How babies talk: Six principles of early language development

Kathy Hirsh-Pasek
Temple University

Roberta Michnick Golinkoff
University of Delaware

This research was funded by joint grants to the authors: From NSF, SBR9615391 and from NIH, RO1HD050199. Thanks to Russell Richie for his help in preparing the manuscript.
In 1995, Hart and Risley’s research became headline news. The language experience of children growing up in poverty was distinctly different than that of children growing up in working class or professional families. On average, the number of words heard per hour in the welfare group was 616 compared to 1,251 in the working class group and 2,153 in the professional group. These findings are dramatic and are consistent with results of recent studies (Hoff, 2002, 2003, 2006a; Arriaga, Fenson, Cronan, & Pethick, 1998; Cristofaro & Tamis-LeMonda, in press). Indeed, even within low-income homes the amount of language directed to children predicts later language output (Vernon-Feagans, Pancsofar, Willoughby, Odom, Quade, & Cox, 2008). Further, disparities in early cognitive development even during infancy grow larger over time. A recent report from the Early Childhood Longitudinal Study (Halle, Forry, Hair, Perper, Wandner, Wessel, Vick & Schultz, 2009) finds that cognitive outcomes that appear at 9 months are already demonstrably larger at 24 months of age.

These findings have enormous implications. First, Hart and Risley (1995) noted that vocabulary assessed at age 3 years significantly predicted scores of language competence at ages 9 and 10 (e.g., PPVT and TOLD). Second, recent results suggest that the amount of early language heard in infancy is strongly related to the speed with which children process language at 18 and 25 months of age (Fernald, Perfors, & Marchman, 2006); to vocabulary size at 25 months; and to linguistic and cognitive skills at 8 years of age (Marchman & Fernald, 2008). These findings hold for children who speak Spanish as well as for those who speak English (Hurtado, Marchman & Fernald, 2008) and appear to be language specific such that when testing processing speed in English/Spanish bilinguals, efficiency in processing English words affects English vocabulary size while
processing speed in Spanish predicts Spanish vocabulary size (Marchman, Fernald & Hurtado, 2010).

Third, not only does early language exposure predict later language ability, but there is also some evidence that it might also be related to brain differences in 5-year-olds. Recent papers by Farah et al. (2006) and Hackman and Farah (2009) suggest that the great disparities in language resulting from SES differences might have neurological consequences. In particular Hackman and Farah (2009) argue “On the basis of our three studies, the effects of poverty were disproportionate for certain neurocognitive systems, including language and executive function …” (p. 66). Though SES did not uniformly predict poor performance in all areas of the brain, middle class kindergarteners did outperform their low SES counterparts in tests that tapped the left perisylvian/language system.

Finally, it is important to note that early language differences are also related to school readiness and to later school outcomes in both reading and math (NICHD ECCRN 2005). For example, a large literature links early language competence to reading readiness in kindergarten and in primary school (National Early Literacy Panel, 2008; Storch & Whitehurst, 2002; Dickinson, Golinkoff & Hirsh-Pasek, 2010; Scarborough, 2001). Indeed, Dickinson and Tabor (2001) found that kindergarten vocabulary and word recognition skills that result from early language input accounted for over two thirds of the variance in fourth grade reading ability (see also Tabor, Porche, & Ross, 2003). Early language ability relates to improvement in the ability to detect separable sounds that feed phonological awareness (Munson, Kurtz, & Windsor, 2005), to vocabulary acquisition, and to grammatical development (Hoff, 2006b) and each of these
language systems is critical to later reading success. Language comprehension
difficulties in children are linked to poor oral language comprehension (Clarke,
Snowling, Truelove, & Hulme, 2010) such that intervention in oral language training was
more positively related to reading outcomes than was training in text comprehension.

The good news that lies beneath these shocking statistics is that the trajectory of
language development is malleable. A number of studies find that the amount of
vocabulary and grammar is related to the amount and quality of the language input (e.g.,
Hoff & Naigles, 2002). Importantly, intervention studies that coached parents of infants
in ways that support language development have been successful. In particular, the
National Early Literacy Panel (2008) reviewed 23 programs in which 18 had measures of
language growth. These programs had effects not only on early vocabulary with effect
sizes of .41, but also on broader language measures with effect sizes of .27. Some of
these programs achieved higher language outcomes by video-taping parents and thus
offering feedback so that they might become more sensitive (PALS; Landry, Smith, &
Swank, 2006).

Classroom interventions with older children show similar progress. Raising
quality in Head Start programs has had positive impact on language outcomes (Zill &
Resnick, 2006; Zill, Sorongon, Kim, Clark, & Wolverton, 2006). These studies suggest
the largest growth for children in the lowest quartile and report larger gains for Hispanic
over pre-k programs in 11 states. On average, creating high quality preschool
environments tends to enrich language outcomes by about 4.8 standard scores across
studies (see also Ross, Moiduddin, Megher, & Carlson, 2008). Perhaps more
impressively, the NICHD Study of Early Child Care and Youth Development (2002) began to unpack some of the mechanisms behind the relationship between high quality care and language outcome. The amount and quality of the input directed to particular children in a child care setting predicts language outcomes.

Finally, in a targeted study of language intervention, Huttenlocher, Vasilyeva and Levin (2002) asked whether the complexity of teacher’s language (multiclause sentences) with 43- to 50-month-old children was related to child outcomes over the course of the school year. The results suggested that while teacher language complexity was not related to child language level at the start of the school year, it was significantly related to children’s grammatical levels at the end of the school year.

These intervention studies give reason to be optimistic, but the findings of an intervention study by Feagans and Farran (1994) suggest that a single intervention may not be enough. Conducting their study at the beginning of the year, they found that experimental children performed better than controls in measures of narrative skill in the fall, but not in the spring. Taking these results with those of the other intervention studies above, it would seem that while language input to low income children portends poorer scores, with sustained intervention we can change the trajectory of development for these children because language development is malleable!

