The Social Dimension in Language Development: A Rich History and a New Frontier

Shannon M. Pruden\textsuperscript{1}, Kathy Hirsh-Pasek\textsuperscript{1} and Roberta Michnick Golinkoff\textsuperscript{2}

\textsuperscript{1} Temple University  
\textsuperscript{2} University of Delaware

Running Head: Social dimension in language development

Corresponding author:

Kathy Hirsh-Pasek  
Department of Psychology  
Temple University, Weiss Hall  
13\textsuperscript{th} & Cecil B. Moore  
Philadelphia, PA 19122

To appear in P. Marshall & N. Fox (Eds.), The development of social engagement: Neurobiological perspectives. This work was supported by NSF grants \# SBR9615391 & SBR9601306 to the second two authors.
“Learning a word is a social act.” (P. Bloom, 2000, p.55)

“Language is social.” (Clark, 2003, p.19)

The past 30 years has witnessed an explosion in research on social/emotional and cognitive development in infants and toddlers (Berk, 2003). Within the realm of social/emotional development, research has demonstrated that infants are capable of social referencing, joint attention, emotion regulation and noting another’s social intent (Baldwin, 1991; Bretherton, 1992; Carpenter, Nagell, & Tomasello, 1998; Corkum & Moore, 1995; Eisenberg & Zhou, 2000; Gergely, Nadasdy, Csibra, & Biro, 1995; Meltzoff, 1995; Murray & Trevarthen, 1985). Within the realm of cognitive development, theories have explored the understanding of infant competencies in numeracy, spatial cognition, object and event perception and early word development (Aguiar & Baillargeon, 2002; Spelke, Katz, Purcell, Ehrlich, & Breinlinger, 1994; Woodward & Markman, 1998; Wynn, Bloom, & Chiang, 2002). Despite the proliferation of scholarship in each domain, this research has largely been conducted on parallel planes, with little attention to how infant and toddler social development impacts upon cognitive growth and how cognitive development might feed social development. Research areas such as social cognition emerged in the 1970s but with a few exceptions (e.g., theory of mind research) most work in this field had a relatively short shelf-life. Recent work on emotion regulation (Eisenberg, 2001; Eisenberg & Zhou, 2000; Fox, 2003; Morris, et al., 2002) and on social prerequisites for school readiness (NICHD, 2003; Raver, 2002) ask anew how our theories of development can move beyond the compartmentalized child to embrace interactions between the cognitive and social self in learning.

One area of research that has struggled with divisions between cognitive and social theories of development is that of language acquisition. Historically, the field of language
development was born from two events: (1) Chomsky’s views on grammar, and (2) the dominance of Piagetian and information-processing models of cognitive development. Chomsky claimed that children were endowed with a universal grammar that made relatively little or no contact with social skills (Chomsky, 1965, 1975, 1981). The social environment provided but a trigger so that children could link internal representations of grammar with the input they heard (see also Pinker, 2002). Information processing models use children’s cognitive competencies (and sometimes innate grammatical structures) as a foundation asking how children compute correlations from the available evidence to derive a more complete grammar (Maratsos & Chalkley, 1980; but see Bates & MacWhinney, 1987 and Elman, 2001 for an alternative view; see also Hirsh-Pasek & Golinkoff, 1996 for a review). These theories make little contact with children’s developing social competence.

Throughout the 1970’s and 1980’s dissenting voices represented the fact that a nativist, modular and even cognitive information processing views of language would be insufficient to account for word and grammatical learning. As Slobin (1985) wrote, “the application of particular operating principles depends on a complex of existing linguistic and extralinguistic knowledge, processing constraints, the structure of social interaction, and the structure of the language being acquired” (p. 1245). To truly understand a complex cognitive behavior such as language, we need to pay attention not only to linguistic mechanisms underlying language acquisition, but also to the social and emotional characteristics of the child in his social environment (Bruner, 1975; Snow & Ferguson, 1977; Snow, 1989).

The area of language development is still largely bifurcated, with researchers emphasizing either the social-interactional (also called social-pragmatic) basis of language or the
child’s innate contributions. There are, however, serious attempts to unite these camps, attempts that could point to new directions in the study of neurological bases for language acquisition.

This chapter is organized in four parts. Part I reviews theories of language development that endorse a social-pragmatic perspective. Part II proposes a hybrid model of language development, called the Emergentist Coalition Model (ECM) that embeds both the pragmatic and cognitive perspectives in a developmental theory of language acquisition. Language development, in this scenario, occurs as a result of the differential weighting of three cues, perceptual/attentional, social, and linguistic. Part III offers evidence for the differential weighting of cues over time. We present several studies (some old and some new) showing that all three cues are used in tandem as the child progresses through early language acquisition. This model has proven fruitful for the study of grammar and early word learning. Finally, in Part IV we project how this more complex model can provide a road map for the neurological underpinnings of language development in order to shed some light on both normal and atypical language development.

Part I: Social Pragmatics in Language Development…A Long Tradition

The role of social and emotional development has been widely recognized within language development and has a long tradition dating back to the 1970’s. Bruner (1975) was among the first to address the role of social development in language learning suggesting that social interaction could be mapped transparently onto linguistic structure, although he subsequently retracted this claim (Bruner, 1983). Nonetheless, the importance of games like “give and take” that infants and toddlers play with their caregivers cannot be underestimated. In these games, where they alternately become the agents or the recipients of an action, the roles that will be expressed in language become revealed. Social interactions also provide a cultural
classroom for language learning by introducing the scripts or routines (Nelson, 1973, 1985; Peters & Boggs, 1986; Snow, 1986; Snow & Godfield, 1983) within which language learning takes place and by introducing the joint attentional episodes that help children find the referent of communication (Adamson, 1995; Tomasello & Farrar, 1986). By way of example, parents participate in common routines with children like the diapering routine or the feeding routine. Within these contexts, parents engage in “proto-conversations” with their infants. Consider the following “proto-conversation” between a mother and her 3-month-old daughter, Ann (Snow, 1977, p.12):

“Ann: (smiles)
Mother: Oh, what a nice little smile. Yes, isn’t that nice? There. There’s a nice little smile.
Ann: (burps)
Mother: What a nice little wind as well. Yes, that’s better, isn’t it? Yes. Yes. Yes.
Ann: (vocalizes)
Mother: There’s a nice noise.”

Western parents guide the infant’s way in the social interaction by showing the baby how to take turns and how to build a relationship (Golinkoff & Hirsh-Pasek, 1999). These “proto-conversations” may be minimal at first, with the adult supplying all the turns. However, as the child grows, the adult “raises the criterion for what counts as a contribution from their infants” (Clark, 2003, p.30).

Nelson (1985), for example, articulated the social-interactional position clearly when she wrote: “Language learning takes place within the framework of social interaction, and the nature of the particular kinds of interaction experienced determines not only the function and context of
the language to be acquired but which segments will be learned first and how these segments will subsequently be put together or broken down for reassembly” (p.109).

L. Bloom’s intentionality model of language development (2000; Bloom & Tinker, 2001) takes Nelson’s position one step further by proposing that language is learned in a social context by an active language learner who is determined to share the contents of their mind. Previous theories emphasizing the role of social and pragmatic cues in language acquisition propose that caregivers structure the learning contexts in relevant culturally specific ways (P. Bloom, 2000).

