From Coo to Code: A Brief Story of Language Development

Julia Parish-Morris  Roberta Michnick Golinkoff  Kathryn Hirsh-Pasek
University of Pennsylvania  University of Delaware  Temple University

January 11, 2012

*The production of this chapter was supported by grants to the second and third authors from the National Science Foundation (SBR9601306 and SBR9615391) and from NIH (RO1HD050199).
1. 100-150 word abstract

Infants begin the process of language acquisition before birth, and their brains appear specially prepared to acquire complex systems of communication. This chapter charts the child’s journey from coo to code, discussing how children find their first words in a cacophony of sound, how they determine what words mean, and what happens when children have variable input or learn more than one language at the same time. The journey is exciting, enlightening, and filled with surprises as we explore the myriad abilities of a child’s developing mind.

2. 5-10 keywords for indexing etc.

language, language development, language acquisition, word learning, child development

3. List of 5 questions for future research

- What is the relationship between children’s early phonological and segmentation abilities and subsequent language acquisition?
- Given that we now know many of the factors associated with language development, can parents be trained to provide the kind of input that will facilitate language development if they do not do so naturally?
- In what ways will the advent of new neurological measures of language functioning shape the questions the field will begin to ask?
- Do children benefit from learning more than one language or are they hindered?
- What is the relationship between infants’ abilities to form nonlinguistic concepts such as path and manner and figure and ground and their subsequent ability to express these in language?

4. Bulleted list of 10 key points to take away from the chapter
• Language is an intricate system of sounds or gestured symbols that people use to communicate with one another.

• Theories of language development are diverse. Five major branches of thought include behaviorist, innatist, cognitive, interactionist (social), and most recently, Bayesian.

• The Emergentist Coalition Model (ECM; Hollich et al., 2000), takes a hybrid approach emphasizing how word learning starts out as an associative process and gradually becomes a process reliant on social and linguistic information.

• Language development starts before birth and is evident shortly thereafter (e.g., attention to prosody)

• At the inception of language acquisition, infants must discriminate between individual sounds (e.g., “pot” vs. “dot”); find larger units such as individual syllables and words through the process of segmentation; and find the meaning these units contain.

• Infant-directed speech is not bad for infants; rather it contributes to their early language learning.

• Infants initially appear specially equipped to learn any language of the world, as demonstrated by their ability to discriminate between speech sounds from a variety of languages

• Infants easily learn more than one language and bilingualism is the norm rather than the exception.

• Infants have been shown to compute “transitional probabilities” between syllables to find those that co-occur frequently in the stream of speech. This contributes to infants’ ability to segment the language stream.

• Verbs are generally more difficult for children to learn than nouns.
By combining words and sounds in a grammatical way, speakers can progress beyond simplistic telegraphic speech (e.g., “apple”) and begin talking about relations *between* referents (e.g., “She ate the apple”).
“How Ann Salisbury can claim that Pam Dawber's anger at not receiving her fair share of acclaim for *Mork and Mindy*'s success derives from a fragile ego escapes me” (cited in Gleitman, 1981).

I. **Introduction**

Our human language ability is impressive, as evidenced by the reader’s ease in understanding the sentence above. How do infants come to understand sentences like these and numerous others they have never seen before? In this chapter, we explore the complex process of language development, starting with the assertion that the task is more difficult than it first appears. We review evidence demonstrating that infants begin the journey toward language before birth, and then discuss the various tools used as they begin to break up the speech stream. The sections that follow describe word learning, grammatical development, and the pragmatics of language. The penultimate portion of this chapter addresses such topics as bilingualism, language in children with autism and the hearing impaired, and the creation of new languages. We propose a theory of language acquisition that unifies evidence from perceptual, social, and linguistic research: The Emergentist Coalition Model (Hollich, Hirsh-Pasek, & Golinkoff, 2000), and discuss future directions in the field of language acquisition research.

The field of language development is rich and varied; although this chapter will give a broad overview of child language, and introduce both classic and new methods and theories in child language research, we unfortunately lack the space to describe each and every landmark study that has been conducted. Elsewhere, however, we have discussed how the advent of various methods changed the field of language acquisition dramatically (Golinkoff & Hirsh-Pasek, 2011). In addition, our examples are often drawn from contemporary, rather than classic
research. This approach allows the reader to delve into the reference sections of recent articles and hone in on classic research as desired.

II. First, what is language?

Language is an intricate system of sounds or gestured symbols that people use to communicate with one another. Children typically begin to use either a spoken or signed system in their first year or shortly thereafter, and begin to understand individual symbols even earlier. Although language learning seems to come quite naturally to most children, it is not a simple task. To become competent speakers of their ambient language, children must master the phonemes (basic sounds in spoken language or basic gestures in a signed language), semantics (meaning), grammar (language structure; e.g., the way gerunds and particles are combined into sentences), and pragmatics (how language is used in social contexts) of that language. To complicate the task, the world’s languages are diverse, using different components of sound to signal meaning. For example, the Chinese word /bai/ can mean “split,” “white,” “swing,” or “defeat,” depending upon how the voice pitch in vowels is varied (Howie, 1976; Luo et al., 2006). Yet other languages are highly inflected, building up meaning by adding inflections (morphology) to a single word rather than using separate words for separate semantic components (Aksu-Koç & Slobin, 1985; MacWhinney & Pléh, 1997). In an agglutinative language like Turkish, for instance, word structure is formed by adding morphemes (suffixes) to root words. To express “The things that are on the bookcase,” four inflections are added to the root kitap (book) to generate a single word kitaplıktakiler (kitap-lik-ta-ki-ler; Tilbe Göksun, personal communication). Other languages using less morphology, such as English, may rely more on word order to convey meaning: Brutus killed Caesar preserves history while Caesar killed Brutus destroys it.
Given this diversity of language structure, how do infants find units (like phonemes, morphemes and words) and how do they figure out how these units combine in the rules of their native language? Over the years, at least five types of language acquisition models have been proposed to explain how children accomplish this task: Behaviorist, innate, cognitive, interactionist (social), and most recently, Bayesian. We can think of these theories as five large branches of a tree, with families of related theories springing from each branch. Although moderate forms of these different theories are not mutually exclusive, and many contemporary theorists draw upon various aspects of each when explaining the course of child language development (e.g., Hollich et al., 2000), there are some fundamental differences in how language acquisition is conceptualized across theories.

According to a behaviorist theory of language acquisition, children start from scratch. They hear language in the world, and acquire language (both words and grammar) based on imitation and experience alone (via positive and negative reinforcement; Skinner, 1957). It has been argued, however, that the input children receive is insufficient to explain how they develop fully generative and grammatical language in such a short period of time (the poverty of the stimulus argument; Chomsky, 1957). There are few adherents of pure behaviorist theories of language acquisition, although an emphasis on frequency and reliance on general purpose learning mechanisms like association can be seen in a number of contemporary connectionist models of language acquisition (e.g., Colunga & Smith, 2005; Elman, 2006, 2008; Elman et al., 1996; Smith, 2000).

The second primary theory of language development is equally extreme, anchored in the ideas of Noam Chomsky (1957), who argued that children must have an innate faculty for language acquisition. This innateness was captured in a metaphorical “Language Acquisition
Device” that children possess at birth, providing them with a set of universal principles of language that are triggered by minimal language experience (for an in-depth discussion of Universal Grammar, see Roeper, 2007). Slobin later modified this position arguing that rather than possessing substantive rules of language at birth, children are born with a set of procedures and inferences that allow them to learn any language in the world (Slobin, 1970).

The cognitive theory of language acquisition was proposed by Jean Piaget, who held that language develops in concert with general cognitive development (Piaget, 1926) and is but another manifestation of the symbolic function children construct in the second and third year of life. On this view, children’s use of language is part and parcel of how they think about the world. For example, if children have not learned the concept “seriation,” their use of comparative words such as bigger or smaller will reveal only partial understanding.

Some, known as interactionist or social theorists, suggest that language exists for the purpose of communication and can therefore be learned only in the context of interpersonal interaction (Levy & Nelson, 1994; Nelson, 2007; Ratner & Bruner, 1978; Tomasello, 2000, 2003). According to this theoretical viewpoint, the language behaviors of adults (e.g., infant directed speech, or IDS, discussed later in this chapter), and the exchange games that children and adults play are specially adapted to support the acquisition process.

Tenenbaum, Xu, and colleagues have recently proposed a Bayesian model of word learning (Frank, Goodman, & Tenenbaum, 2009; Perfors, Tenenbaum, Griffiths, & Xu, 2011). On this theory, children infer the meaning of a word based not on a process of elimination (deductive reasoning) or by association, but rather by a system that rates each possible meaning of a word according to its statistical probability of being correct. Bayesian theory purports to account for a number of conundrums in word learning, including children’s ability to learn a
word based on very few positive exemplars (e.g., a house cat, a toy cat, a cat on TV) without the necessity of negative exemplars (e.g., “That animal (a turtle) is not a cat”; Xu & Tenenbaum, 2007a, 2007b).

This chapter will review the language development literature from all standpoints, examining evidence for how infants and toddlers decipher what was once called “a blooming, buzzing confusion” (James, 1890, p. 492) of sounds and experiences to acquire the richness and intricacies present in all the languages of the world.

III. In the beginning…

By the third trimester, fetuses attend to prelinguistic fundamentals like consonant-vowel patterns and tonality (Lecanuet et al., 1988, 1989), and remarkably, demonstrate experience-related preferences – that is, they show evidence of learning about auditory stimuli. After hearing the muffled sound of their mother’s voice through amniotic fluid near the end of gestation, fetuses react differently to a recording of a stranger’s voice than to a recording of their own mother’s voice. In 2003, Kisilevsky and colleagues explored fetal voice recognition by exposing 60 full-term fetuses (average gestational age: 38.4 weeks) to a 2-minute recording of either their mother’s voice or a stranger’s voice reading a poem. By measuring fetal heart rate before, during, and after the reading, researchers could determine whether infants responded differently to each type of stimuli. Results revealed that infants’ heart rates reliably increased when hearing a recording of their mother’s voice read a poem and reliably decreased when hearing a stranger read the same poem. Thus, infants learn something about human voices prior to birth (Kisilevsky et al., 2003, 2009; see also Smith, Dmochowski, Muir, & Kisilevsky, 2007). Recently, researchers found a similar effect using musical melodies, which may be analogous to the prosody of human voices. Women in their third trimester of pregnancy played a recording of
a descending piano contour to their fetuses for 3 weeks. After the infants were born, researchers found greater heart rate deceleration in response to the familiar contour than to a matched ascending melody (Granier-Deferre, Bassereau, Ribeiro, Jacquet, & DeCasper, 2011). This finding is noteworthy because it suggests that fetuses are not only sensitive to aspects of language prior to birth (such as prosodic contour), but that they retain memories of prenatal experiences up to a month after they are born.

What about the “language” abilities of neonates? Directly after birth, infant behavior suggests that they remember the voices and auditory stimuli heard during gestation. Not only do newborns prefer to hear their mother’s voice over a stranger’s voice immediately after birth (DeCasper & Fifer, 1980), infants also recall particular auditory stimuli heard during gestation. In a classic study of infant memory, researchers measured sucking behavior in response to a speech passage heard in utero (familiar stimuli) versus a novel passage (DeCasper & Spence, 1986). Results revealed that infants preferred listening to the familiar stimuli, although they were not yet born when the speech was heard.

IV. Phoneme discrimination and segmentation of the speech stream

At the inception of language acquisition proper, infants must accomplish at least three tasks. First, they must discriminate individual sounds, such as the sounds that distinguish the words “pot” and “dot”. These sounds are called phonemes. Second, they must use these sounds as the basis for discriminating and segmenting individual syllables and words. Finally, after working on finding the smaller units of their language, children begin to use words to assemble phrases and clauses.

When and how do infants segment floods of sounds into individual units (e.g., phonemes, syllables, words, clauses, phrases), and how do they decide where one unit ends and another unit
begins? Take the case of phonemes. Research suggests that the same phoneme sounds different when couched in different words. This effect, known as coarticulation, belies the naive belief that speech sounds are organized “like beads on a string”. Instead, speech sounds run together during fluent discourse, causing the same phoneme to sound differently depending upon the sounds that come before or after it in a word or phrase. In addition to these issues with coarticulation, there is much variation in how a single phoneme is produced by the same speaker in different contexts (think of your speech when speaking to a baby versus hollering at a ball game), and the same phoneme sounds very different when produced by Dolly Parton versus James Earl Jones (Kuhl, 1979). On top of this, languages do not all employ the same set of phonemes – some phonemes typical of the Tlingit and Hindi languages simply do not exist in English. Variety in the sound of a single phoneme (depending on phoneme location, speaker, rate of speech, etc.) combined with the need to acquire language-specific phoneme repertoires forces infants to rely on multiple cues to discriminate and segment the auditory building blocks of their native language.

What types of information are available to infants to help them discriminate and segment the speech stream? Infants have multiple tools at their disposal: They hear language in a special speech register produced by caregivers around the world when speaking to infants (infant directed speech or IDS), they are born with language-friendly perceptual abilities, they naturally attend to prosodic variability, they are capable of sophisticated statistical analyses, and they harness their knowledge of familiar words like Mommy to learn new words.

