Precursors to language: The first 18 months of life

The field of child language became a highly popular and well-funded research area only after Chomsky (1965) galvanized the public imagination with his proposal of an innately available ‘language acquisition device’ – an image derived from the world of engineering. Chomsky’s now widely known position was based on the premise that, given the lack of evidence of language structure to be gleaned from listening to (error-filled) spontaneous speech, language is far too complex to be learned without some kind of specifically linguistic knowledge being built into the infant brain – that is, without an innate blueprint. In a succinct passage he spelled out his view:

A consideration of the character of the grammar that is acquired, the degenerate quality and narrowly limited extent of the available data, the striking uniformity of the resulting grammars, and their independence of intelligence, motivation, and emotional state, over wide ranges of variation, leave little hope that much of the structure of the language can be learned by an organism initially uninformed as to its general character...The real problem is that of developing a hypothesis about initial structure that is sufficiently rich to account for acquisition of language, yet not so rich as to be inconsistent with the known diversity of language. (p. 58; emphasis added)

Since that time a major concern of many linguists has been to establish, based on analyses of adult languages, what aspects of linguistic structure, or what organizational principles, are so basic as to merit being considered part of that ‘initial structure’, now generally referred to as ‘Universal Grammar’ (or UG; Chomsky, 1981). In the same period the field of child language has come to be increasingly polarized, largely due to differences in focus. In Braine’s inspired formulation,

a primary concern with the...task [of discovering what innate cognitive and linguistic primitives there are] almost inevitably makes one a nativist, and...a concern for the...task [of providing an account of ontogenetic development, including the origin of primitives] almost inevitably makes one dissatisfied with nativism. (1994: p. 10)

On the other hand, many linguists reject the notion that anything specific to language is part of a child’s biological endowment. Reasons for a blanket denial of the validity of Chomsky’s proposal are often rooted in reasoning from evolution, on the grounds that no known processes could have given rise to such prewiring as is implied by UG (this is the complementary concern identified by Braine, that of accounting for the origins – phylogenetic as well as ontogenetic - of ‘initial structure’: Lindblom, MacNeilage & Studdert-Kennedy, 1983); the polarization of research along nativist/non-nativist lines extends to the study of the evolution of language as well, however (see, for example, the range of positions reflected in Hurford, Studdert-Kennedy & Knight, 1998; Knight, Hurford & Studdert-Kennedy, 2000). In fact, linguists who are not primarily concerned with development but who take a more empirical, less mathematically or logically based approach to adult language – e.g.,

1 The same position was repeated more briefly but in similar terms ten years later: 'To come to know a human language would be an extraordinary intellectual achievement for a creature not specifically designed to accomplish this task' (Chomsky, 1975).
the large subfield of Cognitive Linguistics, including Construction Grammar (Langacker, 1987; Goldberg, 1995; Croft, 2001) and Usage-Based Grammar (e.g., Barlow & Kemmer, 2000; Bybee, 2001) – often take a lively interest in the findings of child language researchers, perhaps because of the centrality of the ‘innateness of language’ claim among nativists (see Elman et al., 1996). Among non-nativist approaches to child language development many distinct (but largely complementary) positions can also be identified, referred to by a range of labels such as ‘interactionist’, ‘functionalist’, ‘constructivist’ or ‘emergentist’ (e.g., MacWhinney, 1999; see Braine’s [1994] account of the contemporary irrelevance of a radical empiricist position such as that associated with philosophers of earlier centuries).

In this chapter we will follow Lindblom (1992), who suggested that it is not only more plausible (from an evolutionary point of view) but also a more fruitful research strategy to seek to understand how language could emerge in all normally developing children without any specifically linguistic endowment. We will set out what is known or understood concerning the biological and social foundations of language development, based on both purely observational and experimental studies of infants.

One goal for the field of language development is to arrive at an ecologically appropriate view of the infant and of the environment accessible to the developing infant brain – a brain that is in the process of maturing, learning, and integrating the various patterns or regularities to which it is exposed. An account of infant language learning that aspires to be truly developmental should also avoid predicating preformed adult-like linguistic or communicative goals for the infant. That is, there is no reason to believe that the infant is in any sense seeking to discover the structure of language, nor how to most efficiently communicate information, as sometimes seems to be implied in speech perception studies.

Infants do not really set out to learn language. Instead, they study the movements of faces and voices - the observable displays of talkers - and gradually accommodate to and reproduce these behaviors. (Locke, 1993, p. 8)

Indeed, the infant may be seen to be ‘innately guided’ to direct attention at first in certain biologically specified ways (Jusczyk & Bertoncini, 1988) – ways that are not, however, intrinsically linguistic. As Locke’s words suggest, such ‘guided attention’ leads to noticing sound patterns, rhythms and voices familiar from experience in the womb, to riveting the gaze on caretaker faces, and to rapidly developing familiarity with other sensory experiences available from birth (e.g., the smell of their own mother’s milk, familiar to an infant by 3 days of life at the latest: McFarlane, 1975).