One of the most prolific champions for social mediation in language learning is Michael Tomasello. He and his colleagues hold that early language learning is bathed in a social context (Akhtar & Tomasello, 2000; Carpenter, Nagell, & Tomasello, 1998; Tomasello, 1992). According to Tomasello’s social-pragmatic view, children participate in patterned social interactions or routines that provide a “scaffold” to the pre-linguistic child (Akhtar & Tomasello, 2000; Ninio & Bruner, 1978). Evidence for “scaffolding” comes from studies investigating children’s games like “peek-a-boo” (Ratner & Bruner, 1978), children’s play such as a tea party (Kaye & Charney, 1980), and even studies of picture-book reading (Kaye & Charney, 1980; Ninio & Bruner, 1978; Snow & Godfield, 1983; see Golinkoff & Hirsh-Pasek, 1999 for other examples). A classic example of an adult scaffolding or supporting a child in a language-learning situation comes from Ninio and Bruner’s (1978, pp.6-7) study of book reading:

“Mother: Look!
Child: (Touches picture).
Mother: What are those?
Child: (Vocalizes and smiles).
Mother: Yes, they are rabbits.
Child: (Vocalizes, smiles, and looks up at mother).
Mother: (Laughs). Yes, rabbit.
Child: (Vocalizes and smiles).
Mother: (Laughs). Yes.”

In this example, the mother structures the situation for the child by formatting and framing the interaction. The mother guides her child in the language learning process by initiating exchanges and interactions and elaborating on a topic. Notice that the child is really not saying anything comprehensible. Yet, her contribution to the conversation is honored by the mother as if it were a full-fledged verbal communication.

Tomasello, Akhtar, Baldwin, and P. Bloom claim that language learning is fundamentally social, based on the child’s discovery of social intent (Akhtar & Tomasello, 2000; Baldwin & Tomasello, 1998; P. Bloom, 2000). By noting what people are intending to talk about, children become like Fauconnier’s bricklayer (Fauconnier, 1985). As Fauconnier writes, “a brick could theoretically occupy any position in a wall, but at any stage of the actual building process there is only one place for it to go” (pp. 168-169). L. Bloom uses this as a backdrop to posit a “principle of relevance” that guides active young learners to form correct mappings between world and language through social, linguistic, and perceptual cues. Once you can follow eye gaze and interpret another’s point of view, you have a window onto language use and meaning. Once you can ‘read’ another’s mind, you can determine how the mapping from world to mind works. You can then become an apprentice to a more sophisticated language user.

In a recent review, Clark (2003) explained why social development is an important factor, not to be ignored and underestimated, in language learning: “Infants are born into a social world, a world of touch, sound, and affect, a world of communication. They develop and grow
up as social beings, immersed in a network of relationships from the start. It is in this social setting that they are first exposed to language, to language in use” (p. 25).

From the beginning of life, infants are fundamentally social beings, attending to social information in their environment (Butterworth, 2003; see Golinkoff & Hirsh-Pasek, 1999 and Adamson, 1995, for reviews). Research shows that they attend to social cues, such as eye gaze and pointing in the second 6 months of life (Corkum & Moore, 1995; for a review, see Carpenter, Nagell, & Tomasello, 1998). Thus, the issue at hand is not whether infants are social beings, born into a social world, but whether infants can recruit social cues for the purpose of learning both the rules of language (grammar) and words of their language.

*The Role of Social Development in Children’s Acquisition of Grammar*

The earliest work on the intersection of social development and language came from the “social interactionists” who asked how social input promoted grammatical development. Two lines of research characterized this period of study: (a) the study of infant-directed speech and its impact on language development (Snow & Ferguson, 1977), and (b) the role of corrections on the development of grammar (Bohannon & Stanowicz, 1988).

*“Motherese” or infant-directed speech*

Caregivers around the world talk to their infants in a way very different than they do with adults and even older children (Fernald et al., 1989; Grieser & Kuhl, 1988). “Motherese” or infant-directed (ID) speech is characterized as having a slower rate, an extended frequency range, higher overall fundamental frequency, repeated pitch contours, marked intensity shifts, longer pauses between utterances, word lengthening, and simplified vocabulary (Golinkoff & Alioto, 1995; see Morgan & Demuth, 1996 and Snow, 1995 for related definitions). Research indicates that infants do in fact prefer to listen to ID speech rather than adult-directed speech.
(Fernald, 1985), even as early as at birth (Cooper & Aslin, 1990). According to Clark (2003), ID speech serves three specific functions. First, ID speech is designed to get the baby’s attention. This type of speech, whether it is high-pitched or a whisper, differentiates speech addressed to the child from speech addressed to adults. It says to the baby, “Hey, this talk is for you!” Getting an infant’s attention is the first step on the road to achieving what has been dubbed “joint attention.” Second, ID speech is used to maintain infants’ attention. In English, speakers achieve this goal by maintaining a high pitch and exaggerated intonation (Fernald & Mazzie, 1991; Ratner, 1984). Only when infants are able to maintain attention to the speaker can they learn what the speaker is referring to. Finally, to communicate with a listener who has much less knowledge than the speaker, the speaker must tailor their utterances and make them stand out in the stream of speech. ID speech is generally tailored to the child’s linguistic level though the modifications that are made are different across cultures (Fernald & Morikawa, 1993; Fernald et al., 1989; Schieffelin & Ochs, 1986).

Research has also established that across various cultures, caregivers use ID speech to talk to infants (Fernald et al., 1989; Grieser & Kuhl, 1988) and even to domestic pets (Burnham & Kitamura, 2002; Hirsh-Pasek & Treiman, 1982), but does this type of speech play a specific role in language acquisition? Some suggest that ID speech may help infants isolate words within the speech stream (Golinkoff & Alioto, 1995) and may even provide the foundation for acquiring the grammatical structure of one’s language (Gleitman, Gleitman, Landau, & Wanner, 1988; Gleitman & Wanner, 1982; Morgan, 1986). What little research exists showed that ID speech, at the very least, facilitates various phonological awareness (Karzon, 1985; Kuhl et al., 1997) and perceptual processes (Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989) that might assist learners in the parsing of speech into language-relevant units. In studies of both infant-directed
and adult-directed speech, Hirsh-Pasek and colleagues (1987; Kemler-Nelson et al., 1989) noted that pre-linguistic infants preferred to listen to utterances that were parsed at language-relevant boundaries such as before or after a noun phrase rather than at points within the noun or verb phrases (e.g. “the big house PAUSE” vs. ”the PAUSE big house”). This finding only emerged, however, when the input was in infant-directed speech. Interestingly, prosodic cues offered hints to language structure at the level of both clauses and phrases (Jusczyk et al., 1992). Taken together, these findings suggest that prosodic qualities of motherese might provide infants with cues to units of speech that corresponded to grammatical units of language. These results also lend support to the prosodic bootstrapping hypothesis, which proposes that infants learn about the syntactic structure of their language through the prosodic characteristics of the speech that they hear (Gleitman et al., 1988; Gleitman & Wanner, 1982; see Morgan & Demuth, 1996 for a review).

Recasts, Expansions and Imitation

Once children find a toehold for language through infant-directed speech, one must ask how they arrange words and phrases into grammar. That is, how do they move the units of language around to form the rules for language? Two possibilities were introduced by the social-interactionists in the form of imitation and correction (called “negative evidence”). Perhaps children simply imitate much of what they hear, and perhaps when they generate ungrammatical sentences, they are corrected by benevolent social partners who help them move progressively towards grammatical speech. Although these ideas were appealing, there were several challenges that needed to be addressed. First, some researchers argued that children do not generally imitate what they hear (Brown, 1973). In fact, children generate novel sentences like, “I goed to the store” that have never been said before. Further, in his classic book, Brown (1973) held that
children could not learn language solely through social interaction because parents did not explicitly correct children’s ungrammatical speech. Rather, parents responded to the child’s intended message rather than to the form it arrived in.