*Baby talk is good: Infant-directed speech as a scaffold*

Infants are born with a preference for listening to language over other auditory stimuli, as demonstrated by the finding that newborns prefer listening to human language over an artificial
“language” that mimics some of its properties (Vouloumanos & Werker, 2007) and 3-month-olds prefer human speech over other naturally occurring sounds like macaque vocalizations (Shultz & Vouloumanos, 2010). The same is true in the visual realm: hearing 6-month-olds prefer to watch sign language rather than non-linguistic gesture (Krentz & Corina, 2008). This may suggest that infants are primed, “ready and willing” to learn language in the verbal or gestural modalities – and caregivers typically encourage acquisition by modifying their interactions with infants to facilitate language learning. Originally termed “motherese” (Newport, 1975), this type of modification is now often referred to as child-directed speech (CDS) or infant-directed speech (IDS; for a discussion of infant-directed gesture or “motionese”, see Brand, Baldwin, & Ashburn, 2002; Brand & Shallcross, 2008; Brand, Shallcross, Sabatos, & Massie, 2007). IDS is found in language communities throughout the world, and is used by mothers, fathers, grandparents, aunts and uncles, brothers and sisters, unrelated caregivers, interested bystanders – and even 4-year-olds when they are speaking to 2-year-olds (Shatz & Gelman, 1973). IDS is attractive to children – infants as young as newborn to 4 months of age demonstrate a robust preference for listening to IDS over adult-directed speech (Cooper & Aslin, 1990; Fernald, 1985).

What are the characteristics of IDS? When speaking with children, speakers often modify word order (Narasimhan & Dimroth, 2008), exaggerate semantically meaningful tones (Liu, Tsao, & Kuhl, 2007), speak slowly in shorter phrases, use simpler words with fewer syllables per phrase, speak in higher than average registers, and exaggerate intonation contours by stretching vowels (Fernald, 1992; Fernald & Mazzie, 1991; Fernald & Simon, 1984; Gleitman, Newport, & Gleitman, 1984; Grieser & Kuhl, 1988).
Hearing IDS has been linked to a number of positive outcomes for children, including a larger vocabulary (Gleitman et al., 1984; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Masur, 1982; Rowe, 2008; Tomasello, 1988), which may be due to increased brain activity in response to IDS (Zangl & Mills, 2007), better word memory and categorization skills when hearing novel words in IDS (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 study of Korean-speaking mothers interacting with adults and infants found that mothers speaking to their infants changed their speech in a number of ways that helped children learn the phonological regularities of their language. When speaking with infants, mothers were more likely to use sounds that infants were able to produce themselves and to highlight perceptually salient phonemes and sound patterns (Lee, Davis, & MacNeilage, 2008). Similar results have been found in English, Japanese (Werker et al., 2007), Russian, and Swedish (Kuhl et al., 1997). This type of modification offers children a foothold into the world of words and grammar (Fisher & Tokura, 1996a, 1996b).

In addition to correlations with vocabulary outcome, a recent experimental word-learning study conducted by Ma and colleagues directly compared word learning in IDS and adult-directed speech (ADS; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011). Results revealed that 21-month-olds with small vocabularies relative to their peers did not learn novel words when they were presented in ADS, but did learn them when they words were taught using IDS. Twenty-one-month-olds with larger vocabularies, on the other hand, were equally likely to learn a word in ADS or IDS. Testing 27-month-old children in the same paradigm, the researchers found that they could learn words in both the IDS and ADS conditions. These findings suggest that IDS supports early word learning – especially in toddlers with smaller vocabularies. Ma et al. is one
of the first studies to explore the effect of IDS on vocabulary development within an experimental (rather than correlational) paradigm.

**Hardware included: Infants have language-friendly perceptual abilities**

Child-directed speech helps infants begin to understand language, but is only one component of a complex process that allows children to break up the stream of sounds in their environment. Infants are not blank slates at the inception of language learning and are undoubtedly influenced by eavesdropping on maternal language while in the womb. However, they also appear to possess hardwired perceptual abilities that facilitate phoneme discrimination. For example, infants show evidence of differential responding to categories of consonant sounds at birth (e.g., they react differently to the syllables “ba” and “ga”; Bertoncini, Bijeljac-Babic, Blumstein, & Mehler, 1987), and show categorical responding to different vowel sounds by 6 months of age (Kuhl, 1979). This suggests that infants’ ability to discriminate between sounds emerges concurrently with turning their attention to the problem of segmentation (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005).

The sounds of the world’s languages (from clicking in Khoisan languages to tonal variation in Mandarin Chinese) are remarkably diverse, but infants learn them all. How? Are infants born in Spain inherently prepared to perceive and produce the rolled *rrr*, whereas the infants of Namibia are ready to produce the click consonants of the Khoisan languages?

A large body of research suggests that infants are born sensitive to all of the phonemes in the world’s verbal and sign languages (such as ASL, Tlingit, Swahili, Mandarin, and German; Baker, Golinkoff, & Petitto, 2006; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Jusczyk, 1997; Kuhl, 1993; Streeter, 1976). With increased exposure to a primary language, however, their ability to discriminate non-native phonemes is reduced (Best, McRoberts, Lafleur, & Silver-
Isenstadt, 1995; Kuhl, 2009; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 2005) while they simultaneously improve their discrimination of sounds in their native language (Kuhl, 2009; Kuhl, Tsao, Liu, Zang, & De Boer, 2001). This process happens fairly quickly; evidence suggests that infants narrow in on the phonemic distinctions relevant to their ambient language(s) by approximately 10-12 months (Newman et al., 2006; Tsao, Liu, & Kuhl, 2004) – whether they are learning oral language or sign language (Palmer, Fais, Golinkoff, & Werker, in press).

However, not all infants attend to phonemic distinctions in the same way. For example, better discrimination of the phonemes present in one’s native language at 7 months of age is associated with improved language outcomes (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005), but better non-native phonemic discrimination at 7 months is associated with reduced later language abilities (Kuhl et al., 2005). This suggests that honing in on the correct phonemes for the ambient language and reducing attention to non-native phoneme distinctions results in better language over time (see also Kuhl et al., 2008).

This ability to discriminate non-native phonemes is not completely lost in the first few months of life, however (Best et al., 2001). If 9-month-old infants are exposed to a non-ambient language in a social context for even a brief period of time (e.g., 12 laboratory sessions), the declines that are typically observed can be resuscitated (Kuhl, Tsao, & Liu, 2003). The native-language specialization that occurs over time continues through toddlerhood, as children begin to interpret sounds according to the dictates of their native language by the middle of the second year. For example, toddlers reared in Dutch-speaking environments treat elongated vowels as lexically contrastive (consistent with Dutch) but English-reared children do not (consistent with English; Dietrich, Swingley, & Werker, 2007). Similarly, children learning Mandarin Chinese correctly interpret tonal changes as semantically meaningful whereas English-learning children
do not. Early perception and later language ability in bi- and multilingual children is discussed later in this chapter.

*Semantically meaningful tones*

In addition to the attention-getting prosodic contours of IDS, infants use other types of auditory input to break into the world of language. Infants attend to music and tonality (Saffran, 2003; Saffran, Loman & Robertson, 2000) which is especially important given that many languages of the world use variations in tone to express meaning (Han & Kim, 1974; Gandour, Potisuk, & Dechongkit, 1994; Laniran & Clements, 2003). In Somali, for example, pitch or pitch contour is used to distinguish the meanings of two words that are otherwise phonologically identical. Other languages that use meaningful tones are Mandarin, Navajo, Yoruba, Vietnamese, and Yucatec.

Semantically meaningful tones are produced in different ways by different speakers (Zemlin, 1988), and multiple tones overlap during fluent speech (Xu, 1997). How do infants organize the range of tones they hear and manage to group them into categories? One way to explore this question is through the use of computational modeling, exposing simulated neural networks to a range of stimuli and observing the results of self-organization based on patterns of hierarchical similarities (Chater & Christiansen, 2008). Gauthier and colleagues (2007) used this method to explore how the infant brain captures and categorizes tones. A self-organizing neural network program was given multiple examples of four semantically meaningful tones that are present in Mandarin (Xu, 1997). The tones were produced by speakers with varying fundamental frequencies (i.e., a higher or lower vocal register) and varying pitch contour extrema. At test, it was revealed that the program had organized the input into categories of sound, and that those
categories corresponded to the four semantically meaningful tonal categories of spoken Mandarin (Gauthier, Shi, & Xu, 2007).

**What goes together: Audio-visual synchrony, speech and perceptual-motor activity**

Infants know how to use their eyes and their ears at the same time, and are able to demonstrate this ability in language-learning contexts by 8 weeks of age (Kuhl & Meltzoff, 1982, 1984; Patterson & Werker, 2003). For example, Patterson and Werker (2003) found that 2-month-olds can match auditory vowel sounds to visual face information. By tracking correlations between what they see and what they hear, infants pick up cues that help them to segment the speech stream – even in noisy situations such as when the TV is on or siblings are playing nearby. Research by Hollich and colleagues (2005) revealed that 7.5-month-olds were able to pick out target words from one of two simultaneous (and equally loud) speech streams if, and only if, the target audio was accompanied by a correlated visual display (e.g., a person speaking or an oscillation synchronized to the speech stream). The authors argued that audio-visual correlation (not limited to faces) serves to direct the attention of very young infants to certain sounds, which helps them to segment the speech stream in a noisy context (Hollich, Newman, & Jusczyk, 2005).

The importance of audio-visual correlations to the acquisition of language has likewise been found using neuroimaging methods. For example, a recent study of brain activation in response to speech and non-speech sounds revealed a link between speech perception and motor representations in infancy. Using magnetoencephalography (MEG) to measure responses in different areas of the brain, Imada and colleagues (2006) found a developmental pattern wherein initially, the newborn brain demonstrated an all-purpose activation to any and all sounds, regardless of whether they were syllables, pseudo-speech, or non-speech sounds like music. Six
to 12 months later, however, speech sounds activated an area specially used for perceptual-motor processing (Imada, Zhang, Cheour, Taulu, Ahonen, & Kuhl, 2006). This finding supports a link between speech and perceptual-motor representations in the brain that may be related to the behavioral finding that infants use audio-visual synchrony to segment words from the speech stream.

*All stressed out: Using stress patterns to segment*

Infants are attuned to stress patterns (like the difference between TRICK-y and re-SERVE). By 7.5 and 13.5 months respectively, English-learning infants can use the predominant strong-weak pattern of stress in English to pick nouns (Jusczyk, Houston, & Newsome, 1999) as well as verbs (Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005) out of a fluent speech stream (this is called the trochaic bias). In a language that is *not* characterized by a predominant strong-weak pattern of stress (for example, Canadian French has an iambic (weak-strong) structure), infants segment only those words consistent with the most common stress pattern of their ambient language (Polka, Sundara, & Blue, 2002). Thus, by approximately 9 months of age, infants develop language-specific rhythmic biases that facilitate word segmentation.

*Baby statisticians and rule learners*

Infants demonstrate more than language-friendly perceptual abilities and heightened attention to speech. In fact, infants are remarkable statisticians, and appear to learn rules quite well (Johnson, Fernandes, & Frank, 2009; Marcus, Vijayan, Rao, & Vishton, 1999). The structure of language is conducive to statistical analyses (Christiansen, Onnis, & Hockema, 2009; Mintz, 2006; Onnis & Christiansen, 2008; Soderstrom, Conwell, Feldman, & Morgan, 2009), and significant research suggests that infants exploit these structural regularities when
decoding the speech stream. One groundbreaking study of infant statistical abilities revealed that infants are able to decipher the speech stream by recognizing that sound combinations occurring within words occur more frequently than sound combinations between words (Saffran, Aslin, & Newport, 1996). Using a familiarization-preference procedure (developed by Jusczyk & Aslin, 1995), Saffran and colleagues (1996) familiarized eight-month-old infants with 2 minutes of monotone female speech. The speech stream was a simplified artificial language containing consonant-vowel syllables (e.g., bubidakupadotitabidakubupati), some of which co-occurred 100% of the time (e.g., bida – analogous to a combination of sounds within a word in normal speech, and part of the nonsense word bidaku), and some that occurred together less frequently – around 30% of the time (e.g., kupa – analogous to the sounds that occasionally occur together during the transition between words). For example, in “silly puppy”, the sounds “si” and “ly” have a higher transitional probability than the sounds “ly” and “pu”. A stringent second phase of the study challenged infants by presenting them with nonsense words and nonsense part words from the familiarization stream. Part words were formed by including the last syllable of one word with the first two syllables of another. In this difficult task, infants demonstrated that they had learned the nonsense words by increasing their attention to the part words – indicating that the part words were novel and the words were expected (Saffran et al., 1996). In sum, by capitalizing on the transitional probabilities of phonemes in a speech stream and segmenting those syllable combinations that have a high probability of co-occurring, infants are able to hypothesize which phonemes combine to form words and which do not.

Is the ability to track transitional probabilities specific to language learning? A powerful series of studies suggests that the ability to track transitional probabilities is domain-general (i.e., not limited to linguistic stimuli). In one study, 7.5-month-olds were exposed to the same
statistical learning task used by Saffran et al. (1996) with one difference: consonant-vowel syllables were replaced with musical tones. Results revealed that infants were able to track “tone words” as effectively as they track syllable-words (Saffran, Johnson, Aslin, & Newport, 1999).

Evidence from visual modality likewise suggests that infants’ statistical learning abilities are domain-general. For example, one study familiarized 5- and 8-month-olds to a series of colored shapes in a manner analogous to the original Saffran et al. (1996) statistical learning task. At test, infants differentially responded to trials wherein the statistical probability that certain shapes would follow other shapes had changed from what they had seen during training (Kirkham, Slemmer, & Johnson, 2002). Recently, evidence of statistical learning to isolate events (distinct hand motions) has been shown in 7- to 9-month-old infants (Roseberry, Richie, Hirsh-Pasek, Golinkoff, & Shipley, 2011).