Meeting the goal of arriving at a truly developmental account of language learning will also require advances in our understanding of the nature of infant learning. As Braine (1994) pointed out, ‘the more successful a scientist is in accounting for behaviour in terms of learning, the more powerful the innate learning mechanism that they are forced to posit’ (p. 10). That is, if we refrain from positing innate linguistic principles, we will need to posit, instead, learning principles that are strong enough to result in infants being able to register, represent, and creatively reproduce the complex, multilayered, hierarchical structure of any language(s) to which they may be regularly exposed in interaction with their caretakers.
There have been dramatic changes in our understanding of human learning mechanisms over the last several years (see Ellis, 1994). In particular, research into the separable mechanisms of implicit and explicit learning and access suggests that the infant is indeed equipped with the kind of power that a non-nativist account of language learning would require, in Braine’s terms. Neuroscience has played an important role in these advances, providing a far better understanding of the nature of memory, attention and learning in adults, among other things (see, for example, Rugg, Mark, Walla, Scholerscheidt, Birch & Allan, 1998; Baddeley, Conway & Aggleton, 2001; Ullman, 2001, 2004; Ellis, in press). Experimental studies of infants have been an important source of new understanding as well (Bates & Elman, 1996; Rovee-Collier, 1997; Gómez & Gerken, 2000). It is likely that advances in these two lines of research will increasingly come together in coming years to flesh out an account of language acquisition that requires no innate foreknowledge of the nature of language structure.

Each child’s move into language can be seen as an individual process that constructs or re-constructs the complex, multileveled system of language out of the minimal beginnings proper to a social organism, one designed to mature and learn within the context of the caretaker-child dyad. This chapter is designed to track what we know of the infant’s unconscious and unplanned path toward intentional communication, language comprehension, and speech production. In developing this profile we will provide the newborn with only those resources that have been found to obtain for the dynamically changing states of the human brain over the first year and a half of life. In short, we will start not from the basic premise that ‘all infants come into the world with linguistic skills’ (Pinker, 1994, p. 263) but instead from the working hypothesis that the ‘biological’ or instinctive component in language development is quite general, in accord with Sapir (1921):

Walking is an inherent, biological function of man. Not so language...Speech is a non-instinctive, acquired, ‘cultural’ function. (p. 4.)

Language, on this account, is one developmental consequence of the birth of human infants in a relatively premature state, which leads to their long period of helpless dependence on caretakers. Its ‘construction’ draws on the resources with which infants (and their caretakers) have been endowed over evolutionary time to prepare them to survive in and learn from their environment. Our review of the developmental precursors of language will include both the biological bases and the vocal, cognitive and communicative precursors to the uniquely human ability to serve as both experiencer/listener and communicator/speaker.

The development of linguistic form and function

Language can be seen as emerging out of a dual biological and social foundation in step with more general cognitive abilities (MacWhinney, 1999) – representational abilities that are unique to humans, the only species to have developed symbol use (Deacon, 1997). The basic ‘building blocks’ of language, according to this view, are twofold:

(1) The natural vocal, perceptual and cognitive endowments of the infant which, at birth, are not very different from those of other primates;
(2) the bonding and cultural learning that are the natural consequence of both biological biases and the intense social contact in which human infants are normally immersed over the long period of nurturing and development needed before they can survive on their own.

The past quarter-century or more of empirical study of the newborn and the maturing infant has provided us with a solid foundation for tracing early development. What follows here is an overview of advances that can be taken to move the child toward language learning, as schematized in Figure 1. The figure, which provides approximate chronological age-ranges, is divided into three strands of development in the first year, as follows:

1. those pertaining to FUNCTION OR MEANING (left-most column) trace the route from initial social responses, broad attentional capacities, and communicable needs and feelings to a developing ability to process experience and to initiate and control communicative situations (further distinguishing the ‘child as experiencer’ and the ‘child as actor’ or ‘communicator’);

2. those pertaining to VOCAL FORM (right-most column) trace the transformation of infant auditory biases and vocal capacities into the first recognition and production of word forms or phrases (further distinguishing the ‘child as listener’ and the ‘child as speaker’).

3. those pertaining to the LINKING OF FORM AND FUNCTION (central column) trace the emergence of the ability to grasp verbally encoded meanings and to conceptualize them for expression (‘child as experiencer/communicator’) as well as to distinguish and recognizably reproduce conventional verbal forms (i.e., words and phrases: ‘child as listener/speaker’). These abilities can be seen to emerge from the increasing links between advances in the processing and expression of meaning, on the one hand, and in the response to and production of vocal form, on the other.

Our account will be chronologically organized, but it is important to emphasize that the ages specified provide only a very approximate frame of reference. Individual variability in vocal and language development is considerable (Bates, Bretherton & Snyder, 1988; Lieven, Pine & Dresner Barnes, 1992; Vihman, 1993). Comparative empirical studies of children learning different languages are most informative when based on developmental milestones, such as lexical level, as determined from numbers of words produced in a session (Vihman & Miller, 1988), or cumulative vocabulary, based on diary records or on parental response to a questionnaire such as the MacArthur Communicative Developmental Inventory (Fenson et al., 1993), or syntactic level, as derived from numbers of words or morphemes per utterance, rather than on chronological age (Lieven, 1997). Furthermore, note that the developmental ‘stages’ or ‘phases’ that we trace here are cumulative, not sequential: Each set of attentional or vocal capacities is subsumed in the following set; the older child shows a wider range of abilities and considerable variability, exhibiting behaviours shown by younger infants as well as those of the current phase of development (compare the stages in the ‘emergence of the self’ as described by Stern, 1985).
Early capacities: I. Birth to two months

Birth to two months: The child as experiencer and communicator

A. Function: Infant experiencer

Stern (1985) notes that ‘starting from birth, infants regularly occupy a state called alert inactivity, when they are physically quiet and alert and apparently are taking in external events (Wolff, 1966)’ (p. 39). This state makes it possible for the infant to begin learning about the world through all of his or her senses; events that are regularly repeated will make a particularly deep impression on the developing brain. Beyond this, newborn infants appear to actively seek and evaluate experience (Bruner, 1977, Kagan, Kearsley & Zelazo, 1978), to the extent that the stability of their basic physiological state allows. Stern (1985) suggests that ‘learning itself is motivated and affect-laden’ (p. 42). That is, it is the emotional content of infant/caretaker interactions, the social grounding of the infant’s existence, that supports the learning process.