This paper begins to tell the back-story of language development in an effort to better understand the mechanisms of language growth and hence ways in which we can use current scientific data to best support strong language skills for all children. The story is told in two parts. First, we explore what we know about the progress of language development – reviewing what is obvious to the naked eye (in doctor’s charts) and
importantly what we are learning that is often hidden from view. Second, we suggest 6 principles of language learning that have emerged over the course of the last 2 decades of research. We conclude with suggestions on how these principles can be directly translated for classroom use even with very young children and how the use of these principles in practice might particularly profit low-income children.

A selective review of the evidence on language learning

*Charting Language Development*

*What you see.* The trajectory of language production has been documented. Diary studies dating to the late 1800s from scientists the likes of Darwin (1877) among many others (e.g., Stern & Stern, 1907; Leopold, 1939-1949; Bowerman, 1985, 1990; Dromi, 1987 & Halliday, 1975) relate the momentous climb from the first coos and gurgles of the 3-month-old to the first words at around 1 year of age and multiword speech at around 2 years of age. The milestones are charted in doctor’s offices and in baby books. And to this day, those who examine progress in language development often rely on production as a guide. Figure 1 presents an overview of these productive milestones.

[Insert Figure 1 about here]

*What you do not see.* The real story of language development, however, comes not merely from what children say, but also from what they understand well before they can speak and from the analyses they appear to conduct on language. New methodologies emerged that allow us to see beneath production – or the lack thereof – and to better understand *how* children learn not just *what* they learn.

These methodologies enable us to assemble the puzzle of language development. To learn a language, for example, children must be able to segment the continuous flow
of speech or the “melodies” of speech into sentences, phrases, words and sounds (or handshapes in the case of children learning sign language). They must also parse continuous events into the objects and actions referenced in language. Finally they must map the units of sound onto the objects, actions and events using words and grammar.

What might infants know in each of these areas? Newer methodologies give us some clues. Here we detail just a few of these methods to give a flavor of the tools that allow us entrée into the emerging language system. It is this view of language available through the windows of new methodologies that is shaping the landscape of early language development and offering clues to what types of early intervention might prove most profitable.

*A Methodological Sampler*

Perhaps the earliest breakthrough in language methodology came from a procedure called the *high amplitude sucking paradigm*. Introduced by Eimas, Siqueland, Jusczyk, and Vigorito (1971) this procedure capitalizes on a behavior in young infants’ repertoire. Infants suck on a pacifier that is rigged to electronically measure sucking bursts. Using this method, Eimas et al. (1971) cracked the first part of the language puzzle when they discovered that infants just a few weeks old heard the sounds of language as falling into the same categories possessed by adults. That is, children do not dissect the sounds they hear like physicists who notice every change in amplitude or hertz level. Rather, they lump the sounds they hear into the phonemic categories – b’s and p’s, for example – that correspond to the sounds used by languages.

High amplitude sucking rests on the “orienting response” (Sokolov, 1963) we share with other species to pay attention to new stimuli and cease responding to old ones.
This orienting response was the basis for the habituation paradigm (Bornstein, 1985), another method used to study language acquisition that presents infants with a repeated visual or auditory stimulus. After the stimulus becomes familiar and the dependent variable (e.g., sucking, visual fixation time) is reduced, a new stimulus is presented. If infants’ dependent variable response is heightened, then it is considered evidence that they discriminated between the familiar and new elements. An example of this procedure in action was used to investigate the second part of the language-learning puzzle – the question of whether infants can detect relational components in the events they witness.

Using an offshoot of the habituation response, Golinkoff (1975; Golinkoff & Kerr, 1978) familiarized 14-month-olds infants with filmed, dynamic events in which human actors played different roles in a pushing event (A→B). Would infants watch changes in action roles (B→A) more than a change in the direction of the action across the screen that maintained the same action roles (B<--A)? Action role changes won out suggesting that infants on the verge of language production could segment events into components they would soon encode.

The head turn preference procedure has also made an important contribution to our understanding of early language growth. Pioneered by Fernald (1985), and Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss, and Kennedy (1987), this technique was derived from Colombo and Bundy’s (1981) conditioned head turn procedure. In Hirsh-Pasek et al.’s (1987) version, 7- to 10-month-old babies sat in an enclosure where sound could be emitted from either the left or right sides of the apparatus. Above each speaker, infants saw a red light that would come on with the sound source to familiarize the children with samples of either Natural (with 1-sec pauses inserted at clause or phrase...
boundaries) or Unnatural (with 1-sec pauses within these natural boundaries) passages. The side on which the Natural and Unnatural speech was played for an individual infant was always the same. Here the insight was that infants would turn to one side or the other to activate a stream of speech that they wanted to hear. Duration of infants’ looks toward the speakers revealed a clear preference for the Natural sentences over the Unnatural sentences. This research moved beyond asking whether infants could discriminate individual sounds to ask how they segment fluent speech into the units like clauses, phrases and words.

The Intermodal Preferential Looking Paradigm (IPLP. Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996) enabled us to complete the last part of the puzzle – asking about how preverbal infants might map words and grammar onto objects, actions and events. Derived from Fantz’s (1961) discovery that infants would choose to look at one of two pictures presented side by side, and Spelke’s (1976) extension of the paired comparisons method to dynamic events, Golinkoff et al. (1987) recognized that this method had great promise for testing language comprehension.

The first IPLP studies, tested infants who were expected to show comprehension of nouns, verbs, and word order. For example, Golinkoff et al. tested 28-month-old 2- to 3- word speakers with a pair of videos showing Big Bird tickling Cookie Monster on one screen and Cookie Monster tickling Big Bird on an adjacent screen. The accompanying audio said, “Big Bird is tickling Cookie Monster” (or the reverse). When children’s visual fixation time landed longer on the matching than the non-matching event, it was clear that they could use word order to map to events in the world. To show that they used word order to guide their attention to these events, toddlers needed to a) analyze
who was agent and who patient in each pair of events; b) process the sentences they heard while they watched the events; and c) discern which event in a pair matched the language they were hearing.