In addition to being a domain-general skill, statistical learning ability does not appear to be unique to humans. Hauser and colleagues replicated the results obtained by Saffran et al. (1996) in the cotton-top tamarin, a non-human primate (Hauser, Newport, & Aslin, 2001), although a more difficult test of grammatical structure learning was accomplished only by infants (Saffran et al., 2008). Furthermore, experiments using natural language (rather than a simplified artificial language consisting of nonsense words as in Saffran et al., 1996) have thus far been conducted only with baby humans (Pelucchi, Hay, & Saffran, 2009).

The ability to track transitional probabilities helps infants segment the sound stream into words, with the outcome of this process feeding into word learning in both artificial and natural languages (Graf Estes, Evans, Alibali, & Saffran, 2007; Hay, Pelucchi, Graf Estes, & Saffran, 2011). In one recent study, 17-month-old English-reared infants were exposed to a novel natural language speech stream (Italian) consisting of words with high or low internal transitional
probabilities. When infants were given the opportunity to then map those words onto referents, they were more likely to map words with high transitional probabilities – suggesting that experience with statistical regularity of sounds in a given language plays a role in acquiring word meaning (Hay et al., 2011).

Infants segment the sound stream into words using statistical probabilities, and use those probabilities to map words to individual referents in the world, but what happens when infants need to learn rules that dictate how to put words together? Learning rules for how to properly order words and segment word sequences is a critical step toward language mastery and the use of grammar. However, the task of learning rules is different from tracking transitional probabilities because the latter does not require children to extend their knowledge to new exemplars. That is, in statistical learning procedures, there is a one-to-one mapping between sounds heard (or pictures seen) during familiarization and sounds heard (or pictures seen) at test. In contrast, when infants learn rules for “how things go together” they must be able to generalize an abstract pattern (rule) from one set of exemplars and apply it to another.

The apparently more difficult task of abstracting and extending grammatical rules is accomplished by 7-month-olds in the same way that they accomplish the (relatively) simple task of tracking transitional probabilities between syllables, sounds, or shapes. In 1999, Marcus and colleagues familiarized infants to a sequence of words that followed the rules of an artificial “grammar”. In one condition, syllables appeared in ABA order (e.g., ga ti ga or li na li). In another, they appeared in ABB order (e.g., ga ti ti or li na na). At test, infants heard novel syllables organized in a way that was either consistent or inconsistent with the “grammar” they heard during familiarization (e.g., wo fe fe or wo fe wo). Results revealed that infants preferred to listen to the audio featuring unfamiliar grammar. Thus, infants abstracted a grammatical rule
during familiarization (such as ABA) and were subsequently more interested in hearing syllables organized according to a new rule (such as ABB) at test (Marcus et al., 1999).

Is this algebraic rule-learning specific to language, or is it a domain-general ability like statistical learning? Early studies suggested that infants were only able to learn artificial “grammars” when presented as language (Marcus, Johnson, & Fernandes, 2004). However, subsequent research using category-matched stimuli revealed that 7-month-olds can learn “picture grammar” in the visual domain (Saffran, Pollak, Seibel, & Shkolnik, 2007) and can generalize grammatical rules learned in a speech task to tones, timbres, and animal sounds (Marcus, Fernandes, & Johnson, 2007).

**Familiar words facilitate segmentation**

Children possess a number of perceptual processing tools available to help them segment the speech stream into individual phonemes and words from the bottom up (such as tracking statistical probabilities), but there is also evidence to suggest that they can use top-down lexical information. In fact, lexical knowledge as simple as a child’s own name might provide a useful “wedge” into word segmentation. By 4.5 months of age, infants are attuned to the sound pattern of their own name. Six weeks later, they are able to pick their name out of fluent speech (Mandel, Jusczyk, & Pisoni, 1995). Bortfeld and colleagues (2005) demonstrated that 6-month-olds use knowledge of their own name to take a top-down approach to speech segmentation by showing that infants can segment a novel word from a fluent speech stream when the word is presented after a familiar name (e.g., *Mommy*, or *baby’s name*), but not after an unfamiliar name (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). In addition to the child’s own name or the word *Mommy*, infants can use their knowledge of familiar, frequently occurring function words such as *the*, *her*, and *its* to help define the boundaries between words (see later section on closed-
class words; Shi, Cutler, Werker & Cruickshank, 2006; Shi & LePage, 2008; Shi, Werker, & Cutler, 2006). This indicates that infants have many tools (top-down and bottom-up) at their disposal when cracking the language code. In fact, recent research by Mersad and Nazzi (in press) suggests that infants can use both top-down and bottom-up cues in combination by 8 months of age. When exposed to an artificial language containing words with non-uniform lengths, infants were able to segment artificial words as long as they heard the familiar word “mama” as part of the sound stream (Mersad & Nazzi, in press).

Which words go together?

To become competent speakers, infants must not only learn which parts of fluent speech streams form words, but also which words combine to form larger units of analysis: phrases (e.g., the blue bunny), and clauses (e.g., the blue bunny ate rice). Phrases and clauses are grammatical chunks composed of multiple individual words that are dependent upon one another to express complex meaning.

A number of studies suggest that infants as young as 7 months of age are attuned to natural clause and phrase boundaries. For example, Hirsh-Pasek and colleagues (1987) used a head-turn preference procedure to study whether young infants preferred listening to mothers speaking with pauses inserted either mid-clause (interruption) or post-clause (natural end of the grammatical unit). Results revealed that infants were sensitive to pause placement and preferred listening to pauses that occurred after grammatical units, rather than the middle of those units (Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright-Cassidy, Druss, & Kennedy, 1987). A subsequent study revealed that IDS helps 7- to 9.5-month-olds identify clause boundaries within fluent speech. Using the same paradigm employed by Hirsh-Pasek et al. (1987), infants heard pauses inserted either at the beginning or the middle of a clause spoken in IDS or adult-directed
speech. Only when the audio was heard in IDS did infants show evidence of segmenting fluent speech according to clausal boundaries. The authors concluded, “prosodic qualities of motherese provided Ss with cues to units of speech that corresponded to grammatical units of language” (Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; for a review, see Jusczyk, Hirsh-Pasek, Kemler-Nelson, Kennedy, Woodward, & Piwoz, 1992).

The ability to determine where grammatical units begin and end, coupled with a preference for listening to grammatically intact clauses appears to lead to recognition and segmentation of word sequences – but only when they are heard in the context of grammatically intact clauses (Soderstrom, Nelson, & Jusczyk, 2005). Thus, prosodic consistency and grammatical correctness help children segment multi-word sequences while those same words heard in the context of interrupted clauses are not learned (see Soderstrom & Morgan, 2007, for why this might be the case).

Multi-word word sequences like verb phrases (e.g., she ate) and noun phrases (e.g., the carrot) can be combined to form clauses (e.g., she ate the carrot). What cues do infants use to segment clauses into these smaller grammatical units? One study found that 6- and 9-month-olds familiarized with word sequences preferred listening to those sequences only when they formed a grammatical unit (i.e., a noun phrase) but not when they were presented as a syntactic non-unit (Nazzi, Kemler Nelson, Juczyk, & Juczyk, 2000). A second study demonstrated that 6-month-olds preferred listening to word sequences presented as a phrase and accompanied by a number of prosodic cues to phrase boundaries (Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003). These results suggest that infants are sensitive to smaller grammatical units than clauses (i.e., phrases) and that a confluence of cues, such as prosody, facilitate segmentation.

V. Words found: But what do they mean?
The ability to segment words, phrases, and clauses from a speech stream doesn’t buy infants much in terms of communication unless they know what the words and combinations of words mean – i.e., unless they have a semantic system in place. Semantics is the study of meaning in language – both of individual units and of larger groupings of units. A semantic unit can be a word such as “dog” or the “s” that turns the word for one dog into the word for two dogs. It can also be a noun phrase like “the dog” or a sentence like “the dog slept peacefully.”

When and how do infants build up a semantic system? First, they demonstrate word comprehension. Experimental studies using the head-turn preference procedure and the Intermodal Preferential Looking paradigm (Golinkoff, & Hirsh-Pasek, Gordon, & Cauley, 1987; Hirsh-Pasek & Golinkoff, 1996), revealed that one of the first words infants differentially respond to is their own name. At five months, infants prefer to listen to a voice saying their name rather than a stress-matched word (Mandel et al., 1995) and are able to identify the referents of frequent words like “Mommy” (Tincoff & Jusczyk, 1999). By 8 months, infants demonstrate some understanding of phrases such as “don’t touch” or “give kisses.” However, there are significant differences between an infant’s ability to comprehend words and the ability to produce them. For example, the average 10-month-old’s receptive vocabulary ranges from 11 and 154 words and their expressive vocabulary is nonexistent. This and similar production-comprehension data from 16-month-olds suggests that early word comprehension far exceeds early word production (Fenson, 1994). By around 18 months of age, when children have laboriously acquired approximately 50 words used in a mostly telegraphic way (that is, without grammar; Braine, 1963, Brown, 1973; Grégoire, 1937), some argue that they experience a “word spurt” and begin learning words very rapidly – averaging 8-37 new words per month (Benedict, 1979; Gopnik & Meltzoff, 1987; but see Bloom, P., 2000; Goldfield & Reznick, 1996).
How shall we think about word learning from a theoretical point of view? One theory, the Emergentist Coalition Model (ECM; Hollich et al., 2000), takes a hybrid approach emphasizing how word learning starts out as an associative process (Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006) and gradually becomes a process reliant on social and linguistic information (Golinkoff & Hirsh-Pasek, 2006).

The ECM posits that children employ a confluence of perceptual, social, and linguistic cues to determine the referent of a novel word, and that they weight those cues differentially over developmental time (Hollich et al., 2000; see also Lavin, Hall, & Waxman, 2006). According to this theory, infants begin the word learning task with a heavy perceptual bias; their word learning is contextually bound regardless of word type (e.g., noun, verb, etc.; Behrend & Scofield, 2006; Forbes & Farrar, 1993, Forbes & Poulin-DuBois, 1997; Smiley & Huttenlocher, 1995). By the second year of life, however, they begin to use 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).

Long-debated questions about the mechanisms behind early word acquisition, the reasons why children acquire the words they do, and why some words are learned earlier than others are addressed by the ECM. For example, the ECM holds that nouns are easier for children to acquire because the majority of words in that class are more perceptually accessible than the majority of verbs, even in a language like Mandarin, where the noun-verb discrepancy is less pronounced than in English (Ma, Golinkoff, Hirsh-Pasek, McDonough & Tardif, 2009). Further, basic
perceptual processes such as statistical tracking of cross-situational regularities seem to help infants link nouns to their likely referents (Smith & Yu; 2008), which may also contribute to verb learning (Scott & Fisher, 2011). Later in developmental time, infant word-learners incorporate an understanding of social and linguistic cues into these basic perceptual learning strategies.

*What are children’s first words?*

To answer this question, it is important to note that within the context of language-learning, all words are not created equal. Large studies of early vocabulary composition suggest that certain types of words are more prevalent than others in the productive lexicons of infants and toddlers. One cross-linguistic analysis of vocabulary development revealed that *open class* words such as nouns, verbs, and adjectives are produced much more frequently than *closed class* words like prepositions, determiners, and pronouns – and that this relationship holds for children learning Spanish, Dutch, French, Hebrew, Italian, Korean, and American English (Bornstein et al., 2004). Despite the paucity of closed class words in productive vocabulary, however, laboratory research suggests that children are sensitive to closed class words from an early age (Gerken, 1996; Gerken, Landau, & Remez, 1990; Shi et al., 2006; Shipley, Smith, & Gleitman, 1969). The following sections review research on how children come to understand the meaning of open class words (specifically nouns, verbs, and adjectives) as well as closed class words.

*Open class words*

*Nouns.* Nouns refer to people, places, things or ideas. Proper nouns refer to specific entities (e.g., Martha or Philadelphia), and other nouns refer to classes or categories of objects (e.g., people or city). While it is fairly easy for a child to associate the word “cup” with a single familiar sippy-cup and use the word as a proper name, the ability to attach one word to one
object is insufficient for developing full language. Forming *categories of objects* that can be labeled by a single word is necessary for parsimony; if each and every “cup” had a unique name, language would be hopelessly cluttered. On its face, learning words for categories of objects is deceptively simple; after all, how complex could it be to form a category of “cow” that includes all different cows? Closer examination, however, reveals that the process is far from straightforward. The problem begins with how to determine which entity is being referred to by a given word.

Picture this: a child rides along with her father in the car, passes a farm, and suddenly hears, “Look at the cow!” What might her father be referring to? The barn? The tree near the barn? The large spotty thing eating grass? The *spot* on the large spotty thing eating grass? The question of which part of an object or scene is being labeled when one hears a novel word is called the *indeterminacy of translation* problem (Quine, 1960), and represents a first stumbling block on the path to categorization. If, however, we *assume* that the child is able to pinpoint the cow in the field as the object being referred to, we still cannot conclude that she has formed the category “cow”. She has simply seen one example of a live black-and-white “cow” and may not generalize the word to cows that are brown or cows that are on television. Later that night, however, mother reads a book to the same child and says, “Look at the baby cow. Isn’t he cute?” Five minutes later, the child’s older brother gets scolded and mumbles to mom, “Don’t have a cow.”