In a review of the considerable literature on cross-modal effects in the first month of life, (e.g., Meltzoff & Borton, 1979), Maurer (1993) finds that infants ‘seek out an optimal level of stimulation, summed across all sensory modalities’ (p. 111). Based on several studies, Maurer concludes that the apparent cross-modal matching of early infancy results from a lack of differentiation of the various sensory modalities in the first months, or synesthesia. The ‘optimal level’ sought out by infants is very low at first and is thus quickly reached with exposure to both sound and light, or touch (mouthing) and sight, for example. A principle of attention first to what is already familiar, then to what is novel applies from the earliest experiences (see Hunter & Ames, 1988, for a model and discussion of the move from familiar to novel effects with time of exposure and across developmental time).

Both faces and voices are potent attractors for infant attention, especially when they are in motion (Bushnell, 1979; Girton, 1979; Cooper & Aslin, 1990). A sketch of a face is a captivating visual target from the earliest months of life (Johnson & Morton, 1991); by 6 weeks infants are particularly drawn to study a face that is speaking (Haith, 1980). Findings such as neonatal preferences for faces (Goren, Sarty, & Wu, 1975), early recognition of the mother’s face and voice (Bushnell, Sai & Mullen, 1989; Field, Cohen, Garcia & Greenberg, 1984; DeCasper & Fifer, 1980), and a particular attraction to the eyes (Haith, Bergman & Moore, 1977) have led to the suggestion that humans have a neural specialization for social cognition (Locke, 1993).

B. Function: Infant communicator

Basic, rather global meanings are expressed for the greater part of the first two months, particularly signs of physiological needs and experiences relating to hunger, pain, and fatigue. Distinct categories of affective experience are not easily identified, although in this period infants do provide some sense of their overall ‘hedonic tone’ (globally positive or negative affect: Stern, 1985).
Birth to two months: The child as listener and ‘speaker’

A. Form: Infant listener
The auditory system of the foetus is complete in the last trimester of prenatal development; this permits passive experience of auditory stimuli already before birth, as has been demonstrated in studies based on direct acoustic recordings of the foetal sound environment and auditory responses (Lecanuet & Granier-Deferre, 1993; Querleu, Renard, Versyp, Paris-Delrue & Crispin, 1988). Experimental studies have yielded some of the postnatal familiarity effects that could be expected, given this prenatal experience. Most importantly, infants are found to respond with preferential attention to their mother’s voice over other female voices and to the ambient language over other languages (DeCasper & Fifer, 1980; Mehler, Jusczyk, Lamberz, Halsted, Bertoncini & Amiel-Tison, 1988; Hepper, Scott & Shahidullah, 1993).

B. Form: Infant ‘speaker’
Respiratory and phonatory control is sufficient at birth to permit the expression of pain or distress in cry (Wolff, 1969). In the first two months of life the infant’s expressive repertoire generally remains limited to ‘reflexive vocalizations’, including crying, fussing, and ‘vegetative sounds’ – a reflection of the infant’s immature control over such basic functions as breathing and digestion (Stark, 1980; see also Oller, 1980, 2000). Of these phonation types cry is perhaps the most directly related to speech, in that it provides exercise in the control of breathing, with relatively short in-breaths followed by the long outbreath needed to carry varied sound patterns (Lieberman & Blumstein, 1988).

Birth to two months: Linked form and function
The rise and fall of prosody, or speech melody (termed ‘the mimetic dimension of voice’ by Donald, 1991), can trigger direct and differentiated affective responses in the infant, especially when presented in the exaggerated register of baby talk, or ‘infant-directed speech’ (Fernald, 1991). It has been suggested that prosodic signals have a parallel in nonhuman primate communication systems and may be taken to reflect the most ancient vocal aspect of human language (Fernald, 1992). In terms of our tripartite model, this would provide a foundational capacity: a natural receptive link between (prosodic) form and (affective) function.

Early capacities: II. Two to four months

Two to four months: The child as experiencer and communicator

A. Function: Infant experiencer
Attention to objects is not in the neonatal repertoire but develops rapidly from two months on, when visual orienting and selection of visual foci begin to be coordinated with attention (Posner & Rothbart, 1981). Furthermore, in the early months the infant does not so much direct its attention as have it caught by salient events and entities, especially by faces (Ruff & Rothbart, 1996).

B. Function: Infant communicator
By the third month social meanings are present and even expressible, especially in the form of smiling and brightening of the infant’s face at the sight of a familiar face (a
key ingredient in what Trevarthan [1979] evocatively termed ‘primary intersubjectivity’; see also Stern, 1985). As Fraiberg (1977) put it,

During the first six months, the baby has the rudiments of love language available to him. There is the language of the embrace, the language of the eyes, the language of the smile, [and] vocal communications of pleasure and distress. (p. 29)

It may be worth noting that non-human infant primates and their mothers do not appear to engage in the mutual gaze that is so basic to human bonding (Plooj, 1979).

**Two to four months: The child as listener and ‘speaker’**

**A. Form: Infant listener**

The now well established ability of infants to respond in an experimental setting to change from one isolated speech syllable to another is generally based on experiments with two-month-olds, not newborns; at least one study has shown that the ability to discriminate vowels is available before the ability to discriminate consonants (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy & Mehler, 1988).