When the IPLP showed predicted language comprehension, it became clear that it could be used to assess burgeoning language knowledge. For example, Hirsh-Pasek and Golinkoff (1996) showed that infants as young as 13 months of age recognized that words come in ‘packages’ specifying unique events in the world. When shown a video of a woman kissing a set of keys and holding a ball in the foreground vs. the same woman kissing the ball and dangling the keys in the foreground, babies looked more at the matching event when they heard, “She’s kissing the keys!” This result could only have emerged if infants were processing the sentence as a sentence as opposed to individual words as both videos contained a “she,” the action of “kissing,” and the target item (keys, in this case). By 17 months, infants with as few as 2 words in their productive vocabularies could use word order as a guide to watching the specific events with Cookie Monster and Big Bird (described above) that matched what they were hearing. This latter finding was taken as confirmation that infants not yet speaking were capable of comprehending not only action role relations in language but likely the grammatical categories of subject and object of the sentence. A 3-D version of the paradigm was introduced in 2000 (Hollich, Hirsh-Pasek, & Golinkoff, 2000) and has permitted further studies of early word learning (Pruden, Hennon, Hirsh-Pasek, & Golinkoff, 2006) while teaching us about the processes children use to acquire new words. In short, the IPLP and its extensions enabled researchers to examine how children mapped words onto objects and actions and sentences onto events.
There are a host of other popular methodologies that populate infant and toddler language research (e.g. the switch paradigm, Werker, Cohen, Lloyd, Casasola, & Stager, 1998). And new technologies are now pressing the frontier of our knowledge by exploring brain processes that are enacted when children discriminate words (Mills, Coffrey-Corina, & Neville, 1993; Coffey-Corina, Kuhl, Padden, & Dawson, 2006) or use social support for language learning (Kuhl, 2007; Kuhl & Rivera-Gaxiola, 2008). It is with the advent of these methods that we have been able to survey a new landscape in early language learning – the landscape that lies beneath what we see in the doctors’ charts.

Collectively, these methods opened up vistas of language learning that now allow us to paint a new portrait of what children know well before they can speak. More recently, the findings from these methodologies have also led to longitudinal studies where early competencies in perceptual and social processing for language appear to be related to later language skills. We review selective findings from these methods in the context of the Emergentist Coalition Model of Language, a framework developed as a comprehensive and integrative look at the earliest language growth.

*The Emergentist Coalition Model*

Introduced by Hirsh-Pasek and Golinkoff (1996; Golinkoff, Hirsh-Pasek, & Hollich, 1999; Hollich et al., 2000) the Emergentist Coalition Framework was designed to organize the new findings in language acquisition and to ask how the perceptual, social and linguistic input available to children might be integrated to promote both vocabulary and grammatical growth. The framework makes three assumptions. First, children are sensitive to multiple language inputs at any given time. Second, in the course of
development, these inputs are differentially weighted such that infants and young
children first weight perceptual information that will enable them to parse the incoming
flow of sounds and the visual information in events, only later attending to social cues in
the service of language learning and grammatical information that informs language
learning. Third, principles for language development emerge in the course of weighing
the different sources of information over time.

The Emergentist Coalition Framework answered a call form a number of theorists
looking for a more holistic approach to language development (Bloom, 1993; Thelen &
Smith, 1994; Woodward & Markman, 1998). In fact, in 1996, Katherine Nelson wrote,
“there are no single effective pushes to the developing system but rather a combination of
influences that lead to observable change” (p. 85). To this date, however, most studies of
language growth focus on one aspect of language development – the sounds, the words,
the grammar –rather than on a broader approach. This broader approach also forces us to
look beyond milestones at particular ages and towards the processes the enable language
development. Sternberg (1984) put it best when he wrote:

There are two fundamental questions in developmental psychology. First, what
are the psychological states individuals pass through at different points in their
development? Second, what are the mechanisms of development by which
individuals pass from one state to another? A strong case could be made that the
second question is the more basic one…p. vii

Here we use the Emergentist Coalition Model as a guide to organize and explain new
developments in the various areas of language development that were revealed in the
methods mentioned above. In the next section we turn to a set of principles that can be
derived from these findings. A visual representation of the model can be seen in Figure 2 below.

Add figure 2 about here

*Perception: Detecting the Sounds of Language*

Among the now classic studies was that of DeCasper and Spence (1986) who found that even in the womb, infants begin to listen to and remember speech. These researchers asked mothers of fetuses to read a particular story (e.g., *Cat in the Hat*) two times each day for the last 6 weeks of their pregnancy. After infants were born they either heard that story or an unfamiliar one. Using a sucking procedure, they found that the infants preferred to hear the familiar story indicating that they were listening to language even in the womb. Others have found that newborns prefer listening to human language, spoken or signed, over an artificial “language” that mimics some of its properties (Vouloumanos & Werker, 2007; Krentz & Corina, 2007).

Eimas et al., (1971) demonstrated that shortly after birth infants not only can distinguish a familiar story, but are prepared to discriminate sounds in all of the worlds languages -- an ability that they seem to maintain until approximately 8 months of age when they become less sensitive to sounds not used in the home language as they morphed into language specialists (Golinkoff & Hirsh-Pasek 1999; Kuhl, Stevens, Hayashi, Deguchi, Kiritani, & Iverson, 2006; Werker & Yeung, 2006).

For bilingual children, the ability to discriminate between sounds in non-native languages lasts for a longer time than for monolingual children – a quality that may be adaptive as children sort out which sounds matter for meaning in their languages (Fennell, Byers-Heinlein, & Werker, 2007). Zeroing in on the sounds that comprise the native language (or languages) sets children up for learning words as it narrows down the
range of sounds to consider when novel words are encountered. It may also increase children’s ability to note common sound patterns and, in turn, to attach these sound patterns to referents in the world (Jusczyk, 1997; Werker & Curtin, 2005).

At around 4.5 months of age, using the head turn procedure, Mandel, Jusczyk, & Pisoni (1995) reported that infants recognize the sound pattern of their own name compared to a name with the same number of syllables and stress pattern stress (e.g., LAUren vs HARRy). By 6 months infants use their own name to serve as wedge into the speech stream, recognizing a new word that appears after their name but not after an unfamiliar name. (Bortfeld, Morgan, Golinkoff, and Rathbun, 2005).