The above examples of “cow” make clear that perceptual correlates of the word “cow” are variable (the child may hear “cow” when seeing cows on television, in pictures, or as toys, in addition to live animals). Despite this relatively wide perceptual variability, however, children
somehow manage to form the category “cow” and use the word referentially within the first few years of life. How do they accomplish this task?

Cues to noun meaning are plentiful, and infants capitalize fully on every available hint. For example, syntax, social context, perceptual salience, and comparison provide a number of well-researched cues that infants use to discover meaning and categorize words (Brown, 1957; Fisher, 1994; Gleitman, 1990; Gentner, Loewenstein, & Hung, 2007; Naigles, 1996; Pruden et al., 2006). Theorists have also posited and explored a number of lexical and pragmatic assumptions made by language-learners to guide their hypotheses about the potential meaning of words (Clark, 1995; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Hollich, Golinkoff, & Hirsh-Pasek, 2007; Markman, 1991; Mervis & Bertrand, 1994; Waxman & Markow, 1995).

A classic example of infants employing their knowledge of syntax to determine word meaning comes from Roger Brown’s study of preschool children (1957). Three- to five-year-old children were shown a picture of an action, a substance, and an object (e.g., person kneading dough in a bowl) and given a novel word (sib) in one of three syntactic contexts: object (a sib) action (sibbing) or material (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 dough (substance/material). In this case, preschoolers were able to make inferences about the meaning of a novel word simply by attending to the syntactic context in which it was heard (Berko, 1958; Brown, 1957; Subrahmanyam, Landau, & Gelman, 1999).

Social-pragmatic cues are an undeniably powerful cue to word meaning (Nelson, 2007; Tomasello, 2008). A significant body of 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 language outcome (Brooks & Meltzoff, 2008).

Building on an extensive literature of joint attention research (Adamson & Bakeman, 1991; Adamson, Bakeman, & Deckner, 2005; Adamson & MacArthur, 1995; Tomasello, 1995; Tomasello & Farrar, 1986), recent longitudinal findings illustrate the importance of attention to social-pragmatic cues to infants’ rate of vocabulary development. At age 10-11 months, Brooks and Meltzoff (2008) recorded how well 36 infants followed a person’s eye gaze and attended to an object with that person. Infant vocabulary was measured at 10-11, 14, 18, and 24 months. Growth curve modeling was used to determine the rate at which the children acquired vocabulary. Results revealed that infants who followed a speaker’s eye gaze and fixated on the same object for longer when they were younger had significantly larger vocabularies and a faster rate of vocabulary growth than infants who fixated for less time (Brooks & Meltzoff, 2008).

Infants have access to various social-pragmatic cues to word meaning, including emotional expression (Bloom, L., 1998, 2000; Moses, Baldwin, & Rosicky, 2001). In one study, 18-month-old infants heard an experimenter announce her intention to find a “toma” and observed her searching. An object was found along the way, but the experimenter’s expression of disappointment made it clear that it was not the desired object. A second object was found and the experimenter showed delight. At test, infants were able to identify the target object when asked for it by name, despite having heard the label at the beginning of the search and seeing a different novel object in the meantime. This suggests that 18-month-olds are able to learn a label for a novel object by discerning a speaker’s intention to find a particular object and recognizing the ensuing emotional reactions to success and failure (Tomasello, Strosberg, & Akhtar, 1996).
Some argue, however, that an appeal to social cues is unnecessary to explain children’s word learning, and that a more parsimonious account relies on the presence of general cognitive processes like attention, memory, learning, perception, and action (Christiansen & Dale, 2001; Colunga & Smith, 2008). This *attentional learning account* of language acquisition posits a correlational (connectionist) learning system with online attentional processes that work together with historical experience to determine word reference. According to Smith and colleagues, this process results in systematic biases of a computational nature that are used to guide subsequent word learning (Landau, Smith & Jones, 1988; Imai & Gentner, 1997; Jones & Smith, 2002; Smith, 2003; Soja, Carey & Spelke, 1991; Yoshida & Smith, 2003).

Pragmatic principled accounts of word learning, on the other hand, suggest that children hold certain *pragmatic assumptions* about the meaning of words they are learning. Eve Clark (1993, 1995, 2007) proposed the *principle of contrast*, which states that word-learners assume that new words refer to different (contrasting) forms than words that are already known. This principle is *not* the same as the principle of mutual exclusivity (discussed below) because according to the principle of contrast, the words *dog* and *daschund* can refer to the exact same animal. According to Clark, language-learners assume that if a single animal is referred to both as a *dog* and a *daschund*, the words must have different (non-competing) meanings. In this case, *dog* refers to the larger category in which *daschund* is a subcategory and *Charlie* is the individual exemplar. Clark further proposed the *principle of conventionality*, which states that word meaning is consistent and agreed upon by members of a language community (Diesendruck & Markson, 2001). This principle, in combination with the *principle of contrast*, prevents redundancy and helps early learners determine word meaning (Clark, 2007; Diesendruck, 2005).
Lexical principles that steer children’s word learning have likewise been proposed, two of which are the whole-object assumption (Behrend, 1990) and the assumption of mutual exclusivity, (Markman, 1991, 1994). When presented with a new label, lexically unprincipled word learners would be faced with an unmanageable plethora of hypotheses for word reference. However, when children approach the problem armed with constraints on possible word meaning, the task is simplified. In the cow example above, for instance, children operating under the whole-object assumption will not attach the label cow to a portion of the animal such as the horn or the hoof – rather, they will automatically assume that the word refers to the animal in its entirety. This assumption reduces the number of hypotheses for word meaning that children must consider, and has been shown to be robust. Although a majority of prior work on children’s use of the whole-object assumption focused on infants 18 months of age and older (Kersten & Smith, 2002; Saylor, Sabbagh, & Baldwin, 2002), a recent study revealed that 12-month-olds effectively utilize the whole-object assumption to guide word learning, even when faced with objects composed of salient, detachable parts (Hollich et al., 2007).

The assumption of mutual exclusivity states that children approach word learning with the expectation that different words mean different things. For example, young word-learners asked for a blicket will almost always attach the label to a novel object when it is presented alongside a familiar object that the children already have a name for, such as a shoe (Markman & Wachtel, 1988). The stability of this assumption was demonstrated in a recent study pitting social cues against the assumption of mutual exclusivity. Twenty-four 3- to 4-year-old children saw an experimenter point to (study 1) or look at (study 2) a familiar object that the child already had a label for (such as a crayon) while simultaneously asking for the blicket. Even when provided very obvious cues to a speaker’s intention that the child give the familiar object (conveyed by
such cues as looking at/pointing to the object while requesting the *blicket*), children nonetheless reached for a novel object and handed it to the experimenter (Jaswal & Hansen, 2006). This suggests that when given a new word, children are reticent to attach the new word to a familiar object when a novel referent is available. How does this work in the case of bilingual development, when children are necessarily given two names for a single object (e.g., *dog* and *perro*)? We explore this interesting question later in the chapter during our discussion of word learning in bilingual children.

**Verbs.** Children use many of the same cues described above to determine verb meaning. As with nouns, linguistic cues (Fisher, 2002; Fisher & Song, 2006; Golinkoff & Hirsh-Pasek, 1996; Naigles, 1996), social cues (Akhtar & Tomasello, 1996; Childers & Tomasello, 2006; Tomasello & Akhtar, 1995), perceptual cues (Brandone, Pence, Golinkoff, & Hirsh-Pasek, 2007; Maguire, 2004; Naigles & Kako, 1993), and various combinations of these cues differentially weighted over time (Hollich et al., 2000) are used to determine the referent of a verb (for a review, see Golinkoff & Hirsh-Pasek, 2008). The principles of contrast/conventionality, mutual exclusivity and whole-object reference have likewise been applied to learning words for actions (Clark, 2003, 2007; Golinkoff, Hirsh-Pasek, Mervis, Frawley, & Parrillo, 1995; Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996; Merriman, Evey-Burkey, Marazita, & Jarvis, 1996; Merriman, Marazita, & Jarvis, 1993).

However, the task of verb learning differs from noun learning in significant ways. Unlike nouns, which are mastered relatively quickly by toddlers (Markson & Bloom, 1997; Woodward, Markman, & Fitzsimmons, 1994), verbs are harder to learn (Gentner, 1982; Gillette, Gleitman, Gleitman, & Lederer, 1999; Golinkoff & Hirsh-Pasek, 2008) and are less often produced (even when they are comprehended; Goldfield, 2000). Infants across many languages have fewer verbs
than nouns in their early vocabularies (Bornstein et al., 2004), and although word-action
associations have been found in infants as young as 18 months of age (Casasola & Cohen, 2000),
many fail to demonstrate novel verb learning in laboratory tasks until much later (Imai et al.,
2008). It has been argued that verbs are harder to learn than nouns because the typical referent of
a verb is less concrete, imageable, individuable, and/or clearly shaped than the typical referent of
a noun, which makes it more difficult to map (Gentner & Boroditsky, 2001; Maguire, Hirsh-
Pasek, & Golinkoff, 2006; Parish-Morris, Pruden, Ma, Hirsh-Pasek, & Golinkoff, 2010). Indeed,
Ma and colleagues (2009) found that a word’s imageability, or the capacity of a word to call up a
mental image, is related to the age of acquisition of nouns and verbs in English and Chinese. As
expected, nouns had higher imageability ratings than verbs across both languages, which may
explain the predominance of nouns in young children’s vocabularies and is consistent with the
predictions of the ECM (i.e., that early word learning is driven by perceptual salience; Ma et al.,
2009; McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2011).

Given the challenges of learning a word that has a potentially ephemeral and dynamic
referent (Genter, 1982), children must fully utilize all available cues to verb meaning – and may
lean especially heavily on linguistic cues. The use of these cues in the service of word learning is
referred to as syntactic bootstrapping (Gleitman, 1990; Naigles, 1996; Naigles, Gleitman &
Gleitman, 1993). Syntactic bootstrapping occurs when children use their knowledge of abstract
linguistic structure (such as agent-patient relationships and argument number; Fisher &
Snedeker, 2002; Hirsh-Pasek & Golinkoff, 1996) to guide their interpretation of a novel verb,
and is employed by children learning diverse languages like Mandarin, Kannada, and French
A recent study of syntactic bootstrapping replicated and extended Golinkoff et al., (1996), and Naigles (1990) by exploring children’s use of word order as a cue to verb meaning (Gertner, Fisher, & Eisengart, 2006). Two-year-olds were shown a split screen with a bunny performing a novel action on a duck on one side and a duck performing a novel action on a bunny on the other side (see Figure 1).

**Figure 1.** Example of study stimuli (used with permission from Gertner et al., 2006).

Children heard audio containing a novel transitive verb (e.g., *look, the bunny is gorp*{ }*ing the duck!*). Gaze duration data revealed that toddlers looked longer to the side of the screen
depicting a scene consistent with the word order of the audio (e.g., the screen where the bunny was acting on the duck), revealing that they were able to use syntactic information to determine verb meaning. This finding has been replicated in subsequent research, and the effect remains when the first label (e.g., bunny) is replaced with a pronoun (e.g., he’s gorping the duck!), and when the subject and object are people rather than animals (Gertner & Fisher, 2006). In fact, recent research suggests that toddlers as young as 19 months of age can use syntactic cues (such as word order) to determine the meaning of a novel verb (Yuan, Fisher, & Snedeker, in press).

Children’s use of social cues in determining the meaning of a novel verb has frequently been studied in terms of understanding the goals, intentions, desires, and perceptions of their communicative partners (Meltzoff, Gopnick, & Repacholi, 1999). For example, a looking-time study by Poulin-Dubois and Forbes (2002) revealed that 27-month-olds are able to use their understanding of intentions to differentially label actions that an actor produced purposefully versus accidentally (e.g., topple versus knock over). A subsequent live-action study revealed that 27-month-olds, but not 21-month-olds, were able to learn names for similar looking actions that differed only in the nature of the actor’s intention (Poulin-Dubois & Forbes, 2006). Thus, children are sensitive to the intentions of others by 2 years of age, and are able to map words based on those intentions in the face of perceptually similar referents.

A recent series of studies exploring the role of perceptual versus social cues to verb meaning sheds light on the relative importance of social and attentional information to the learning process (Brandone et al., 2007). Thirty-two 22-month-olds were taught a label for one of two available actions. Either the labeled or the unlabeled action produced a result (e.g., a light or a sound). At test, infants saw a video depicting the labeled action on one side of a split screen, and the unlabeled action on the other side. Results of a looking-time analysis revealed that 22-
month-olds could only learn a word for an action when the speaker was naming the action and the action produced a result. When the speaker named an action that did not produce a result, children disregarded the social cues to reference and failed to learn a word. A second study revealed that 22-month-olds could not learn the label for an action when both possible actions had equally salient results. Finally, 34-month-olds in a third study managed to overcome the lure of a perceptually salient result and follow speaker cues to attach a word to the result-less action (Brandone et al., 2007). Object-related perceptual cues can also make it difficult for children to attend to an action, as demonstrated by the finding that 3.5- to 4-year-olds are more likely attend to the object in a verb-learning task when the object in the scene is novel. When the objects are familiar, however, children correctly direct their attention to the action (Kersten & Smith, 2002).

