed and syllable-timed languages is difficult and perhaps culturally inured. Lyrical music, sung by those whose accent is apparent in speech, disappears when singing, revealing that timing is everything.

Often it is the language that changes, not the speakers, especially when there are large numbers of Global English speakers as in subcontinent India. Patel suggests that studies of the differences in rhythm between Singapore English and British English, for example, might show evidence of the putative theory as reflected in changes to durational differences in vowel and consonant patterns--its rhythmic classes. This paper's proposition is that lyrical music is an easy mechanism for encoding these rhythmic classes for the novice, even if they lack true periodicity in stress, but are simply perceived to have regularity. Our mental topography is built upon perceptions of reality, mediated by our current abilities and focus. Thus, stroke damage to the visual system can cause visual recall dysfunction, much like attentional load causes texting and driving disasters. Therefore, the sense that song lyrics are closer to speech rhythms is nonsense, but the relevance to constraints and rules of perception is important. The music does not stop, even if the singer pauses. Synchrony will reassert itself without mindful effort.

Speech rhythm is important because it goes to the heart of learnability theories. Babies have to be able to extract word edges and adults have to overcome phoneme deafness (Patel, 2008, p. 138) and possibly stress deafness caused by their sense of diminished self. But those very non-existent speech rhythms (remember it is duration, not periodicity) will eventually lead the learner to the detection of phrase boundaries in connected speech. Humans are able to detect slight hesitations within phrases to provide sense. For the L2 learner, the implicit parsing strategies used for the L1 are automatically applied to the target language (Patel, 2008, p. 148).

Although they have dissimilar syntaxes, language and music share many other faculties, especially acoustical production and perception. Testing of music/language brain function correlates can take two forms: Either lyrical music or melodic patterns without lyrics. Typically, experiments test for brain function based on electrical measurements revealed by anomalies such as 'sour' notes or anticipatory sequential notes. Incongruence testing, a staple of language testing, is also a baseline for a number of testing paradigms for pre-linguistic infants and animals. Obviously, animal survival depends on a developed ability to distinguish and deliver fight or flight reactions. For music, incongruence is a "lack of membership in the prevailing scale" (Patel, 2008, p. 201). Membership is learned by exposure. Several problems arise in this testing paradigm because of the human capacity to learn implicit patterns from a very few instances. A neonate's exposure to sound systems predisposes it to expectations of specific sorts. Like all animals, we seek regularity, not necessarily symmetry, and prefer the familiar. The three basic ways that humans draw inferences from their environment are resemblance, cause and effect and contiguity under safe conditions (Patel, 2008, p. 336). For the preyed-upon, being able to scan the horizon and pick out shapes that are anomalies; being able to deduce natural laws like gravity; and finally, being able to distinguish close or distance sounds and patterns of regularity (like hooves versus wind) are all selective for survival.

Melodic pitch contour is the part of speech production we notice in Valley Girl uptalk, where every statement sounds like a question with rising intonation at the end. In fact, all languages follow acceptable patterns, and among age cohorts, there is standardization. So, there must be both implicit and explicit controls on these very same aspects of speech. Tone-deaf persons seem to be able to speak and hear reasonably well, but their singing can be painfully unaware: They are unable to detect 'sour' notes. Through his hypothesis, Patel (2008) explains that 'tone deafness' is a result of a deficit in pitch direction recognition, obscured in speech by surrounding context.

Detect or Regret

Song, with its greater dependence on continuous pitch and direction, does not just pitch shift, nor does it rely solely on change detection. Patel submits a hypothesis based on short-term memory deficit for pitch patterns, using neuroanatomical studies. Since control of pitch direction is a major contributor to fluency in L2 English acquisition, this paper suggests pitch is amenable to explicit learning through song. Just as Valley Girls can code switch from dizzy blonde to academic English by changing pitch direction, so too can L2 learners acquire pitch control for language through lyrical music, and real word edges like *"isa," "wanna," "hafta"*.

Experiments designed to establish empirical evidence for specific theories often confound the effects of one kind of activity with others. Since language is the most complicated interpersonal technology that we have,[4] focus must be narrowed to the task effects when examining/testing/developing maps of brain activity (Hickock & Poeppel, 2007, p. 393-402). All the sensorimotor systems are in play with language, and even deaf or blind persons who have missing systems seem to have brain territory recruited for language performance in other modalities. Speech processing requires a perception side for the incoming acoustical signal prior to phonemic decoding. For speech recognition to happen, there needs to be a computation function prior to the interface with the lexicon. This interface between the morphemic structure and lexical access is language-specific, and for multi-linguals, addressable consciously. That this might be true, think of standing in a crowded polyglot public conveyance. As one listens in turn to neighboring conversations, each can be switched on and off. Hickock and Poeppel, (2007) believe that there is a specialization for language in the human brain, because of our ability to pick out man-made sounds. We are experts at picking out voices, which is to be expected, but can also focus on specific musical instruments, or even seat of the pants 'tuning' of automobile engines by mechanics, for example. They surmise that lexical items must have some property "that sets them apart from other auditory information." (p. 396). This allows them to develop a model of speech processing that separates the sound stream

[4] Music is slightly less complicated as a technology because it lacks semantic content.

from the speech recognition stream; that is, processing for comprehension separated from articulatory networks. In their theory, new vocabulary entails deposition of a "sensory representation...that codes the sequence of segments or syllables" (p. 399). Thus auditory and motor sequence interactions become conjoined like Siamese twins.

The inseparable nature of voice sounds and motor skills takes place in the pre-lingual period for infants, and can be accomplished by adult language learners through non-passive interaction with the sound systems of the target language. Hickock and Poeppel's (2007) theory nicely complements this paper's proposition that lyrical music can serve as an aid to language acquisition because it describes how the sounds of words have a life of their own, connected to their own network. Rather than just shared syntactic systems for melodic music as in Patel's (2008) theory, lyrical music must also share those specialty functional areas of the brain that process the language sound stream. It might be obvious, but articulatory actions associated with singing, humming, whistling (or sisseling, the whistle associated with a 'su' phoneme, versus whistling associated with a 'wh' phoneme) all can have a 'thinking for speaking' component when recalling a specific lyrical piece. Except for singing, all other background musical activities, such as drumming one's fingers, etc., can be purely rhythmic, even atonal, like a cat's purr. Even infant cries have rhythm and have a sound envelope similar to the speech in their environment.

A question for future research: What sounds can be produced in and around the human mouth (besides imitative animal calls) which can be successfully separated from the sound stream of voices, for which we are exquisitely developed to recognize?

As soon as those sounds are attached to words, even nonsense words, we can recognize the sound as human, and often, with very few iterations, who the speaker is. (Scat singing is an example.) Disembodied radio voices suddenly attached to TV faces can be disconcerting, but henceforth, we cannot hear the same speaker without calling up an image. Using the face recognition system as an example for adult learning, many people instantly forget a new introduction's name, even if they repeat it with polite phrases such as "Pleased to meet you, Robert," and again on parting. This phenomenon can be explained by the lack of salient features to tag the explicit name memory to the implicit face memory that is stored away without effort. Weeks later, the face is recalled automatically, but the name is an arbitrary word like a foreign language vocabulary word, quickly forgotten. What is missing is a classification system, which faces have, but names do not, resulting in naming conventions in many languages, such as 'von' and 'Ibn' which attaches the name to places or tribes.

This is the basis of theories of multimodal learning: Yes, it is possible to learn a reified fact about something, and regurgitate it for an exam, or rehearse it for a speech act. But real learning, native-like L2 learning, requires a kind of experiential learning that is embedded in several layers of intersecting webs of sensorimotor perception leading to cognitive activity. This is not a new idea in linguistics. Edward Sapir (1921) considered along with Franz Boaz, fathers of modern linguistics, said that language could not be localized is a specific brain area, but rather had recruited existing functional parts of the brain, both newer cortex areas and deeper, older, more primitive areas. Paths of association and shifting networks, between "...all possible elements of consciousness on the one hand and certain selected elements localized in the auditory, motor, and other cerebral and nervous tracts on the other" (p. 4) describe current thinking, in a continuity that spans a century.