**B. Form: Infant ‘speaker’**

Vocal development is rapid in the first few months (Oller, 1980, Stark, 1980; Roug, Landberg, & Lundberg, 1989; Vihman, 1996, ch. 5). By two to three months oral vowel-like sounds are produced, often accompanied by incomplete velar closures effected by contact between the infant’s relatively large tongue and the soft palate (these are the ‘cooing’ or ‘goo’ sounds of popular description). The most speech-like sounds produced in this period are oral vowels; these can sometimes be elicited by adult production of similar sounds (‘vocal imitation’: Uzgiris, 1973; Kuhl & Meltzoff, 1988; Papousek & Papousek, 1989) and these are the vocalization types most likely to dominate in turn-taking episodes in which an adult times his or her speech contingently with that of the infant (Bloom, 1975, 1988, Bloom, Russell & Wassenberg, 1987).

**Two to four months: Linked form and function**

Among the reflexive or vegetative sounds generally reported for the first few months are grunts, the product of physiological changes associated with effort. In particular, Stark (1993) described such grunts in the context of *en face* communication, as a result of the physical effort involved for the infant in maintaining the head erect. The same involuntary vocal production has been noted in association with other postural adjustments, such as reaching (Trevarthan & Hubley, 1978) or crawling (Stark, Bernstein, & Demorest, 1993). This can be seen as another foundational sound-meaning link in production, though it is neither intentional, communicative, nor arbitrary – all defining characteristics of adult language.

**Early capacities: III. Four to six months**

**Four to six months: The child as experiencer**

The ability to recognize objects appears to depend at first on seeing them in motion; such recognition is evidenced by four months, whereas stable objects do not appear to be discriminated (under experimental conditions) until some months later (Kellman,
1993). At about 5-6 months infants replace their earlier focus on persons with visual and haptic exploration of objects (seeing and touching; alternate looking and mouthing is the typical pattern: Ruff & Rothbart, 1996). In part the developmental change is no doubt related to increased motoric capacities; additionally, however, there is a natural progression from stimulus-bound reflexive responses (0-3 mos.) to a period of ‘magnetic’ social interaction (3-6 mos), in which the infant is riveted to the caregiver’s face in frequent en face interaction. Over this time the child comes to be highly familiar with the caregiver’s face and expressive style; it is likely that this familiarization process is a factor facilitating the shift away from intense dyadic interactions, making it possible for the child to direct attention to other objects and events (Malatesta & Izard, 1984).

Four to six months: The child as actor
In this age period a sense of agency begins to form as well (Stern, 1985), as infants prepare motor acts before actually completing them, shaping hand opening and finger position prior to grasping an object, for example (Bower, Broughton & Moore, 1970). The sense of agency, as Stern spells it out, involves

three possible invariants of experience: (1) the sense of volition that precedes a motor activation, (2) the proprioceptive feedback that does or does not occur during the activation, and (3) the predictability of consequences that follow the activation...The presence of the motor plan as it exists in mind allows for the sense of volition or will. (p. 76f.)

Nevertheless, reaching and grasping are manifested in clumsy attempts at best at this age; they will be performed smoothly only from about six months on (Thelen, Corbetta & Spencer, 1996). According to Bruner (1973), this is the result of the initial action components being reorganized into a higher-order routine after a considerable period of practice.

Four to six months: The child as listener and ‘speaker’
A. Form: Infant listener
Experimental work has shown that in this period infants are sensitive to the prosody of clausal units, showing a ‘preference’ to listen to coherent clauses over clauses with arbitrarily placed pauses (Jusczyk & Kemler Nelson, 1996), although this is true only in child-directed speech (Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989). This sensitivity to well-formedness in the prosody of long units of speech (which are presumably well beyond the infant’s capacity for understanding) is an extension of the early ‘training’ in prosodic patterns experienced in the womb; prosody is the first aspect of language accessible to the infant.

B. Form: Infant ‘speaker’
In what has been termed the ‘expansion stage’ (Oller, 1980) the child explores the potential of the vocal tract, which has begun to take on a more adult-like structure (Kent, 1992; Vihman, 1996, ch. 5). Typical vocalizations include extremes of loudness and of pitch, resulting in ‘squeals’, ‘growls’, ‘yells’ and whisper, as well as friction noises, nasal murmurs and trills of various kinds.
First advances: Six to nine months

Six to nine months: The child as experiencer

Significant neurologically driven changes in attention have been noted in the second half of the first year. According to Ruff and Rothbart (1996), an early maturing ‘orienting/investigative’ attentional system comes to be supplemented by a later maturing system that underlies goal-oriented attention and control of complex activity. This attentional system is related to the development of frontal cortex and permits planned, goal-directed behavior. In part, the change appears to involve the addition of inhibitory processes, permitting greater flexibility in the voluntary deployment of attention (Tipper, 1992).

In this period the interpersonal ‘attunement’ or ‘primary intersubjectivity’ of the early months gradually begins to be transformed into a more outward-directed social interaction, involving adult, child and an object or event of mutual interest. Butterworth (1995) proposed that joint attention has its origins in an ‘ecological mechanism [which] enables a “meeting of minds” in the self-same object’ (p. 32):

What initially attracts the mother’s attention and leads her to turn is also likely, in the natural environment, to capture the attention of the infant...It is as if the change in the mother’s gaze serves as an orienting signal specifying the direction for the infant to look while the interesting object completes the communicative link with the adult to specify the position at which to look. (p. 32).