**Adjectives.** Children begin to produce words referring to object properties (e.g., color, material, size) early in development (for example, toddlers might say big dog; Fenson et al., 1994). Although children aged three and younger typically prefer to map novel words to an object’s basic-level kind (Markman, 1989) rather than a property (Hall, Waxman, & Hurwitz, 1993), infants can be compelled to attend to object properties under certain conditions. In a test of adjective learning, 21-month-olds were shown three objects, two of which shared a property (color or texture; Waxman & Markow, 1998). In one condition, all of the available objects belonged to the same basic level category (e.g., 3 different horses). In another, the object that was described belonged to a different basic level category than the two other objects (e.g., 1 green horse (training object) and 2 keys – one green (target) and one yellow). Half the subjects heard one object described with an adjective (e.g., this is a citron one), and were asked at test, “Can you find another citron one?” The other half heard, “Look at this one,” and were asked at test, “Can you find another one?” Results revealed that 21-month-olds were able to select an
object based on property when all of the objects belonged to the same basic level category, but not when they belonged to different basic level categories (Waxman & Markow, 1998; Klibanoff & Waxman, 2000). Importantly, infants saw only one exemplar of an object possessing a particular property before they were asked to select an object based on that property.

What is the role of object familiarity in adjective learning? As with verbs (Kersten & Smith, 2002), there is evidence to suggest that familiarity with the object whose property is being labeled facilitates mapping (Hall, Waxman, & Hurwitz, 1993; Mintz, 2005). One way to explain this “familiarity effect” is to remember that infants have a default preference to attach novel labels to novel objects (assumption of mutual exclusivity; Gelman & Markman, 1985). If the object already has a name (e.g., cup), infants operating under the assumption of mutual exclusivity will search for another aspect of the object to attach the label to – such as a property (Gelman & Markman, 1985; Waxman, 1990).

In a study that encouraged object property categorization by providing multiple exemplars, Waxman (1999) found that infants as young as 12-13 months of age differentially responded when an object was referred to using a noun versus an adjective. In a novelty-preference paradigm, infants were familiarized to four different toys that had one common property: either texture or color. In one familiarization condition, infants heard the experimenter refer to each object using an adjective (e.g., this one is blick-ish); in the other, they heard the experimenter refer to the four objects using a noun (e.g., this one is a blick). At test, infants were given the opportunity to interact with two objects, neither of which had been seen during familiarization. One object had a novel shape and a novel color (new property), while the other had a novel shape and the same color (old property). Results revealed a significant preference for the object with the new property, but only when children heard an adjective during
familiarization. This suggests that 1-year-olds attend more to property differences when they hear an adjective than when they hear a noun during familiarization with objects, which is consistent with research suggesting that infants are sensitive to linguistic cues to grammatical category from a very early age (Morgan & Demuth, 1996).

Waxman’s (1999) study of the role of adjectives in directing infants’ attention to the properties of an object did not directly assess adjective learning per se, because the dependent variable in that study was increased attention rather than proactive response to questions by an experimenter. To address this issue, Waxman and colleagues conducted a live-action study wherein children were familiarized with four same-colored objects each described as “blickish”. Children were given an opportunity to compare “blickish” to “not blickish” (Waxman & Klibanoff, 2000) and then explicitly asked to retrieve “the blickish one” from two novel objects at test. Results revealed that 14-month-olds were able to select a new object based on certain property information (color) but were unable to do so when texture was the common property, demonstrating the fragility of adjective learning at 14 months (Booth & Waxman, 2003).

Mintz and Gleitman (2002) examined the effects of multiple exemplars and lexical specificity on 2- to 3-year-olds’ ability to learn a novel adjective. Children saw three different objects that shared a single property. A puppet described each of the three objects using a novel adjective (e.g., “See this? This is a stoof horse.”) or neutral audio (e.g., “See this? Look at this!”). At test, children were asked to give the puppet one of two test objects. The first test object matched the training exemplars in object kind but not property. The second matched the training exemplars in property but not kind. Results revealed that when children heard the puppet describe the three training objects using an adjective (but not with neutral audio), they were able to abstract the common property and select the property-matched target at test. Interestingly,
children did not abstract the common property if object familiarity was not highlighted lexically, or if they saw only one exemplar (Klibanoff & Waxman, 2000). That is, when children heard audio like “See this? This is a stoof one,” during training, they no longer reliably selected the property-matched object at test. Thus, 2- to 3-year-olds appear to need both lexical specificity and multiple exemplars to learn a novel adjective (Mintz & Gleitman, 2002). Adjective learning and extension is challenging even for older children. In fact, Song and colleagues (2010) illustrated that 3- and 4-year-olds were not yet at ceiling in learning and extending new adjectives on a forced choice match-to-sample task in both English and French, where the adjective appears after the noun.

Social and linguistic context also contribute to children’s adjective acquisition (Diesendruck, Hall, & Graham, 2006). In 2002, Akhtar described a novel object using a novel word (e.g., daxy) to 2.5- and 3.5-year-olds. In one condition, she set up a conversational context wherein an object’s shape was the relevant feature of interest by describing two other objects using shape words (such as round and square) before describing the novel object as daxy. In a second condition, she set up a context wherein an object’s texture was the conversationally relevant feature by describing two other objects as smooth and fuzzy before using the word daxy. Children were asked to extend the adjective daxy to a new object. Results revealed that children who learned daxy in a shape-relevant context extended daxy in terms of shape, where as children who learned daxy in a texture-relevant context extended the word as if it referred to an object’s texture (Akhtar, 2002). This suggests an important role for conversational relevance to children’s mapping and extension of novel adjectives in the real world.

In sum, adjective learning appears to benefit most when multiple cues to word meaning are available – perceptual, social, and linguistic. Children are more likely to learn a novel
adjective when it labels a perceptually salient feature seen across multiple exemplars (Klibanoff & Waxman, 2000), when social information such as conversational relevance is available (Akhtar, 2002), and when they hear lexical/grammatical information that narrows the range of possible referents for the novel word (such as a familiar object’s name or morphological inflection; Behrend, Harris, & Cartwright, 1995; Mintz & Gleitman, 2002).

Closed class words

At the earliest stages of language acquisition, closed class words such as prepositions (e.g., in, on, to, for), determiners (e.g., a, an, the, this, that), conjunctions (e.g., and, but, or), classifiers (used to modify nouns in languages like Japanese and ASL), pronouns (he, she, you, I, they), and morphological inflections are generally not present in children’s productive vocabularies (Braine, 1963; Brown, 1973). However, a significant body of research suggests that infants are sensitive to various closed-class words before they are produced, and that sensitivity to those words facilitates segmentation and categorization of open-class words.

Closed class words occur frequently in spoken language, but the set of closed class words is small and stable (i.e., new closed class words are rarely added to the set, whereas new open class words like the verbs google or blog are regularly added to the lexicon; Kucera & Francis, 1967). In almost all languages, grammatical words differ predictably from content words in length, acoustics, and phonology (Shi, Morgan & Allopenna, 1998). These acoustical and phonological differences are so pronounced that newborns show evidence of perceiving open (lexical) versus closed (grammatical) class words categorically (Gerken, Landau, & Remez, 1990; Shipley, Smith, & Gleitman, 1969). Using a habituation paradigm, Shi and colleagues (1999) exposed 1- to 3-day-old infants to a list of words that were either grammatical or lexical. Once infants had habituated, an experimental group heard a list of words of the other type (i.e., if
they heard grammatical words during habituation, they heard lexical words at test, and vice versa). A control group heard a new list of words of the same type (i.e., if they heard grammatical words during habituation then they heard new grammatical words at test, or lexical words during habituation and new lexical words at test). Results revealed that infants reliably increased their attention when the type of word (grammatical or lexical) changed, but not when the list of words changed but word type remained the same (Shi, Werker, & Morgan, 1999). The authors concluded that newborns use a multiplicity of perceptual cues to reliably distinguish grammatical from lexical words, and that this ability may give them a foothold into developing syntactic categories.

Infants may be able to discriminate between open and closed class words at birth by using perceptual cues, but when and how do they begin to use closed class words as tools to aid in language acquisition? To determine when infants start to recognize functors in a phonetically specific way, researchers presented infants with nonsense functors and actual functors and asked which type infants preferred listening to (Shady, 1996; Shafer, Shucard, Shucard, & Gerken, 1998). By matching real and nonsense functors on perceptual elements such as prosody and vowel shape, Shi, Werker, and Cutler (2006) conducted a particularly stringent test of functor recognition. Eight-month-olds, 11-month-olds, and 13-month-olds heard a nonsense functor (such as ris) paired with a pseudo content word (such as breek), and an actual functor (such as the) paired with a pseudo content word in a habituation paradigm. Results revealed that 13-month-olds reliably preferred listening to real functors, 11-month-olds showed a trend toward preferring real functors, and 8-month-olds had no preference. This suggests that infants possess a phonologically precise grasp on functors after the first year, and that functor knowledge develops gradually between 8 and 13 months of age.
Functors are specifically segmented by English-learning infants at 13 months, but year-old children have already had extensive exposure to closed class words like his and her. What role do such frequent functors play in segmenting content words from fluent speech? Recent research suggests that English-reared infants as young as 13 months of age use frequently heard functors (such as the) but not infrequent ones, to segment nouns from a speech stream (Shi, Cutler, Werker & Cruickshank, 2006). This finding is not specific to English. In fact, when a language has more distinct functors, as in French, infants as young as 8 months of age are able to use frequent functors, but not nonsense or infrequent functors, to segment content words (Hallé, Durand, & Boysson-Bardies, 2008; Polka & Sundara, 2003; Shi & LePage, 2008). A similar ability to use functors to segment and categorize new open class words has been found in infants learning German (Höhle & Weissenborn, 2003; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004) and French (Van Heugten & Shi, 2009). In a cross-linguistic demonstration of functor use in word learning, Lew-Williams and Fernald (2007) showed that Spanish-reared children can use articles in Spanish (distinguished by gender) to quickly find referents for words, a skill that undoubtedly contributes to the rapidity of sentence comprehension.

By 18 months of age, infants respond more quickly to objects when their labels follow a determiner (e.g., “Look at the dog!”) than when they do not (e.g., “Look. Dog!”) (Fernald & Hurtado, 2006). In a recent exploration of determiner sensitivity in 18-month-olds, Kedar and colleagues showed pictures of familiar objects presented in various linguistic contexts (Kedar, Casasola, & Lust, 2006). Infants heard audio asking for the objects using an actual determiner (e.g., Can you see the book?), a nonsense determiner (Can you see el book?), a conjunction (Can you see and book), or no closed-class word at all (Can you see _ book?). Results revealed that 18-month-olds distinguish between different types of function words (determiners versus
conjunctions) and have expectations for how those words work in sentences. Children were quicker to look at the correct object when the label was preceded by the correct function word (i.e., *the*) than when it was preceded by a conjunction, nonsense determiner, or no word at all (Kedar et al., 2006).

How sensitive is children’s sentence comprehension to grammatical articles, and when does comprehension become robust to errors? Zangl and Fernald (2007) found that 18-month-olds’ sentence comprehension was compromised when they heard a familiar noun preceded by a nonsense article, but that 34-month-olds’ comprehension was unaffected. The sentence comprehension of 34-month-olds worsened, however, when the noun was unfamiliar to the children and preceded by a nonsense article (Zangl & Fernald, 2007). This suggests that children’s ability to disregard nonsense functors improves over time, and is robust to the extent that the noun in the sentence is familiar. In the case of novel nouns, however, 3-year-olds become confused and less able to comprehend the sentence – their processing is disrupted by the presence of a nonsense functor.

*Classifiers* are a type of closed-class lexeme whose purpose is to classify nouns according to meaning (e.g., person + animate classifier = animate person). English happens to be an *inflected* language (discussed below) rather than a classifier language, but many languages of the world use classifiers to modify the animacy status, size, shape, structure, numeracy, possession, relation, and/or location of noun referents (Aikhenvald, 2003). Examples of such classifier languages include the American, Arabic, Australian, Austrian, British, Brazilian, Croatian, German, Hong Kong, Irish, Korean, Mexican, South African, Swedish, Swiss German, Turkish, and Taiwan Sign Languages (Milkovic, Bradaric-Joncic, & Wilbur, 2007; Morgan & Woll, 2007; Perniss, 2007), Chinese (Li & Thompson, 1981; Wei & Lee, 2001), Cantonese
(Erbaugh, 2001, 2002; Tse, Li, & Leung, 2007), Japanese (Uchida & Imai, 1999; Yamamoto & Keil, 2000), Mayan languages like Jakaltek and Akatek (Zavala, 2000), Austronesian languages such as Fijian (Lichtenberk, 1983), and Papuan languages such as Waris (Brown, 1981).

Although the features of the world that are classified appear to be common across languages (e.g., animacy, shape, size, numeracy), systems of classification are not universal. That is, different languages carve up the world in different ways, and emphasize certain features over others. For example, in Japanese, the same classifier is never used for both animate and inanimate objects. In Mandarin, on the other hand, it is possible to use one classifier for a set that includes both animate and inanimate objects (Uchida & Imai, 1999). Given the cross-linguistic variability of classifier systems and internal exceptions to the rules of classification, how do children begin to understand and acquire classifiers? The answer is: with difficulty. In fact, the majority of published research suggests that 4- to 6-year-old Thai learning (Carpenter, 1991), Chinese learning, and Japanese learning children have trouble comprehending – much less producing – classifiers (Uchida & Imai, 1999; but see Szeto, 1998 and Tse et al., 2007 for evidence that Cantonese-speaking children produce general-purpose classifiers at younger ages). Nonetheless, when compared to the young age at which children are able to fast-map nouns to objects (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992), classifiers are acquired relatively late in development.

