The Language Niche

Helena H. Gao

Nanyang Technological University

John H. Holland

University of Michigan

Abstract

The term *niche* is widely used, with varying precision, in disciplines ranging from ecology to marketing. Here, 'niche' is used to designate a complex flow of signals and interactions between language users. Basic patterns of these interactions persist through time, a bit like the vortices that persist around rocks in a white water stream. Viewed in this way, language and language acquisition depend upon persistent, dynamic social interactions that exploit the language niche.

1. Introduction

Human language acquisition is in many ways remarkable. The newborn receives an incredible amount of stimulation: In the retina alone there are roughly 1,000,000 light sensitive neurons, each firing dozens of times a second. Inevitable variation in incoming light levels means that no firing pattern ever exactly repeats – there is perpetual novelty in the input. Similar comments apply to other sensory modes. To handle this perpetual novelty the newborn must extract and respond to regularities and patterns, treating similar patterns as "equivalent". Then the infant must learn from social interactions to make utterances that produce predictable reactions in other humans. As time goes on the infant learns to combine these utterances to produce "sentences", allowing increasingly subtle descriptions and, even, conjectures.

There is some wired-in help in this task, supplied by a long evolutionary history. Infants can imitate facial gestures between 12 and 21 days of age, an age much earlier than predicted by stage development theory (e.g. Jean Piaget). Such imitation implies that human neonates equate their own behaviors, which they often cannot see, with gestures they see others perform (Gopnik and Meltzoff 1997; Melzoff and Moore 1977; Meltzoff and Williamson 2013). But how does the newborn go on from there to make sense of the torrent of novel input? In particular, how does the newborn travel the long distance from very limited initial abilities to full language acquisition? Though we have large collections of relevant data, we have little theory of the dynamics of this process.

Our objective in this paper is to suggest a way to bridge the gap between linguistic data and a theory of the dynamics of language acquisition. There are substantial differences between the suggested bridge and previous attempts to come to a broad understanding of language acquisition. A brief examination of these previous attempts will help explain why we consider these differences important:

1.1 Behavioral Theories

Behavioral theories became known in the early and mid 1990s. Chomsky's critique of Skinner's Verbal Behavior (Chomsky 1959; Skinner 1957) led to behaviorists' approach becoming increasingly known and used as a point of reference. Behaviorists proposed that human verbal behavior is a form of operant conditioning subject to the same controlling variables as any other operant behavior (Skinner 1957). That is to say, children learn language like they learn other complex behavior, through principles of operant conditioning (reinforcement and imitation). The role of parents as models of other complex behavior was emphasized but relatively little attention was paid to children. The fact that children are able to produce perfectly understandable original sentences and phrases without reinforcement and imitation does not seem to support the behaviorist view. This is because there is little evidence that parents use strictly structured techniques to teach their children at home; yet almost all children learn their native language without much effort within their first three years of life. Also, research has shown that parents seldom correct children when they produce any ungrammatical utterances but only focus on the meaningfulness of their spoken messages (Brown and Hanlon 1970) and that new grammatical rules, such as tenses and plurals, are typically used spontaneously with no evidence of imitation before their school age (Bloom et al. 1974).

A post-Skinnerian account of human language and cognition called the Relational Frame theory proposed by Hayes, Barnes-Holmes, and Roche (2001) explains the origin and development of language competence and complexity based on Skinnerian behaviorism and argues that children acquire language through interacting with the environment. Under the concept of functional contextualism in language learning, predicting and influencing psychological events, such as thoughts, feelings, and behaviors are regarded as important variables that can manipulate contextual learning. Their empirical studies show that a learning process which they call "derived relational responding" appears to occur only in humans possessing a capacity for language (Hayes et al. 2001). This suggests that children learn language via a system of inherent reinforcements rather than purely depending on innate abilities.

Findings of various studies make a behavioral theory unlikely adequate to explain some critical features of child language acquisition. Thus, this approach has not been widely accepted in either psychology or linguistics. In the past decade, however, there has been a re-thinking of behaviorists' approach and behaviorist models are used today in empirical studies (Ramscar and Yarlett 2007; Roediger 2004).

1.2 Nativist Theories

Nativist theories were started by the work of Chomsky (1957). Chomsky's critique of Skinner's Verbal Behavior (Chomsky, 1967) leads to an even more abstract analysis of language acquisition. His *Logical Structure of Linguistic Theory* (Chomsky, 1955, 75) is based on his analysis of a context-free grammar that he extends with transformational rules to account for the productivity or creativity of language. Chomsky emphasized that it is something innate that allows children to acquire language effortlessly. To account for the fact that the language input that children receive is usually too complicated and ambiguous for them to distinguish grammatical rules, Chomsky suggested that besides the surface structure of a language, a second structure of language, the deep structure, exists. It is the underlying meaning of language, a "species-specific characteristic…latent in the nervous system until kindled by actual language use." (Sacks 1989, 81). Chomsky believed that humans possess an innate neural device, the language acquisition device (LAD). It is this device that imposes order on incoming stimuli. This nativist idea of Chomsky has influenced most linguistic theorists' understanding of the nature of child language acquisition. His well-known "universal grammar" (UG) and "language acquisition device" (LAD) have guided much of linguistic research for decades. However, the UG/LAD approach left unexplained the mechanisms whereby an infant actually acquires a language. Later, more empirical approaches along these lines, such as Relational Frame Theory, Functionalist Linguistics, Social Interactionist Theory, and Usage-based Language Acquisition, still relied on an innate, "wired in" grammar.

1.3 Social Interactionist Theories

These theorists take an interactive perspective and see the social environment as playing a more important role than innate "wiring". The innate learning mechanisms proposed by the nativist theorists and the domain-general set of learning devices as proposed by the behaviorists are well accepted by the social interactionist theorists, but the aspects of the environment, especially parents, are regarded as being specially important. Jerome Bruner (1983) is the leading theorist who holds this view. He believes that language is presented to children with a selection of the contents for the child's current abilities by the people around them. By doing so, children are provided with the best possible chance of learning. This social-pragmatic view of language acquisition argues that "children's initial skills of linguistic communication are a natural outgrowth of their emerging understanding of other persons as intentional agents" (Carpenter et al. 1998:126). Tomasello (2003) also argues that language acquisition is based on more primitive social processes, such as shared joint attention. Language thus is considered as a social-cognitive tool, user-based, learned through functional distributional analysis, and used to manipulate other people's attention.

In psychology, Jean Piaget's experimental study of child cognitive development revealed stage-development in children. He found that a child's intellectual skills in the first 2 years of life depend upon sensori-motor experiences, not words and symbols (Piaget, 1954, 1962). The wider cognitive system then makes it possible for children to develop a series of rules for language. The constraints of sensori-motor experience lead to similarities in children's progressive acquisition of language. Following Piaget, psychologists, as well as linguists such as Melissa Bowerman (1990, 2004), Elizabeth Bates and her colleagues (Bates 1975, 1999; Bates and Goodman 1997, 1999), and Jean Mandler (1998, 2004), made data-based assumptions that there could be many learning processes involved in language acquisition. For example, empirical studies on child-directed speech have provided ample data to such a nature of children's learning. Child-directed speech, or infant-directed speech, which was originally termed by Snow (1972) as *motherese*, is found to be a simpler and more redundant speech than the speech that mothers or others use with older children. More specifically, mothers are found to typically use high-pitched tones, exaggerated modulations, simplified forms of adult words (, many questions, and many repetitions (Cooper and Aslin 1990, 1994; Fernald, 1992; Fernald and Mazzie 1991, Fish and Tokura 1996; Hoff, 2001; Karzon 1985; Kitamura et al. 2002; Kuhl et al. 1997; Masataka 1996, 1998; Moor et al 1997; Stern et al 1982; Trainor et al 2000). Some researchers regard this special form of adult speech to young children as a reflection that there is some innate language-transmittal mechanism in adults' brains. For example, Bruner (1983) suggested that adults' brains have a special mechanism that allows them to respond to young children with an automatic change of speech to a more understandable form, which called Language acquisition support system (LASS). These studies show that although infants are biologically prepared to acquire language, social-emotional context provides an equally important foundation for language development. The social interactionist approach thus compromises between "nature" and "nurture" in their theoretical arguments.