In response to Brown’s claims, some social-interactionists sought to demonstrate that infant speech was only an approximation of adult speech, and that partial if not complete imitation did have a role to play in later language learning (K.E. Nelson, 1989). Further, while parents do not respond explicitly to ungrammatical speech (e.g., telling the child to say, “I goed” as “I went” instead), they do respond implicitly by repeating more ungrammatical utterances or by expanding these utterances in ways that made them grammatically correct. By way of example, Bohannon and others demonstrated that adults directly respond to ungrammatical utterances in their children’s speech by using expansions, repeats, recasts, and requests for clarification (Bohannon & Stanowicz, 1988; Chouinard & Clark, 2003; Demetras, Post, & Snow, 1986; Farrar, 1992; Hirsh-Pasek, Treiman, & Schneiderman, 1984). Consider the following anecdote from Clark (2003, p.427) in which 30-month-old Abe asks his father for milk:

“Abe: Milk, milk.
Father: You want milk?
Abe: Uh-huh.
Father: Ok. Just a second and I’ll get you some.”

The father reformulates the child’s utterance by filling in the missing terms needed to make the utterance more complete and grammatically correct. These contingent replies to ill-formed child utterances are different from the types of responses used to respond to well-formed utterances. The use of different reply-types to the children’s utterances might indicate to children whether their utterance was ill-formed or well-formed. Chouinard and Clark (2003) argue that
when children receive negative evidence, they are presented with two forms: Their own ill-formed utterance and the adult’s well-formed utterance, both expressing the same underlying meaning. Chouinard and Clark write, “Since these two forms do not contrast in meaning (they express the same intention), the one that is conventional has priority” (p. 643). As a result of hearing a well-formed utterance that expresses their intention, children “defer to the adult speakers, the experts on the conventional forms for expressing specific meanings” (p.643). Taken as a whole, these studies show that adults do correct and reformulate children’s errors using various reply-types.

Of course, corrections and reformulations are useful to children only if they can detect the corrections and if they are prevalent enough in the input for children to make use of them. Shipley, Smith, and Gleitman (1969) reported just this finding in young children’s responses to well-formed versus ill-formed commands. Choiunard and Clark (2003) further investigated children’s ability to detect adults’ corrections. Longitudinal data, collected from five young children, showed that children do attend to adult reformulations. On average, adult corrections of ill-formed sentences were detected between 10% and 50% of the time. Thus, studies do demonstrate that implicit correction is available. Whether these corrections are frequent enough to assist children in grammatical learning (see Hirsh-Pasek et al., 1984 or Pinker, 1984) is still at issue.

At the grammatical level, then, there is reason to believe that infant-directed speech might capture infant attention (Fernald, 1985), and might carve the perceptual flow into language relevant units like clauses and phrases (Hirsh-Pasek et al., 1987; Kemler-Nelson et al., 1989; Morgan, 1986). Further, social feedback through expansions and recasts might alert children both to the fact that their own utterances were not “up to snuff” and to ways in which their
utterances could be repaired (Choiunard & Clark, 2003). While this might not solve the
grammatical learning problem by itself, it certainly highlights potentially important roles for
social input in grammatical learning.

The Role of Social Information in Word Learning

As documented above, researchers in the 1970’s and 1980’s focused on the role that
social information might play in grammatical development. In contrast, scientists in the 1990’s
turned their attention to the role of social input in word learning.

On the surface, word learning seems like a very easy problem to solve. At the simplest
level, a child need only attach a word to an object in plain view. Upon closer inspection,
however, word learning proves to be a difficult task. To learn a word, infants need to (a) segment
the continuous sound stream into units; (b) discover a world of objects, actions, and events, and
figure out how those are divided into meaningful units to find the natural “carving joints”; and
(c) map the word onto the correct referent in the world.

Quine (1960) highlighted the difficulty of word learning in his well-known vignette on
word reference. He wrote about a linguist who sees a rabbit scurrying by while hearing a native
exclaim “gavagai!” Quine argued that the word “gavagai” could refer to an indeterminate
number of possible referents including the whole rabbit, rabbit’s ears, or rabbit’s hopping. If
Quine’s example captures the word-learning problem, then we must ask how children ever learn
what a word refers to. Three major theories of word learning arose to debate this question: (1) the
constraints/principles view; (2) the domain-general view; and (3) the social-pragmatic view. In
fact, these theories can largely be distinguished by whether they embrace or reject the Quinean
conundrum as a foundational assumption.

Constraints/Principles Views
The constraints/principles view adopts Quine’s view of the problem space. Under the constraints/principles view, word-to-world mapping is under-determined, so human minds must be equipped with constraints or principles that narrow the search space (Golinkoff, Mervis, & Hirsh-Pasek, 1994; Woodward & Markman, 1998). Children approach word learning biased to make certain assumptions over others for what a word might mean. For example, Markman’s principle of mutual exclusivity (N3C; Markman, 1989; see also Merriman & Bowman, 1989, and Golinkoff et al., 1994) proposes that children hypothesize that an object can have only one label. As a consequence, children assume that an unfamiliar name refers to an unfamiliar object rather than to an object that is labeled. In the last 20 years, numerous constraints and principles have been proposed to address the word-learning problem (see Golinkoff et al., 1994 and Woodward & Markman, 1998, for reviews).

**Domain-General Views**

The constraints/principles approach spearheaded an early and popular view of word learning, but has not gone unchallenged. Both the domain-general and the social-pragmatic views reject Quine’s assumptions and offer alternatives to the constraints approach. Proponents of the domain-general view like Smith (1995, 1999, 2000), Samuelson and Smith (1998), and Plunkett (1997), suggest that word learning can be accounted for through “dumb attentional mechanisms” such as perceptual saliency, association, and frequency. Samuelson and Smith (1998), articulated this position when they wrote that “general processes of perceiving, remembering, and attending when placed in the word-learning context may be *sufficient* in and of themselves to create children’s smart word interpretations” (p. 95). Abundant evidence is available to support these claims. Smith, Jones, and Landau (1996), for example, showed that young children’s extensions of novel names to novel objects was influenced by the relative
salience to perceptual features rather than to conceptual knowledge such as functional information. However, when children were asked to judge the similarity of the objects they used conceptual knowledge to group the objects. Smith et al. suggest young children’s initial interpretations of novel words are indeed guided by a “dumb attentional mechanism.”

Social-Pragmatic Views

The third dominant view, the social pragmatic alternative also dismissed Quine’s conundrum as a basis for word learning. Researchers in this camp argued that children were socially gifted from an early age and that their social precociousness would allow them to learn words as apprentices to the master word users around them. The last 20 years has revealed that infants have remarkable social skills from an early age. Infants’ have the ability to follow eye gaze (see Butterworth, 1995 for a review; for more recent work see Brooks & Meltzoff, 2002 and Woodward, 2003) to pointing (Butterworth & Itakura, 1998; see Butterworth, 2003 for a review; for more recent work see Woodward & Guajardo, 2002), to enter into joint attention with another (see Moore and Dunham, 1995 for a review), to use a significant others’ facial expressions for information (see Feinman, 1992 for a review), and to even imitate another’s intended, but never completed, actions (Carpenter, Akhtar, & Tomasello, 1998; Meltzoff, 1995; Meltzoff & Brooks, 2001). A voluminous literature appeared linking these social competencies to word learning and later language learning.