Perhaps the most celebrated of the perception studies launched a new and critical field relevant to later intervention. Conducted by Saffran, Aslin and Newport (1996), infants listened to two minutes of artificial speech created out of nonsense syllables that were strung together and presented in monotone with no temporal breaks between the syllables. Within the string were embedded three-syllable nonsense “words” that always appeared together. Thus, Saffran et al. manipulated the transitional probabilities between syllables such that some syllables were more likely to occur together. Could 8-month olds learn that “pidago” was a “word” but “pidare” was not? With only 2 minutes of exposure to the artificial syllable string, infants showed a preference for syllables that had been part of “words” over syllables that were not part of “words.”

This study, using the headturn preference procedure, made clear that infants bring computational abilities to the language-learning task and act as statisticians on the input before it has meaning to them. One consequence of this finding is that infants need to
hear enough input on which to perform the statistics – a point that we will return to under principles of language learning.

Importantly, another now classic study revealed that infants not only learn the statistics relevant to word learning, but also for grammar learning. Marcus, Vijayan, Rao and Vishton (1999) found that infants detect algebraic patterns like ABA (e.g., ga ti ga) vs ABB (ga ti ti). Impressively, 7-month olds who became familiarized with the ABA sequence recognized that it was the same as one that included novel syllables in the same ABA pattern (e.g. li na li) and different than a sequence with the same novel syllables in a different pattern (e.g., li na na or ABB). Infants not only calculate statistics over particular elements in a sequence, but also abstract patterns in the language data they hear.

A final note on some of the work that has investigated perceptual properties of language is that infants tend to pay more attention to the speech they hear when it is in infant-directed speech than in adult-directed speech. Characteristically, when adults are speaking to infants rather than to adults, they often modify word order (Narasimhan & Dimroth, 2008), exaggerate semantically meaningful tones (Liu, Tsao, & Kuhl, 2007), speak more slowly, use simpler words with fewer syllables per phrase, use higher than average frequency, and exaggerate intonation contours by stretching vowels (Fernald, 1992; Fernald & Mazzie, 1991; Fernald & Simon, 1984; Gleitman, Newport, & Gleitman, 1984; Grieser & Kuhl, 1988).

This change in register has many positive outcomes for children, including larger vocabularies (Gleitman et al., 1984; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Masur, 1982; Rowe, 2008; Tomasello, 1988), better word memory and categorization
skills (Singh, 2008), improved word segmentation (Thiessen, Hill, & Saffran, 2005), and better speech perception – especially in the area of phonology (Liu, Kuhl, & Tsao, 2003; Tsao, Liu, & Kuhl, 2004; Werker et al., 2007). A recent paper provides the first clear demonstration that 21-month-old infants learn words more readily under infant-directed than adult-directed speech (Ma, Golinkoff, Houston, Hirsh-Pasek, in press).

*Perception: Detecting Objects and Actions in Events*

In the first year, infants are not only statisticians, but physicists. A rich literature reviews their attention to properties of objects as early as 4-months of age (Spelke, 1990; Baillargeon, 1987; Baillargeon, Li, Ng, & Yuan, 2009; Carey, 2009). Among the discoveries are that infants under 6 months mentally represent absent objects (Baillargeon, 1987), expect unsupported objects to fall (Needham & Baillargeon, 1993), and have different expectations for solids and liquids (Hespos, Ferry, & Rips, 2009). Further, infants as young as 4 months can detect individual objects in a spatial array (Spelke, Kestenbaum, Simon, & Wein, 1995), and by 12 months, infants can individuate these objects on the basis of object kind information alone (e.g., if a blue ball emerges from behind a screen and returns behind it, and then a red ball emerges from behind the same screen and then returns, 12-month-olds understand that they have seen two different objects; Xu & Carey, 1996; Xu, Carey, & Welch, 1999). Learning about objects and their properties, and being able to categorize them, is critically related to word learning and might even be related to the vocabulary spurt (Gopnik & Meltzoff, 1987). More recently, researchers are also asking about how infants might parse the events that they witness into conceptual foundations for learning action words and relational terms like prepositions and verbs (Song, Golinkoff, Hirsh-Pasek, under review). To learn a verb like
‘march,’ for example, English-reared infants must differentiate the act of ‘marching’ from, say, ‘hopping’. They must recognize that the manners of two actions – how the person moves (e.g. marching vs hopping) are different and that they are distinct even if they travel along the same path. To use the linguistic terminology, these manners of motion are carried out by an agent who serves as the figure of the action, and they always occur against some ground (running on the beach vs the street).

What is interesting is that learning verbs is not only contingent upon dissecting events in ways that will be labeled by language, but also in packaging the components together in language-ready ways. For example, “marching” is encoded as a verb in English “march into the class” whereas in Turkish “sınıfa yürüyerek girdi” – means go into the class marchingly” surfacing as an adverb – if at all. Spanish patterns in the same way as does Turkish with the path information conflated in the verb and the manner or how the action is performed in an adverb. Notice that bilingual and ESL children will have to learn that the packaging of event components required in their two languages is different. This might pose a particular problem in the learning of relational terms vs nouns especially for ESL children who already have an entrenched first language.

As a backdrop to asking how children might master relational terms like verbs, some are now investigating when children zoom in on the foundational semantic properties of events like manner and path and how attention to these event components might wane or wax when the processing of events meets up against the native language (see Göksun, Hirsh-Pasek & Golinkoff, 2010 for a review). Briefly the findings suggest that 7-month-olds can discriminate between different paths (over/under) and manners (spinning vs toe touching, Pulverman, Golinkoff, Hirsh-Pasek, & Sootsman Buresh,
2008; Song et al., under review) and can categorize these properties of events by the beginning of the second year of life (Pruden, Hirsh-Pasek, Maguire, & Meyer, 2004). Similar results were obtained from Spanish-reared and Mandarin-reared infants (Pulverman, Chen, Chan, Tardif, & Meng, 2007; Pulverman et al., 2008).

Infants viewing events are also sensitive to properties like containment (English “in”) and support relationships (English “on”; Choi & Bowerman, 1991; Casasola, Cohen, & Chiarello, 2003; McDonough, Choi, & Mandler, 2003; Hespos & Spelke, 2007), figures and grounds (Göksun et al., 2010), and the source and goal of the same action (e.g. the bird flew from the tree (source) to the bush (goal), Lakusta, Wagner, O’Hearn, & Landau, 2007; Lakusta & Carey, 2008). And as described earlier, infants seem to be able to tell the difference between agents and patients of action in events (Golinkoff, 1975; Golinkoff & Kerr, 1978).