Six to nine months: The child as listener and ‘speaker’

A. Form: Infant listener

Six-month-olds have been found to listen with longer attention to lists of words in their native language (English) when it is contrasted experimentally with a prosodically quite different language (Norwegian), even when low-pass filtering is used to remove much of the segmental patterning from the acoustic signal; however, the contrast between English and a prosodically similar language, Dutch, was appreciated only at 9 months, and only on the basis of the full acoustic signal (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993). In short, infant knowledge of the prosodic patterning of the ambient language continues to develop over the first six months of life, although knowledge of segmental patterns (patterns involving particular consonant and vowel sequences) begins to show significant advances only in the second half of the first year.

B. Form: Infant ‘speaker’

The chief milestone in production in the first year is the sudden emergence, at about six to eight months, of the rhythmic speech-like consonant-vowel syllables known as ‘canonical babbling’ (Oller, 1980). It is a striking shift, readily observed by anyone looking for it who is in day-to-day contact with the infant. Furthermore, the change is robust, being found with little variation in age of onset even under a range of challenging conditions (including socio-economic deprivation, prematurity, or both: Oller, 2000). The emergence of easily recognized babbled syllables with adult-like timing in the middle of the first year appears to be maturationally based and fits into a
broader framework of rhythmic motoric advances that occur around that age (Thelen, 1981; Kent, 1984; Iverson & Fagan, 2004).

**Attainments of the first 9 months**

From a being whose attention during waking hours could seem to be very largely occupied by the struggle to handle the stresses of physical existence in the world outside the womb the infant has become a social being, capable of giving and receiving communicative signals and of voluntarily directing his or her attention, whether to persons, objects or events (although coordination of these various sources of stimulation is not yet generally possible). The infant whose disorganized movements led to no predictable effects has now come to move rhythmically and to grasp objects at will. Rhythmic vocal production has also succeeded the far less speech-like sound-making of the first months; this new capacity adds significantly to the infant’s perceptual experience, as his or her own adult-like syllables are now available for matching against the rapidly changing sound patterns of input speech.

In short, by 9 months the basic foundations have been laid for the infant to attend to and represent the world presented to his or her senses (social and material objects and events), to develop an understanding of the intentions of others and to produce the meaning-bearing phonetic forms of speech. These three strands will prove crucial for making the transition into symbolic or referential language use, but the tools are not yet fully formed at this point. They will be transformed by use or practice and experience over the three to six months that follow. The timing of the development of each strand differs from one child to the next, however, the ‘tipping over’ into the next ‘phase’, the transition into language proper, being dependent on the fulfillment of threshold criteria in each separate domain of development.

**Bringing the strands together: Nine to twelve months**

**Nine to twelve months: The child as expericer**

Butterworth (1995) and his colleagues found the first step in the establishment of joint attention to be in place in 6-month-olds, the youngest infants they tested. Using a more stringent operational definition, however, Corkum and Moore (1995) found that 10-month olds were the youngest infants to show spontaneous joint attention in the laboratory, though many eight-to-nine-month olds learned to follow the experimenter’s head turn toward an attractive moving object in the course of the 28-trial experiment.

Comprehension of manual pointing has been reported for the same age (Murphy & Messer, 1977; Leung & Rheingold, 1981). However, a longitudinal study found that no infants understood pointing before 12 months, and more than half showed understanding only in the second year (15 months: Desrochers, Morissette & Ricard, 1995). Finally, a longitudinal study of 28 infants engaged in spontaneous play in the home found that over the period 6-18 months infants showed increased episodes of ‘passive joint attention’ with mother (in which both parties ‘are actively involved in the same object, but the baby evidences little awareness of the other’s involvement or
even presence’; Bakeman & Adamson, 1984, p. 1281). Note, however, that it was only at 15 months that infants spent more than 5% of their time in ‘coordinated joint attention’ with mother (meaning that ‘the infant is actively involved with and coordinates his or her attention to both [mother] and the object...’, p. 1281).

The importance of the shift to an understanding of intentionality in others – or the ability to infer such intentions and to use them as a guide to understanding others’ behavior, including language behavior – is emphasized throughout P. Bloom’s engaging review of How Children Learn the Meanings of Words (2000). The timecourse for the development of this ability is not considered there in any detail. Nevertheless, it is clear that there is a need for development, in interaction with social as well as sensorimotor experience; ‘theory of mind’ (Bretherton, McNew & Beeghly-Smith, 1981) is not an ‘instinct’ available from the start. Bloom provides descriptions of ingenious experiments designed to pin down the nature of the child’s intentional attributions, showing, for example, that infants expect humans but not robots to have ‘intentional states’ (see Johnson, Slaughter & Carey, 2004). We have outlined a continuum in the emergence of evidence of this understanding, from a capacity to be entrained in joint attention, to comprehension of head turns, to stable participation in bouts of joint attention, to comprehension of pointing, to routine joint engagement with mother and an object. This continuum characterizes the child’s ‘first advances’ in the processing of meaning in what we may call the pre-symbolic period.

Nine to twelve months: The child as communicator

At about the same age at which a reliable capacity for joint attention begins to emerge (that is, when infants respond by following an adult head turn: 9-12 months), social referencing is also reported, in which infants will turn to look toward an adult under conditions of uncertainty (Feinman, 1992). Similarly, by this age an infant will turn to look back at the adult for visual feedback after following the direction of a point, showing ‘a deliberate attempt to validate whether the joint attention has been achieved, that is, whether the focus of attention is being shared’ (Stern, 1985, p. 129). Bates, Camaioni and Volterra (1975) described for this same time period the emergence of ‘proto-communicative behaviors’, in which infants begin to use both objects to gain adult attention and the help of adults to obtain objects. In the absence of language use infants have already begun to share their attentional states, their intentions, and affective states as well. Trevarthan and Hubley (1978) used the expression ‘secondary intersubjectivity’ to characterize these advances in dyadic communication. The spontaneous integration of attention to a social partner and an object of mutual interest – the manifestation of joint attention in the fullest sense – will become routine only later, however.