In addition to closed class words and/or classifiers, many languages contain a set of units called morphological inflections (examples from English are the plural -s, the gerund -ing, and the past tense -ed). These units alter the meaning of the open class words to which they are bound by changing tense, number, aspect, etc., and, like function words, are often omitted by very young children (Armon-Lotem & Berman, 2003; Berko, 1958; Brown, 1973; Matthews &
Theakston, 2006; Valian & Aubry, 2005). Although children do not evidence *explicit* semantic understanding of some inflections (such as the impact of the plural -s) until the late preschool or early elementary years (Johnson, de Villiers, & Seymour, 2005; de Villiers & Johnson, 2007), many researchers have shown that infants (some as young as 16 months old) are able to distinguish correct from incorrect applications of the plural -s (Fenson et al., 1994; Mervis & Johnson, 1991; Soderstrom, 2008; Tomasello, 1992), especially when the root word is familiar (Soderstrom, White, Conwell, & Morgan, 2007). Furthermore, children can use grammatical plurality to guide their search patterns (Wood, Kouider, & Carey, 2009), spontaneously produce the plural -s by 23 months (Tomasello & Olguin, 1993), and use it to distinguish between generic and specific meaning by 2 years of age (Gelman & Raman, 2003). Likewise, the progressive -ing is used relatively early in development (Brown, 1973; DeVilliers & DeVilliers, 1973). Infants show a preference for sentences that contain the correct usage of -ing as an aspectual inflection by 18 months in the Head Turn Preference Procedure (Santelmann & Jusczyk, 1998) and the Intermodal Preferential Looking Paradigm (Golinkoff, Hirsh-Pasek, Schweisguth, 1998; Hirsh-Pasek, 2000; Hirsh-Pasek & Golinkoff, 1996). Productively, 2-year-old toddlers use verbs inflected with -ing (Akhtar & Tomasello, 1997; for further discussion of the progressive verb form in English, see Wagner, 2001; Weist, 1991; Weist, Lyytinen, Wysocka & Atanassova, 1997; Weist, Atanassova, Wysocka & Pawlak, 1999). The past tense -ed is another common morphological inflection that is often treated appropriately by 2-year-olds (Hohenstein & Akhtar, 2007) but which may also be over-regularized by children learning the rule (e.g., the past tense of run may be called runned by a child instead of ran; Cazden, 1969; Marchman, 1997; Marchman, Plunkett, & Goodman, 1997; Marcus, 1996; Mervis & Johnson, 1991; Plunkett & Marchman, 1993).
Statistical regularities in morphological marking have the potential to contribute to early word *categorization* (e.g., a word ending with *-ing* is probably a verb; Bedore & Leonard, 2000; Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995). In fact, a recent study by Soderstrom and colleagues (2007) revealed that 16-month-olds use regularities in morphological information (along with word order) to distinguish different grammatical classes, such as noun versus verb (Soderstrom, White, Conwell, & Morgan, 2007). This use of morphological marking to distinguish word classes has been found cross-linguistically, as in the case of distinguishing male- and female-gendered nouns in Russian (Gerken, Wilson, & Lewis, 2005).

Children use morphology as a cue to word class membership, but can they use morphosyntactic information to “bootstrap” them into word meaning (e.g., *morphological bootstrapping*; Behrend et al., 1995)? Significant research across multiple languages suggests that the answer is yes; children use such cues to infer information about the meaning of novel words (e.g., gender cues in Spanish; Lew-Williams & Fernald, 2007). A study of English-learning children found that the presence or absence of the morphological inflections *-ing* and *-ed* strongly influenced preschooler’s extension of the verbs to new instances (e.g., verbs learned with the *-ing* ending were preferentially extended to ongoing actions with similar manner, whereas verbs learned with the *-ed* ending were more often extended to actions with a similar end-point or result; Behrend et al., 1995; Carr & Johnston, 2001). Similarly, Göksun and colleagues found that Turkish-reared children enact sentences in a more causal way when one of the noun phrases is accompanied by an accusative morpheme (explicitly designating it as the patient of an action). Two- to 5-year-old children were asked to use toy animals to act out sentences containing words like *pig dog push*. In one condition, the second noun phrase was accompanied by an accusative morpheme (*pig dog-ACC push* – meaning “the pig pushed the
dog”). In the other condition, both nouns were bare. Results revealed that subjects enacted the actions in a more causative way (e.g., making one animal physically push the other) when the verb was accompanied by an accusative marker than when it was not, suggesting that the presence of a morphological inflection influenced their interpretation of verb meaning (a process the authors term *morphosyntactic bootstrapping*; Göksun, Küntay, & Naigles, 2008).

Noun learning is also facilitated by the presence of morphological cues, as demonstrated by a recent exploration of the plural -s using a preferential looking paradigm. Thirty-month-olds were shown two pictures. One picture depicted a novel object A, and the other depicted a group of novel objects B (all novel objects B were identical, but different than the picture of the single object A). Rather than providing rich syntactic information as well as morphological information (as in some other studies of morphological cues to word meaning; Booth & Waxman, 2003; Kouider, Halberda, Wood, & Carey, 2006; Waxman & Booth, 2001, 2003), the authors provided only information about word class while manipulating the presence or absence of a plural inflection. Results revealed that infants reliably increased their looking to the correct referent at test when they were trained in the plural condition but not the singular condition, suggesting that the presence of an inflection helped them to infer the meaning of the novel word (a process the authors term *inflectional bootstrapping*; Jolly & Plunkett, 2008).

Further evidence of the relation between understanding morphological marking and improved word learning has been shown cross-linguistically in both correlational and longitudinal studies. For example, children’s understanding of morphological inflections like -ed and -ing at 2.5 and 3 years of age predicts language ability two years later (Lyytinen & Lyytinen, 2004), and even after controlling for a number of factors (such as phonological awareness), morphological awareness predicts a significant amount of variance in children’s long-term

VI. How do words go together? Grammatical development

After infants segment words and use a variety of cues to discover which part of speech they are and what they mean, the next step is to figure out how to combine those words to express complex ideas.

Why is grammar important?

Grammar allows language-users to create phrases and sentences out of lexical items. By combining words and sounds in a grammatical way, speakers can progress beyond simplistic telegraphic speech (e.g., “apple”) and begin talking about relations between referents (e.g., “She ate the apple”). Syntax is critical because it enables a speaker to do more than describe objects, actions, and events; it allows them to offer a point of view, or perspective, on those events. In addition, certain lexical items can modify or modulate other items when combined with one another (e.g., “very tiny” or “ridiculously big”).

Syntactic rules dictate the way words or gestures can be combined to form a well-constructed phrase. Diversity in language structure includes variation in morphology, word, order, classifiers, etc. to express meaning. For example, Japanese has an extensive classifier system for modifying nouns, whereas English does not. Likewise, pronouns are typically used when communicating in English (e.g., in order to convey that a female went to the store, a person often says, “She is going to the store,” but in Spanish, the she is often dropped and a person would simply say, “Va a la tienda” which roughly translates to “___ is going (or goes) to the store.” Information about the subject of the sentence is typically drawn from the context in which
the sentence is uttered. However, even in Spanish, information about the subject is available through the verb conjugation. To say, *she, he, or it* is going to the store, the verb is conjugated as “va.” To say *they* are going to the store, the verb is “van.” To say *we* are going to the store, the verb is “vamos.” In this way, Spanish does not require that listeners infer every bit of information about the subject from pragmatics alone, even though Spanish is generally considered a “pro-drop” language (i.e., a language that permits the dropping of the noun arguments that accompany a verb, as in ‘John kissed Mary’ where John and Mary are arguments of the verb “kiss”). Japanese is an even more dramatic example of a pro-drop language, in that pronouns are almost never used – to use them is not grammatically incorrect *per se*, but sounds very odd to a native speaker.

Word order is another component of grammar that differs across languages. In English, for example, sentences typically take the subject-verb-object (SVO) form [e.g., “The *frog* (subject) *ate* (verb) *the fly* (object)]. To say, “Ate the frog the fly” is incorrect because English does not support the VSO form. Other permutations of the more common SVO and VSO forms exist in other languages, but orders such as OVS are rare (although it was used in the artificial language *Klingon*, courtesy of Star Trek). Word order knowledge is fairly robust, as young children exposed to one type of word order in the ambient language will not accept as grammatical sentences with different word orders (Akhtar, 1999; Matthews, Lieven, Theakston & Tomasello, 2005; Matthews, Lieven, Theakston, & Tomasello, 2007).

In an early exploration of children’s ability to use syntactic information (in this case, word order) to determine reference, 17.5-month-old infants were shown two scenes. In one scene, Big Bird was tickling Cookie Monster. In the other, Cookie Monster was tickling Big Bird. Toddlers were asked to look at where Cookie Monster was tickling Big Bird. Results of a
looking time analysis revealed that young toddlers used word order information to correctly infer which scene was the “correct” scene (i.e., which scene matched the sentence). This ability suggests that infants used an abstract grammatical rule (SVO order in English) to determine which character was the subject (or agent) and which was the object (or patient) of a transitive action (Hirsh-Pasek & Golinkoff, 1996; Naigles, 1996). A more recent study tested the ability of slightly older toddlers to use word order to determine the referent of a novel word embedded in a transitive sentence. Results of a looking-time study revealed that 21- and 25-month-olds looked significantly longer at a transitive scene in which the subject of the target sentence was the agent and the object was the patient (Gertner, Fisher, & Eisengart, 2006; see also Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008; Fisher, 2002). This suggests the ability to use word order information to determine the referent of a novel verb, and that 1.5-year-olds have some understanding of “agent” and “patient” (Golinkoff, 1975; see Brandone et al., 2005, and Fernandes, Marcus, Di Nubila, & Vouloumanos, 2006, for a discussion of abstract syntactic knowledge of intransitive structure).

In 2003, Seidl, Hollich, and Jusczyk demonstrated that infants as young as 15 months of age show some understanding of a deep syntactic structure long believed to be acquired only after age three (Ervin-Tripp, 1970; Tyack & Ingram, 1977). The case of Wh-questions (e.g., *What did Sally do to Jim?* or *Who did Sally say Jim hugged?*; Chomsky, 1981) has been difficult to assess using standard production methods, but was successfully explored using a preferential looking paradigm. Fifteen- and 17-month-old infants were shown scenes of familiar objects knocking into one another, and then heard audio asking them to look at one of the two objects. The test question was formed as either a subject-question (e.g., *What hit the keys?*) or an object-question (e.g., *What did the keys hit?*). Results revealed that 15-month-olds correctly looked to
the subject in the subject-question condition but not to the object in the object-question condition, whereas 20-month-olds looked correctly in both the subject and the object conditions. This suggests that infants are able to understand syntactic structure much earlier than suspected (i.e., 20-month-olds demonstrated a reliance on Wh-movement rather than simply word order information), and that this ability develops over time (Seidl, Hollich, & Jusczyk, 2003).

Evidence that children may have some innate knowledge of grammatical structure (or, minimally, are equipped to learn grammar with relative ease) comes in part from research showing that the extent of children’s early grammatical understanding cannot be deduced from the input they receive (Chomsky, 1957). However, other researchers argue for a computational model of grammar acquisition (Elman, 2006; Elman et al., 1996) and still others argue that training studies demonstrate an effect of learning and social experience on the acquisition of syntactic structure (Abbot-Smith & Tomasello, 2006; Childers & Tomasello, 2001; McClure, Pine, & Lieven, 2006; Tomasello, 2003).

VII. Social cues, enter stage right: The role of pragmatic development

In addition to perceptual and syntactic cues to word reference, children have access to a plethora of social cues. However, despite the use of infant-directed speech, the language heard by infants tends to be incomplete and referentially ambiguous. For example, an infant might hear an older sibling say, “I’m hot and sweaty,” to which a parent could gesture toward the kitchen and respond with, “There’s a juice box in the refrigerator.” In this case, understanding what the child means by “I’m hot and sweaty” and what the parent is trying to convey by “There’s a juice box in the refrigerator” requires young children to take into account a number of cues that can be gleaned neither from a strict semantic interpretation of each sentence individually nor by an analysis of the syntactic structure of the interchange. Rather, understanding the communicative
nature of the question and answer depends upon children’s ability to access social-pragmatic/contextual cues to meaning, such as where the mother gestured when she offered the juice box, and that a person tends to feel thirsty when they are hot and sweaty. As children acquire language, they will implicitly act in accordance with a set of communicative principles and maxims (such as being relevant, being cooperative, telling the truth, speaking parsimoniously and completely) that ordinarily dictate verbal interaction (Grice, 1957, 1975; Kasher, 1977; Sperber & Wilson, 1986).