Looking at the biologically wired-in and experiential aspects from a neurolinguistic perspective, Locke (1997) proposed a theory of neurolinguistic development involving four overlapping and interactive stages: vocal learning, utterance acquisition, analysis and computation, interaction and elaboration. Experimental studies conducted by other researchers have shown that the early phases of children's development do require social cognition for language development (DeCasper and Spence 1986; Fernald 1992; Locke 1993; Mehler et al. 1988; Spence and Freeman 1996). In addressing the relation between lexical and syntactic development, Elizabeth Bates and her colleagues (Bates and Goodman 1997, 1999) have shown that the relation between vocabulary size and syntactic complexity is stronger than the relation between age and syntactic complexity, indicating that age alone is unlikely the only underlying factor for language development. This finding allowed Bates and her colleagues to be in a good position to argue against Chomsky's LAD hypothesis. Their data show that children may not need a "grammar module". Grammatical rules emerge from a need to organize the vocabulary when it becomes big enough to be sorted, categorized, and used with certain rules. What Bates and her colleagues propose is a connectionist approach in which grammar emerges out of the acquisition and analysis of individual words (Elman et al. 1996). They emphasize the dynamic interaction between a brain with innate processing constraints and a language-speaking world. This is different from the nativist view that different aspects of language development are relatively independent.

More recent approaches emphasize that, with built-in learning mechanisms, child language emerges through imitation and social interaction (Bates 1999; MacWhinney 2004; Snow 1999;

Tomasello and Bates 2001; Tomasello 2003). For example, emergentist theories, represented by MacWhinney's competition model (1987), argue that language acquisition emerges from the interaction of biological pressures and the environment through a cognitive process. These theories emphasize that nature and nurture need to be jointly involved to trigger language learning in a child.

1.4 Statistical Learning

Based on the experimental findings of connectionist models, statistical learning theories of language acquisition (Creel et al 2004; Saffran et al 1996; Saffran et al. 1997) proposed that statistical learning is a basic mechanism of information processing in the human brain and many of the language aspects, especially for bootstrapping early language skills and competence at a later age, can be explained by statistical properties of the language input. For example, in the phonetic category development children are born with the capability to distinguish any native and non-native phonetic contrasts (Aslin et al. 1981; Best et al. 1988; Eimas et al. 1971; Werker and Tees 1999). They can also discriminate between speech sounds that belong to the same native category but differ in the acoustic domain. However, by the end of the first year, infants' ability is restricted to their own native language sounds, but not the non-native contrasts (Sheldon and Strange, 1982; Werker et al. 1981; Werker and Tees 1984). However, infants' acquisition of native phonetic categories is not a simple task. It is found that the development continues at least to the second decade of life (Hazan and Barret 2000). Evidence suggests that the phonetic categories are formed by statistical properties of the speech that the infant is being exposed to (Maye et al. (2002). Empirical studies of children's learning of words and syntax also show the statistical learning principles that contribute largely to infants' language acquisition (Houston and Jusczyk 2000;

Johnson and Jusczyk 2001; Maye et al. 2002; Mintz et al. 2002; Mills et al. 2004; Newport and Aslin 2004; Saffran; 2001; Saffran et al. 1996; Saffran et al., 2008; Singh et al. 2008; Smith and Yu 2008; White and Morgan 2008; Wonnacot et al. 2008).

Each of these approaches provides a useful perspective on language acquisition, but there is a considerable gap between the observations they explain and the mechanisms that supply the dynamics of language acquisition. An historic example points up the difficulty of relating observations to dynamics: For millennia humans accumulated observations of the movement of inanimate objects, ranging from falling stones and flying arrows to the movement of planets. Fallacious "laws" were common: "a moving object always comes to rest", "heavier objects fall faster", and so on. It took Newton's theory to bring unity to this vast array of observations. It is unlikely that simple laws will encompass language acquisition, but even a rudimentary theory should offer advantages in, say, second language teaching and automatic language translation.

2. Consciousness

In a newborn, language development certainly depends upon an expanding consciousness, but "consciousness", like "life" or "mind", is difficult to define precisely. Still, for "life" and "mind" there are well-developed sciences, biology and psychology respectively, so we should not be too quick to dismiss an approach centering on "consciousness". Relating consciousness to language is, of course, not new:

"He gave man speech, and speech created thought, which is the measure of the universe." Percy Bysshe Shelley (1792 - 1822)

"Speech was given to man to disguise his thoughts. [La parole a ete donnce a l'homme pour deguiser sa pensee.] " Charles Maurice de Talleyrand-Perigord (1754 - 1838)

"Speech was given to the ordinary sort of men, whereby to communicate their mind; but to wise men, whereby to conceal it." Robert South (1634 - 1716)

"Men use thought only to justify their wrong doings, and employ speech only to conceal their thoughts. [Ils ne se servent de la pensee que pour autoriser leurs injustices, et emploient les paroles que pour deguiser leurs pensees.] " Francois Marie Arouet Voltaire (1694 - 1778)

In these quotations, consciousness is implicitly discussed as thought expressed in language. Even further back, in Plato's time, there was a general agreement that one can only speak of what one is consciously aware.

Personal Construct theory (Kelly 1955/1991) defines human consciousness as undergoing both conscious and unconscious processes. It postulates human cognition as starting from unconscious processes, or "low levels of cognitive awareness". These fundamental concepts of consciousness form the basic understanding of how human beings develop as social beings.

Nowadays, however, linguistic theories rarely touch upon consciousness. Chomsky's (1965) theory of universal grammar holds that all children are born with an innate grammar based on mental transformational rules, enabling the production of sentences not previously heard. In that theory there is no discussion of what mental activities impel a child to create a novel sentence. A fortiori, the theory makes no provision for consciousness as an activator of transformational rules.

Recent cross-disciplinary research does, however, open a channel for researchers to discuss consciousness in relation to language. Peter Carruthers (1996, 2000), for instance, argues that much of conscious thinking takes place in natural language. On this view, a better understanding of consciousness contributes to a better understanding of language and vice versa. Certainly, consciousness does not exist on its own; its existence becomes observable only when other mental activities are actively involved, such as conscious body movement vs. unconscious body movement, conscious utterance vs. unconscious utterance, and conscious thought vs. unconscious thought. Therefore, consciousness is to a large extent revealed by other types of mental phenomena that together make one's mental life as a whole.

3. Levels of Consciousness

According to Zelazo's (2004) Levels of Consciousness theory (LoC), the characteristics of children's development of consciousness undergo various levels of what Kelly termed "conscious processes" before the child reaches full linguistic capability. The age-related increases in LoC are related to the quality of experience, the potential for recall, the complexity of children's explicit knowledge structures, and the possibility of the conscious control of thought, emotion, and action.

The basic assumption of the LoC model is that children's consciousness has several dissociable levels of consciousness – information can be available at one level but not at others. This differs from models mainly based on adult data that distinguish between consciousness and a meta-level of consciousness (e.g., Moscovitch 1989; Schachter 1989; Schooler 2002). With the LoC model, consciousness is hierarchically arranged, and it is possible to observe the level at which consciousness is operating in specific situations.

The approach we take here examines mechanisms (behavioral traits) that generate the behaviors at different levels of consciousness, relating different levels of consciousness to well-known transitions in physiological and mental control as the newborn develops. Following Braitenberg's (1984) *Vehicles, Experiments in Synthetic Psychology*, we will select mechanisms at each level that mesh smoothly with the mechanisms adduced for earlier levels. We delimit five levels, starting from capabilities that are pre-primate – a wired-in evolutionary heritage – and ending with capabilities that characterize human language production – learned through interaction with other language-competent humans. The term "levels" immediately suggests a progression from level to level, and a corresponding dynamics.

We are particularly interested in an approach that moves up levels of consciousness by forming new rules that give the learning infant increasing *autonomy* – the ability to have ongoing interior activity that is modulated, but not determined, by current stimuli. This developmental approach thus relates language acquisition to a newborn's increasing autonomy as it gains experience through social interactions. Such autonomy gives the ability to plan ahead, anticipating future effects of current actions. In language acquisition, this ability allows the learning agent to determine what future utterances will fit grammatically with current utterances.