What is particularly exciting about these findings is that infants are not only processing language-relevant information in the fluid events that they witness, but they seem to start with a set of universal divisions in events. As children learn their native language they allocate attention to different aspects of the event and thus reframe the event in ways that conform to their native language (see Göksun et al., 2010 for a review). In some ways this shift is analogous, though not homologous, to the shift from a universal set of sounds discriminated by all infants that becomes refined and tailored to the native language with language experience (Hespos & Spelke, 2004).

In sum, then, infants in the first year of life already have important analytic tools that allow them to perceive the sounds of speech and the elements in events that will prepare them for language learning. Recent research suggests that these perceptual
beginnings might predict to later language competence. Speech segmentation ability measured in children younger than one year of age is related to language ability and cognitive abilities 3 to 5 years later (Newman, Ratner, Jusczyk, Jusczyk & Dow, 2006; Höhle, van de Vijver, & Weissenborn, 2006; Tsao, Liu, & Kuhl, 2004). And the ability to detect phonemes present in one’s native language at 7.5 months of age is also predictive of later language abilities, but the ability to detect phonemes from other languages is negatively correlated with language outcome (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Tsao, Liu, & Kuhl, 2004; for a review, see Kuhl, 2009).

This attention to individual differences in performance during infancy and its relationship to later language learning is relatively new. Research is only beginning to link early studies in object and event perception with emergent language. Already, studies like Marchman and Fernald’s (2008) are using this kind of work to suggest that the amount of language directed to infants and the ability to segment the language stream is predictive of later differences in processing speed and word learning. While these are but correlations, if they are borne out they have enormous consequences for children growing up in poverty who might not hear enough language to do the statistical analyses needed to foster a strong foundation for language learning.

Mapping Sounds and Events onto Words and Grammar

Much of the research on early language development centered on early perception. Another body of research, however, focuses on how children map words to world and grammar to events (see Woodward & Markman, 1998; Lavin, Hall & Waxman, 2006; Waxman & Lidz, 2006; Golinkoff et al., 2000; Bloom, 2000; Yuan & Fisher, 2009). As many have pointed out, the task of mapping words to world is
deceptively simple. A child who sees Quine’s (1960) proverbial rabbit hopping by while a linguist utters “gavagai” might think the word refers to the whole rabbit, the rabbit’s ears, or color of the rabbit ears, or the rabbits hopping. How is a child to know?

Researchers have proposed a number of answers to this question ranging from those who argue for perceptual salience or attentional cues as the link between word and meaning (Landau, Smith & Jones, 1988; Imai & Gentner, 1997; Jones & Smith, 2002; Smith, 2003; Soja, Carey & Spelke, 1991; Yoshida & Smith, 2003; Smith & Yu, 2008; Hills, Maouene, Maouene, Sheya, & Smith, 2009), to those who hold that infants have certain cognitive constraints that guide them to one interpretation for a word over another (e.g., children assume the word maps to the entire rabbit, Markman, 1991, 1994; Kersten & Smith, 2002; Saylor, Sabbagh, & Baldwin, 2002; Golinkoff, Hirsh-Pasek, & Mervis, 1994; Hollich, Golinkoff, Hirsh-Pasek, & Arnold, 2007), to those who suggest that children attend to social cues in the environment that support word mapping (Nelson, 1996; Adamson & Bakeman, 1991; Adamson, Bakeman, & Deckner, 2005; Adamson & MacArthur, 1995; Tomasello, 1995; Tomasello & Farrar, 1986). Research has been marshaled to support each of these claims.

The ECM, however, suggests that rather than endorse one theory to the exclusion of the others, this research can be unified in a developmental model of word learning. Hollich et al. (2000), for example, presented infants with two toys in a live adaptation of the preferential looking paradigm. The experimenter looked at and trained children on either the name of the interesting toy (Coincident Condition) or looked at and labeled the boring toy (Conflict Condition). While hiding behind a barrier that displayed both toys side by side, the experimenter then asked the child “Where is the modi? Can you find the
modi?” At 10 months of age, infants formed associations between the object they preferred (the interesting toy) and the word that they heard – regardless of whether the interesting or the boring toy was labeled by the experimenter (Pruden, Hennon, Hirsh-Pasek & Golinkoff, 2006). By 12 months of age, infants learned the label in the coincident condition but not in the conflict condition. That is – they started to notice the social cues but could not recruit social cues in the service of word learning. By 24 months of age, toddlers could ignore the perceptually interesting object to map the label onto the boring object that was designated by the experimenter. They used social cues like eye gaze to interpret the speaker’s intent. A reliance on social over perceptual cues in word learning is critical as we do not necessarily label the most interesting objects in our immediate environment.

The role that social cues play in word mapping has been extensively studied. Research on communicative intentions suggests that children figure out the referent of a word by attending to what a person means to convey by word or action (Baldwin & Moses, 2001; Birch, Vauthier, & Bloom, 2008; Diesendruck, Markson, Akhtar, & Reudor, 2004; Liszkowski, Carpenter, & Tomasello, 2008; Tomasello, Carpenter, Call, Behne & Moll, 2005) and that attention to social cues like eye gaze direction predicts later language outcomes (Brooks & Meltzoff, 2008). Research also demonstrates that children learn words best during episodes of joint attention or when the adult looks at the object that the child is focused upon rather than trying to draw the child to the object they find interesting (Adamson & Bakeman, 1991; Adamson, Bakeman, & Deckner, 2005; Adamson & MacArthur, 1995; Dunham, Dunham & Curwin, 1993; Tomasello, 1995; Tomasello & Farrar, 1986). Tomasello, Carpenter, Call, Behne, & Moll (2005) argues
that these moments of joint attention are also episodes of joint intention where the child tunes into the intentional goals of the speaker. With this window into the other person’s mind (what Lois Bloom (2001) called the “principle of relevance”), children become apprentice word learners to seasoned speakers.