What extra-linguistic social cues do children exploit to determine reference in an ambiguous situation, and when do they begin to use them to learn language? A preponderance of evidence suggests that children are sensitive to cues like a speaker’s goals and intentions (Baldwin, Baird, Saylor, & Clark, 2001; Behne, Carpenter, Call, & Tomasello, 2005; Buress & Woodward, 2007; Carpenter, Akhtar, & Tomasello, 1998; Csibra, Bíró, Koós, & Gergely, 2003; Gergely, Nádasdy, Csibra, & Bíró, 1995; Johnson, Booth, & O’Hearn, 2001; Luo & Baillargeon, 2005; Meltzoff, 1995; Parish-Morris, Hirsh-Pasek, Hennon, Golinkoff, & Tager-Flusberg, 2007; Saylor, Baldwin, Baird, & LaBounty, 2007); a speaker’s true and false beliefs (Csibra & Southgate, 2006; Southgate, Senju, & Csibra, 2007); a listener’s knowledge and perceived speaker knowledge (Birch et al., 2008; Bowler, Briskman, & Grice, 1999; Moll & Tomasello, 2006, 2007; Nurmsoo & Robinson, 2009; Onishi & Baillargeon, 2005; Sabbagh & Baldwin, 2001; Sabbagh, Wdowiak, & Ottaway, 2003; Scofield & Behrend, 2008); joint and triadic attention (Adamson, Bakeman, Deckner, & Romski, 2009; Baldwin, 1995; Carpenter, Nagell, & Tomasello, 1998; Cleveland & Striano, 2007; Morales et al., 2000a; Mundy et al., 2007; Mundy & Newell, 2007; Striano, Chen, Cleveland, & Bradshaw, 2006; Striano, Stahl, Cleveland, & Hoehl, 2007; Tomasello, 1995); gestures (such as pointing or touching; Brand et al., 2002;
Hollich et al., 2000; Moore, Angelopoulos, & Bennett, 1999; Woodward & Guajardo, 2002); eye gaze (Bloom, L., 2000; Booth, McGregor, & Rohlfing, 2008; Brooks & Meltzoff, 2005; Farroni, Massaccesi, Pividori, & Johnson, 2004; Morales et al., 2000b; Phillips, Wellman, & Spelke, 2002; Reid & Striano, 2005; Reid, Striano, Kaufman, & Johnson, 2004; Woodward, 2003) and whether the speaker is reliable or trustworthy (for a review, see Koenig & Harris, 2007).

The direction of a speaker’s gaze provides important information about word meaning. In fact, infants are able to learn some words based on eye gaze direction alone (Baldwin, 1991, 1993; Hollich et al., 2000; Houston-Price, Plunkett, & Duffy, 2006). In addition to helping children learn words in the moment, the ability to follow eye gaze has been linked to productive vocabulary over time. In one particularly striking study, Brooks and Meltzoff (2008) found that 10- and 11-month-olds’ ability to follow an adult’s gaze to an object predicted the speed of their productive vocabulary growth through age 2.

Despite evidence suggesting heavy reliance on eye gaze direction for word learning, data also support the notion that children use this cue prudently. For example, Nurmsoo and Bloom (2008) found that 3- to 4-year-olds rely on eye gaze as a cue to word reference only when it is relevant to the context as a whole. Rather than simply using eye gaze as an invincible cue to word reference, children are able to take into consideration linguistic and intentional information as well, and ignore eye gaze information when necessary (Nurmsoo & Bloom, 2008). Similarly, children will ignore information about word reference when it is presented by an ignorant speaker, but learn a word from a knowledgeable speaker (Birch et al., 2008; Sabbagh & Baldwin, 2001; Sabbagh et al., 2003; Scofield & Behrend, 2008). This suggests that by the preschool years, children flexibly and wisely exploit a wealth of pragmatic cues in their quest to determine word meaning.
Thanks to other types of pragmatic acuity such as the ability to determine a person’s goals and intentions, young children can learn words in ambiguous situations where eye gaze may not be a sufficient cue to word meaning. For example, children learn names for objects that are hidden from view at the time of labeling (Baldwin, 1993), have never been explicitly labeled (Akhtar & Tomasello, 1996; Tomasello & Barton, 1994), or that are labeled based on the false belief of a speaker (Carpenter, Call, & Tomasello, 2002; Happé & Loth, 2002). One study of false-belief tracking revealed that that 3- to 5-year-olds are better at tracking the false beliefs and intentions of a speaker in the context of word learning than in a typical false belief task that does not have a word learning component (Happé & Loth, 2002). This result implies that there may be special significance associated with taking another person’s perspective (i.e., having a “theory of mind”; Baron-Cohen, 1995) in word learning (Parish-Morris et al., 2007; Patael & Diesendruck, 2008).

Importantly, recent research suggests that word learning – at least in the earliest years – requires a social, interactive and contingent environment. It is not enough, for example, to just hear language or words spewing forth from a television. In fact, findings by Kuhl, Tsao, and Liu (2003) and Roseberry, Hirsh-Pasek, Parish-Morris and Golinkoff (2009) suggest that children under the age of three are unlikely to learn words from mere exposure on television. Surprisingly, however, children presented with novel verbs on video chat did learn new words. In a video chat format, contingent language was delivered where the input was responsive to the child and 2-year olds were able to master the new words just as they had in a live, face-to-face condition (Roseberry, Hirsh-Pasek & Golinkoff, in preparation). Thus, language learning requires sensitive and responsive conversations with children where language input is tailored to the interest and timing of the child’s attention.
VIII. Topics of special interest

*Word learning and thinking: Is cognition affected by language or vice versa?*

Do differences between languages (such as linguistic representations of space and time that highlight different aspects of the world, or the presence/absence of a classifier system) influence the way speakers think? The *Sapir-Whorf hypothesis* (Whorf, 1939) or theory of *linguistic relativity* holds that language shapes thought; an idea that has drawn both support (Bloom, 1981; Bloom & Wynn, 1997; Boroditsky, 2001; Casasanto, 2008; Casasanto & Boroditsky, 2008; Cifuentes-Férez & Gentner, 2006; Dessalegn & Landau, 2008; Feist & Gentner, 2007; Gentner & Christie, 2008; Lakusta, Wagner, O’Hearn, & Landau, 2005;) and opposition (Bloom & Keil, 2001; Heider, 1972; Heider & Oliver, 1972; January & Kako, 2007; Papfragou, Li, Choi, & Han, 2007; Tse & Altarriba, 2008; Pinker, 1994). Some theorists have argued for the existence of “culturally specific” cognitive styles, and indeed, support for this position has been found (e.g., Choi, Nisbett, & Smith, 1997; Nisbett, 2003; Nisbett, Peng, Choi, & Norenzayan, 2001; Saalbach & Imai, 2007; Varnum, Grossmann, Katunar, Nisbett, & Kitayama, 2008). Others argue that the truth lies somewhere in the middle (Gleitman & Papafragou, 2005).

The connection between cognition and spatial language has been extensively researched (Li & Gleitman, 2002; Majid, Bowerman, Kita, Haun, & Levinson, 2004). For example, a large body of evidence suggests that relational language facilitates relational thinking (Feist & Gentner, 2007; Loewenstein & Gentner, 2005), and that there are language-specific ways of thinking about color (Gilbert, Regier, Kay, & Ivry, 2006; Kay & Kempton, 1984; Kay & Regier, 2007; Roberson, Davidoff, Davies, & Shapiro, 2004; Roberson, Davies, & Davidoff, 2000, 2001; Winawer et al., 2007; Witthoft et al., 2003), time (Boroditsky, 2001; Casasanto & Boroditsky,
2008; January & Kako, 2007; Núñez & Sweetser, 2006), animacy (Anggoro, Waxman, & Medin, 2008) and number (Bloom & Wynn 1997; Dehaene, Izard, Spelke, & Pica, 2008; Gelman & Gallistel, 2004; Ji, Zhang, & Nisbett, 2004; Miller, Major, Shu, & Zhang, 2000; Pica, Lemer, Izard, & Dehaene, 2004; Spelke & Tsivkin, 2001). Recent research with Turkish-reared infants and Spanish- or English-speaking adults confirmed that the peculiarities of one’s native language direct attention to aspects of a scene that are linguistically relevant (Cifuentes-Férez & Gentner, 2006; Göksun, Hirsh-Pasek, & Golinkoff, 2010).

What about language-specific spatial concepts? Analogous to the initial ability of prelinguistic infants to distinguish sounds from all of the world’s languages (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Jusczyk, 1997; Kuhl, 1993; Streeter, 1976), it appears that prelinguistic infants are capable of distinguishing conceptual categories like “a tight-fitting relationship between two objects, one of which is contained by the other” that are used in languages other than their ambient language (Hespos & Spelke, 2004; Göksun et al., 2010). Furthermore, in the same way that infants gradually specialize and “zoom in” on the phonemes used in the ambient language (Best, McRoberts, Lafleur, & Silver-Isenstadt, 1995; Werker et al., 1981), some suggest that all infants start with the same conceptual bases – e.g., a pool of conceptual options from which a few are highlighted as relevant an individual child’s linguistic environment (McDonough, Choi, & Mandler, 2003). Over time, the highlighted conceptual distinctions are retained and strengthened, whereas irrelevant or unlexicalized distinctions are disregarded and ultimately not attended to by adults (Hespos & Spelke, 2004).

In 2004, Hespos and Spelke explored this hypothesis by looking at the conceptual distinctions made by Korean- and English-reared infants and adults. In Korean, a lexical distinction is made between tight-fitting containment relations (such as a cylinder fitting tightly
into a similar-sized cylindrical container) and *loose-fitting* relations (e.g., a cylinder fitting into a container with a significantly larger diameter) that is not made in English. In English, tight-fitting and loose-fitting instances of containment are both described by a single word: *in*. Hespos and Spelke demonstrated that 5-month-olds in both languages distinguish the Korean categories *tight-fit/loose-fit* (suggesting a universal conceptual base) whereas adults categorize the same relation in language-specific ways (Korean-speaking adults categorize according to *tight-fit* and *loose-fit* whereas English-speaking adults do not). These results support the hypothesis that while infants may attack the issue of language learning with the same initial stock of conceptual distinctions, their default concepts are gradually shaped by the lexical peculiarities of their native language (Hespos & Spelke, 2004). Göksun, Hirsh-Pasek, and Golinkoff (2010) developed this thesis further and extended it to a wide range of language-relevant concepts. They suggested that infants start with language-general nonlinguistic constructs that are gradually refined and tuned to the requirements of their native language. In effect, maintaining their sensitivity to some relational distinctions while dampening other distinctions, depending on how their native language expresses these constructs (Göksun et al., 2010).

*The case of bi- or multi-lingual language acquisition*

What happens when children are exposed to more than one language from birth, or when they begin learning a new language at the start of elementary school? Do they learn the languages at the same time, and are they aware that they are learning more than one? Do they make mistakes and mix languages together? Do they think in both languages? Does learning two or more languages slow down their acquisition of *all of them*? Is competence in a “native language” compromised if children learn a new language later in the toddler years? What are the
effects of bilingual language acquisition on other areas of cognition? These questions and many others have been extensively researched, and answers are beginning to emerge.

In the past, some believed that raising a child in a bilingual environment would result in a slow learner not competent in any language. According to this line of thought, children younger than age three would not recognize that they were hearing two different languages, and would be confused and delayed while trying to sort out the input (Redlinger & Park, 1980; Vihman, 1985; Volterra & Taeschner, 1978). However, children raised bilingually do appear to understand that they are learning more than one language (Genesee, 1989; Genesee, Nicoladis & Paradis, 1995; Lanza, 1992; Meisel, 2001), and brain activity during linguistic tasks shows that non-identical systems process the different languages of bilingual toddlers (Conboy & Mills, 2006). Learning two languages simultaneously does not appear cause a delay in acquisition per se. Evidence suggests that children exposed to two languages from birth master both languages in the same order and with the same timing as their monolingual peers (Holowka, Brosseau-Lapré & Petitto, 2002; Oller et al., 1997; Pearson, Fernandez, & Oller, 1993; Petitto, Katerelos, et al., 2001; Petitto & Kovelman, 2003; but see Fennell, Byers-Heinlein, & Werker, 2007) and may demonstrate a “bilingual advantage” in the area of phonetic processing (Bialystok, Majumder, & Martin, 2003, Norton, Baker & Petitto, 2003; but see Bosch, Costa, & Sebastián-Gallés, 2000; Sebastián-Gallés & Bosch, 2005). Does a child have to produce words in the second language to benefit from early experience with two languages? Research suggests that simply overhearing another language regularly during the early years can have a significantly positive effect on an adult’s ability to produce phonemes correctly when learning to speak the overheard language later in life (Au, Knightly, Jun, & Oh, 2002).
What are the larger effects of bilingualism on children’s development? According to the literature, bilingualism has both positive and not-so-positive effects on language and cognitive outcomes (Petitto & Dunbar, 2004; Werker & Byers-Heinlein, 2008). These effects are modulated by language experience and the age at which each language was acquired (Petitto & Dunbar, 2004). Although bilingualism can be achieved by older learners, acquiring a second language after age 5 is difficult if the child is exposed to the second language in a classroom only; optimal outcomes occur when children hear a new language in multiple contexts (e.g., at home, in school, at the playground, at the store; Kovelman & Petitto, 2003; Petitto, Kovelman & Harasymowicz, 2003). Adult bilinguals tend to retrieve words in their dominant language more slowly than monolinguals (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Kaushanskaya & Marian, 2007), and bilingual children tend to have smaller vocabularies in each language than their monolingual peers (Bialystok & Feng, 2010; Mahon & Crutchley, 2006; Oller & Eilers, 2002), although this gap narrows over time. Consistent with this finding, recent research by Hoff, Core, Place, Rumiche, Senor and Parra (2011) suggests that for a matched sample of high-SES bilingual and monolingual children aged 1 to 2-and-a-half years, the total vocabulary of the children is equivalent. That is, whereas monolingual children have all of their vocabulary in one language, bilingual children merely distribute vocabulary across two languages.