An important part of our outlook for generating new rules is the idea that new rules can be formed by combining building blocks extracted from rules already established. This procedure involves operations similar to those used by breeders in crossbreeding to get more sophisticated varieties. The ontogeny of language is reflected, then, in the mechanisms that transform each model into the next model in the sequence. As a bonus, many of the mechanisms, and the resulting changes, can be reinterpreted as relevant to the phylogeny of language. By taking this dynamics into account we hope to arrive at new experiments and, even, new approaches to teaching language.

4. Agent-based models of *LoC*

Because this approach concentrates on the social nature of language, the exposition uses agent-based models. Each agent is defined by a set of IF/THEN rules that respond to external and internal signals. As an agent learns, through interactions with its physical environment and other agents, new rules are discovered and extant rules are modified. In effect, the agent's rules amount to hypotheses at different levels of precision, with the rules being confirmed (or disconfirmed) as the agent gains experience. The different levels of precision will be related to different levels of consciousness.

Increasing consciousness in these agent-based models arises from mechanisms that increase the recirculation and feedback of signals within the agent. This increasing recirculation allows responses that are only modulated by exterior stimuli. We claim that such autonomy is a *sine qua non* for language production. We will illustrate the claim with a series of rule-based models that serve as simple examples of the progressive autonomy that accompanies successive levels of consciousness. One advantage of this agent-based approach is that it lets us employ methods used in the general study of complex adaptive systems (*cas*) – a point we will elaborate later.

Based on the central claim, LoC models make the following assumptions of the age-related changes in executive functions as children develop:

LoC 0 Unconscious activities: Inherited ('wired in') cognitive abilities.

Pre-primate precursors to language acquisition:

(i) Ability to imitate utterances and gestures.

(ii) Ability to distinguish between objects and actions.

(iii) Awareness of a mutually apprehended salient object or action.

(iv) Basic learning procedures (akin to Hebb's learning rule).

LoC 1 Minimal consciousness: Innate reinforcement of repeatable activities, ranging from repetition of sounds and motions to actions that produce innate rewards.

Example: Directed motion of hand across visual field (a precursor to gesture).

LoC 2 Stimulus-response (conditioned) consciousness: Labeling from semantic long-term memory.

Example: Utterances that cause innate rewards (such as causing T to smile).

LoC 3 Simple recursive consciousness: Use of labels (utterances) to cause others to act (such as causing T to fetch bottle).

Example: Utterances that lead to food acquisition when food visible.

LoC 4 Extended recursive consciousness: Use of labels to cause others to act when object is not present.

Example: Food acquisition when food not visible.

LoC 5 *Self-consciousness: Use of labels facilitate planned sequences of action (e.g. sequenced utterances), characterize mental activities of others, to look ahead and explore alternative courses of action*.

Example: Distinguish between two similar objects using a sequenced pair of utterances.

Newborn babies are assumed to experience LoC1, the simplest, but still conceptually coherent, consciousness that accounts for the behavioral evidence (cf. Armstrong 1981). LoC1 is unreflective, present-oriented, and makes no reference to a concept of self. So, in LoC1, the infant is conscious of what it sees (the object of experience), but not of seeing what it sees, let alone that its 'self ' is seeing what it sees. As a consequence, it cannot recall seeing what it saw. LoC1 infant behavior lasts till the end of the first year.

Numerous new abilities appear within months of birth, such as producing repeatable utterances, using objects in a functional way, pointing proto-declaratively, and searching flexibly for hidden objects. The infant works through conditioned consciousness, LoC2, to a new form of consciousness – recursive consciousness (LoC3), observed in 2-year-olds. LoC3 is marked by two signs: (i) the existence of a perceptual experience and (ii) the ability to label using semantic long-term memory. For instance, if a 1-year-old toddler says 'dog', it is assumed that the infant successfully combines a perceptual experience with a label from semantic long-term memory. The existence of labeling makes the contents of minimal consciousness perceptible and recoverable, and thus provides an enduring trace of the experience of certain content. This trace lasts long enough to be deposited into both long-term memory and working memory. The contents of working memory (e.g. representations of hidden objects) can then serve as goals to trigger action programs indirectly so the toddler is not restricted to responses triggered directly by minimal and conditioned experience of an immediate present stimulus (LoC1 and 2).

When provided with a pair of rules to use in a choice situation, 3-year-olds, in contrast to 2-year-olds, do not perseverate on a single rule. However, there are still limitations on 3-year-olds' executive function, as seen in their perseveration in the Dimensional Change Card Sort (see below), which represents the integration of two incompatible pairs of rules into a single structure. The Dimensional Change Card Sort requires children to adopt an even higher level of consciousness, extended recursive consciousness, LoC4. Evidence indicates that LoC4 first emerges around 4 years of age, together with a range of meta-cognitive skills studied under the rubric of 'theory of mind' (Frye, et al. 1995; Perner and Lang 1999; Carlson and Moses 2001).

William James (1901) suggested that an understanding of consciousness would provide the key to intellectual accomplishment. A child's awareness of self, then, is seen as a major developmental transition. It begins at the end of the second year. Linguistically, this transition is marked by a child's first use of personal pronouns; cognitively, it is marked by their self-recognition in mirrors; and emotionally it is marked by their display of self-conscious emotions such as shame (Tracy and Robins 2004). This transition is marked by another level of consciousness, referred to as self -consciousness, LoC5.

According to Zelazo et al (2007), increases in level of consciousness are brought about by the re-processing of experienced information via neural circuits in prefrontal cortex.
The potential for recall, the complexity of knowledge structures, and the possibility of action control, all increase. Reprocessing adds depth to subjective experience because more details can be integrated into the experience before the contents of consciousness are replaced by new environmental stimulation. Each degree of reprocessing causes information to be processed at a deeper, less superficial level, which increases the likelihood of retrieval (Craik and Tulving 1975). So the more complex knowledge structures increase the scope of cognitive control. This advance, in turn, moves consciousness away from stimuli and responses, making possible more de-contextualized discursive reasoning.

The infant's increasing autonomy shows itself first in conditioning that provides the infant with responses based on experience, then in anticipation and short-term planning and, finally, in internal models that allow planning and lookahead (as in playing a game of chess). Each of stage is closely related to progression in control of utterances and, ultimately, in organized, meaningful sequences of utterances. In short, the approach outlined here claims that this increasing autonomy marks progressive changes in linguistic ability.

5. Experiments

The need for different LoC's for a knowledge/action dissociation is illustrated by experiments done by Zelazo and his group of researchers with different versions of the Dimensional Change Card

Sort (DCCS). In a typical experiment, children are shown two target cards (e.g. a blue rabbit and a red car) and asked to sort a series of bivalent test cards (e.g. red rabbits and blue cars) according to one dimension (e.g. color). Then, after sorting several cards, children are told to stop playing the first game and switch to another (e.g. shape), 'Put the rabbits here; put the cars there.'). Regardless of which dimension is presented first, 3-year-olds typically continue to sort by that dimension despite being told the new rules on every trial (e.g. Brooks et al. 2003; Frye, et al. 1995; Jacques, et al. 1999; Perner and Lang 2002). By contrast, 4-year-olds recognize immediately that there are two sets of rules for the game and a switch of rules is needed regarding shape or color.

The example shows that the two groups of children, 3-year-olds and 4-year-olds, are at different stages of conscious development. For 3-year-olds, understanding one set of rules at a time, doesn't seem to be a problem. But, once a new set of rules is introduced, children at this age are found to be unable to switch back and forth. They seem to be unable to represent the rules at a level of consciousness that allows a deliberate decision to follow either the pre-switch rules or the post-switch rules. For 4-year-olds, the two ways of construing the stimuli are perceived as a reflection of multiple perspectives on the situation. Such understanding allows an integration of the different rules into a more complex rule structure. Zelazo (2004) makes this point clearly.