Evidence from both typical and atypical populations suggests that early access to and experience with social cues is crucial to long-term language outcome, and that certain measurable social behaviors in the first year of life are predictive of later language ability (Markus, Mundy, Morales, Delgado, & Yale, 2000; Morales, Mundy, & Rojas, 1998). For example, Morales and colleagues (1998) found that 6-month-olds’ ability to follow a person’s gaze to an object (considered an early indicator of joint attentional skills) correlated positively with receptive/expressive vocabulary size at 12, 18, 21, and 24 months (Morales, Mundy, & Rojas, 1998). Further studies revealed that infants’ responding to various parental and experimenter bids for joint attention (e.g., pointing, looking, touching) is related to later language ability (Morales et al., 2000a, 2000b). In fact, populations that have less access to social cues tend to have poorer language abilities. In one study by Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff and Tager-Flusberg (2007), children diagnosed on the autistic spectrum engaged in a series of tasks designed to measure their access to social intent. Their attention to social intent predicted 68% of the variability in their vocabulary outcomes.

Finally, mapping comes not only from attention to perceptual salience and social intent, but also through a sensitivity to the grammatical structure of the language. By the second year of life, children use both social and linguistic cues to word meaning, such as eye gaze direction, information about the goals and intentions of others (Baldwin &
Moses, 2001; Brooks & Meltzoff, 2008; Hollich et al., 2000), tense agreement and morphology (e.g., a word ending with -ing is probably a verb and therefore labeling an action or event; for a review, see Fisher & Song, 2006). A classic example of infants employing their knowledge of syntax to determine word meaning comes from a study by Roger Brown (1957). Three- to five-year-old children were shown a picture of an action, a substance, and an object (e.g., person kneading a confetti-like mass in a bowl) and given a novel word (sib) in one of three syntactic contexts: as a noun (a sib), a verb (sibbing), or as a quantifier (some sib). Depending upon the syntactic context in which children heard the novel word, they interpreted it to mean either the bowl (noun), the process of kneading (verb), or the confetti-like mass (substance/material). Preschoolers made inferences about the meaning of a novel word from attending to the syntactic context in which it was heard (Brown, 1957; Naigles, Gleitman, & Gleitman, 1993; Subrahmanyam, Landau, & Gelman, 1999). Recently, Song et al., (2010) showed that 3-year-olds (English reared and French-reared) can use the morphology of a novel adjective or verb with no other cue to decide the meaning and part of speech of a new word.

The use of these grammatical cues in the service of word learning is referred to as syntactic bootstrapping (Gleitman, 1990; Naigles, 1996; Naigles, Gleitman & Gleitman, 1993). The term refers to when children use their knowledge of abstract linguistic structure (such as the fact that transitive verbs may be causal) (Fisher & Snedeker, 2002; Hirsh-Pasek & Golinkoff, 1996) to guide their interpretation of novel words. Some argue that given the ephemeral nature of verbs and the fact that verbs are generally harder to learn than nouns (Gentner, 1982; Gentner & Boroditsky, 2001; Imai et al., 2008), syntactic bootstrapping might be even more critical for verb learning than for noun
learning. Studies suggest that even toddlers as young as 21 months of age can use syntactic cues such as the arrangement and number of a verb’s noun arguments to determine something of the meaning of a novel verbs (Fisher & Song, 2006; Lee & Naigles, 2005, 2008; Lidz, Gleitman, & Gleitman, 2003; Naigles & Lehrer, 2002; Gertner, Fisher, & Eisengart, 2006).

In a stunning demonstration of the role grammar can play in verb learning, Gillette, Gleitman, Gleitman, and Lederer (1999) conducted their “human simulation” study. Researchers asked adults to guess the words used by mothers captured on videotape interacting with their children. Hearing a “bleep” that was overlaid onto the video sequences, adults guessed the word from pretaped maternal speech delivered to 18- to 24-month-old infants. Adults correctly guessed the missing nouns in 45% of the cases; their score for verbs was only 15%, and for mental verbs, 0%. When syntax was systematically added verb identification improved substantially. Snedeker, Li, & Yuan (2003) replicated this result with Mandarin and English speakers hearing Mandarin samples, and Piccin and Waxman (2007) replicated the finding with 7-year-olds.

The findings on mapping from word to world have several important lessons for those interested in supporting strong language growth. First, children will notice perceptually interesting objects and actions and will attempt to attach labels to those actions. Second, if we are sensitive to what they find interesting and label those objects and actions, children will engage in joint attention and master more vocabulary. As toddlers (Akhtar, Dunham & Dunham, 1991; Masur, 1982; Tomasello & Farrar, 1986; Dunham, Dunham & Curwin, 1993) they can also follow our gaze to new objects and those who do so more effectively become better word learners. Word learning and
sentence learning occurs in conversations where children and adults are focused on objects, actions and events. Third and critically, this research suggests that offering one word at a time is not supportive of language learning. Rather, exposing children to full sentences and rich language best promotes word learning as vocabulary and grammar acquisition are reciprocal processes.

Implications for the Classroom and for Intervention

The research presented thus far is but a fraction of that available in the literature. It does, however, reveal the significance of some of the new findings emerging in infant language development along with a sense of how much development goes on beneath the surface of what we see day in and day out. The new findings also give us direction on how we can intervene in the process to ensure positive language trajectories for young children. Moving back to the basics, we have seen that language learning can be distilled into three main tasks: a) finding the units of speech (or handshapes) that will become the sounds, words, phrases and sentences, b) finding the units in the world (objects, actions and events) that will be labeled by language, and c) mapping between word and world in ways that move from speech to meaning and meaning to speech in the native tongue (or tongues). With these basics in mind, we move from research findings to some speculation and offer 6 principles of language learning that are consistent with the literature and that might guide practice in the living room and the classroom.

Principle 1: Children learn the words that they hear most. As Neuman and Dwyer (2009) suggested, “Talk may be cheap but it is priceless for young developing minds” (p. 384). Support for this principle comes from the Hart and Risley (1995) study, a number of correlational studies (Hoff, 2006a; Hoff & Naigles, 2002; Hoff-Ginsberg, 1991;
Tamis-LeMonda & Bornstein, 2002), and the fact that amount of language exposure has long range consequences for later language and reading levels (Walker Greenwood, Hart, & Carta 1994; Weizman & Snow, 2001; Clarke et al., 2010; Marchman & Fernald, 2008). Saffran et al.’s (1996) findings begin to shed light on why the amount of language emerges as critical for young learners. Language must be heard in sufficient quantity for the computational brain to do the calculations necessary to derive the units of language.