Using speaker intent to learn words. Children learn words by reading the social intent of their mentors in this world (P. Bloom, 2000). When an adult says “dog” while looking at a dog, infants follow the gaze of the speaker, interpret speaker intent, and are literally guided to the correct word meaning. Thus, by virtue of being a social animal, language comes for free.
Among the first to demonstrate the link between social engagement and language acquisition were Tomasello and Farrar (1986; also see Tomasello & Todd, 1983) who examined the role of joint attention in early language development. They found that extended periods of joint attention between the adult and child serve as a scaffold for linguistic interactions. Adult references to objects that were already in children’s focus were positively correlated with children’s later vocabulary (also see Dunham & Dunham, 1992). In addition, words were learned better for objects that were in the child’s focus than those that were not.

More recently, studies investigating the role of social input in word learning have explored children’s ability to pick out the adults’ intended referent using pragmatic cues, such as eye gaze, head direction, body posture, and voice direction (Baldwin, 1991, 1993; see Baldwin & Moses, 2001 for a review). For example, Baldwin (1991, 1993) showed that 18- to 19-month-olds followed a speaker’s eye gaze upon hearing a novel label. Rather than linking the new label to an object occupying their own focus, these children used the speaker’s line of regard to guide their inferences about word meaning (also see Dunham, Dunham, & Curwin, 1993). Further, a study by Baldwin and colleagues (Baldwin, Bill, & Ontai, 1996) demonstrated that even infants as young as 12 months actively monitor a speaker’s face for information about the referent of a word. By age 2, children demonstrate their sophistication by learning a novel word in an intentional context, but not an accidental context (Diesendruck et al., 2004).

These studies demonstrate the utility of social information and social intent in the learning of object labels. Research also documents that infants recruit social information to learn action words or verbs (Akhtar & Tomasello, 1996; Poulin-Dubois & Forbes, 2002; Tomasello & Akhtar, 1995; Tomasello & Barton, 1994). Poulin-Dubois and Forbes (2002), for example, demonstrated that toddlers use social cues such as a speaker’s gesture and eye gaze to learn novel
action words. Just as in object word learning, Tomasello and colleagues illustrated that infants were able to learn novel action words even when the referent was absent at the time of naming (Akhtar & Tomasello, 1996; Tomasello & Akhtar, 1995). Moreover, Tomasello and Akhtar (1995) showed infants use pragmatic information to indicate what kind (an object or an action) of referent a label is referring to.

Where emotion meets word learning. Finally, children learn words not only by joint attention or by attending to speaker intent, but also by being attuned to emotional valence in the word-learning environment. For example, Mumme, Fernald, and Herrera (1996) found that children were very sensitive to the emotional affect of the speaker. When infants heard their mother utter a sentence like “oh, how frightful,” without a matching facial expression, 12-month-old infants looked at their mothers longer and showed an increase in negative affect. Facial emotional signals did not elicit the same responses suggesting that vocal expressions may be more powerful emotional signals relative to facial expressions (Fernald, 1992).

Akhtar and Tomasello (1996) demonstrated how emotional expression relates to word learning. In their study, 2-year-olds watched as a speaker announced her intention to find a toy in a bucket, saying, “Let’s go find the toma!” After retrieving the first toy from the bucket, the speaker displayed disappointment. Upon revealing the second toy, the speaker displayed pleasure and terminated her search. Infants immediately assumed that the object eliciting pleasure was the “toma”. In a replication of this study, Tomasello, Strosberg, and Akhtar (1996) noted that even 18-month-olds use speakers’ emotional expression and actions to determine an adult’s referential intentions.

Finally, L. Bloom investigated how emotional expression by the child affected word production across time (Bloom & Capatides, 1987). Early in language production, children
cannot express affect and talk at the same time. Instead, their affect becomes neutral before they speak and is not displayed again until after they produce a word. Those who are less expressive emotionally are more linguistically competent. Further, children tend to say “newer” words using a neutral affect, while those words said with emotional expression are generally familiar (L. Bloom, 1998). The coordination of emotion and language, while understudied at the current time, is a fruitful area for future research.

The evidence that social and emotional forces are important to language learning is compelling. What is less clear is just how far the social information can and does take the learner (Golinkoff, Hirsh-Pasek, Bloom, Smith, Woodward, Akhtar, Tomasello, & Hollich, 2000) in the learning of grammar and words. Some hold that those endorsing the social-emotional perspective have overstepped the limits of the theory. In grammatical learning, for example, this challenge comes from Hoff and Naigles (2002) who find that while children use social information to inform language learning, by age 2, syntactic frames are the predominant source of information about word meaning and grammatical form. Supporting the plausibility of a syntactic bootstrapping hypothesis, Hoff and Naigles argue that the grammatical input to children provides structural cues to language learning (also see Naigles & Hoff-Ginsberg, 1995).

In word learning, the challenge comes in two forms. First, some argue that the social information for the learner might not be social at all. Rather, the perceptual/attentional changes that accompany social cues may be the real data for word learning (Samuelson & Smith, 1998; but see Diesendruck, Markson, Akhtar, & Reudor, 2004). Samuelson and Smith (1998), for example, modified a study by Akhtar, Carpenter, and Tomasello (1996) to show that children could solve the word-learning task by relying exclusively on attentional cues, rather than on social cues. Hoff and Naigles (2002) support a different but related argument when they stated,
“the usefulness of social-pragmatic bases of information for child word learners around 10 to 15 months of age is due to maternal sensitivity, not children’s social-cognitive abilities…even though children over 18 months of age can use speaker intent, it appears not to be a particularly important source of information for actual word learning” (p. 428). Complexity in the sentences that children hear might account for more of variance in word learning than does social input. Under these scenarios the social input to children works through mediating cues like perception and grammar to build the child’s lexicon.

Second, some researchers challenge whether children can really recruit social intent in the service of word learning prior to 18 months of age (Hoff & Naigles, 2002; Hollich, Hirsh-Pasek, & Golinkoff, 2000). While there is some data to suggest that infants are sensitive to social intent during the latter half of the first year (Carpenter, Akhtar, & Tomasello, 1998; Gergely et al., 1995), few studies have discovered infant ability to use social intent in word learning prior to 18 months of age. If children are not using social intent to drive word meaning from the outset, then other strategies for word learning must be in play when children learn their very first words (but see Akhtar & Tomasello, 2000).

At present, each of the camps – be it the researchers emphasizing constraints, those espousing a domain-general approach, or the social pragmatists – seeks dominion over the theoretical landscape. Each, however, also pays lip service to the fact that word learning will not be possible without input from the social world, the perceptual world, and perhaps some linguistic principles that guide word learning. Even two of the major proponents of the social-pragmatic view argue that language learning is a complex process that requires a more comprehensive explanation. Baldwin and Tomasello (1998) wrote that language learning “…requires an explanation encompassing both its social and cognitive roots” (p. 19). And L.
Bloom (2000) has long argued for considering a whole, active child in explaining the language acquisition process. She noted, “The acquisition of language is, itself, embedded in other cognitive, social, and emotional developments that occur at the same time. Efforts to explain word learning, therefore, must involve broad principles that account for both developmental process and changes in behavior over time” (p.19).