Many evident precursors to language have been noted in the last trimester of the first year. The child’s intent to communicate can first be inferred from the timing and use of gesture and vocalization, with pauses for the expected response, gaze alternation between adult and object, ritualization of the form of communicative gesture or

---

2 Means ranged from 19% to 23% in such a state, out of the 10-minute periods of recorded mother-infant observation
3 More formally, social referencing is ‘a process in which one person utilizes another person’s interpretation of the situation to formulate her own interpretation of it’ (Feinman, 1992, p. 4).
vocalization, and apparent frustration when communicative efforts fail (Harding &
Golinkoff, 1979). The basic communicative gestures - pointing, ‘showing’ (holding
an object out to an adult), and ‘giving’ (often with immediate retrieval) - are regularly
observed in individual infants before consistent word use has been established
(Vihman & Miller, 1988), typically between 10 and 14 months.

Pretend play demonstrates the child’s ability to construct and express meanings
through action, and provides a window on the process by which mental life gradually
expands beyond the constraints of immediate perceptual context (Rovee-Collier,
1990, 1997). The simplest kind of pretend play, typically seen by about 9 months,
involves relatively ‘concrete’ or ‘literal’ activities that mimic the child’s experiences,
using real objects or small replicas (pretend to drink from an empty cup, put brush to
hair, push toy car).

Nine to twelve months: The child as listener and ‘speaker’

A. Form: Infant listener

The attentional shift described by Ruff and Rothbart for visual attention has its
counterpart in a particular category of auditory perception, although the mechanism
remains unexplained. As demonstrated in a landmark study by Werker and Tees
(1984) and later replicated by Best (1994) using a different methodology, the
unbiased capacity for speech sound discrimination exhibited in the first months of life
is replaced at about 10-11 months by a narrowing of attention to the sounds contrasted
in the native language. It should be noted that this change occurs at about the same
time as the first evidence of infant familiarity with the segmental patterns of the
native language (Jusczyk, Frederici et al., 1993, 1994); the fact that attention to
segmental patterning in input speech is found only after the infant develops the ability
to produce adult-like segmental patterns (with canonical babbling) is probably not
coincidental – an idea to which we will return later.

An equally striking advance occurs in relation to word forms at about the same age: In
an experimental setting devoid of any relevant situational cues both French (Hallé &
Boysson-Bardies, 1994) and English 11-month-old infants (but not 9-month-olds: Vihman,
Nakai, DePaolis & Hallé, 2004) have been found to respond with longer
looks to lists of words likely to be familiar from the home than to phonotactically
comparable word forms that they were highly unlikely to know (e.g., bottle, nappy,
thank you vs. nettle, whacky, juncture). It is important to emphasize that this is word
form recognition, not comprehension. An Event Related Potentials study (ERPs)
using the same basic rationale – presentation of isolated word forms, out of context –
found evidence of discrimination of familiar from unfamiliar word forms at the same
age (Thierry, Vihman & Roberts, 2003). The time course for the differentiated brain
response to familiar vs. unfamiliar word types after onset of the stimuli was so short
(250 milliseconds) as to indicate that word form recognition must be automatic, i.e.,
an implicit brain response. This may be a useful corrective for the tendency to refer to
infant responses in the head-turn paradigm as ‘preferences’, with its implication that
the infant makes a voluntary ‘choice’ of what to attend to.

B. Form: Infant ‘speaker’

The first adult-like production pattern, canonical babbling, involves simple
‘mandibular oscillation’ (jaw opening and closing), typically with labial or dental stop
closure and low central vowels filling in the supraglottal phonetic ‘content’ for the
rhythmically carried CVCV skeleton or ‘frame’ (MacNeilage & Davis, 1990, Davis & MacNeilage, 1995). Once this articulatory framework is in place, infants are free to develop ‘favorite’ sound patterns, or ‘vocal motor schemes’, defined as ‘generalized action patterns that yield consistent phonetic forms (Piaget, 1952)’ (McCune & Vihman, 2001, p. 673). Children show significant differences in the number of such schemes developed within the first months after beginning canonical babbling; regular production of a variety of different consonants, in particular, appears to constitute the best prelinguistic index of phonological advance (Stoel-Gammon, 1992) and has been proposed as one of the ‘control parameters’ determining the timing of language onset (McCune, 1992).

It is likely that the development of vocal motor schemes (VMS) is also a factor in the narrowing of attention to native language contrasts at this time. Since the child necessarily is most familiar with his or her own articulatory output, the speech-like sound patterns represented by the child’s vocal motor schemes may serve as a kind of ‘top-down’ articulatory filter on the language patterns heard (Locke, 1986; Elbers & Wijnen, 1992; Vihman, 1991, 1993); those speech patterns would then be likely also to play a role in early segmentation of the adult speech stream (see Cutler, 1996, and Aslin, Woodward, LaMendola & Bever, 1996 for suggestions as to the other available cues). As the adult vocal patterns (words and phrases) associated with situations of interest in everyday life come to be increasingly salient to the child in the latter part of the first year (recognition of word forms: Vihman et al., 2004; Vihman et al., in press), an unconscious matching process between own vocalizations and adult input can be expected to lead to incipient familiarity with the repertoire of native language phones or phonemes; unfamiliar patterns, including phonetic sequences that have no close equivalent in the native language, would no longer hold the child’s attention.