This relationship between adult input and child output not only appears in home environments but also in studies of childcare and early schooling (Hoff, 2006b; Hoff & Naigles, 2002; Hoff-Ginsberg, 1991; McCartney, 1984; NICHD Early Child Care Research Network, 2000, 2002, 2005). Research also finds that children learn not only language that is directed to them, but also from overheard speech (Akhtar, 2005; Weizman & Snow, 2001). In short, providing environments filled with opportunities to hear and respond to language supports children’s growing vocabulary and grammar.

**Principle 2: Interactive and responsive rather than passive contexts favor language learning: Social interaction matters.** Amount of language heard is important, but amount alone will not support language competencies. For example, it is not enough to just hear language or words spewing forth from a television. Findings by Kuhl, Tsao, & Liu (2003) and Roseberry, Hirsh-Pasek and Golinkoff (2009) suggest that children under the age of three are unlikely to learn language from mere exposure on television (see also DeLoache et al., 2010). Language learning requires sensitive and responsive conversations with children where language input is tailored to the interest and timing of the child’s attention.
Adults who take turns in interactions with young children, share periods of joint focus, are sensitive and responsive to children, and express positive affect provide children with the scaffolding needed to facilitate language and cognitive growth (Bronfenbrenner & Morris, 1998; Clarke-Stewart, 1973; Bradley et al., 1989; Howes, 2000; Katz, 2000; Tomasello & Farrar, 1986; Landry, Smith, Swank, Assel, & Vellet, 2001; Wakschlag & Hans, 1999; Tamis-LeMonda & Bornstein, 2002; Hirsh-Pasek & Burchinal, 2006). This link has been observed in childcare homes and relative care as well as in center care (Kontos, Howes, Shinn, & Galinsky, 1997; NICHD ECCRN, 1998; Clarke-Stewart, Lowe Vandell, Burchinal, O'Brien, & McCartney, 2002).

A longitudinal study that examined teacher-child conversations when children were four, controlling at age three for children’s language ability (i.e., the mean length of their utterances), parental income, education, and home support for literacy (e.g., reading), found that higher quality conversations and richer vocabulary exposure during free play and group book reading were related to children’s language, comprehension, and print skills at the end of kindergarten (Dickinson, 2001a; Tabors, Snow, & Dickinson, 2001) and fourth grade (Dickinson, 2001a; Dickinson & Porche, in press; Tabors, et al., 2001). Parental sensitivity across time also fosters change in child outcomes (see Bornstein & Tamis-LeMonda, 1989; Landry, Smith, Swank, & Miller-Loncar, 2000; Landry et al., 2001; NICHD ECCRN, 2003; Tamis-LeMonda & Bornstein, 2002). Finally, using the NICHD Study of Early Child Care and Youth Development, Hirsh-Pasek and Burchinal (2006) noted similar relationships with children in child care settings. The moral of this story is that language competencies grow when we remain
conversational and attuned to children’s focus of attention and interests during the course of the infant, toddler and preschool years.

Why might social interactive settings be critical for language growth? Contingent interaction not only facilitates advanced babbling in infants (Goldstein, King, & West, 2003), but also helps older children come to discern adults’ intentions (Baldwin & Moses, 2001). Contingent responsivity is crucial for language growth (Roseberry, 2010), possibly because it engages the child and mirrors the child’s responses. It is interesting in this light that children under 3 have trouble learning language from television (Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009) but not from video chats (Roseberry, 2010).

**Principle 3:** *Children learn words for things and events that interest them.* In 2000, Lois Bloom wrote about the “Principle of Relevance,” commenting, “Language learning is enhanced when the words a child hears bear upon and are pertinent to the objects of engagement, interest and feelings…” (p.19). In fact, the evidence suggests that younger children readily assume that words map onto objects they find interesting (Hollich et al., 2000). The joint attention literature attests to the fact that parents who talk about what children are looking at have more advanced vocabularies. The corollary is also true: parents who try to redirect children’s attention and label objects not of interest have children learn fewer words (e.g., Dunham, Dunham & Curwin, 1993; Hollich et al., 2000; Golinkoff, 1981).

When children are interested and actively engaged in activities they also learn language from their peers. Dickinson (2001b) noted that the amount of time 3-year-olds spent talking with peers while pretending was positively associated with the size of their
vocabularies two years later, when they had begun kindergarten. And Bergen and Mauer (2000) found that 4-year-olds’ play, in the form of making shopping lists, and “reading” storybooks to stuffed animals, predicted both language and reading readiness after the children had entered kindergarten (Nicolopoulou, McDowell & Brockmeyer, 2006).

**Principle 4. Children learn words best in meaningful contexts.** Sparking interest is often a first step in meaning making. People learn best when information is presented in integrated contexts rather than as a set of isolated facts (Bartlett, 1932; Tulving, 1968; Neisser, 1967; Bruner, 1972; Bransford & Johnson, 1972). Words connected in a grocery list are better remembered than the same list of words presented without context.

Meaningful connections between words develop in conversations and in studies that use thematic play as a prop for language development. Christie and Roskos (2006), for example, find that children who learn connected vocabulary for the category of building (words like hammers, hard hats, screw drivers and tool belts) better remember and use these words than do children who do not learn in this more integrative way.

Meaning comes not only by connecting words with experiences, but also by hearing them used in sentences within narratives. Play can often provide a rich context for those narratives. New research by Han, Moore, Vukelich, and Buell (in press) finds that children who are given an opportunity to use vocabulary in a playful context learn it better than those who learn only under explicit instruction. As Neuman and Dwyer (2009) pointed out, experimental research on comparing vocabulary learning in meaningful versus less meaningful contexts is scant. Yet, correlational studies in language, play, and memory research converge to suggest that teaching vocabulary in integrated and meaningful contexts enriches and deepens children’s background
knowledge and hence their mental lexicons (Hirsh-Pasek, Golinkoff, Berk, & Singer, 2009).