Throughout the history of language development, then, be it in the study of grammar or word learning, scientists have made enormous progress determining the role of social input and engagement in language learning. The role of social information weaves a constant thread throughout the fabric of the study of language development. Yet, that thread has not been fully integrated into the tapestry of the science. In the last five years, there has been some movement towards a more integrative theory. Perhaps the question that we need to ask is which components of which theories govern children’s grammatical and word learning at different points in developmental time. This would provide us with an integrative and truly developmental theory. It would consider multiple perspectives to solve the complex problem in language development and would afford a role for social, perceptual, and cognitive inputs into the system. This kind of approach is not new (L. Bloom, 1993, 2000). As L. Bloom (1993) argued “cognitive developments bring the infant to the threshold of language only in conjunction with other developments in expression and social connectedness” (p.52). More recently, Woodward and Markman (1998) echoed the same sentiment when they wrote “word learning depends on an ability to recruit and integrate information from a range of sources” (p.371). The complexity of language learning requires a model that takes into account the findings from not only the social-pragmatic literature, but other literatures as well (e.g. constraints/principles, associationist; Hollich et al., 2000).
Part II: The Emergentist Coalition Model

The coalition models of grammatical and word development (Hirsh-Pasek & Golinkoff, 1996; Hollich et al., 2000) offered one attempt to provide an integrative theory that used multiple cues and strategies to break into language learning. For the purposes of this paper, we only talk about one of these models, the Emergentist Coalition Model (ECM) that attempts to provide such a theory for word learning. Like it’s predecessor for grammar, the ECM is an empirically based systems approach that incorporates aspects of the perceptual, social, and linguistic theories of learning. Thus, it embraces the role of social input but sees the role of social input in language learning as interactive with other inputs across time. The ECM rests on three central tenets: (1) that children utilize multiple cues, attentional/perceptual, social, and linguistic, that are always available to them in the language-learning situation; (2) that children’s ability to use these cues changes over time as does the weighting of these cues; and (3) that children develop word learning principles of language over time through these differential weightings (Hollich et al., 2000).

*Children are Sensitive to Multiple Cues*

Figure 1 graphically depicts the multiple inputs that are available for during the language learning process.

-Insert Figure 1 about here-

Among the earliest cues used in the service of word learning are attentional cues, such as perceptual salience, temporal contiguity, and novelty. Studies show that infants, well before the age of language acquisition, detect and link arbitrary relations in intermodal displays (Bahrick, 1983; Gogate & Bahrick, 1998). Literature on word learning also finds the novelty of an object draws children’s attention. In the presence of a word, children will map a novel word to a novel
object (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992). As noted above, research also demonstrates that infants are sensitive to social cues, such as eye gaze, pointing, and speaker intent (Butterworth & Itakura, 1998; Carpenter, Nagell, & Tomasello, 1998). This vast literature finds that infants can follow another’s eye gaze and point. Finally, infants also exploit linguistic input to help them find the words in the speech stream and to identify their part of speech (Hirsh-Pasek et al., 1987; Kemler-Nelson et al., 1989; Morgan, 1986). The question for language researchers then becomes not whether children are sensitive to various cues in the word-learning environment, but rather whether and when children exploit these various cues to learn a word?

**Children Differentially Weight Cues Over Time**

The ECM posits that while these cues are available to children at all times they are not equally utilized in the service of word learning. *Not all cues are created equal* (Hirsh-Pasek, Golinkoff, & Hollich, 2000) over the course of development. By way of example, the model posits that infants at the very beginning of word learning will rely more heavily on perceptual salience than on social or linguistic cues. When given a choice between attaching a novel word to a boring object that an adult is looking at and a more perceptually interesting object, children at the very early stages of word learning (10-12 mos.) should rely on perceptual salience rather than on social cues (e.g., adults’ direction of gaze or pointing). After they can interpret social intent of the speaker, they will realize that a label refers to the object or action that the speaker has in mind. Indeed, as apprentices to master language users, they might even come to over-rely on social information about word learning. Finally, as suggested by Hoff and Naigles, (2002; Naigles & Hoff-Ginsberg, 1995) children should come to use cues from the sentences they hear in the language to predict word meaning. At around two years and later, children comprehend sophisticated language structure and can use this structure to bootstrap word meaning. Indeed,
many theories of verb development rely quite heavily on the child’s burgeoning ability to detect linguistic structure (e.g., transitive & intransitive) in the input (Fisher, 2002; Lidz, Gleitman, & Gleitman, 2003; Naigles, 1990, 1996).

**Principles of Word Learning are Emergent**

The third assumption of the model holds that principles for word learning are emergent. Only through the combination of all sources of information can the process of word learning proceed. Lexical principles are the products not the engines of lexical development. They develop from immature principles (e.g., the label goes with the most interesting thing in the environment) to more mature principles (e.g., the label stands for what the speaker has in mind) only as their capabilities develop and they obtain more “practice” in learning words (see also Smith, 2000; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002 for arguments that word learning strategies change with practice). Early versions of the principles are subsequently refined to allow for faster and more precise word learning. For example, children do not start word learning with the novel name - novel category principle (N3C). Research shows that this principle is not in place until after the vocabulary spurt (Mervis & Bertrand, 1994).

Research from our laboratories has empirically assessed the assumptions of the ECM within the context of the principles of reference (that words map onto the child’s representation of objects, actions and events) and extendibility (that words map onto more than one exemplar; see Hollich et al., 2000). Three hypotheses are borne from the ECM: (1) children learning their first words (at 10- or 12-months of age) would be informed by multiple cues, of an attentional, social, and linguistic nature; (2) perceptual salience would be more heavily weighted than social cues for the novice than for the expert word learner; and (3) word learning principles develop along a continuum from immature to mature such that children are first attracted by what is most
salient to them, and only later note what is important to the speaker. As they break through the language barrier, children are guided (though not completely) by associationist laws i.e., perceptual salience. As they mature into veteran word learners, they are guided (though not completely) by social-pragmatic strategies—speaker intent.

Part III: Evidence for the Emergentist Coalition Model

Investigation of a hybrid model like the ECM demanded experiments that could trace development of the principles of reference and extendibility from their immature to their mature states. To assess the principle of reference, we examined whether infants would attach a label to both interesting and boring objects. We reasoned that a child with an immature principle of reference might attach a novel label to the perceptually interesting object - regardless of which object an adult was labeling. The child with a mature principle, on the other hand, should overcome the salience of the object in favor of relying on the speaker’s social cues to what is being labeled. To assess the principle of extendibility, we first asked whether infants would extend a label for a given object to one that differed only in color from the original exemplar. We then put infants in a very difficult task and asked whether they would use social information to extend that label to an object that bore no resemblance to the original object. After all, beanbag chairs and dining room chairs bear little resemblance to one another and yet they are both called, “chairs.” Children who fail to extend a label or who will only extend the label to close perceptual relatives possess an immature principle of extendibility. Alternatively, children who trust a social mentor extending a label in the face of contrasting perceptual cues are operating with a mature principle of extendibility (Maguire, Hennon, Hirsh-Pasek, & Golinkoff, submitted).

Examination of the emergentist coalition theory required a method that could be used equally effectively with children in the age range of interest (10 to 24 months) and one that
would enable researchers to manipulate multiple cues (attentional, social, and linguistic) and their interactions. The Interactive Intermodal Preferential Looking Paradigm (IIPLP) provided this new method (e.g., Hollich et al., 2000). Based on the “Intermodal Preferential Looking Paradigm” (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996) used to study lexical and syntactic comprehension, Baldwin’s (1991) “bucket task,” and Fagan’s (1971; Fagan, Singer, Montic, & Shepard, 1986) infant intelligence test, the method allows for the study of multiple cues to word learning in the first two years of life. The physical setup is depicted in Figure 2.