In an experimental study designed to test the idea of an articulatory filter on speech perception, 12.5-month-old infants acquiring either English or Welsh were found to be influenced by their own vocal production patterns in listening to lists of nonwords: The nonsense disyllables featuring a consonant that was novel to the infant (one rarely produced) elicited significantly longer looking times (or headturns in response to the stimuli) than disyllables featuring well-practiced consonants, or VMS (the sounds tested were matched for adult input frequency in each language; the findings were based on correlation of individual looking times in response to the nonwords with individual child production frequency: Vihman & Nakai, 2003). The results demonstrated an effect of the infant’s experience with his or her own vocal production on perception of input speech. However, what was found was a novelty effect, not a familiarity effect. This may reflect the fact that the infants tested were already over a year old: By 12 months it is likely that they were so familiar with the consonants they themselves produced most frequently that relatively unpracticed consonants held their attention longer.4

The first imitative learning is reported for the end of the first year (Lewis, 1936), including the imitation of word forms (Tomasello, 1995) – although only a subset of children are found to imitate to any appreciable extent (REF). The child will now

---

4 See Vihman et al., in press, in which the effect size for a differential response to familiar vs. unfamiliar words rises month-by-month from 9 to 11 months and then declines at 12 months.
have been babbling, or practicing the production of adult-like syllables, for several months (McCune & Vihman, 2001); the range of possible patterns may be sufficiently varied to make adult identification of such imitated forms a likely outcome when they are attempted.

**Nine to twelve months: Linked form and function**

*A. Infant listener and experiencer*

For most children receptive word knowledge begins in this period, although in diary accounts word comprehension with the support of context is reported anecdotally as early as six months. Experimental studies place the beginnings of rapid learning of new words (‘recognitory-comprehension’) at about 12-14 months (Oviatt, 1980; see also Huttenlocher, 1974 and Benedict, 1979, for careful observational studies), with a sharply rising word-learning curve over the period 9-17 months.

*B. Infant ‘speaker’ and communicator*

A number of observers of infants in this age range have noted the use of ‘protowords’, or vocal forms used for a consistent function or meaning but with no apparent adult model (see the review in Vihman, 1996, ch. 6). Typically, the forms are simple consonant-vowel shapes, with glottals ([i] and [h]) and oral or nasal stops filling the consonant slot. Meanings are simple as well: broad markers of focus of attention or attempts to share interest or make a request; also expressions of emotion - excitement, disgust - or accompaniments to the child’s activity (another kind of focus), or simply conversational fillers (Vihman & Miller, 1988).

It has been observed that grunts commonly cooccur in this period with moments of focussed attention, involving quieting of the body, concentrated visual and manual exploration of an object or looking toward an event (Vihman & Miller, 1988; McCune, Vihman, Roug-Hellichius, Delery & Gogate, 1996; compare the criteria for coding focused attention in Ruff & Capozzoli, 2003, p. 879). These minimal vocal expressions, brief monosyllables of the form [V], typically with a neutral vowel [i] or a syllabic nasal, are produced quietly, with no apparent communicative intent; they appear to be related to the effort grunts observed contemporaneously as well as earlier and may be a direct physiological product of the ‘effort of attention’ (Porges, 1992; Richards & Casey, 1992).

The landmark event most enthusiastically heralded by families is the first identifiable word use. Those children who produce adult-based word forms in appropriate contexts in this period generally do so only in a restricted sense, however: The first words are characteristically ‘context-limited’, dependent on perceptual memory for their elicitation, narrowly attached to a single exemplar of a noun category (e.g., *kitty*), like a proper name in adult language, or involving ritualized responses embedded in adult/child routines (‘What does the doggie say? Woof!’; see Volterra, Bates, Benigni, Bretherton, & Camaioni, 1979; Vihman & McCune, 1994; McCune & Vihman, 2001). Thus the most noticeable seeming discontinuity in vocal expression, the first word, is heavily embedded in the social and action context of the child’s life and has few of the distinguishing features of adult language. It does not appear to be symbolic, nor is it part of a system of contrasting forms or meanings. The transition to symbol use or reference, which depends upon further advances in representational capacity and, in some cases, on maturation of the child’s understanding of the
potential of vocal signals for intentional expression, constitutes the second significant discontinuity in the infant’s communicative development.

**Transition to language: twelve to 18 months**

Bloom (2000, ch. 2) raises two basic questions about word learning: (1) Why do children start learning words at about 12 months and (2) why does the rate of lexical learning increase over time? For Bloom, word learning means having

1. a certain mental representation or concept
2. that is associated with a certain form…

Two things are involved in knowing the meaning of a word – having the concept and mapping the concept onto the right form (2000: p. 17).

Bloom makes no distinction here between receptive vs. expressive word learning – although the evidence generally adduced for or against the existence of a ‘lexical spurt’, for example, depends on word production, not comprehension. It is evident that word comprehension has already begun well before 12 months, although a notable acceleration is evident at about that age, when word production is also typically reported.

In answer to the first question Bloom emphasizes the development of an understanding of ‘referential intent’, or what could broadly be termed ‘pragmatic understanding’ (‘theory of mind’ is the grander-sounding expression that is currently fashionable). This would place a lower limit of about 10-12 months on the first word learning. Bloom also briefly considers a role for phonetic or phonological development and for advances in the domain in memory (‘perhaps very young children cannot store arbitrary form-meaning correspondences for long enough to be of any use’: p. 45).