Kandel and Hudspeth (2013) in summarizing the history of knowledge of the brain, from a current perspective, moving from older view of discrete areas to an interlaced capacity, declare "all cognitive abilities result from the integration of many processing mechanisms distributed in several regions of the brain." (p. 17). Besides providing for redundancy, and allowing recovery from injury, massively parallel processing allows switching channels as more important input is parsed. This is close to how we build modern enterprise wide databases: Transactional processing in each module adheres to a posting schedule that returns the greatest knowledge about the enterprise at any point in time, while conserving computing resources. This reflects the distributed nature of the processing while allowing for integrated functions that include data from sub-processes or sub-modules. This could produce a strobe-light effect, but the brain integrates the snap-shots and smooths the herky-jerky input into a continuous stream.

Think with Your Eyes

One of the curious things about domesticated dogs is that they watch your eyes when you speak, just as do human infants. But this is not true for autistic youngsters; their gaze is fixated on mouths. Motorcycle riders are admonished to look ahead to where you want to go, not down at the road in front of you. You and your bike will follow your head's lead. So too, older people or people recovering from injuries and 're-learning' how to walk are advised not to watch their feet, but keep their eyes a few steps ahead and let their kinesthetic sensibilities to do the walking, while their eyes probe the near distance for oncoming issues. Thus, motor activity runs its own course, offloading the automatic part of the activity to routines in other modules, rather than attempting to explicitly control everything. Recitation of poetry or singing of songs aids language learning by eliminating conscious choices and re-directing attentional loads. Songs also have the added advantage of endless repetition, like earworms, caught in an endless loop, sub-vocal and sneaky.

Future Research

Despite Patel's evidence of functional brain correlates for syntax in music (typically chords or sour notes) and language syntax (typically morphological rules like verb agreement and word order) he cautions that empirical evidence of agrammaticism in Broca's aphasia plus musical disharmony is tentative at best in determining structural knowledge (implicit rules) in the two domains. Brain lesion studies cannot be definitive because, among other things, damage in one patient will seldom be replicated in another (Patel, et al. 2008, p. 776-789). What Patel suggests is that there is a need to study the relationship between musical harmonic dysfunction and specific language dysfunctions, and whether there are mutual deficits (because both are implicitly acquired) in "other measures of language function," (Patel, et al. 2008, p. 788), presumably also implicitly acquired like morphemes.

This paper argues for a structural relationship as the underpinning of all syntax, whether musical, visual, tactile or lingual. This somewhat rigid structural system is balanced against an implicit reliance on the right fit, with a large aural component, and a visual component. We are constantly scanning the visual field, even when blind. Speech is ultimately an implicit sound pattern recognition cued to the visual space. For social animals, that visual space is faces, especially the eyes of others, and their voices. Rather than saying that humans have an innate learning ability for language, perhaps what we should be saying is that we humans have an innate detector for anomalies in our sensorimotor faculties. For language, we are prenatally attuned to the voice of our mothers, and then extend this to other senses by maturational experiences. But that might entail a social explanation for language acquisition, bringing us full-circle to the nature vs. nurture epistemological debate that the philosophy of language has become ensnared. Indeed educational philosophies are caught up in the same web, relying on externalism as determining our mindfulness.

CHAPTER IV

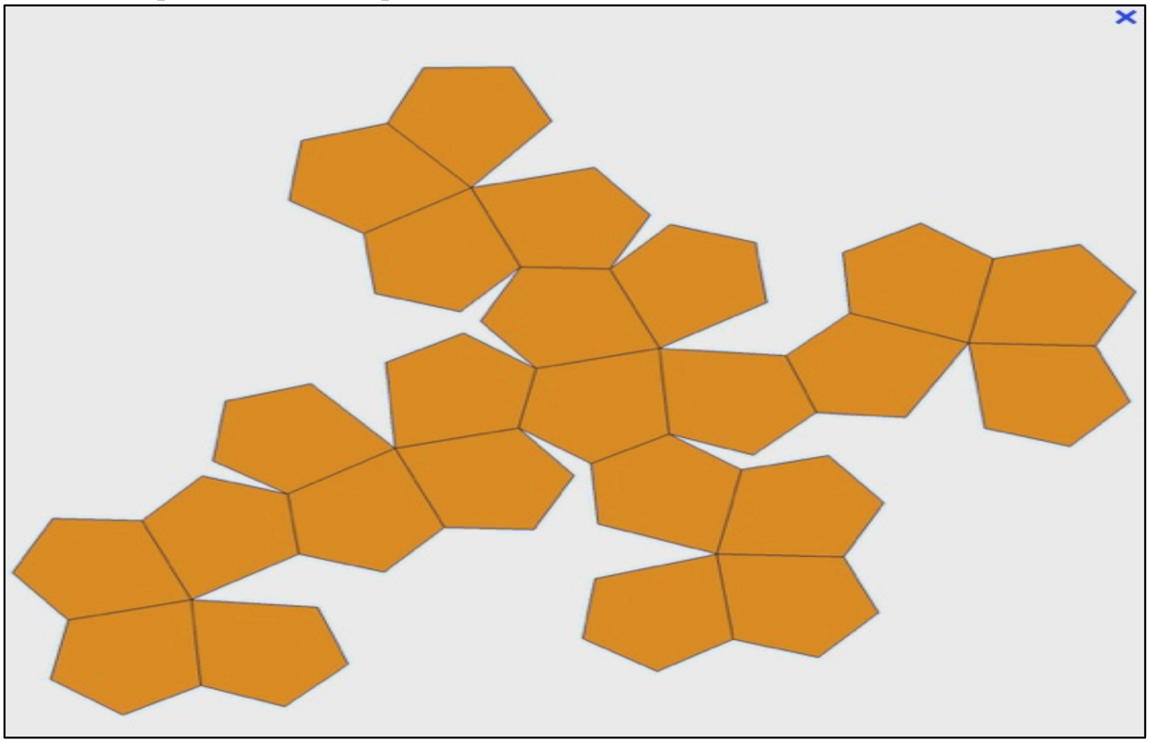
L2 ACQUISITION AS IMPLICIT, BUT AMENABLE TO STRATEGIC ALLIANCES

Language Learning Strategies

Fast mapping is a language strategy utilized by early learners. However, it has a downside. According to Twomey, Horst and Morse, (2013), for 24 month old early learners at least, recall after five minutes is virtually nil (p. 175-176). For example, given three fruits, an apple, and orange and a quince, and asked which is the quince, the early learner can easily identify which is the quince, if the learner is already familiar with apples and oranges, but at the same time be unable remember the label quince when asked what it is just five minutes later. This applies to L2 adult learning, too, especially when the conversation partner supplies the missing word to move the communication forward. Since the word is not remarked, it is swiftly lost. Learning can occur if interest is engaged by directly labeling and comparing quince qualities to the apple and evoking other senses; its fuzzy, waxy exterior; its aroma, etc. This requires repeating sub label similar characteristics to the known item, whose characteristics have been observed, but not perhaps previously labeled. Then, the early learner must use the new word make it more permanent. Its connection to the fruit node in the semantic library is strengthened by packing the new item into place with more subcategories, providing a recursive route for mnemonic memory and recall for the fast mapped item, per the model of stranding.

Edges join, reinforcing and strengthening the connectivity of the network, or are at least momentarily activated in association areas because of close neighborhood ties.