There are numerous examples of age-related changes in children's ability to disengage from a compelling construal of a situation. For example, children become more likely, over the course of the second year, to perform pretend actions (e.g., talking on the telephone) with pretend objects (e.g., a spoon) that bear little physical resemblance to the real objects. They also become more likely to perform pretend actions without objects altogether (e.g., Werner and Kaplan, 1963; DeLoache 1995). Similar changes continue into the preschool years (O'Reilly 1995; Overton and

Jackson 1973). As these changes occur, there are complementary changes in children's ability to resist responding with actions suggested by the objects (e.g., putting the spoon into the mouth). (Elder and Pederson 1978; Pederson et al 1981.)

This general development pattern—from stimulus-dependent to cognitively-controlled activity – we identify with changes in level of consciousness. In the A-not-B task (Piaget 1954; Marcovitch and Zelazo 1999), 9-month-old infants watch as an object is hidden conspicuously at one of two locations, and then they retrieve it. When the object is then hidden at a new location, 9-month-olds are likely to search for it at the first location. Older children, rather than responding in a perseverative, stimulus-bound fashion, evidently use an updated representation of the object's location to guide their search.

Gradually, linguistic meaning comes to dominate sensori-motor experience, as described by Vygotsky (e.g., 1978) and Luria (e.g., 1961). A preliminary study of 3- to 5-year-olds' flexible understanding of the adjectives "big" and "little" (Gao and Zelazo 2008) provides a good example. When shown a medium sized square together with a larger one, 3-year-olds had little difficulty answering the question, "Which one of these two squares is a big one?" However, when the medium square was then paired with a smaller one, and children were asked the same question, only 5-year-olds reliably indicated that the medium square was now the big one. This example shows an age-related increase in children's sensitivity to linguistic meaning when it conflicts with children's immediate experience. It reveals that interpretation becomes progressively decoupled, from perseverative stimulus properties.

Increasing sensitivity to linguistic information is also seen in children's difficulty in interpreting ambiguous adjectives. Preliminary research (Gao et al. in preparation) indicates that 3-year olds shown medium-sized pictures of a rabbit and a bear, have no difficulty identifying the bear as "a big animal." However, when 3-year-olds are shown a big picture of a rabbit and a small picture of a bear typically point to the rabbit as the "big animal". 4-year-olds seem to sense the ambiguity in the questions; they typically hesitate and then reply in an inconsistent fashion. By 5 years of age, however, children often ask, "What do you mean? The animals here in the picture or real animals?", and they are more likely to point to the bear. Older children are increasingly likely to use verbal input to restrict their attention to the appropriate aspects of a situation (Ebeling and Gelman 1988, 1994).

Language plays a causal role in helping a child to attain higher levels of consciousness. Previous research (e.g., using the Dimensional Change Card Sort and measures of children's theory of mind; Frye et al 1995) suggests that 4-year-olds are capable of considering two incompatible perspectives in contradistinction, even if they do not always do so. In another study, (Jacques and Zelazo 2005) 4- and 5-year-olds were presented with the Flexible Item Selection Task. On each trial of this task, children were shown sets of three items designed so that one pair matches on one dimension, and a different pair matches on a different dimension (e.g., a small yellow teapot, a large yellow teapot, and a large yellow shoe.)

Children were first told to show the experimenter two things that go together in one way (Selection 1), and then asked to show the experimenter two things that go together in a different way (Selection 2). To respond correctly, children must represent the pivot item (i.e., the large yellow teapot) first according to one dimension (e.g., size) and then according to another (e.g., shape). Four-year-olds generally perform well on Selection 1 but poorly on Selection 2, indicating that they have difficulty thinking about the pivot item in more than one way—they have difficulty disengaging from their initial construal of the item. However, when children were asked to label the basis of their initial selection (e.g., when they were asked, “Why do those two pictures go together?”), their performance on Selection 2 improved substantially. This was true whether children provided the label themselves or whether the experimenter generated it for them.

In terms of the LoC model, children stepped back from the Selection 1 perspective when it was labeled, reflecting on it at a higher level of consciousness. That higher level of consciousness transforms the child’s initial perspective (seeing the objects in terms of size) from a subjective frame into an object of consideration. It puts psychological distance between the child and the perspective allowing the child to adopt an alternative perspective (seeing the objects in terms of shape).

Another study done by Gao et al (under review) examined 3- and 4-year-olds’ interpretation of the relative adjectives "big" and "little". In Experiment 2 of the study, children were asked to interpret adjectives pairs with respect to a medium-sized stimulus that was compared either to a smaller stimulus or to a larger one. On each trial, children were presented with three stimuli that varied only in size, and the experimenter pointed to two of the stimuli, asking children to indicate which one was big or little. On interference trials, children were required to interpret the medium stimulus as either big or little despite interference from the presence of a bigger or littler stimulus, respectively. On switch trials, children were required to interpret the medium stimulus in a way that differed from the previous trial (e.g., they were required to interpret it as big on the switch trial when they had interpreted it as little on the previous trial). (see Figures 9.1, 9.2, and 9.3)

Figure 9.1: A non-interference trial with the question (with the experimenter pointing at the two in the circle): "Which of the two chicks is the big one?"

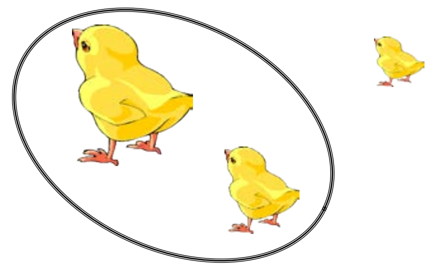


Figure 9.2: An interference trial with the question (with the experimenter pointing at the two in the circle): "Which of the two chicks is the little one?"

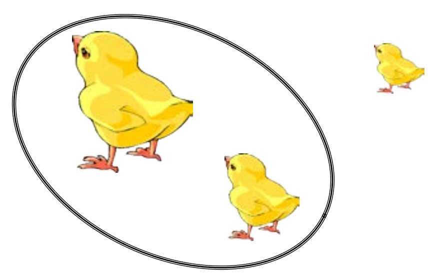
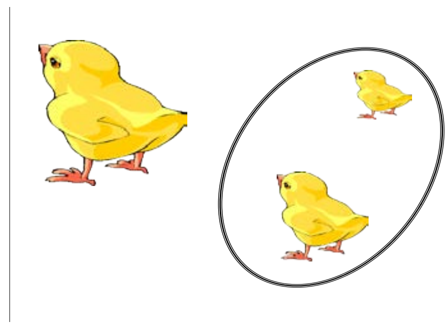


Figure 9.3: A switch trial with the question (with the experimenter pointing at the two in the circle): "Which of the two chicks is the big one?"



By examining performance on both interference and switch trials, it was possible to explore whether young children's difficulty was with switching per se, or whether children also had difficulty attending selectively to the interpretive context in the presence of interference from the stimulus display. To examine further (and in a different way) the effect of stimulus support, children were assessed in one of two conditions, one in which there were blue borders around the two stimuli in question, and one in which there were not. It was assumed that children might be easily distracted by multiple stimuli displayed at the same time and thus black borders were used as a condition under which attention was expected to be drawn to particular stimuli. It was also assumed that linguistically explicit cues would further confine children's attention in a physically and perceptually distracted context and thus verbal feedback were to be given spontaneously when a child failed a trial the first time.

Results of this study first verified that even the younger children were able to understand the basic concepts of *big* and *little*. When the ambiguous stimulus (the medium sized one) became the switching target in both the interference and switch conditions, however, children had to interpret this single stimulus as *big* in one sense (e.g., in relation to the small stimulus) and then as *little* in another sense (e.g., in relation to the big stimulus). Results of the interference and switch trials confirm that children's interpretation of adjectives is sensitive to context from a young age, but also reveals an age-related improvement in flexible adjective interpretation during the preschool years.

The major issue addressed is the apparent disparity between prior findings showing that even 2-year-olds can flexibly interpret these words, and a large literature showing young children's difficulty with switching tasks (e.g., the Dimensional Change Card Sort). The experiments provide a resolution to this disparity by showing that children fairly readily switch interpretations when the context clearly signals a shift (e.g., by changing which items are available to compare), but have more difficulty switching when they need to ignore competing cues that are present in the context. Furthermore, switching is more difficult for younger than older children.