Further, many studies show that children learning two languages do just as well or better than children learning only one language. When the methods used to test bilingual children are appropriate and not merely the same as those used with monolingual sample, some of the advantages of learning two languages are unveiled (Sebastian-Galles, 2010). For example, children and adults raised with two languages have certain cognitive advantages over the
monolingual population (Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Craik, & Luk, 2008; Bialystok, Craik, & Ruocco, 2006; Bialystok, Craik, & Ryan, 2006; Craik & Bialystok, 2005; Emmorey, Luk, Pyers, & Bialystok, 2008). Bilingual children demonstrate advanced inhibitory functioning and can control their attention to misleading cues in laboratory tasks (Martin-Rhee & Bialystok, 2008). Bilingual children are more likely to be able to ignore obvious perceptual cues (such as the figure of a duck) to see the alternate figure of a bunny in the same picture (Bialystok & Shapero, 2005). They are more likely to correctly draw the line in Piaget’s water level task by ignoring the perceptually salient slant of the container and attending to the effect of gravity on the surface of the water (Bialystok & Majumder, 1998; see Figure 2). All of these tasks require executive control; an area in which bilingual children and adults have been shown to excel (Costa, Hernández & Sebastian-Galles, 2008; Martin-Rhee & Bialystok, 2008) in certain aspects more than others (Carlson & Meltzoff, 2008), although the effect may disappear once SES and ethnicity are tightly controlled (Morton & Harper, 2007). Older bilingual children show flexibility in the area of purportedly universal word-learning principles like mutual exclusivity (Baldwin, 1992; Markman & Hutchinson, 1984; Markman & Wachtel, 1988), which allows them to entertain the possibility that one object is labeled by more than one word (as in different languages; Davidson & Tell, 2005). Early-bilingual children (defined as having learned two languages prior to age three) demonstrate reading abilities comparable to monolingual children, and perform significantly better on phonemic awareness tasks (Kovelman, Baker, & Petitto, 2008). In sum, learning two languages has pros and cons, but does not damage a child’s outcome overall – bilingualism may even improve certain aspects of cognitive functioning.

_The predictive power of early sounds and behaviors_
Can tests administered during the first months of life measure a child’s ultimate language outcome? Is it possible that a baby’s first noises and behaviors are a window into vocabulary development and reading ability years later? According to numerous researchers, the answer to these questions is a very qualified yes. Studies of early production (i.e., babbling), perception (e.g., distinguishing bottle from bobble), and social skills (e.g., eye gaze following) suggest that there are, indeed, links between early and later development.

The noises infants make when they begin to experiment with language-like sounds are referred to as “babble” and represent a universally important communicative milestone (Oller, 2000). Because babbling is pervasive and relatively consistently found in the typical population, much of the research connecting babble quality to language outcome has been conducted in the context of language disorders or delay (Davis & Velleman, 2000; Fasolo, Majorano, & D’Odorico, 2008; Hall, 2003; Locke, 1989; Maassen, 2002; Oller, Eilers, Neal, & Cobo-Lewis, 1998; Oller, Eilers, Neal, & Schwartz, 1999; Stoel-Gammon & Otomo, 1986; Whitehurst et al., 1991). For example, a retrospective correlational study based on parent report revealed that children later diagnosed with suspected apraxia of speech also babbled less as infants (Highman, Hennessey, Sherwood, & Leitão, 2008). An analysis of home videos demonstrated that children diagnosed later with Selective Language Impairment (SLI) engaged in less canonical babbling than undiagnosed age-mates (Magaldi, 2008). Although the mere presence of babbling is predictive of language outcome, the quality of babbling is predictive as well. For example, phonetically complex babbling is positively associated with better global language outcomes in toddlers implanted with cochlear hearing devices (Walker & Bass-Ringdahl, 2007). If children are “poor babblers,” they can be encouraged to engage in more babbling practice. Social cues affect babbling in typically developing infants, such that the quality and quantity of babbling
improves in the context of contingent social interactions (Gros-Louis, West, Goldstein, & King, 2006).

Certain perceptual skills measured in infancy are also predictive of language outcome. For example, the electrophysiological activity of 71 newborn brains in response to speech stimuli discriminated those same children 5 years later on the verbal subtest of the Stanford-Binet (Molfese & Molfese, 1997), and brain activity in toddlers can be used to predict reading and language outcome in the elementary years (Molfese et al., 2008; Molfese, Molfese, & Molfese, 2007). Other research has shown that 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, Bernstein, Ratner, Jusczyk, Jusczyk & Dow, 2006). In addition to the predictive power of segmentation skills, infants’ ability to discriminate the sounds of their native languages has been linked to later linguistic skills (Höhle, van de Vijver, & Weissenborn, 2006; Tsao, Liu, & Kuhl, 2004). Kuhl and colleagues found that 6-month-olds who were better at discriminating native-language vowel sounds (such as ee and oo) also had superior receptive/productive vocabulary skills and improved phrase understanding over the next 18 months. Interestingly, the ability to detect phonemes present in one’s native language at 7.5 months of age is predictive of later language abilities, but the ability to detect phonemes from other languages is negatively correlated with language outcome (Kuhl et al., 2005; Tsao, Liu, & Kuhl, 2004; for a review, see Kuhl, 2009). This suggests that infants whose natural perceptual tendencies are compatible with the characteristics their native language may have an easier time acquiring words at the outset.

Vocabulary size and speed of reaction time in infancy (e.g., quickly looking at the cup when asked “Where’s the cup?”) are also strongly predictive of language outcome (Bernhardt, Kemp, & Werker, 2007; Marchman & Fernald, 2008). In a longitudinal study, Fernald and
colleagues (2006) found that measures of grammar and vocabulary from 12 to 25 months were correlated with vocabulary size at age two years (Fernald, Perfors, & Marchman, 2006). The original longitudinal study was extended in a 6-year follow-up, wherein measures of vocabulary were collected from the same children at age eight. Results revealed a strong relation between performance on a vocabulary task in toddlerhood and later language and cognitive ability (Marchman & Fernald, 2008).

Early social behaviors are tightly linked to language development (Brooks & Meltzoff, 2008), and come in a variety of forms (e.g., eye gaze following, pointing, joint attention, understanding intentions/goals; Carpenter, Nagell, & Tomasello, 1998). 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 et al., 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, 2000a, 2000b). Brooks and Meltzoff also found that a combination of infant behaviors (eye gaze following in conjunction with infant vocalization; length of gaze duration) at 10-11 months old predicted vocabulary comprehension at 24 months (Brooks & Meltzoff, 2005, 2008).

What happens to language outcome when children have reduced access to social cues? The case of children with autism is instructive, as these children are characterized in part by
impaired social functioning (e.g., difficulty with eye gaze, pointing, understanding others’ intentions, attention to faces, responding to bids for joint attention; Adamson, Bakeman, Deckner, & Romski, 2009; Baron-Cohen, 1995; American Psychiatric Association, 2000; Dube, MacDonald, & Mansfield, 2004; Grice et al., 2005; Parish-Morris et al., 2007). Although language outcomes in children with autism are heterogeneous and range from non-verbal to above average, even very high functioning children with autism nonetheless experience difficulty with the pragmatic (social and practical) aspects of language (Kjelgaard & Tager-Flusberg, 2001; Lord & Paul, 1997; Rapin & Dunn, 2003; Tager-Flusberg, 1996; Volden, Coolican, & Garon, 2009). What are the characteristics of those children who do acquire a vocabulary? Research suggests that the presence and quality of various social abilities (e.g., intentional understanding, responding to bids for joint attention) are key predictors of language outcome in children with autism, above and beyond the effects of IQ, mental age, or initial language ability (Adamson et al., 2009; Parish-Morris et al., 2007; Siller & Sigman, 2008). These results support the hypothesis that the capacity to access social cues is fundamentally important to successful language acquisition.

*Children can create a language: Natural experiments in language invention*

Language is such a resilient aspect of human nature that children, when faced with a situation in which language is not provided, will produce their own (Goldin-Meadow, Mylander, & Franklin, 2007). For example, deaf children born to hearing parents do not have access to auditory language input, and instead invent a “homesign” system of gestures to communicate with family members (Goldin-Meadow, 2007; Tervoort, 1961). These homesign systems become more complex with individual usage, but they never develop the characteristics of a full language (e.g., complex syntax, morphology, generativity; Tervoort, 1961). There are situations, however,
in which a full and independent language develops in a community of deaf people. Two recent examples come from a small Bedouin community in Israel and a school for the deaf in Nicaragua (Aronoff, Meir, Padden, & Sandler, 2008; Sandler, Meir, Padden, & Aronoff, 2005; Senghas, 2003, 2005; Senghas & Coppola, 2001; Senghas, Kita, & Özyürek, 2004; Senghas, Senghas, & Pyers, 2005).

Historically, the deaf population in Nicaragua was isolated and each individual communicated almost exclusively with hearing family members. In the 1970’s, however, deaf people from all over Nicaragua were pooled together in a newly established school for the deaf. At this school, they spontaneously began to create their own sign language (Senghas, 1995). Senghas and colleagues followed different cohorts of children as they entered the school over a number of years, and observed the evolution of Nicaraguan Sign Language (NSL) as it was passed from cohort to cohort. Interestingly, as NSL developed key features of language (such as syntax, spatial morphology, discrete and recombinable parts), those features were only learned and expanded upon by subsequent cohorts, and did not “filter up” to the first signers (who communicated via less evolved versions of NSL; Senghas & Coppola, 2001; Senghas, 2003; Senghas et al., 2004). Thus, language evolution was sparked by each new group of children entering the school, who took the existing form of NSL and developed it to another level; each time attaining more and more properties of an established language. These results are especially intriguing because they provide evidence to suggest that in a span of 20-30 years, children are capable of naturally creating well-formed languages whose basic foundations correspond to those of languages that developed over centuries. Over a slightly longer period of time in Israel, Al-Sayyid Bedouin Sign Language (ABSL) developed within a smaller deaf population (Aronoff et al., 2008; Sandler et al., 2005). Although ABSL exhibits some characteristics of a fully
developed language (e.g., word-order grammar; Sandler et al., 2005), it lacks others (e.g., spatial morphology; Aronoff et al., 2008). The reason for the apparently slower development of ABSL (70 years) versus NSL (30 years) has been hypothesized to be because new languages evolve most quickly when the language community is large (NSL has more speakers), when exposure to the sign language begins early, and when the language is transmitted to many subsequent generations or cohorts (Senghas, 2005).

Better late than never? When language input is delayed

Is there a “critical period” in development during which children must learn language? If this window of opportunity is missed, will the child ever acquire language? A large natural study of the critical period hypothesis recently resulted from advances in cochlear implant technology (see also Maurer & Lewis, this volume, for a discussion of critical periods).

It used to be that children born deaf remained so for the rest of their lives, or received some variable assistance from a hearing aid. Now, however, the technology of cochlear implants (CIs) drastically improves the hearing of many infants within the first few years of life. Recent research suggests that when children are implanted with CIs at or before 12 months of age, they catch up to their hearing peers by age two years (Tait, De Raeve, & Nikolopoulos, 2007). Another study of 76 children implanted with CIs showed that the language scores of children implanted before 24 months of age were to equal to the scores of hearing age-mates by the time they were 4.5 years old, which indicates a steep growth curve for children using CIs (Nicholas & Geers, 2007). Numerous studies demonstrate that as the age of implantation goes down, language proficiency goes up (Holt et al., 2004; Kirk et al., 2002; Miyamoto et al., 2003; Miyamoto et al., 2008; Svirsky, Teoh, & Neuburger, 2004). What does this evidence suggest about the existence of a critical period after which language cannot be acquired? As with bilingual development,
research on children using CIs suggests that although language can be learned after the toddler years, that earlier is better.

**X. Conclusion**

The roots of language development are in place before birth, and the journey from “coo to code” is multi-faceted. Children are born with a preference for listening to language, and are equipped with language-friendly perceptual abilities and statistical processing skills. They are attuned to social cues such as eye gaze, pointing, goal-orientation, and referential intent. They acquire linguistic rules in the absence of explicit instruction, and are able to learn more than one language simultaneously. When placed in a situation where fully developed speech input is unavailable, they even create their own language. The field of language development has made enormous progress in understanding the processes involved in language acquisition despite the fact that many of the children we study have yet to say a single word. The advent of new methodologies such as the non-nutritive sucking paradigm, the head turn preference procedure, and the intermodal preferential looking paradigm have offered us a window into how hard infants are working at discovering the units and meaning of their language even in the first year of life. The future is promising as newer methods such as near infrared spectroscopy (NIRS) are coming on line (Chen, Vaid, Bortfeld, & Boas, 2008) and we are beginning to ask important longitudinal questions about the relation between early behavior and later language ability. Language acquisition is an area of human development that inspires awe as infants continue to surprise us with their remarkable ability to learn.


*Infancy, 5*, 39–60.


Göksun, T. Personal communication, May 1, 2009.


Göksun, T., Roeper, T., Hirsh-Pasek, K, & Golinkoff, R.M. (in press). From noun phrase ellipsis to verb phrase ellipsis: The acquisition path from context to abstract reconstruction. In J.
Harris and M. Grant, (Eds.) University of Massaschusetts Occasional Working Paper in Linguistics 38: Recent Work on Ellipsis. Amherst, MA: GLSA.


(Eds.), Beyond names for things: Young children's acquisition of verbs. (pp. 185-222). Hillsdale, NJ: Erlbaum.


for Research in Cognitive Science summer workshop, Philadelphia, PA.


*Language Sciences, 10*, 69-88.


*Joint attention: Its origins and role in development* (pp. 103-130). Hillsdale, NJ: Lawrence Erlbaum.


Figure 2. Piaget’s water level task. It is correct to draw the water level as horizontal (parallel to the table). Bilingual children perform significantly better on this task than monolingual children (Bialystok & Majumder, 1998).