Figure 1
Stranded Explanation for Simple Network and Recursion



Note: This stranded model illustrates how tension is increased at the node edge allowing salient features from one strand to 'stretch' to attach to another strand. © S. Stephens

Domain General Acquisition Processes

Arguing for domain-general processes to explain infant language acquisition, in contrast to the innate module theory of nativists like Chomsky, Walk and Conway (2013) looked to evidence from differential results on language skills for children with cochlear implants (CI), compared to control groups and norms of hearing children. Two domain-general abilities, auditory and serial order processing were used to verify what was 'missed' by not being in tune with environment. One of the earliest windows for language acquisition is phoneme discrimination. Current accepted theory is that acquisition begins pre-birth; by three months, the infant is strapped rather rigidly into its native language straightjacket, rejecting other phonemes as possible clues. In a review, Kuhl, (2004, p. 831) calls it "the brains commitment to the statistical and prosodic patterns". Walk and Conway's theory, supported by their own and other studies, seems to implicate phonemic learning as having a narrowing and focusing effect, with a precise short window of opportunity, almost as if maturation were on a forced march which heaves unutilized bits into the mental dustbin in its headlong rush to lighten the cognitive load. The implication for the late L2 learner would be to spend as much time (all the time) listening and producing the sounds of the new target language. Instead of babbling, (immature vocalizations), late learners have the advantage of mimicking actual songs, poetry, prayers, etc., which can perhaps excavate those lost bits from the dustbin of history. In fact, adults seem to better utilize structured sequential pattern learning when it is provided orally, in preference to written learning. Deprived of the early sequential input based on language, deaf CI children may exhibit deficits in motor tasks seemingly unrelated to language that requiring sequencing. Hand-eye coordination seems to be inextricably connected to the ears and mouth. According to Walk and Conway (2013), domain-general sequential learning is strongly implicated in language deficits, specifically "learning the rules governing word order." (p. 365). The so-called language universals are no more than the infant's brain responding to environmental clues to interpersonal meaning through language parsing. "[D]ifferent domains of sequence processing require the extrapolation of different types of dependencies that in turn lead to the development of different types of language skills" (Walk & Conway, 2013, p. 353).

Language acquisition moves from discrimination of phonemes to expectation of phoneme order and then to parsing individual words before moving on to word order. Last, functors, requisite morphemes and progress from single words to multiple words develop (discussed elsewhere in this paper.) As a strategy, Walk and Conway (2013) support the essential idea that sequence learning has a fundamental role, and that it is domain-general, part of hearing and thus environmentally correlated.

Late L2 learners might be comforted by knowing that deaf and hard of hearing youngsters also have difficulty in simple Hebbian learning. But CI kids do eventually catch up. Lists of items, such as vocabulary, are supposed to be amenable to learning by repetition. For L2 learners, this usually means solitary practice supported by flash cards and drills involving writing. But for CI youngsters, this kind of learning is significantly impaired, and seems to correlate to smaller vocabularies, as well as degraded responses to learning non-verbal stimuli that require repetition. Because there are two serial order mechanisms for learning, the adult L2 has got an advantage over the deaf learner, even if the adult L2 learner also struggles to hear, differentiate or articulate some of the phonemes for proper phonemic sequence processing. The adult has intact domain-general tools for language acquisition, which the CI child is still developing. "[P]robabilistic sequence learning underlies syntax acquisition...whereas repetition sequence learning underlies lexical development" according to Walk and Conway (2013, p. 362). Recovery from the deprivation of inputs for the CI children seems to be ameliorated, if not almost completely overcome with time and exposure, and so it also improves (eventually) for L2 adults. Walk and Conway suggest further studies to longitudinally examine the deficits due to delayed probabilistic sequence learning and repetition learning on vocabulary and syntax in CI children. This writer suggests the same should be done with L2 adults, testing improvement to predictions of rhyme, meter, etc., since "an individual's ability to make sense of structured sequences directly predicts language competencies in both typically developing adults...and hearing-impaired children" (Walk & Conway, 2013, p. 361).

Implicit Versus Explicit Learning

Theories of serial order processing through repetition sequence and probabilistic sequence learning as separate and distinct mechanisms must be contrasted with theories of explicit and implicit learning, a major thread in learning theory. (Morgan-Short & Ullman, 2013). Walk and Conway (2013) reference Ullman's (2004) separation of syntax from vocabulary learning via procedural/implicit versus declarative/explicit memory functions, but position themselves as agnostics, believing that explicit and implicit sequence learning are false dichotomies. Rather, they implicate both implicit and explicit memory uses for probabilistic and repetition sequence learning (Walk and Conway, 2013, p. 362). And, as it turns out, Morgan-Short and Ullman (2013) would agree. It is all a matter of proficiency. In a more recent study comparing event-related potential brain responses (ERPs) of L1 and L2 learning, Morgan-Short and Ullman tried to disambiguate the problem of explicit and implicit learning by using a contrived language with natural language syntax. Tested were the native-like neural responses to ungrammatical expressions and the similar neural responses for high-proficiency L2 learners, who in this experimental paradigm learned names of game pieces and descriptors, both adjectival and adverbial along with a few verbs and case determiners. They concluded that only implicit training resulted in the native-like ERP signature for ungrammatical expressions. What is not known is how the learner actually attains high proficiency. What Morgan-Short and Ullman (2013) speculate is that though the route to L2 proficiency is not known, high proficiency L2 learners should exhibit similar brain function to those with native proficiency. Not solved is the pedagogical problem of immersion versus explicit grammar teaching/learning. What Morgan-Short and Ullman are ultimately arguing for is that there is a profound difference in brain activity as a result of implicit training versus explicit, and that implicit function using language is more native-like than calling up declarative, explicit knowledge. L1 learning mechanisms are so efficient, L2 learners would be smart to emulate them. This is not a declaration of war on explicit grammar, or meta-language explanations. As their ERP studies revealed, high proficiency L2 brain function is much like L1 brain function. This only reinforces the imperative to study the mechanisms of early language acquisition in infants. Further, they found that there were brain function differences between implicit and explicit learning are not reflected in outward behavioral differences in production. Morgan-Short and Ullman's (2013) study using ERPs to examine implicit training to acquire a small artificial language concludes that "...syntactic processing can come to rely on the same biphasic mechanisms found in L1: rule-governed automatic structure building, which may involve the engagement of procedural memory, followed by controlled structural reanalysis, accompanied by an increasing demand on working memory" (p. 942-943). Their observed increased demand on working memory is viable because units are consolidated into larger and larger chunks, finally becoming real phrases which role off the tongue like political clichés: "The fact of the matter is...."

One of the widely used tests for L2 proficiency is the ability to mimic longer and longer sentences. Elicited imitation (EI) is based on the limited number of items which can be held in working memory (phrases, words, syllables), but which can be increased by off-loading to habit or predictable chunks. Thus, memory is increased with added knowledge, or by firmly attaching a grammatical item to its most likely niche. Unlike spontaneous speech, which was the basis for most of the original studies in the 1970s, EI allows targeted testing and can be standardized between subjects, languages and age groups. Also, because response time is very limited, mnemonic devices or other 'lifelines' cannot be utilized by savvy test subjects.

A well-known issue for all experimenters is bias produced by the test/observation itself. By directing a floodlight on a solitary, isolated activity, the activity may become enhanced, exaggerated, or worse yet, a mockery of real activity. Many studies show explicit grammar training, as an example, is more effective than implicit rules acquisition, but there are skeptics. The phenomenon may just be the effect of "reliance on domain-general attentional mechanisms rather than the syntactic or lexical/semantic processing that is typical for native speakers and, apparently for implicit learners of an L2." (Morgan-Short & Ullman, 2013, p. 943). In general, procedural memory, as opposed to semantic memory, seems to be more 'sticky'. It is invoked through motion or volitional direction towards a goal, that which is a feature of language production from an embodied cognitive viewpoint. Figurative language, especially metaphor, lyrics, proverbs, etc., also have an automaticity, achieved through repetition and ubiquity. Morgan-Short and Ullman have provided a glimpse of a tedious path for L2 learners from low-proficiency language skills reliant on declarative memory to a smooth passage supporting lexical/semantic processes revealed by native-like ERPs with high-proficiency and reliance on procedural memory. (Morgan-Short & Ullman 2013, p. 944). Procedural memory, (implicit) or how we interact with each other, is recognized for its automaticity. Procedural memory also has a large component of motor memory, and motor memory is always sequential. This may explain why song can animate persons suffering from dementia. A common example of memory that has become proceduralized is when you cannot tell someone your password, but need a keyboard for your fingers to 'remember' it.