Using the LoC model to interpret findings, we can argue that the 3-to-4 year olds performed above the level of the recursive consciousness (LoC3), since they all passed warm-up trials and non-interference trials. The existence of a perceptual experience and the ability to label using semantic long-term memory are obvious. However, the results of the interference and switch trials show that most children were unable to make the switch flexibly. In particular, children's performance was worsened when scaffolding was removed. For example, children performed better when the relevant items were highlighted (e.g., edged with black borders) than when they were not. Children failed to step back from the perspectives of the interference and switch trials when asked to compare only two instead of three items. A higher level of consciousness did not exist that would have allowed the children to transform their initial perspective (seeing the items in comparison in terms of normality) into a flexible or dynamic perspective to interpret the comparison in terms of relativity. This study addresses a rich question about the relation between language processing and cognitive flexibility, and executive functions in general. What is more interesting in this study is that the specific tasks given to the children allowed us to see that the cognitive flexibility in processing both perceptual stimuli and linguistic cues require different levels of consciousness.

These outcomes make us think that a higher level of consciousness allows for both greater influence of thought on language and greater influence of language on thought. On the one hand, it allows for more effective selection and manipulation of rules (i.e., it permits the control of language in the service of thought). On the other hand, it allows for top-down structuring of interpretive frames. Top-down structuring permits children to respond more appropriately to linguistic meaning despite misleading context – it allows language to influence thought. As a result language and thought become increasingly intertwined in a complex, reciprocal relation. This reciprocal relation can be seen in the growing richness of children's semantic understanding and increasing subtlety of their word usage. Consider, for another example, a child's developing understanding of the semantics of the verb "hit". Children first understand "hit" from its use to depict simple accidental actions (e.g., an utterance by a child at 2;4.0: "Table hit head"; Gao, 2001:220). Usage is initially restricted to particular contexts. Eventually, however, reflection on this usage allows children to employ the word in flexible and creative ways (e.g., "I should hit her with a pencil and a stick", uttered metaphorically by the same child at 3;8.6; Gao 2001:219). As Tomasello (2000) explains, such restricted productivity requires engagement in the processes of analogy-making and structure-mapping. A child's linguistic constructions depend on a first step of imitative learning, with some understanding of functional roles. A process of analogy-making then takes place to get to first order constructions and then later to get to higher order constructions.

These findings bring forth a central claim of the LoC framework: Increasing linguistic ability is the result of recursive processing, whereby the content of consciousness at one level is compared to other content at that same level.

6. An Agent-based Model of LoC

Now our objective is to incorporate these observations into a theoretical framework that suggests the mechanisms underlying these LoC transitions, opening the way to new experiments. Specifically, we want to examine mechanisms that use social interactions to build new LoC on top of levels already acquired. Though the model is constrained by facts and observations, it does not try to supply parameters, such as statistical parameters, for prediction of data; it is an exploratory model.

To emphasize the role of social interaction we will use an *agent-based model* (Holland 1995). In an agent-based model two or more agents interact through an exchange of signals, learning new behaviors as they adapt to each other. Agent-based models have been used to describe interactions in systems as different from each other as the immune system (where the signals are proteins) and markets (where the signals are buy and sell orders).

Here we will use *rule-based, signal-processing* agents (Holland et al. 1986), with rules of the form

IF (signal x is present) THEN (send signal y).

Signals x and y could be utterances, gestures, or visual input. The kinds of signals processed determine a rule's level of performance, so that we can typically associate certain kinds of rule conditions with the LoC involved. In the following examples, T ("teacher", e.g. the mother) stands for a competent adult that regularly interacts with the infant L ("learner"). For example, a simple rule for L might be,

IF (T lifts milk bottle) THEN (say "milk").

Signals can also serve to coordinate internal process, in which case they have no intrinsic meaning, serving much like the un-interpreted bit strings that coordinate instructions in a computer program. Each agent has many rules and, indeed, many rules can be active simultaneously. This simultaneous activity is roughly the counterpart of the simultaneous firing of assemblies of neurons in the central nervous system (Hebb 1949).

[In the rules that follow, <action> denotes an overt action caused by a particular signal.]

Typical rule at *LoC 0* [Unconscious activities]:

IF (T utterance) **THEN** (<imitate utterance>)

(Note that L will use limited current abilities to attempt match.

E.g. T-utterance "Gloria" can become L-utterance "Do-ee".)

Typical rule at *LoC 1* [Minimal consciousness – innate reinforcement].

IF (hand in vision cone) **THEN** (<move hand right>)

Typical rule at *LoC 2* [Stimulus-response with labeling from long-term memory].

IF (milk bottle present) **THEN** (<utterance “milk”>)

(Note that there will often be correlations between recurring patterns in the environment, such as actions and objects. These correlations can be exploited through conditioning.)

Typical rule set at *LoC 3* [Simple recursive consciousness; e.g., using utterances to cause others to act].

IF (milk bottle present) **THEN** (<utterance “milk”>)

T fetches milk bottle.

IF (milk bottle at mouth) **THEN** (<consume milk>)

Typical rule set at *LoC 4* [Extended recursive consciousness; e.g., using utterances to cause others to act on objects *not* present].

IF (hungry & no food visible) **THEN** <“milk”>

T fetches milk bottle.

IF (milk bottle at mouth) **THEN** (<consume milk>)

Typical rule set at *LoC* 5 Self-consciousness [Planned sequences of action].

IF (red ball present and blue ball present) **THEN** (internal signal x)

IF (internal signal x & red ball desired) **THEN** (internal signal y & <“red”>)

IF (internal signal x & internal signal y) **THEN** (<“ball”>)

(Note that this set of rules only allows the object word “ball” to be uttered *after* the modifier “red” – a simple form of proto-grammar.)

To *learn* in this rule-based context, the agent must have the ability to modify its signal-processing rules. Such rule-modifying, learning abilities are innate capacities supplied by evolution. Learning abilities can also be expressed as rules – think of Hebb's (1949) learning rule in neuro-psychology – so it is important to distinguish these meta-rules for learning from the signal-processing rules that are the grist for the meta-rules. In agent-based models, the meta-rules are unchanging and common to all agents.

The LoC models described here are based on meta-rules that are demonstrably available to pre-primates. That is, the meta-rules are not language specific. There are two general learning tasks the agent must be able to carry out:

(i) Credit-assignment.

As an agent interacts it must be able to decide which of its rules are helpful and which are detrimental. At higher LoC the agent must even be able to determine which early-acting, stage-setting rules make possible later beneficial outcomes. (As an example, consider the sacrifice of a piece in a game like checkers in order to make a triple jump later.) The credit-assignment learning process assigns strengths to the rules. A rule's strength reflects its usefulness to the system, useful rules having high strengths. Rules then compete to control the agent, stronger rules have a better chance of winning the competition. In effect, the rules in this system are treated as hypotheses to be progressively confirmed or disconfirmed. (See Holland 1998, chapter 4).

(ii) Rule discovery.

Once rules have been rated by credit-assignment, it makes sense to replace rules that have little or no strength by generating new rules (hypotheses). Random generation of new rules is not an option here; that would be like trying to improve a computer program by inserting random instructions. Instead, newly generated rules must somehow be plausible hypotheses in terms of experience already accumulated. (See Holland 1995, chapter 2).

Requirement (ii) leads us to the final topic of this section, *building blocks* (Holland, 1995, chapter 1 ff). Building blocks (*generators* in mathematics) have a familiar role in the sciences, best exemplified by the building block hierarchy of the physical sciences – the quark / nucleon / atom / molecule / membrane /… hierarchy. Selected combinations of building blocks at one level form the building blocks of the next level. For example, selected nucleons can be combined, much like children's building blocks, to yield the 92 atoms of classical physics. The atoms can in turn be combined to yield a vast array of molecules; the laws that constrain the combination of atoms were originally set out in the periodic table of the elements. Each level of the hierarchy can be understood in the same way. For spoken language there is a similar a phoneme / word / sentence hierarchy. A grammar specifies the laws that determine how words can be combined to yield sentences. As with atoms, a relatively small number of words, under the compact rules supplied by a grammar, can yield a vast array of sentences.