**Principle 5. Vocabulary learning and grammatical development are reciprocal processes.** The amount and diversity of verbal stimulation fosters earlier and richer language outcomes in terms of both vocabulary and grammar (Beebe, Jaffe & Lachman, 1992; Hart & Risley, 1995, 1999; Snow, 1986; Huttenlocher, Haith, Bryk, Seltzer & Lyons, 1991; Tamis-LeMonda, Bornstein & Baumwell, 2001). Vocabulary and grammar are not divorced. They feed one another. Dixon and Marchman (2007), for example, argue from a large sample of children ages 16-30 months (N=1461) that words and grammar are “developing in synchrony across the first few years of life” (p. 209). This relationship between grammar and vocabulary learning is also celebrated in research with bilingual children. Conboy and Thal (2006) find, for example, that toddlers’ English vocabulary predicted their English grammar and the reverse, and their Spanish vocabulary predicted their Spanish grammar.

Why might language environments that offer more sophisticated language support learning? First, syntactic bootstrapping, which assists children in narrowing word meaning, can only be useful if children hear many sentences. Upon hearing “John is blicking Mary,” for example, the child may assume that John did something to Mary – an assumption not licensed by the sentence, “John and Mary are blicking.” Second, observing the diverse linguistic contexts in which words are used helps children detect nuances in word meaning (Gillette, Gleitman, Gleitman, & Lederer, 1999; Naigles, 1990). Finally, it is worth noting that oral language measured as both vocabulary and grammar (NICHD ECCRN, 2005) will be crucial for early literacy (Clarke et al., 2010).
**Principle 6: Keep it positive.** One of Hart and Risley’s startling findings was that lower income children are more likely to hear prohibitions (e.g., “Don’t touch that!”) than to hear what they called “affirmations” (e.g., “That’s an interesting toy”). Prohibitions are not only more negative in tone, but they serve as conversation closers. In a lovely illustration, Chase-Lansdale and Takanishi’s (2009) opened a recent report entitled, *How do families matter?*, with a vignette they called “three mothers and an eggplant.”

The first mother wheels her shopping cart down the produce aisle, where her kindergartner spots an eggplant and asks what it is. The mother shushes her child, ignoring the question. A second mother, faced with the same question, responds curtly, ‘Oh, that’s an eggplant, but we don’t eat it.’ The third mother coos, ‘Oh, that’s an eggplant. It’s one of the few purple vegetables.’ She picks it up, hands it to her son, and encourages him to put it on the scale. ‘Oh, look, it’s about two pounds!’ she says. ‘And it’s $1.99 a pound, so that would cost just about $4. That’s a bit pricey, but you like veal parmesan, and eggplant parmesan is delicious too. You’ll love it. Let’s buy one, take it home, cut it open. We’ll make a dish together.’

Rather than closing off the conversation, the third mother affirms the child’s interest, speaks in full sentences, and continues the conversation in a way that builds vocabulary and grammar. The research suggests that when we expand on our children’s language and when we ask questions rather than simply giving directives we talk more and we create a climate that spurs language growth. Continuing the conversation increases the amount of talk, uses language in a social context, builds on children’s interest, makes language meaningful, and generates more complex language samples.
Taken collectively, the 6 research-derived principles of language development offer a way to alter the trajectory of a child’s language development. The principles dictate a kind of pedagogical approach that yields optimal language growth. The topics under discussion can vary with language or culture. Strong language competence is as likely in conversations about family heritage as it is in discussions of cooking, drawing, or building a fort. Teachers and parents can confidently give children a rich language base by applying the principles in areas that are of interest to them and their children. The trick is to start the conversation and keep it going. As David Dickinson has said, “Strive for five,” meaning 5 back and forth turns with the child. When conversations are only one-side prohibitions or one-word answers, children are not hearing the language they need to fuel their language-learning engine. Nor are they being sufficiently exposed to the concepts language encodes.

Final thoughts. Language is a core mental ability that serves as the foundation for communication, academic learning, and social navigation. Children growing up in poverty are at a significant disadvantage for developing these language competencies (e.g., Hart & Risley, 1995). Given the centrality of language and communication for school readiness and academic success, this research has sent ripples to the highest levels of government. In the last 15 years, researchers have not only replicated these findings, but also expanded them, suggesting that the lack of positive language input might have lasting effects on children’s mental processing speed (e.g., Marchman & Fernald, 2008), their later reading abilities (e.g., Clarke et al., 2010) and possibly in brain functioning (e.g., Farah et al., 2006).
Research in the last decade, however, also offers an antidote. Language learning is malleable and we can change young children’s trajectories. We now know enough about the course of language learning to make the leap from the pages of scientific journals into application. Using a smorgasbord of the latest research, we offered 6 principles that distill the science into a roadmap for practitioners who work with young children and their families. If we are to narrow the achievement gap between rich and poor, it is imperative that we build strong language competencies in all children and that we put our science to work.


Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus


Figure 1: Milestones of language development

- **Birth**: 0-3 months: coos
  - 6-9 months: babbling
  
- **1 year**: 3-6 months: coos and laughs
  - 9-12 months: first word, points
  - 12-18 months: 2 words per week, 50 words at 18 months
  
- **2 years**: 18-24 months: naming explosion ("whas sat?")
  - 500 words; asks questions; uses wh- words, past tense, and other grammatical morphemes

- **3 years**: 800 words; uses contractions; narrative skills develop; more complex (e.g. multiclause utterances)

- **4 years**: 4-5: 2000 words; can make up stories
  - 1st grade: 11,000 words
  - 3rd grade: 20,000 words
  - 5th grade: 40,000 words

Figure developed by Russell Richie.
Figure 2: The coalition of cues available for establishing word reference and utilized differently across developmental time. Children shift from Phase I, a reliance on attentional cues such as how compelling an object is (perceptual salience) and the coincident appearance of an object and a label (temporal contiguity), to Phase II, a greater dependency on social and linguistic cues like eye gaze and grammar. By 12 months, dependence on Phase I cues has begun to wane and shift to the social cues in Phase II.