Morgan-Short and Ullman suggest that there is a path for L2 learners from tedious high-stress, low-proficiency declarative memory supporting lexical/syntactic processes to a state of native-like ERPs with high-proficiency and reliance on procedural memory (Morgan-Short & Ullman, 2013, p. 944). That path out of declarative memory purgatory depends upon the goal: Short-term explicit grammar learning, soon to be forgotten, or a longer-term implicit grammar acquisition, such as that obtained through language immersion (Morgan-Short et al., 2012, p. 15). Since we know that lyrical music and poetry were universally used to both entertain and pass the great stories of myth and legend between generations, why have we given the method short shrift in academic English programs in the era of books and Google?

Babbling

Accent amelioration is possible, even for late learners, but most researchers agree there is probably a limit. The phenomena of universal babbling[5] in both bird and human infants suggests to researchers that there is a window of opportunity for infants to reject non-conforming sounds and for caregivers to enforce conforming sounds, at least in humans. Just how the infant acquires the sounds of its surroundings and discards trials, is detailed in a recent research article by Lipkind, et al. (2013), concerned with the babbling phenomenon in birds and infants.

If parsing is a major milestone for the language-prone infant, then babbling must surely be another. Productive capacity is probably based on imitative abilities as well as motor control, and usually observed by caregivers when the infant is five to seven months old. But much has already biased the learning of the infant in its environment. So, although the act of babbling has many similarities across languages, such as snipping of unused phones happening at same time as combinations of new phones are experimented with, it should be emphasized that these babblings are language specific. Lipkind, et al. (2013), studying finches, and comparing them to utterances of pre-lingual infants found in the CHILDES database, discovered that the phone bigrams

[5] Known as subsong in birds. Birds raised in isolation still produce partial songs, leading researchers to believe that there must be a component of feedback, as is true for human infants.

were stepwise, that is chained, but never simply intercalated, even as they seemed random. Birds could be taught new phones, but they did it first like AB->ABC but not like ABC -> ACB (Lipkind, et al. 2013, p. 107). Phones are added to the head or to the tail in reduplicative sequences or to replicate bits of environmental sounds. As the infant babbles, a simple motor network is built. The agonizing wait from those first syllables to real words is explained by the network itself, which has to integrate new 'pairwise vocal transitions' by first building connectors and amending the new sound sequence on the fly for sounds that are not usually sequential. Development of the human infant language or songbird's repertoire has multiple sensitive periods, and depends upon the synchrony of sight and sound modules in dynamic interactions. Lakshmi Gogate and George Hollich (2013) characterize this as a "complicated cascade of outcomes each of which builds upon and results from the complex interactions that take place between the organism and environment (Gogate & Hollich, 2013 p. 30). This streams theory of sensitive periods easily accommodates the dynamic neural field theory (DNF) model of in-line cognition as described by contributors Samuelson, Spenser and Jenkins in Gogate and Hollich's (2013) book on trends in word learning. These two models, DNF and cascading streams of sensitive periods support this paper's stranding model in which sensory streams constantly re-work the pathways like the forces building an alluvial flood plain. Plasticity is always guaranteed, and memory is distributed among multi-modal pathways. To extend the cascade and alluvial flood plain metaphor further, it is not just the existing streambeds that determine where the floods go. The pattern also depends on the amplitude and frequency and each network is idiosyncratic.

Hearing dominates infants' learning giving way to visual at about four years; adults exhibit visual dominance and visual bias (Gogate and Hollich, 2013, p. 36-37). Intersensory redundancy allows children to learn language, even with sight or hearing deficits. For adult L2 acquisition, past the brain maturation age, irrespective of theories of sensitive periods, immersive learning provides the similar redundancy effects, retraining L1 mappings to accommodate new stimuli. "Language development is much like a river;" the more you know, the more you can learn (Gogate & Hollich, 2013, p. 43). Like most sensorimotor language models, their model is based on invariance detection, presumably after or co-terminus with familiarity of phonemes through labeling to the arbitrary word-referent basis of language. Invariance is a domain general learning mechanism, and is utilized by all organisms. Babbling, so far unique to birds and babies, is thus a necessary feedback loop for both hearing and articulation, and necessarily precedes whole-word production. Karaoke sung in another language is a modern version of adult babbling.

Mirror Neurons

The possibility of mirror neurons was probably never in doubt to modern cognitive scientists, even as observant animal behaviorists recorded examples of mimicry in the beasts. This writer first encountered the mirror neuron's powerful influence when surreptitiously watching a moose mother and calf browsing on the roadside; somehow forgetting myself, I began munching in synchrony with the bulky pair. The ubiquity of expressions such as aping someone, or monkey see, monkey do, though somewhat pejorative, reflects a common view of human behavior that mimicry is a normal activity. Experiments on monkeys, at the single neuron level[6] have indeed confirmed the existence of mirror neurons that have the effect of causing the firing of the same motor neurons associated with real activity when the monkey observes another animal doing something familiar such as grasping a cup. We all know the contagion of the yawn reflex, which is intra-species. As a human learning mechanism, mimicry along with its siblings pantomime, universal facial expressions, and posture contagion all contribute to unconscious implicit learning. That may be why whole language learning within a natural milieu can be more successful over rote memorization of discrete items. Subtle clues are observed, but perhaps not consciously noted. Repeated, these gestalts become part of the learner's repertoire. Learning in situ and oriented by a panoply of sensorimotor inputs allows recall to run along various strands, those alternative routes that support language, but are not traditionally thought of as central to language.

Are the mirror neurons turned off in adult learners? In order for the motor neurons to work for articulation, the language learner needs to already have in place motor control for those vowels and consonants. The learner has to be able to distinguish the sound in the first place, and also has to have the multiple combinatory phonemes stored and retrievable. Individualized feedback through both sight and sound ensures efficient learning. Professional speech coaches and TESOL instructors alert new learners to the general errors made in some of the larger speech families like the *r* and *l*

---

[6] Most notable is research done by Giacomo Rizzaolatti in the mid-1990s. Probes were permanently inserted into the monkey's brain. Such experiments are not considered ethical, today.

consonants for Japanese natives, but ultimately, success is goal driven. Adult learners, rather than lacking in brain plasticity, may consider nuances of articulation not necessary for communication or integration, and resist repetitious practices as wasteful of their time; they are not trying to 'fit in'. As stated earlier, accent will not be self-remediated, even through years of exposure, if it not an existential threat. Accent learning/remediation requires both sight and sound feedback, ideally in synchronous articulation with the tutor/expert.

Mirror neurons fire automatically in persons trained in piano, stenography or Morse code, when, for example they see or hear other individuals performing or speaking. Their hands covertly move in synchrony with the incoming sound stream. These specialized movements were developed over time and with considerable practice, leaving behind permanently enlarged pathways in the brain. Songs, poetry and language prosody generally do the same thing, but perhaps with less explicit practice and repetition because of a phenomenon called 'earworms' whereby a person is plagued by repetition of a catchy phrase, over and over. At the extreme, it might be related to paranoia, even hearing voices, but in the normal course of living most people have experienced phrases or bars of music that seem to stick around, sometimes for days. This annoying repetition of catchy phrases and melodies, rather than a negative syndrome has been harnessed by second language learners through listening to the repeated playing of popular music, like the Beatles, and finally parsing and then singing along with their favorites. How does this work, and under what conditions? It has been a central theme of this paper that principles of synchrony, periodicity, and mimicry along with pitch discrimination, innate mechanisms of brain modules designed by natural selection, which have allowed humans to develop an exogenous tool called language. Any activity that enforces and enhances these innate structures will reinforce language learning. Music, both lyrical and melodic, and speech prosody in the form of poetry provide strategic alliances for reinforcement of these structural patterns.

CHAPTER V

MORPHEMES PLUS FUNCTORS: KEY TO L2 MASTERY

Overview of Language Spurt and Brain Reorganization

Morphemes and functors are described as integral for an explanation of the expansive growth after a period of latency in which gestures and phonemes are tested through mimicry. This model of language growth is applicable to infants and adult L2 language acquisition and is a major assumption of this paper. New language learners must establish a taxonomic network of usage-based routes along which meaning is communicated to themselves and others.