An important advantage of building blocks in the study of LoC is that the building blocks occur as repeated patterns in the ever-changing torrent of sensory input. That is, the building blocks provide repeatable experiences in a perpetually novel environment. The human face provides a clear example of the extraction and combination of building blocks to provide simple descriptions of complex objects. Indeed the highly variable pattern that we call a “face” can be represented by the combination of just a few building blocks, as exemplified by the “smiley face” emoticon ☺. By adding a few more building blocks we can describe an astonishing array of individual faces. To see this possibility, divide the face into 10 “features”: “hair style”, “forehead shape”, “eye shape and so on (see Figure 9.4). Allow 10 alternatives for each feature. There are thus 100 building blocks in total. By selecting one alternative for each feature, we can form a complete face. 10 billion distinct faces can be formed in that way. Other important sensory inputs can be treated similarly. Moreover, these building blocks can be arrayed in a LoC hierarchy, similar to the hierarchy in the physical sciences.

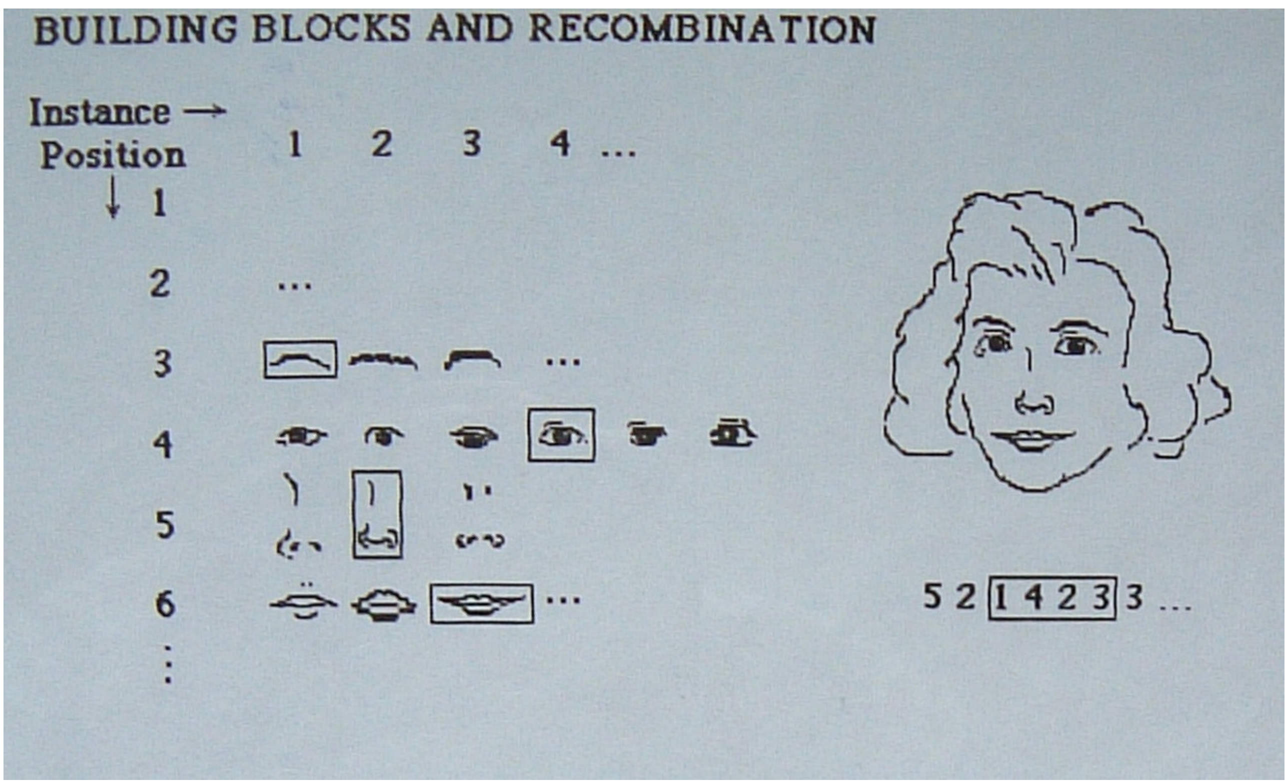


Figure 4. Building blocks.

When a building block is repeatedly associated with a rewarding event, such as food or a mother's smile, it becomes a sampled regularity that is associated with valuable experience. From the sampling point of view, the building block's reliability is continually tested under the credit assignment procedure. A confirmed building block becomes a plausible hypothesis when combined with other similarly distilled building blocks.

The random variation and imitation that accompany the earliest levels of consciousness provide a random sampling that helps uncover the most primitive building blocks, say phoneme-like utterances. Plausible new conditions and rules are generated by *recombining* building blocks already confirmed. The procedure is much like the crossbreeding of good plants (or animals) to get better plants. There is a substantial literature, centering on *genetic algorithms* (Holland 1995), that discusses the production of new rules in agent-based models via the crossing of extant rules. There is not space here to discuss genetic algorithms in detail, but it is a well-established procedure.

The meta-rules for credit assignment and rule discovery allow the neonate to achieve a gradual increase in control, corresponding to increasing LoC. The process begins with the acquisition of repeatable sound and gestures. Sounds and gestures reinforced by T become the building blocks for more complex utterances and gestures. For example, various combinations of 2 utterances, 3 utterances, and so on, provide substantial refinements in expression and meaning. The utterance "milk", at the child's single utterance stage, can have a variety of meanings: "give me some milk", "look at the milk bottle over there", and so on. Combining the two utterances "give" and "milk" greatly reduces this ambiguity. In mathematical terms we refine a broad equivalence class into a set of smaller, more informative sub-classes.

In this way, selected combinations of building blocks at one LoC become the building blocks for the next level. Building blocks, like grammars, offer combinatoric possibilities: a large variety of useful or meaningful structures can be constructed from a small number of building blocks. Moving up the LoC hierarchy thus becomes a much more efficient process than trying to "establish" a monolithic rule for each possibility at the highest LoC.

The discussion of credit assignment above pointed up the importance of strengthening "stage-setting" rules. Stage-setting, provided by the self-consciousness of Loc5, is the very essence of planning and autonomy. Autonomy of this kind requires that the agent be able to explore alternatives internally, without taking overt action. Rules so organized constitute an *internal model*. An agent with an internal model can internally explore possible sequences of action until a sequence leading to a desired outcome is located. Because the agent is dealing with sequences not yet executed, the "stage-setting" act in a sequence may not be obviously related to the last act in the sequence. For instance, in checkers, an agent may give up a piece to accomplish a later triple jump, or, in language, the agent may utter the first word of a sentence in preparation for a later utterance that will clarify the communication. Once the first act in the sequence is executed it causes a change in the environment and constrains what act(s) may be taken next. For instance, the rule set for the LoC5 example exhibits a case where the first rule executed produces an utterance which constrains what kind of utterance may follow. The study of internal models in a rule-based system is presented in *Induction* (Holland et al. 1986).

7. Summary

This paper is closely related to the manifesto produced by the "The Five Graces Group" at the Santa Fe Institute (Five Graces Group 2010). In particular, it explores the idea "that patterns of use strongly affect how language is acquired, used, and changes over time." Here we explore the effect of perceptual constraints and social motivation upon the newborn's increasing autonomy, described as "levels of consciousness" The resulting interlocked set of models does *not* try to parametrize the data presented here – the models are exploratory. Nevertheless, the models are constrained by this data and they do suggest how the data might arise. Indeed, the models are meant to suggest further experiments that will clarify both the data and the models.

Because this approach highlights the effects of social interaction on the acquisition of language, the treatment uses a *complex adaptive systems* (*cas*) framework, as suggested in the Five Graces manifesto. That is, we consider models in which multiple agents interact and learn from each other. *Cas* emphasize dynamics, and the agents therein rarely settle down to a static equilibrium. This suggests that each agent will develop its own idiolect and that agents that interact regularly will have many common constructions in their idiolects, a familiar finding of modern linguistics.

It is our intention to build executable versions of these models, testing them and extending human experiments in ways that will suggest improvements to the models. Above all, we would like to provide a strong, testable theoretical base for the "levels of consciousness" approach to language acquisition.