- Insert Figure 2 about here –

Infants are seated on their blindfolded mother’s lap facing the experimenter and our testing apparatus. After pre-exposure to the toys – familiar toys on some trials and novel toys on others – the toys are fixed with velcro onto one side of a two-sided black board that can be rotated so that the toys can go in and out of view for a specified period of time. The experimenter hides behind the board while children are inspecting the toys and during test trials. Coding is done off-line from video taped records and inter-rater reliability has consistently been very high.

Using this apparatus, it is possible to examine word learning in a controlled setting. Familiar object trials allow us to ask whether the child can “play our game.” The use of unfamiliar, novel objects permits exploration of the cues and combinations of cues that children use to guide word learning across development. The logic of the design (Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996) is that children should look more at an object that “matches” the linguistic stimulus than at an object that does not match. Thus, the dependent variable is visual fixation time to the target (named) object versus to the unnamed object.
Validation of the method comes from the familiar trials. Children at three ages were tested: 12- to 13-month-olds just at the beginning of word learning; 19- to 20-month-olds who may or may not have yet experienced a vocabulary spurt; and 24- to 25-month olds who typically have sizeable production vocabularies. In over 23 experiments children demonstrated the potency of the method by looking significantly more at the target item than at the non-target item in the familiar condition when an item was requested (Hollich et al., 2000). Evaluation of the hypotheses comes from children’s responses to novel stimuli. Using this method, we were able to explore how infants move from immature to mature principles of reference and extendibility and to examine the hypotheses that form the foundation for the emergentist coalition model. We were able to better understand how and when children used social information to buttress word learning.

*Evidence from the Studies on the Principle of Reference*

Reference, or the assumption that words refer, is the most basic of the word learning principles. Do infants assume that a word refers to an object, action, or event? How do they decide which object, action or event should receive a label when one is offered? To investigate these questions, conditions were created in which multiple cues were available to children but were sometimes placed in conflict. In what we called the *coincident* condition, we labeled the novel toy that coincided with children’s preferences – the interesting toy. In the *conflict* condition, we labeled the novel toy that did not coincide with the children’s preferences – the boring toy. We reasoned that learning the word in the coincident case should be easy for children because all of the “cues” – attentional, social and linguistic – were in alignment. In contrast, learning a novel word in the conflict condition should be more difficult because the coalition of cues is not acting in concert.
The experiment was conducted in four phases. First, children were given the opportunity to explore both the interesting and boring toys. Second, children participated in a salience trial in which they saw both the interesting and boring toy mounted side by side on the board. Third, in the labeling phase, the experimenter captured children’s attention, displayed both toys and labeled the target five times with a novel word (e.g., “danu”). In the coincident condition, the experimenter looked at and labeled the interesting toy; in the conflict condition, she looked at and labeled the boring toy. Finally in the test trials, the experimenter, now hiding behind the board, asked for the object that was labeled during training, once again getting the child’s attention first. For example, she might say, “Eve, where’s the danu?” If children learned the name of the correct toy, they should look more to the target than at the non-target (see Hollich et al., 2000, for details).

Do children understand that words refer? What cues do they use to determine the referent of a word? The participants in this study were 32 children at each of 12, 19 and 24 months of age. At all three ages, there is evidence that children detected the range of cues available. Importantly, even the 12-month-olds detected the social cue of eye gaze. They could not, however, use the social cues to inform word learning when they were in conflict with perceptual salience. The 12-month-old children learned the name of the object only in the coincident condition, as several further studies indicated (Hollich et al., 2000). The 19-month-olds learned the names of the objects in both conditions, but were still influenced heavily by perceptual salience. Even the oldest group who learned the names of the novel objects in both the conflict and the coincident conditions still showed the effects of perceptual salience by looking much longer at the target object in the coincident condition than in the conflict condition. This suggests that children were lured by the perceptual salience of the interesting toy, but were able
to overcome it when the boring toy was the focus of the experimenter’s attention. In short, these data suggest that infants with an immature principle of reference are more dominated by perceptual salience than are their counterparts with a more mature principle of reference. Nineteen- and 24-month old children recruit the speaker’s social intent when mapping word to object.

In light of these data, we conducted studies with 10-month-old infants to see whether children who were just beginning to acquire a comprehension vocabulary, operated like the 12-month-olds who were starting to produce language. Results from this age group suggest that 10-month-olds are even more bound to perceptual salience than the 12-month-olds. They demonstrated a clear preference for the interesting toy even in the conflict condition, suggesting these children ignore the presence of conflicting social cues. They apparently assume that labels “go with” interesting rather than boring objects, regardless of what the speaker is labeling (Hirsh-Pasek, Hennon, & Golinkoff, in preparation)!

What we see in the data is a clear pattern that changes over time, such that infants become increasingly less dependent on perceptual cues and more dependent on social cues to determine reference. Such data speak to both the associationist and the social-pragmatic theorist. The associationist position would predict that children would form a mismapping between the interesting object and the label in the conflict condition. If the 10-month-old data stand, then these data fit this prediction – but only for the very youngest children. Yet, by as early as 12 months of age, children with only 3 words in their productive vocabularies are already demonstrating some sensitivity to social information in a word learning task. These children, at the cusp of word learning, learned the novel labels only in the coincident condition. In the coincident condition, the experimenter labeled the object that the babies were most interested in.
For these babies, learning took place when the cues coincided. However, when multiple cues failed to coincide in the conflict condition, infants showed little evidence of word learning. They wanted to look at the interesting object despite the fact that the experimenter persisted in labeling the boring object. Though they looked at the interesting object much more than the boring object in the conflict condition, they did not falsely conclude that the novel label was attached to the interesting object. Even 12-month-olds were sensitive to the fact that the experimenter was looking elsewhere and not labeling the interesting object. These children are able to use multiple cues for word learning, but for learning to occur, the cues had to overlap. While the 12-month-olds are not social pragmatists yet, they also defy the predictions made by the associationistic camp. Only a hybrid theory that talks about attention to multiple cues and differential attention to these cues over time, can account for the data.

The ECM can and has been readily applied to verb learning. Using a similar design, but using the standard video version of the preferential looking paradigm, Pence et al., (2003) tested 21- to 24-month-old children’s ability to attach a novel verb to a natural action and noted a similar pattern of results using the live demonstration of an interesting action and a boring action (see Pence et al. for condition demonstrating unequal salience of actions).

**Evidence from Studies on Extension**

Our research not only examines the use of social information in determining word reference, but also word extension. That is, what cues do children use to extend a label for an object to another perceptually similar and dissimilar object (Maguire et al., submitted)? As in the study of reference, we reasoned that very young children might accept the same label for objects that are perceptually similar only later relying on speaker knowledge – social cues – to extend the label to objects that are perceptually distinct. Indeed, this is exactly what Maguire et al found.
Thirty-six children at each of ages 12, 19, and 24 months of age participated in this research. The experimenter offered a novel label for objects that differed only in color (three lemon juicers), that shared perceptual features for one dominant part (see Table 1) or that had no features in common.

-Insert Table 1 here-

The results suggested that infants had no trouble extending a novel label to objects that shared shape but not color, or even shape in a dominant part. In stark contrast, we witnessed a developmental progression in children’s willingness to accept the same label for objects that did not have any properties in common. Only 24-month-olds were willing to learn a common label for two objects that appeared to be very distinct. Only 24-month-olds were willing to “trust” the social mentor.