Setting the Stage

Roger Brown (1973) developed a five-stage language acquisition model for infants through about three years. Published as *A First Language*, it was originally meant to be the first of two volumes, with Stages I and II in the first volume, with Stages III through V to follow. Only the first book was published, mainly because, for the latter stages, data were too difficult to collect and collate. The outpouring of child language after Stage II was just overwhelming. Somehow, having reached Stage II and on the way to Stage III, child speech takes an exponential leap in size and complexity. That turning point has been a common thread throughout this paper. The data collected were almost entirely from three youngsters, known as Adam, Eve and Sarah. Brown (1973) included comparisons of his subjects to those of other contemporaneous researchers as both a foil for his unique morphemic stages theory, but also to elucidate contrary theories, especially where the 'live' data failed to support, or indeed enforced various other theories. One of the outcomes of the study was the discovery of 14 requisite forms that are not only minimal forms, but also easily detected as being obligatory in the particular context/utterance. As these forms modulate meaning beyond the expression of solitary nouns or verbs, they are also notable for being (mostly) monosyllabic, not stressed, and unlike their content cousins, members of small closed classes (Brown, 1973 p. 12).

These stages are roughly based on mean length of utterances, with cut-offs between stages at the point a majority number of utterances falls above the arbitrary Stage I, (one word), Stage II, (two words), etc. Multiple word utterances were also roughly correlated to age/maturation. It is this paper's proposition that there is a similar stage transition in adult L2 acquisition tied to a similar brain re-organization, allowing lexicon to conjoin syntax to provide meaning and lead to productivity and the super growth seen in vocabulary acquisition beyond these two stages. The requisite morphemes produce connections between subject/object and action/being that enhances contextual memory. Syntax is not necessary until two-word utterances, i.e., order plus inflections.

Stage I for Brown (1973, p. 75) is notable for its lack of functors,[7] (grammatical relationships), which he lists as inflections, auxiliary verbs, articles,

---

[7] Functors and grammatical morphemes are both terms for minimal units, sometimes used interchangeably; Brown (1973) uses both.

prepositions and conjunctions. This results in infant speech that is sometimes described as 'telegraphic', only essential nouns and verbs. Even though it includes the beginning use of functors, Stage II is variable and not linear in direction, depending on the construction. Order is arbitrary, and gestures not gratuitous. Some functors, as single words, are utilized in Stage I: These are pronouns such as *I, me,* and *my*, as well as locatives such as *here, there,* and demonstratives *there, this, that*, and demands for *more* and *'nother* (from Brown, 1973 Table 13, p. 84). The putative explanation for these early words is based on their semantic salience to relations among things rather than modulation of meaning, regardless of their frequency in surrounding adult speech. This relation among things for Stage I is therefore based on agent-object type propositions (Brown, 1973 p. 122 & p. 134). The modulation of meaning has to wait for an understanding of the "things and processes" resulting in naming before tuning (Brown, 1973 p. 253).

Brown (1973) was not the first to notice that there seemed to be an order for the acquisition of these functors or morphemes, but he was the first to bring the rigor of detailed analysis to the field, which is still being tested and debated forty years later, and now applied to L2 acquisition. The basic order may have to be slightly modified based on the L1 and target L2, but along with universals in grammar, there appears to be consensus that there is a regular order of functor/morpheme acquisition in both adults and infants. The question is why? It cannot just be brain maturation driven, since it applies to infant and adult language (as I am asserting) and it is not amenable to explicit training. It is something about the building of the language network(s) that drives this regularity of order of acquisition, which will be explored in this chapter.

Order of Acquisition

According to R. Brown (1973), morpheme acquisition seems to follow a natural order (in English, at least), which does not appear to be altered by instruction. After acquisition of these 14 morphemes, a huge increase in the scale of language output brings the learner into the mainstream of communication. Networks as models were built in earlier chapters in this paper to model how both order of acquisition and nodes or hubs (Ullman's (2005) substrate) operate with words and classes. Now we direct attention to morphemes and their role in brain reorganization that must occur before language can become a tool rather than an impediment to expressing meaning. Ullman (2005) says there is a "distinct set of links …expected among neurocognitive markers of grammar (across subdomains, including morphology and syntax), motor and cognitive skills, and procedural memory in brain structures" (p. 150). Among these links are the 14 requisite morphemes for early language learning. The simplified 5-node structure of Barabási, and colleagues (2003, 2013) is assumed as a model plus stranding as a result of multiple sets of input is utilized in this instance to depict parallel but distinct systems.

Comparing Brown's (1973) list of obligatory morphemes in Table 1, to the highest usage words from the CHILDES database, in Table 2, draws attention to the heavy use of pronouns. We can see why these early morphemes are a necessary part of minimal participation in social interaction. While it is not directly suggested that these morphemes are necessary for 'thinking' in English (or their equivalent in other languages), they certainly are necessary for communication, which is the only way a child can acquire a language.

Table 1
Order of Acquisition of English Morphemes in obligatory context

Rank	Morpheme
1	Present progressive (-ing): Kitty is sleeping
2,3	in, on: Kitty is in the box. The box is on the floor.
4	Plural (-s): The cats have stripes.
5	Past irregular: Kitty went bye-bye
6	Possessive (-'s): Kitty's box is dirty.
7	Un-contractible copula (is, am, are): Here are the cats.
8	Articles (a, the): The kitty is here
9	Past regular (-ed): The cat played with me.
10	Third person singular (-s): She wants, she licks
11	Third person irregular: Kitty does a bad thing.
12	Un-contractible auxiliary (is, am, are): She was biting me.
13	Contractible copula: That's my kitty. Who's that?
14	Contractible auxiliary: She's playing. She's sleeping.

Note: Table built from morphemes described in Roger Brown's (1973) seminal work, *A First Language*: *The early stages.*

The pronoun usage reflects both the early understanding of polysemy (She has a name, and it's Kitty) and its integral place in language's syntax order of acquisition. It is possible, however to over-generalize when applying these selfsame forms to L2 adult language acquisition, since an older person has more control over intention, and a more established sense of self, outside of language. Anecdotally, the most frustrating part of early L2 experience is not being able to talk about oneself to others, and sounding like a simpleton when one tries. Table 2 is built from the CHILDES Parental Corpus that consists of utterances of parents, caregivers, and experimenters extracted from corpora in the original CHILDE's database. Although, strictly speaking, not all of these utterances are infant-directed; they form a representative sample of the speech that children are typically exposed to (e.g., dinner table talk, activities of free play, and storytelling).

Table 2
Infant-directed highest use words

you	know	a	go
it	see	I	yeah
to	can	that	did
what	like	is	we
and	here	in	me
do	put	on	get
oh	for	that's	want
this	now	have	ok
your	not	don't	all
no	with	there	up
he	she	are	
one	what's	right	
of	the	it's	

Note: Data from the CHILDES database (MacWhinney, 2000).

Theories of Classification

These early high use words and early morphemes intersect with theories of classification in building the first networks of language for infants and possibly for L2 adults. Going through the list of who, what, when, why or how for interpersonal caregiving should direct attention to the large number of pronouns which are necessary for interpersonal chat. Infants first use the present progressive *-ing* form because of its noun-ness. They are just emerging from a world of words dominated by naming things. This affords them the ability to assert any proposition, be it a demand or an effort to join in the 'conversation' with caregivers. Their language is periphrastic before it becomes inflective. Then, with the prepositions of *in* and *on*, they can place the objects within their personal space in relation to themselves. Grammatical case, indicating possession is the next big step, reflecting their animal competitive instincts, similar to gulping food to prevent others from stealing it. This brat behavior is essentially a survival technique used to assert oneself in a social group. It applies to the L2 adult in a similar fashion, which is why immersion learning is essential to promote these buried selfish imperatives. Listed below are the early morphemes translated into social imperatives, still maintaining Brown's (1973) order:

1. Assertion of fact about something happening of mutual interest

2/3. Where you should look, put or go

4. Is it one or many?

5. Is it now past, but still relevant?

6. Assert ownership, my relation to the assertion

7. Existence, existential angst

8. Is it of general or particular interest?

9. Is the past still relevant (extension of 4, above)

10-14. Statements of fact about ownership/kinship

What we see in the list above is the skeleton of a conversation between people, either strangers or intimates. The social imperative is what drives language acquisition, and the same order is probably universal because social neediness is universal.