References

Armstrong, D.M. 1981. What is consciousness? In *The Nature of Mind and Other Essays*, ed. D. M. Armstrong. Ithaca, New York: Cornell University Press.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., and A. J. Perey. 1981. Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Development* 52:1135–1145.

Bates, E. 1975. Peer relations and the acquisition of language. In *Friendship and peer relations: The origins of behavior*, eds. M. Lewis and L. Rosenblum, 259–292. New York: John Wiley and Sons.

______. 1999. Nativism vs. development: Comments on Baillargeon & Smith. *Developmental Science* 2 (2): 148–149.

Bates, E., and J. Goodman. 1997. On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real–time processing. *Language & Cognitive Processes* 12 (5/6): 507–586. [redacted version reprinted in Essential readings in language development, eds. M. Tomasello and E. Bates. Oxford: Basil Blackwell, 2000].

______. 1999. On the emergence of grammar from the lexicon. In *The emergence of language*, ed. B. MacWhinney, 29–79. Mahwah, NJ: Lawrence Erlbaum.

Best, C. T., G. W. McRoberts, and N. M. Sithole. 1988. Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English–speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14:345–360.

Bloom, L., Hood, L., and P. Lightbown. 1974. Imitation in language development: If, when and why. *Cognitive Psychology* 6:380–420.

Bowerman, M. 2004. From universal to language–specific in early grammatical development [Reprint]. In *The Child Language Reader*, eds. K. Trott, S. Dobbinson and P. Griffiths, 131–146. London: Routledge.

———. 1990. Mapping thematic roles onto syntactic functions: Are children helped by innate linking rules? *Linguistics* 28:1253–1290.

Braitenberg, V. 1984. *Vehicles: Experiments in Synthetic Psychology*. MIT Press.

Brown, R., and C. Hanlon. 1970. Derivational complexity and order of acquisition in child speech. In *Cognition and the Development of Language*, ed. J. Hayes, 155–207. New York: Wiley.

Bruner, J. 1983. *Child's Talk: Learning to Use Language*, New York: Norton.

Carpenter, M., N. Akhtar, and M. Tomasello. 1998. Fourteen– to 18–month–old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development* 21:315–330.

Carruthers, P. 1996. *Language, Thought, and Consciousness: An Essay in Philosophical Psychology*. Cambridge University Press.

Carruthers, P. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge University Press.

Chomsky, N. 1975. *The Logical Structure of Linguistic Theory*. New York: Plenum Press.

———. 1967. Review of Skinner's Verbal Behavior. In *Readings in the Psychology of Language*, eds. L. A. Jakobovits and M. S. Miron, 142–143. Prentice–Hall.

———. 1965. *Aspects of A Theory of Suntax*. MIT Press.

———. 1959. Review of Verbal Behavior, by B.F. Skinner. *Language* 35: 26–57.

———. 1957. *Syntactic Structures*. Mouton.

———. 1955. *Logical Syntax and Semantics. Language*. January–March.

Cooper. 1994. Developmental differences in infant attention to the spectral properties of infant–directed speech. *Child Development* 65:1663–1677.

Cooper, R. P., and R. N. Aslin. 1990. Preference for infant–directed speech in the first month after birth. *Child Development* 61:1584–1595.

Creel S. C., E. L. Newport, and R. N. Aslin. 2004. Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning,Memory, and Cognition* 30:1119–1130.

DeCasper, A. J., and M. J. Spence. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development* 9:133–150.

Deloache, J. S. 1995. Early understanding and use of symbols: The model model. *Current Directions in Psychological Science* 4:109–113

Ebeling, K., and S. Gelman. 1994. Children's use of context in interpreting "big" and "little." *Child Development* 65:1178–1192.

———. 1988. Coordination of size standards by young children. *Child Development* 59:888–896.

Eimas P. D., E. R. Siqueland, P. Jusczyk, and J. Vigorito. 1971. Speech perception in infants. *Science* 171:303–306.

Elder, J. L., and D. R. Pederson. 1978. Preschool children's use of objects in symbolic play. *Child Development* (49) 2:500–504.

Elman, J. L., E. A. Bates, M. H. Johnson, A. Karmiloff–Smith, D. Parisi, and K. Plunkett. 1996. *Rethinking Innateness: A Connectionist Perspective on Development.* Cambridge, MA: MIT Press.

Fernald, A. 1992. Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In *The Adapted Mind,* eds. J. H. Barkow, L. Cosmides, and J. Tooby, 391–28. New York: Oxford University Press.

Fernald, A., and C. Mazzie. 1991. Prosody and focus in speech to infants and adults. *Developmental Psychology* 27 (2): 209–221.

Frye, D., P. D. Zelazo,, and T. Palfai. 1995. Theory of mind and rule–based reasoning. *Cognitive Development* 10:483–527.

Gao, H. 2001. *The Physical Foundation of the Patterning of Physical Action Verbs: A Study of Chinese Verbs*. Lund, Sweden: Lund University Press.

Gao, H. H., and P. D. Zelazo. 2008. Language and development of consciousness: Degrees of disembodiment. In *Developmental Perspectives on Embodiment and Consciousness*, eds. W. F. Overton, U. Mueller, and J. Newman, 225–246. Mahwah, NJ: Erbaum.

Gao, H. H. , P. D. Zelazo, D., Sharpe, and A. Mashari. under review. Beyond early linguistic competence: Development of children's ability to interpret adjectives flexibly.

Gopnik, A., and A. N. Meltzoff. 1997. *Words, Thoughts, and Theories*. Cambridge, Mass.: Bradford, MIT Press.

Hayes, S. C., D. Barnes–Holmes, and B. Roche, eds. 2001. *Relational frame theory A Post–Skinnerian Account of Human Language and Cognition*. New York: Kluwer Academic / Plenum Publishers.

Hazan V., and S. Barret. 2000. The development of phonetic categorization in children aged 6–12. *Journal of Phonetics* 28:377–396.

Hebb, D. O. 1949. *The Organization of Behavior*. Wylie.

Hoff, E. 2001. *Language Development.* 2nd edition. Belmont, CA: Wadsworth.

Holland, J. H. 1998. *Emergence: From Chaos to Order*. Addison–Wesley.

———. 1995. *Hidden Order: How Adaptation Builds Complexity.* Addison–Wesley.

Holland, J. H., K. J. Holyoak, R. E. Nisbett, and P. Thagard. 1986. *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, MA: Bradford Books/ MIT Press.

Houston, D. M., and P. W. Jusczyk. 2000. The role of talker specific information in word segmentation by infants. J*ournal of Experimental Psychology: Human Perception and Performance* 26:1570–1582.

Jacques, S., and P. D. Zelazo. 2005. Language and the development of cognitive flexibility: Implications for theory of mind. In *Why Language Matters for Theory of Mind*, eds. J. W. Astington and J. A. Baird, 144–162. New York: Oxford University Press.

Jacques, S., P. D. Zelazo, N. Z. Kirkham, and T. K. Semcesen. 1999. Rule selection and rule execution in preschoolers: An error–detection approach. *Developmental Psychology* 35:770–780.

Johnson, E. K., and P. W. Jusczyk. 2001. Word segmentation by 8–month–olds: When speech cues count more than statistics. *Journal of Memory and Language* 44:548–567.

Karzon, R. G. 1985. Discrimination of polysyllabic sequences by one– to four–month–old infants. *Journal of Experimental Child Psychology* 39:326–342.

Kelly, G. A. 1955/1991. *The Psychology of Personal Constructs Vol.1: A Theory of Personality* New York, US: Norton. 2nd printing: 1991, Routledge, London, New York

Kitamura, C., C. Thanavisuth, S. Luksaneeyanawin, and D. Burnham. 2002. Universality and specificity in infant–directed speech: Pitch modifications as a function of infant age and sex in a tonal and non–tonal language. *Infant Behavior and Development* 24:372–392.

Kuhl, P.K., J.E. Andruski, I.A. Chistovich, L.A. Chistovich, E.V. Kozhev–nikova, V.L. Ryskina, E.I. Stolyarova, U. Sundberg, and F. Lacerda. 1997. Cross–language analysis of phonetic units in language addressed to infants. *Science* 277(5326):684–686.