Taken together, this research on reference and extendability of words suggests that children are aware of the multiple cues for word learning, but that the potency of these cues changes over time. For word learning, social cues are important for young children because they allow them to tap into the vast resources of adult competencies. While some challenge the universality of the naming explosion (Godfield & Reznick, 1990, 1996), it is at least interesting that the ability to use social cues in word learning seems to come on line at around the time that children begin to amass a large number of words in their vocabulary.

Part IV: Implications of Emergentist Coalition Model

Our research suggests that children use multiple cues, at different points along their journey, to learn words. These data further suggest that the myopic views of word learning offer only incomplete snapshots of behavior. If it is the case that we can find testable hypotheses in an integrated view, then it should be possible to apply the theory broadly and gain new insights into
the ways in which social cues feed into language development as broadly conceived. Two areas that might prove particularly fruitful for study would be the study of language disabilities, and the study of the neurological correlates of early language development. What would happen, for example, if one of these critical forms of input (perceptual, social, and linguistic) was unavailable? How might these children and adults coordinate the cues differently, in ways that optimize learning and processing? Further, if language development is about the development of, use of, and coordination of multiple cues across time, than any disruption in the processing of those cues should differentially impair the acquisition and processing of language in both comprehension and production at different points in time.

Neurological studies could also provide insight into how social information and language interact. First, if areas of the brain that control social and emotional development are compromised, one should see concomitant deficiencies in the development of early language ability, especially at around 18 months to two years of age, when social information in language acquisition becomes more prominent. It would also be interesting to perform longitudinal studies in typically developing populations looking at both behavioral and neurobiological correlates. There should not only be biological evidence that underlies normal growth, but also some biological underpinnings for the coordination of social, perceptual and linguistic cues across time. While these are but speculative ideas, the twinning of behavioral and neurobiological measures could offer rich hypotheses about understanding the mechanisms behind language development.

**Atypical Language Development**

As a comprehensive theory of language development, the ECM allows us to explain the problems observed in normal, as well as atypical populations. For example, we can ask if
language impairments seen in atypical populations come from a dampening of one of the input cues (e.g., social cues) or from a delay in the accessibility of that cue. Language impairments that relate to the use of social information in autism may be best explained using our model.

Baron-Cohen et al’s (1997) work, for example, suggests that some of the key social cues used in language are severely compromised in autistic children. They found that only 29% of autistic children were able to use a speaker’s direction of gaze to infer the referent of a novel word, while over 79% of normal children and 70% of mentally handicapped children were able to use a speaker’s eye gaze in a word learning situation. Further, many describe autism’s primary impairment as an inability to understand and interpret another’s social intent (Baron-Cohen, 1995; Baron-Cohen, Baldwin, & Crowson, 1997; Tager-Flusberg, 1999a, 2001). If autistic children do not orient to naturally occurring social stimuli, such as hearing their name called (Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998), and they show diminished evidence of joint attention (Baron-Cohen, Tager-Flusberg, & Cohen, 2000), autistic children should show a different and deficient trajectory of language development (see Lord & Paul, 1997 and Tager-Flusberg, 1999b, 1999c for reviews). Indeed, they do! Research in this area shows that lower-functioning autistics either have echolalic speech or never acquire any expressive language at all (mutism; Bailey, Phillips, & Rutter, 1996). Those who exhibit echolalic speech often repeat parts or all of what a speaker said (Fay, 1993). Higher-functioning autistic children acquire language, but also have both delayed and deviant language development. While these children show vocabulary growth and understand word meanings similar to mental-aged matched children (Tager-Flusberg, 1985; Tager-Flusberg et al., 1990), they show delays in prosody, syntactic development, and pragmatics. Mundy and colleagues (Mundy, Sigman, & Kasari, 1990; Mundy, Sigman, Ungerer, & Sherman, 1987) show that less impaired autistic children who engage in
joint attention develop language skills far superior to those that do not engage in joint attention. Research also shows that some autistic children exhibit a regression that entails loss of previously acquired language and social skills, including loss of word use, following directed eye gaze and orienting to one’s name (Goldberg et al., 2003). More than half of the autistic children who showed evidence of word loss in their sample also showed a loss of social skills. These losses were evident in the 2nd year of life - an age associated with developments in social intention.

Using a model like the ECM offers a slightly different way to frame the problem. If autistic children have trouble using social information in word learning or grammatical learning, do they compensate by using information from either the perceptual or grammatical input at their disposal? Given our model about how social information contributes to the process of word learning we might expect (1) that word learning might be slower during the second half of the second year, when social information comes on line for word learning, and (2) the word learning that is achieved might be accomplished through different strategies than those used by typically developing children. As our model predicts, these other means might be perceptual, attentional, and/or grammatical elements.

There is evidence to suggest a decrease in the rate of word learning for autistic children in the second year (Goldberg et al., 2003; Lord & Paul, 1997). Recent research in our laboratory directly addressed the second prediction that autistic children come to rely on perceptual strategies as a means to lexical learning (Hennon, Hirsh-Pasek, & Golinkoff, 2003). This research examined 3- to 7-year-old autistic children’s language development and showed that these children do not to have the ability to use social cues in word learning. Rather, these
children learned words through increased attention to perceptual salience, responding more like typical one-year-olds in their language processing (Hollich et al., 2000).

Thus, by looking at atypical populations, one can begin to test not only the ECM Model of language learning, but can better specify the role and timing of social information in language acquisition. To fully take advantage of this strategy, it would be important to mount longitudinal studies of atypical children who are reported to have particular deficits and to not only chart their language development, but also pinpoint the ways in which particular processes come on line and interact in the course of language development.

**Neurobiological Correlates of Language Acquisition**

There is a rich literature on neurological correlates of language (see Bates, Thal, Finlay, & Clancy, in press, for a recent review). For example, research on the neurological correlates of lexical learning demonstrates that by 20 months of age processing of linguistic stimuli (both familiar and unfamiliar words) is limited to the temporal and parietal areas of the left hemisphere (Mills, Coffey-Corina, & Neville, 1993; although see Bates et al., 1997 and Neville & Mills, 1997). Some researchers have proposed two postnatal correlates of language development that may underlie the major language milestones. The first of these proposals explores the changes in frontal activity in the brain in 8- to 10-month-olds -- a period during which we see a growth in word comprehension and communication, and social development such as imitation, intention and joint referencing. Chugani, Phelps, and Mazziotta (1987) demonstrated an increase in frontal lobe metabolism in 9- to 12-month-olds which is thought to be caused by a burst in synaptogenesis. Such findings led to the proposal that the frontal lobes “come on line” in the first year of life and coincide with language milestones such as word comprehension. More recently, Clancy and Finlay (2001) argue that the behavioral developments seen between 8 and 10 months
of age are a result of the social and cognitive systems reaching “a certain critical level of organization” rather than a sudden maturation of the frontal lobes (p. 324). The second of these proposals explores the changes in the brain correlated with a burst in vocabulary and grammar between 16 and 30 months of age. There appear to be parallels between synaptogenesis and the bursts in both vocabulary and grammar, but little research has explored these parallels. However, some argue that these bursts in language development do not depend entirely on synaptogenesis. Rather, “the compelling parallel between the language burst and the synapse burst may represent a mutually beneficial relationship, but not a crucial and direct relationship of cause and effect” (Clancy & Finlay, 2001, p.324). Research on the neurological correlates of language acquisition that takes a developmental perspective is just beginning to emerge. Even less of that research is relevant to the ways in which social information might support or interact with language development. In this sense, studies that examine correlations between neurological and behavioral developments might offer insight into the ways in which social information is utilized in word learning and grammatical development.