Humans classify their inner worlds through language, unlike our companion species, dogs, who probably live in a world of odors. It is also probable that dogs classify odors, like we classify words, into busy Hub Centrals, because memory depends on collection, consolidation and recall. Although the basis for a dog's taxonomy of odors is unknown, it is likely to be similar to how humans order their networks, because cognition is universal, modular and ontogenetic. We classify by our personal relationships with an item, ignoring or forgetting details, or tags, when they are not primary to that relationship. Saliency depends upon usage, at the individual level, and since we are social, words and favored phrases have viral lives because we value synchronicity with our cohorts.

Word order, an early syntactic tool utilized by children, cannot be unequivocally inferred from one-word utterances. However, by end Stage I, at about 15-18 months of age, with mostly one, sometimes two or three word utterances in 'conversations' with caregivers, word order is inputted by interlocutors and seen as semantically intentioned because of circumstances (Brown, 1973, p. 150). What is also inferred is that control of language is preceded by actual expression; implying that language, words and syntax are co-dependent and developed in parallel, with production a lagging indicator of thought. The premise is that the more complex syntax will be delayed in relation to the very simplest syntax. For English, that is the S-V-O concept, reduced by Hudson's Word Grammar to the simplest noun naming phrase '*isa*' as in the phrase "This isa kitty." This form is dispensed with, but implied by the child once she has mastered the art of pointing. The only issue for Brown (1973) is the question of whether noun contentives need to be learned separately for role as agent, patient or beneficiary.

We have the benefit of the last forty years of cross cultural studies and descriptive linguistics. Additionally, newer methods of anatomical and functional studies of the brain have shown the re-purposing of older motor control areas for newer technology of language. This has been revealed by methods such as fMRI studies. These include merely watching persons in a single activity, as well as 'sensing' the activity through other modalities causing excitation in similar areas of the brain. Brown presciently connects Piaget's sensori-motor stage to his language developmental Stage I characterized by semantic intent through use of single words and imputing word order. Before an infant can point, she looks at the coveted object. Remember, Stage I is characterized by egocentric pronouns, sometimes a few locatives corresponding to early manual manipulation, grasping, and locomotion. Brown (1973) makes an interesting observation about the basis of language as communication when he says that children do not seem to be concerned with making themselves understood. Rather they *assume* this to be the case (p. 164). This assumption is probably gradually reversed in the instance of chronic child neglect or impaired reciprocity in autism.

Overcoming Thresholds

Threshold models in social and epidemiological networks refer to the hurdle to growth that a fad, social movement, virus, etc., must overcome in order to impact its environment; that is, influencing others. Passing that critical threshold is a milestone for the exponential increase in any network model, be it neuronal or created by humans. The breakout moment and its aftermath is sometimes referred to as a cascade. Barabási (2003) declares that the concept is integral to all modern theories of diffusion, but not without caveat, since some scale-free networks seem to operate outside the paradigm. It turns out that hubs, more highly connected individuals, words, or carriers of pathogens can make or break passing the threshold hurdle resulting in the cascade. This connectivity is also the Achilles heel of networks, as seen in human degenerative diseases of the neuronal networks.[8]

The question under scrutiny: Is this syntactic and social model of language, with its earliest hubs the 14 requisite morphemes described by Roger Brown in 1973, really the basis of the language surge seen in Stage II and Stage III? If that were true, we would then in the final stages of Alzheimer’s disease expect to see a gradual return to a lower state of unconnected words. Can there be a correlation between the early words of children and the last words of dementia? One observation in Alzheimer's and other degenerative diseases of the brain is that singing, even whole verses, lasts beyond the ability to communicate in spoken language. The constituents of music that are not necessary to speech may be what provides the handle for grabbing the song, full-blown, from the shattered connections of language syntax necessary for 'normal' retrieval from the lexicons. While there is currently no evidence for Brown's 14 obligatory morphemes marching in reverse; the lyrical music 'hook' is interesting because it suggests an

[8] See Granovetter's (1978) social science description of crowd theory for a readable version of the mathematics behind thresholds and the effects of minor perturbations caused by single individuals.

alternative source of memory not associated with the language constituents of syntax and lexicons. It also suggests the primary reason for pursuing this paper's line of speculation for lyrical music as an aid to second language acquisition. Prosody, in all its forms lies alongside the recognized channels of language acquisition, memory and recall, sharing resources, but with some independence. This independence is possibly traceable to prosody's more secure attachment to limbic and scattered motor areas. As the dismemberment of the syntactic network occurs, vertices and shortcut edges that provide quick access to the lexicons are no longer available, making the network "more homogeneous" as in a random network with degrees of separation greater than three and wiping out the hubs or classes upon which polysemy depend.

The cascading phenomenon upon mastery of the requisite morphemes results in increased word usage, once one acquires the means to construct simple sentences. The motive power lies with the ability to direct attention, distinguish between objects, and assert ownership as seen in the list of social imperatives developed from the 14 requisite morphemes established by Brown (1973). Cascade events seem to depend on what Watts (2003) characterizes as being exposed to numerous early adopters, not just a single new connection. Now that the early learner is attuned to syntax matrices, she is flooded with additional lexical content that she could previously safely ignore. Early learners are more easily influenced because of local stability, rather than simply being well connected.

Further inquiries will be needed to document the connection between the small networks of early learners to the minimal lexicon necessary for language acquisition for L2s (that is, lowest lexical threshold for obligatory morphemes). Early learner lexicons are dominated by nouns and imperatives, usually single verbs. The morphemes and word order impose structure. How delayed conceptual mastery might modify morpheme acquisition for L2 adults is important since we know that conceptual delays affect youngsters' language acquisition. These issues are applicable to models of network robustness. At its lowest level, robustness simply requires an alternative path for continuing the same process when an obstacle is encountered such as not enough information upon which a decision can be made. Synonyms are lexical alternates, and robustness is built into the lexical network by being able to call and retrieve the alternates as appropriate. Watts suggests that in crowded networks, a local shortcut strategy for the nodes that communicate the most with each other reduces the overall load and congestion. In this small world, it cannot be done randomly, hence the need for a directed network of small node clusters with few hierarchies, which he calls the ‘local team’. This new model is multi-scale, which can minimize network failures, probably equivalent to Barabási’s ‘modules.’ Imagine the 'local team' as a cluster of the 14 or 40 or 240 morphemes necessary for child speech or L2 adult learner: Shortcuts are initiated based upon the language stream, either output or input. These shortcuts are templates where appropriate words are slotted. Even for the experienced language user, garden path sentences lead one astray because language, like music is anticipatory, and anticipation always eliminates some alternatives in preparation for the most likely next step, tone, word, phrase, etc.

The assumption of the model after cascade is that the language spurt (that period just after two and three word speech described as Stage I and Stage II by Brown

(1973)) confirms that the child or adult L2 learner now has the capacity to organize new knowledge and extend the rules of syntax to allow exponential increases in vocabulary as well as control over subtle social interactions. Comparing Brown's (1973) list of obligatory morphemes to the highest usage words from the CHILDES database, we can see why these words in particular are a necessary part of participation in social interaction. While proof is lacking that these functors, pronouns and simple commands are necessary for 'thinking', they certainly are necessary for hierarchical activity, as in master-slave, tutor-learner, leader-follower. It is possible, however to over-generalize obligatory morphemes as applied to L2 adult language acquisition, since an older person has more control over intention.