Locke, J. L. 1997. A theory of neurolinguistic development. *Brain and Language* 58:265–326.

———. 1993. *The Child's Path to Spoken Language*. Cambridge, MA: Harvard University Press.

Luria, A. R. 1961. *Speech and the Regulation of Behaviour*. London: Pergamon Press.

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

———. 1987. Toward a psycholinguistically plausible parser. In *Proceedings of the Eastern States Conference on Linguistics* (*ESCOL* 1986), S. Thomason ed. Columbus, OH: Ohio State University.

Majid, A., M. Bowerman, S. Kita, D. B. M. Haun, and S. C. Levinson. 2004. Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8 (3): 108–114.

Mandler, J. M. 1998. Babies think before they speak. *Human Development* 41:116–126.

Marcovitch, S., and P. D. Zelazo. 1999. The A–not–B error: Results from a logistic meta–analysis. *Child Development* 70:1297–1313.

Masataka, N. 1996. Perception of motherese in a signed language by 6–month–old deaf infants. *Developmental Psychology* 32:874–79.

———. 1998. Perception of motherese in Japanese sign language by 6–month–old hearing infants. *Developmental Psychology* 34:241–46.

Maye J., J. F. Werker, and L. Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82:B101–B111.

Mehler, J., P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini, and C. Amiel–Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29:143–178.

Meltzoff, A. N., and M. K. Moore. 1997. Explaining facial imitation: A theoretical model. *Early Development and Parenting* 6:179–192.

———.1977. Imitation of facial and manual gestures by human neonates. *Science* 198:75–78.

Meltzoff, A. N., and R. A. Williamson. 2013. Imitation: Social, cognitive, and theoretical perspectives. In *Oxford Handbook of Developmental Psychology*, ed. P. R. Zelazo, 651–682. NY: Oxford University Press.

Mills, D. L., C. Prat, R. Zangl, C. L. Stager, and H. J. Neville. 2004. Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14– and 20–month–olds. *Journal of Cognitive Neuroscience* 16:1452–1464.

Mintz, T. H., E. L. Newport, and T. G. Bever. 2002. The distributional structure of grammatical categories in speech to young. *Cognitive Science* 26:393–424.

Moore, D. S., M. J. Spence, and G. S. Katz. 1997. Six–month–olds' Categorization of Natural Infant–Directed Utterances. *Developmental Psychology* 33:980–989.

Moscovitch, M. M. 1989. Confabulation and the frontal systems: strategic versus associative retrieval in neuropsychological theories of memory. In *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving*, eds. H. L. Roediger and F. I. M. Craik, 133–60, Erlbaum.

Newport, E. L., and R. N. Aslin. 2004. Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology* 48:127–162.

O'Reilly, A. W. 1995. Using representations: Comprehension and production of actions with imagined objects. *Child Development* 664:999–1010.

Overton, W. F., and J. Jackson. 1973. The representation of imagined objects in action sequences: A developmental study. *Child Development* 44:309–314.

Pederson, D., A. Rook–Green, and J. Elder. 1981. The role of action in the development of pretend play in young children. *Developmental Psychology* 17:756–95.

Perner, J., and B. Lang. 2002. What causes 3–year–olds' difficulty on the dimensional change card sorting task? *Infant and Child Development* 11:93–105.

———. 1999. Development of theory of mind and executive control. *Trends in Cognitive Sciences* 3:337–344.

Piaget, J. 1962. *The Language and Thought of the Child*. London: Routledge & Kegan Paul. (*Le Langage et la pensée chez l'enfant*. 1923.)

———. 1954. *The construction of reality in the child.* New York: Basic Books. (*La construction du réel chez l'enfant*. 1950, also translated as *The child's construction of reality*. London: Routledge and Kegan Paul, 1955.)

Ramscar, M., and D. Yarlett. 2007. Linguistic self–correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science* 31:927–960.

Roediger, R. 2004. What happened to behaviorism. *American Psychological Society* 17, Presidential Column.

Sacks, O. 1989. *Seeing Voices: A Journey into the World of the Deaf*. New York: Harper Perennial.

Saffran, J. R. 2001. Words in the sea of sounds: the output of infant statistical learning. *Cognition* 81:149–169.

Saffran, J. R., R. N. Aslin, and E. L. Newport. 1996. Statistical learning by 8–month old infants. *Science* 274:1926–1928.

Saffran, J. R., M. Hauser, R. Seibel, J. Kapfhamer, F. Tsao, and F. Cushman. 2008. Grammatical pattern learning by human infants and cotton–top tamarin monkeys. *Cognition* 107:479–500.

Saffran, J. R., E. L. Newport, R. N. Aslin, R. A. Tunick, and S. Barrueco. 1997. Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science* 8:101–105.

Schooler, J. W. 2002. Re–presenting consciousness: dissociations between experience and metaconsciousness. *Trends Cognitive Science* 6:339–44.

Sheldon, A., and W. Strange. 1982.The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3:243–261.

Singh, L., K. S. White, and J. L. Morgan. 2008. Building a word–form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development* 4:157–178.

Skinner, B. F. 1957. *Verbal Behavior*. Acton, MA: Copley Publishing Group. ISBN 1–58390–021–7

Smith, L., and C. Yu. 2008. Infants rapidly learn word–referent mappings via cross–situational statistics. *Cognition* 106:1558–1568.

Snow, C. E. 1999. Social perspectives on the emergence of language. In *The Emergence of Language,* ed. B. MacWhinney, 257–276. Mahwah, New Jersey: Lawrence Erlbaum.

———. 1972. Mothers' speech to children learning language. *Child Development* 43:549–565.

Spence, M. J., and M. S. Freeman. 1996. Newborn infants prefer the maternal low–pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development* 19:199–212.

Stern, D. N., S. Spieker and K. McKain. 1982. Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology* 18:727–35.

The Five Graces Group [C. Beckner, R. Blythe, J. Bybee, M. H. Christiansen, W. Croft, N. C. Ellis, J. H. Holland, J. Ke, D. Larsen–Freeman, and T. Schoenemann 2010. Language is a complex adaptive system. *Language Learning.*

'The Five Graces Group' Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen–Freeman, D., Schoenemann, T. 2009. Language is a complex adaptive system. Position paper, *Language Learning* 59, Supplement 1:1–27.

Tomasello, M. 2003. *Constructing a Language: A Usage–based Theory of Language Acquisition*, Harvard University Press.

Tomasello, M., and E. Bates, eds. 2001. *Language Development: The Essential Readings.* Oxford: Basil Blackwell.

Tracy, J. L., and R. W. Robins. 2004. Putting the self into self–conscious emotions: A theoretical model. *Psychological Inquiry* 15:103–125.

Trainor, L. J., C. M. Austin, and R. N. Desjardins. 2000. Is infant–directed speech prosody a result of the vocal expression of emotion. *Psychological Science* 11:188–95.

Vygotsky, L .S. 1978. *Mind in Society*. Cambridge, MA: Harvard University Press.

Werker J. F., and Tees, R. C. 1999. Influences on infant speech processing. Toward a new synthesis. *Annual Review of Psychology* 50:509–535.

———. 1984. Cross–language speech perception: Evidence from perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49–63.

Werker, J. F., J. H. Gilbert, K. Humphrey, and R. C. Tees. 1981. Developmental aspects of crosslanguage speech perception. *Child Development* 52:349–355.

Werner, H., and B. Kaplan. 1963. *Symbol Formation: An Organismic – Developmental Approach to Language and the Expression of Thought*. New York: Wiley.

White K. S., and J. L.Morgan. 2008. Sub–segmental detail in early lexical representations. *Journal of Memory and Language* 59:114–132.

Wonnacott, E., E. L. Newport, and M. K. Tanenhaus. 2008. Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology* 56:165–209.

Zelazo, P. D. 2004. The development of conscious control in childhood. *Trends in Cognitive Sciences* 8:12–17.

Zelazo, P. D., H. H. Gao, and R. Todd. 2007. Development of consciousness. In *The Cambridge Handbook of Consciousness*, eds. P. D. Zelazo, M. Moscovitch, and E. Thompson, 405–434. New York: Cambridge University Press.