Not only is there beginning research on neurological development and language, but there is also nascent research on areas of the developing brain related to social development. Studies on joint attention by Mundy, Card, and Fox (2000) found that 14- and 18-month-old infants who had greater left-frontal and left- and right-central cortical activity were more likely to initiate joint attention episodes. Frith and Frith (2001) investigated the biological basis of social interaction and demonstrated that a network of areas in the brain including the medial prefrontal and temporal cortex are involved with representing one’s own and other’s mental states. These areas of the brain may be involved with aspects of social interaction such as representing intention and the goals of actions. Such studies provide evidence for neural
processes associated with language development and social development (see also Mundy & Acra, this volume). They do not, however, directly address the underlying neural processes that may be important to the linkage between language and social development.

Is there any relation between the brain areas involved in social development and those involved in language acquisition? This is a central question that could help us better understand the mechanisms behind language learning. Mundy, Fox, and Card (2003) are among the first to ask this question by studying whether neurological underpinnings for joint attention are related to early aspects of language development. According to Mundy and colleagues “some element of the commonly observed linkage between initiating joint attention and language may involve basic infant brain-related maturation processes” (p.52). Their research indicates that there is in fact a relationship between brain maturation and joint attention skills on developing lexical abilities. Specifically, later language ability was predicted by both EEG coherence measures and joint attention.

To date, these studies of language development and social cognition are exploratory. They are already, however, beginning to chart neurological foundations for the social cognition that might mediate language development. From the rich data on typically developing children, several areas of study should prove even more profitable as researchers begin to chart this domain. First, given the behavioral data described in Part II of this paper, continued exploration of the neurological bases for social constructs like joint attention and social intent would be the best starting points for an understanding of how social cues mediate early lexical learning. There are clear behavioral operational definitions of these constructs that can be used to isolate incidents of these behaviors and to look for neurological underpinnings of these behaviors. There is also a wealth of data relating the use of social intent and joint attention to concurrent and later
language skills (e.g., Carpenter, Nagell, & Tomasello, 1998; Tomasello & Farrar, 1986; Tomasello & Todd, 1983).

Looking ahead, it would probably be worthwhile to design not only cross-sectional studies of neurological and behavioral development in social cognition and language, but also longitudinal studies. Perhaps social and cognitive developments in the brain jointly pave the way for later language skills (Clancy & Finlay, 2001). Another possibility is that there may be timing shifts such that neurological changes that feed the development of social intent enable concomitant changes in language that allow for behaviors like the so-called naming explosion (Mills, Coffey-Corina, & Neville, 1997) where children move from learning around 2 new words a day at roughly 19 months of age, to almost 9 new words a day at 23 months of age (Bloom, 1973; Carey, 1978; Nelson, 1973; but see Godfield & Reznick, 1990, 1996 for alternative view).

Not only would it be interesting to examine the ways in which social behaviors come on line to support language development, but it would also be interesting to more closely examine the ways in which varied behaviors are integrated towards a complex task like language learning. For example, the ECM suggests that perceptual cues give way to social cues as the dominant roadway for language learning. Is there neurological support for this behavioral evidence? Can we examine the ways in which children come to coordinate various sources of information towards an end? Some have argued that children with particular disabilities like autism might have access to independent sources of information, but that they cannot coordinate the information towards a goal like language learning (Frith, 2003). Similarly, the ECM model suggests that principles of language learning emerge in the coordination of perceptual, social and linguistic cues as children move towards language competence.
In short, behavioral data on developing social and language cues offers a looking glass for neurological researchers. As we better understand the neurological foundations for social behaviors and for language development in their parallel planes, we can begin to examine how areas might jointly support both developments. How might the absence of abilities in one area challenge the other? How might the brain compensate for the lack of social cues in language development? Are there specific areas of the brain that specialize in the coordination of cues towards a complex behavior like language development? Looking at neurobiological data in tandem with behavioral accounts will surely help us move from mere descriptions of what the brain does to explanations of why there appears to be a social dimension that might mediate early lexical and grammatical acquisition.

Conclusions

It is becoming increasingly clear that social and cognitive developments are inextricably intertwined. Nowhere has this story been more forcefully told than in the area of language development where the “interactionists” have struggled to find equal footing with the nativists to account for the acquisition of words and grammar. Language is learned within a social context and certain social developments seem pivotal to the language growth. In the area of word learning, the ability to use joint attention and to interpret social intent are perhaps necessary ingredients. In grammar, children may also benefit from attention to infant directed speech and to the recasts and expansions that might offer informal language lessons for the naïve grammarian.

Language is a complex cognitive process that no doubt will be served by multiple processes and behaviors. Recently, a number of researchers have tried to develop complex, interactive models that capture that complexity. These models go beyond evidence that social information plays some role in particular parts of language development to ask how, social
processing, in tandem with other kinds of processing yields emergent language behavior. The ECM is one such account.

In this chapter, we have argued that accounts like the ECM can be used to re-evaluate our understanding of the relationship between social cognition and language in typical populations. The model might help us to better understand how these behaviors interrelate in atypical populations and it might chart directions for the new frontier of research in neuropsychology and language.

Approaches like the ECM moves beyond searching for a single-factor, “smoking gun”, explanations for language development. Instead, they encourage us to take multiple, interacting factors into account as we unpack complex behavior. They force us to ask not only what biological foundations might separately underlie aspects of language, but how these factors mediate the path towards language competence as they separately develop, come on line in ontogeny, and coordinate with other abilities (social, linguistic, & perceptual input).
References


it rooted in intentional understanding or a result of simple orienting? Paper presented at
the International Conference on Infant Studies, Providence, RI.

Baldwin, D.A., & Moses, L.J. (2001). Links between social understanding and early word

understanding. In E. Clark (Ed.), Proceedings of the Twenty-ninth Annual Child
Language Research Forum: Vol. 29 (pp.3-23). Cambridge, UK: Cambridge University
Press.

MA: MIT Press.

speaker’s direction of gaze strategy to crack to code of language? Child Development, 68,
48-57.

Perspectives from developmental cognitive neuroscience. New York: Oxford University
Press.

MacWhinney (Ed.), Mechanisms of language acquisition. Hillsdale, NJ: Lawrence
Erlbaum Associates.

neural correlates. In I. Rapin, & S. Segalowitz (Eds.), Handbook of Neuropsychology:


Bretherton, I. (1992). Social referencing, intentional communication, and the interfacing of
minds in infancy. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 57-77). New York: Plenum Press.


acquisition (pp. 136-164). New York: Oxford University Press.


prosodic cues in motherese might assist language learning. *Journal of Child Language*, 16(1), 55-68.


(Eds.), *The teachability of language* (pp. 263-310). Baltimore: Brookes.


<table>
<thead>
<tr>
<th>Object</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Similar original category member: purple</td>
</tr>
<tr>
<td></td>
<td>Similar original category member: red</td>
</tr>
<tr>
<td></td>
<td>Similar original category member: green</td>
</tr>
<tr>
<td></td>
<td>Similar possible extension object: same-handle</td>
</tr>
<tr>
<td></td>
<td>Similar possible extension object: juicer</td>
</tr>
</tbody>
</table>
Figure Captions

*Figure 1.* Children are sensitive to multiple inputs, including perceptual salience, and social and linguistic cues, during the language learning process.

*Figure 2.* Interactive IPLP: The child sits on his or her parent’s lap in front of the flip board. The experimenter stands behind the flip board. A hidden camera behind the curtain records children’s looking preferences toward two objects on a display board.