Zipf's (1949) thesis of 'least effort' as applied to language acquisition might be neatly applied to Brown's (1973) 14 obligatory morphemes as the minimal hurdle necessary for learners to achieve communicative sufficiency. Once having achieved minimal sufficiency, L2s can happily ignore accent in spoken language, because communication does not happen in the space between people, but in their respective brains. A contrarian might submit that Zipf's law should be characterized not as an explanation of language networks, but an artifact of the arithmetical distribution of members of any large class, which leaves us with the issue that network theory, though elegant, might be insufficient to describe language acquisition, which appears to be tightly constrained by morphemes and word order but not constrained for content words. All native language speakers, child or adult, regardless of the relative size of vocabularies, reflect this enigma: their ability to proficiently handle language without instruction. As a recurring theme, schoolchildren and L2 adults can recover from their initial small vocabularies, but the learning or enrichment cannot stop, and has to include their entire environment, since language is the tool used to access other people and conceptualize the environment outside the mind. Speaking the second language at school or work, and going home to a first language environment may result in delayed L2 language growth.

Nowadays, most educators and parents are aware of the special effects of the sensitive period for language acquisition, both in first and multiple languages. These sensitive periods of massive neuronal growth, both in children and adults are a result of the brain's reaction to new stimuli. Whether there is a language spurt as claimed by some, or whether it is a relatively smooth upward curve over time, it is proposed that all new learners' language acquisition follows similar paths which can be approximated by models based upon network theories of growth and consolidation. Solé, Corominas-Murtra, Valverde and Steels (2010 p. 23) found that for children, “the observed transition occurs in parallel with the emergence of functional particles and inflectional morphology.” The child's lexicon transitions to a scale-free small world (characterized by short paths). It is structured by implicit acquisition of the constraints afforded by word order, grammatical morphemes and relational modifications. This structure imposes its constraints without the child having to 'know' rules. Solé, et al. (2010) suggest that this phenomenon, the huge increase in language abilities, reflects a general observation of what happens in random networks when they transition from many small networks to a larger hub dominated network. What has snapped into place along this minimal linguistic network? Most probably some form of hierarchy (hubs), radicals (roots) plus inflections, agency/act, and temporal/spatial relationships. These special connections prime the infant for the cascade by providing 'sticky' places for the new linguistic elements to reside.

All of these fundamental connections between the infant and caregiver worlds are set in place with 14 morphemes and a lexicon of less than 100 words. How does the new network provide the "tools" necessary to expand? Houston-Price and Law (2013, p. 84-85) describe these tools as threefold: Associative and/or invariance detection, social/linguistic clues in the environment, and finally heuristics or biases developed through exposure and usage. Thus, the totality of experiences with words provides the momentum behind the language spurt.

Network Growth

Vocabulary growth is dependent upon successful deployment of a category or class utility. For the purposes of this paper, this clustering utility has been defined as, at the lowest level, a node in the instance of a radical (root) or as a hub in the sense of a non-specific essence or generic thing. This might be a bundle of salient features that can be amended with education or experience. According to Smith, et al. (2002), various experiments on children trained with a shape bias before categorizing novel objects were able to increase their vocabulary two-hundred fold over non-trained children. This supports the network cascade theory through build-out of category/class nodes, which are prepped and ready for new words at that crucial intersection of Brown's (1973) Stage II.

Sight is integral to language acquisition. Priming by shape both alerts and directs the learners' attention. How the lack of this bias affects language learning in some children is yet to be established. Some children are unable to be primed to notice similarities in shape because their object input is not high in shape similarity, or the categorization goes unremarked or unnoticed by the caregiver. This affects the infant's ability to categorize, and co-locate synonyms and similar features through reorganizing her lexicon. It is not clear how ultimate vocabularies are affected from the lack of discriminatory shape biases and when catch-up (and at what level) is achieved. Variability in infant/child receptive and productive vocabulary and individual rates of acquisition are so great that the shape bias reflects but one thread in the greater web of experience and maturation, but one which is the most widely tested and measured, because it is easy and choices are usually unequivocal. The literate adult L2 learner, especially when tackling variable alphabet shapes such as initial capitals, joining and termination, as well as commercial and advertising stylistics could explicitly utilize this particular bias for learning.

What is currently missing from claims for shape bias, noun bias (or verb bias in Korean) is conclusive evidence of input being the greatest determiner of the bias. Are biases towards particular language acquisition tools a result of caregivers or are they universal? Annette Henderson and Mark Sabbagh (2013) review the literature and conclude there is yet to be an integration between the studies which show that more input creates larger vocabularies and the cognitive studies of learning strategies. They call for an integrated framework in order to better reveal how experience determines lexical development. (Henderson & Sabbagh 2013, p. 124). An embodied approach to examining infant word-learning requires not just gaze and the words spoken, but interaction of all of the social cues between infant and caregiver.

Concluding Remarks

A lot is known about adult behavior, especially micro movements of eyes and body orientation. Some movements are not consciously controlled, unless one is a professional gambler or actor. But, animals, especially social animals watch and are aware of very subtle body movements and posture. It is a survival technique, and for dogs may have co-evolved with human society. Vulnerable infants probably do not lack this ability to learn cues, turning to speakers and fixating on faces. Perhaps we should extend the learning propinquity to the larger social environment of first the caregiver, and then the immediate objects, especially those within sight, touch and hearing. What Yu and Smith (2013) have shown with detailed recordings of interactive play with 15-20 month olds and their mothers is that the toddlers narrow the focus and simplify the noisy background by paying close attention through manipulating objects and bringing them closer, a "one-object view"(p. 147). Mothers lean in and provide labels, carefully enunciating. This does not happen when lessons are presented by DVD or tapes. There is a strain of passivity in language training through non-human tutorials that cannot be bridged by most adults, much less infants parked in front of devices. This may not be the case with MOOCs. Though adults learn by utilizing conscious strategies and tools, the drop-out rate reflects the basic premise presented here: Humans learn best through exposure to humans. Learning to extend memory through note taking, mnemonic devices and social support through gesture, motherese and lyrical songs, for example, provide the repetition, feedback and engagement that youngsters and adults seem to need. So how does this human interface work?

Language prosody, especially in exaggerated infant-directed speech promotes word-boundary awareness, while it grabs and holds infant attention. Infants, even in utero, prefer their mother's voice over their father's or other women, *if* motherese is used. Flat intonation by mother evokes no preference. Saliency seems to be embodied in a highly charged prosody for the infant with bodily movement creating a synchrony, thus preventing attenuation. Pitch is not sufficient to attract attention and reveal preference. "Rather, the (positive) affective properties of speech directed to infants interact with the tendency to exaggerate pitch contours, driving infants' preference..." (Bortfeld, Shaw & Depowski, 2013 p. 159). Interaction, even pre-verbal, seems to encourage verbal skills, unlike watching TV, video, or simply listening to background speech. The equivalents of motherese for adults are incantatory exhortations, (enshallah, gesundheit) beloved poetry, psalms and songs. The principles remain the same; synchrony, repetition, rhythm, variable pitch and social engagement. Synchrony through mimicry cannot be emphasized enough in the learning of language.

<u>Accomplishments</u>

The core arguments of this paper surrounding the nexus of music and language was based upon a literature search that supports the four topics as detailed in the Introduction to this paper: How language and music basic research supports by overlap theories of language and music syntax processing. Networks, as described by modern graph theory, and brain modularity theories were shown as supporting language acquisition hypotheses. Finally, there was a synthesis of theories as applied to the acquisition of implicit syntax and language readiness for adult second language learners through an adaptation of Brown's (1973) requisite morphemes to the imperatives of social interaction, a driver for both infants and second language learners.

Looking Ahead

There were some issues with older experimental studies of brain function. Specifically, women and men have different processing schemas as reflected in the amount of bilateral functions required for certain activities. Studies must have like groupings of both declared sexes. Left-handedness and self-reported dyslexics have significantly different brain functioning and probably should be included in subgroups, or excluded. Because studies suggest bilingual lexicons are called up, unbidden, in word association tasks, it must be said that bilinguals process language differently than monolinguals, even if switching registers seems fairly facile. There was even speculation that music like language is identifiable as being French or German because of the patterns of intonation, quashed by Patel, (2008). Therefore, studies might have to be excluded for which there are identifiable ‘nationalities’ for the tunes used. For these and other reasons, older studies, even when simply assigning functions to regions, often must be discarded. However, there are some interesting studies that deliberately use music and dyslexics and persons with brain lesions, useful because of what they predict for the entire learner continuum, but more longitudinal work must be done. The history of music and L2 memory, cognition and enhanced learning outcomes should be recalled for what went wrong in the past: Was there a leap to practice, before the science could be reliably replicated? What were the assumptions, test methods, practices: Do they really fit into current theories that have more ‘science’ behind them?

REFERENCES CITED

References Cited

Barabási, A-L. (2003). *Linked: How everything is connected to everything else and what it means for business, science, and everyday life*. New York, NY: Plume.

Barabási, & Ravasz, (2008). Hierarchical organization in complex networks. APS/123-QED. Web. 12 Nov. 2012

Bortfeld, H., Shaw, K., & Depowski, N. (2013). The miracle year: From basic structure to social communication. In Lakshmi Gogate & George Hollich (Eds.) *Theoretical and computational models of word learning: Trends in Psychology and Artificial Intelligence* (153-171). Hershey, PA: IGI Global.

Brown, R. (1973). A first language: The early stages. Massachusetts: Harvard University Press.

Carrus, E., Koelsch, S., & Bhattacharya, J. (2011). Shadows of music-language interaction on low frequency brain oscillatory patterns. *Brain & Language, 119,* 50-57. doi:10.1016/j.bandl.2011.05.009

Ettlinger, M., Margulis, E., & Wong, P. (2011). Implicit memory in music and language. *Frontiers in Auditory Cognitive Neuroscience 2,* 1-10. doi: 10.3389/fpsyg.2011.00211

Groussard, M., Viader, F., Hubert, V. Landeau, B., Abbas, A., Desgranges, B., Platel, H. (2009). *Neuroimage, 49*, 2764-2773. doi:10.1016/j.neuroimage

Gogate, L. & Hollich, G. (2013). Timing matters: Dynamic interactions create sensitive periods for word learning. In Lakshmi Gogate & George Hollich (Eds.) *Theoretical and computational models of word learning: Trends in Psychology and Artificial Intelligence* (pp. 28-48). Hershey, PA: IGI Global.

Ghoshal, G. Chi, L. & Barabási, A-L. (2013). Uncovering the role of elementary processes in network evolution. *Scientific Reports 3*: 2930. doi: 10.10.1038/srep02920

Granovetter, M. (1978). Threshold models of collective behavior. *The American Journal of Sociology,* 83, No. 6, pp. 1420-1443

Hall, E. T. (1959) *The silent language*. Garden City, NY: Doubleday.

Henderson, A. & Sabbagh, M. (2013) Learning words from experience: An integrated framework. In Lakshmi Gogate & George Hollich (Eds.) *Theoretical and computational models of word learning: Trends in psychology and artificial intelligence* (pp. 109-131). Hershey, PA: IGI Global

Hudson, R. (2007) *Language networks: The new word grammar*. Oxford: Oxford University Press

Jackendoff, R. & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition, 200*, 33-72. doi:10.1016/j.cognition. 2005.11.005

Judde, S., Rickard, N. (2010). The effect of post-learning presentation of music on long-term word-list retention. *Neurobiology of Learning and Memory, 94,* 13-20. doi:10.1016/j.nlm.2010.03.002

Kandel, E., Schwartz, J. Jessell, T. Siegelbaum, S. and Hudspeth, A. (Eds.) (2013). *Principles of neural science, Fifth Edition*. New York, NY: McGraw-Hill.

Koda, K. (2004). *Insights into second language reading: A cross-linguistic approach*. New York, NY: Cambridge University Press

Koelsch, S. (2011). Toward a neural basis of music perception-a review and updated model. *Frontiers in Psychology, 02:110*, 1-20. doi:10.389/fpsyg. 2011.00110

Kwon, E-Y. (2005) The 'Natural Order' of morpheme acquisition: A historical survey and discussion of three putative determinants. *Working Papers in TESOL & Applied Linguistics, 5*(1) 1-21

Legg, R. (2009). Using music to accelerate language learning: An experimental study. *Research in Education, 82,* 1-12

Lipkind, D. Marcus, G., Bemis, D., Sasahara, K., Jacoby, N., Takahasi, M. Suzuki, Feher, O., Ravbar, P., Okanoya K., Tchernichovski, O. (2013) Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature 498*, 104-108, doi:10.1038/nature12173

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. *Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Montag, C., Reuter, M. & Axmacher, N. (2011). How one's favorite song activates the reward circuitry of the brain: Personality matters! *Behavioural Brain Research, 225,* 511-514. doi:10.1016/j.bbr.2011.08.012

Morgan-Short, K. Steinhauer, K., Sanz, C. & Ullman, M. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience, 24*(4). 933-947.

Morgan-Short, K. Finger, Ingrid, Grey, S. Ullman, M. (2012) Second language processing shows increased native like neural responses after months of no exposure. *PLoS ONE 7*, e32974.

Patel, A. (2008). *Music, language and the brain.* New York, NY: Oxford Univ. Press

Patel, A. (2010, June 16). *Music, biological evolution, and the brain.* Retrieved from: Connexions Web site: http://cnx.org/content/m34255/1.7/

Pinker, S. (2009). *How the mind works*. New York, NY: WW Norton.

Rinaldi, A. (2009). Speak to me, melody: Music's biological roots and its relationship with language under scrutiny. *European Molecular Biology Organization, 10,* 1294-1297. doi:10.1038/embor.2009.241

Ravasz, E. & Barabási, A-L. (2003). Heirarchical organization in complex networks. *Physical Review E 67*, 026112-l. doi:1063-651X/2003/67(2)/026112(7)

Sapir, E. (1921) *Language, an introduction to the study of speech*. Kindle Edition.

Samuelson, L. Spenser, J. & Jenkins, G. (2013) A dynamic neural field model of word learning (pp. 1-27) in *Theoretical and computational models of word learning: Trends in Psychology and Artificial Intelligence*, Gogate, L. & Hollich, G. (Eds.) Hershey, PA: IGI Global.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2007). Songs as an aid for language acquisition. *Cognition ,106*, 975–983. doi: 10.1016/j.cognition.2007.03.005

Smith, L., Jones, S., Landau, B., Gershkoff-Stowe, L. & Samuelson, L. (2002) *Psychological Science, 13*, 13-19

Solé, Corominas-Murtra, Valverde & Steels (2010). Language networks: Their structure, function, and evolution. *Complexity 15*, (6). 20-26. doi: 10.1002/cplx.20305

Twomey, K., Horst, J. & Morse, A. (2013) An embodied model of young children's categorization and word learning. In L. Gogate & G. Hollich (Eds.). *Theoretical and computational models of word learning and artificial intelligence* (pp. 172-196). Hershey PA: IGI Global.

Ullman, M.T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult language acquisition: Methods, theory, and practice.* (pp. 141-178). Washington, D.C.: Georgetown University Press.

Watts, D. (2003) *Six Degrees of Separation: The Science of a Connected Age*. New York, NY: W.W. Norton.

Walk, A., & Conway, C. (2013) Two distinct sequence learning mechanisms for syntax acquisition and word learning. In L. Gogate & G. Hollich (Eds.). *Theoretical and computational models of word learning* (pp. 350-369). Hershey, PA: IGI Global.

Wong, P., & Ettlinger, M. (2011). Predictors of spoken language learning. *Journal of Communication Disorders, 44,* 564-567. doi:10.1016/j.jcomdis

Yu, C. & Smith, L. (2013) A sensory-motor solution to early word-referent learning. In Lakshmi Gogate & George Hollich (Eds.) *Theoretical and computational models of word learning: Trends in Psychology and Artificial Intelligence* (133-152). Hershey, PA: IGI Global.