To account for the increase in rate of word learning, Bloom mentions practice – possibly including phonetic practice, along with maturation of the faculties of memory and attention. Another factor in the increase in learning rate must necessarily be the increase in size of the child’s existing lexicon: Knowing some words facilitates learning more. This is true for at least two reasons: (1) Learning a few often used content words will aid segmentation of adult input, putting function words into relief as the ‘noisy bits’ in between known words (Brent & Siskind, 2001); (2) as has recently come to be understood, familiarity with phonotactic structure – a cumulative effect of lexical learning – supports new word learning (Storkel, 2001; Edwards, Beckman & Munson, 2004).

On the whole, however, Bloom leaves these questions open and further concludes that we have no real idea why there should be individual differences in language learning, nor even what the correlates of relatively fast or slow word learning may be. We will return to these issues later but will consider his first question in the remainder of this chapter.

We agree with Bloom and others that an understanding of communicative intentions is one of the key strands that must be in place before the shift to referential language use (see also Vihman, Macken, Miller, Simmons & Miller, 1985; McCune, 1992) – but this pragmatic prerequisite is only one of at least three critical strands (McCune
A second strand is the phonetic prerequisite: The child must be capable of producing consistent vocal forms at will – and production of a remembered word form must be sufficiently stable and target-like to lead to adult recognition of the child’s intention. Arguably, the limitations derive not so much from motor or articulatory capacity – which is less of a hurdle in the case of sign than spoken language (Meier & Newport, 1990) – as from phonological memory, or the capacity to hold in mind a novel sound sequence (or, in the case of sign, a sequence of manual postures and movements: see Mayberry, 1994) long enough to make a successful match with one of the motor patterns in the child’s production repertoire.

Finally, the child must have an understanding of and a capacity for representing symbols (Werner & Kaplan, 1963/1984; Bates, 1979). It is necessary first to come to recognize that a linguistic unit is a symbol, and to appreciate that this means that the arbitrary form-meaning relationship embodied in such a unit links recurrent phonetic sequences (word forms) not to specific referents but to a category of potential referents. That is, any one of a number of doggies, or depictions of doggies, may be intended when someone makes verbal reference to : This requires generalization of the child’s understanding of reference (Barrett, 1995). Additionally, in order to show productive symbolic or referential language use the child must be able to access a situationally appropriate form outside of any priming context (Vihman & McCune, 1994); this is the decontextualization requirement (Barrett, 1995). This in turn presupposes the capacity for voluntarily accessing or calling up memories of arbitrary form-meaning associations - which in itself requires a qualitative shift in remembering. These advances together can be taken to constitute the representational prerequisite.

The essence of reference can be stated simply for our purposes: A particular vocal pattern is conventionally linked with a particular intended meaning or referent. Recognition of this link has been termed the ‘nominal insight’ (Dore, 1978; McShane, 1979; Kamhi, 1986). Despite the fact that not all – or perhaps not even most – children show an identifiable ‘spurt’ in their learning rate, something like a ‘naming insight’ can be inferred from the qualitative change that can be observed in children’s use of words, or means of accessing words and phrases, in the period in which referential language use is first identified.

In what follows we will pursue the division into advances in function, form, and the link between the two that we show in Figure 1, but in each case we will combine the roles of recipient (experiencer, listener) and expressor (communicator, speaker), as the child is now able to move fluently between these roles. Word form recognition and speech production have become linked and what is learned (through listening) or practiced (through speaking) will affect both receptive knowledge and the representations that underlie productive expression. In concluding the chapter we will return to and elaborate briefly on the development of the three strands proposed above.

---

5 P. Bloom (2000) is undoubtedly right to insist that the often cited ‘lexical spurt’ has little general empirical support: See Ganger & Brent, 2004, who report finding only 5 such profiles out of 35 children followed longitudinally.
Twelve to 18 months: The child as experiencer/communicator

Representational prerequisites established
The development of the representational capacity needed to free the child from dependence on perceptual 'reminders' to call up previous experiences is most readily observed in spontaneous pretend play (McCune-Nicolich, 1981b; McCune, 1995). The earliest such play is seen toward the end of the first year, when infants engage in single play acts that reflect everyday activities. It is only in the second year that we find the child moving from single acts involving the self to activities involving others (feed Teddy bear, brush mother’s hair). In this same period we can sometimes observe the child enacting such activities in a more abstract form, using 'substitute objects' (putting a shoe in the oven to bake [Velten, 1948]; watering flowers with an empty milk carton).

McCune (1995) applies the term ‘combinatorial symbolic games’ to the still more sophisticated type of pretend play typically not seen before 14 to 18 months. This play level, in which the same pretend scheme is enacted with more than one recipient (feed mother, then doll) or a sequence of acts is performed with the same recipient (put blanket on doll, rock doll), appears to reflect the child’s achievement of the minimal representational capacity needed for the transition to symbolic use of language.

Twelve to 18 months: The child as listener/speaker

Phonetic prerequisites established
At some point in the period 9 to 16 months most – although not all – children develop skill in producing at least two or more supraglottal consonants (typically, stops, fricatives, or nasals). McCune and Vihman (2001) found that achievement of this minimum level of phonetic skill, identified on the basis of consistent use in babbling or words in at least three out of four consecutive monthly recording sessions in the home, was critical to the onset of referential word use. They argued that ‘stable production control of two or more different consonants allows the child more readily to attend to and recall adult word forms and their associated meanings across different contexts’ (p. 680f.). Facility in the production of a range of distinct sound patterns has been found to promote the consolidation of new word learning in adult second-language learners and in older children (Gathercole & Baddeley, 1989, 1990; Ellis & Beaton, 1993; Storkel, 2001). It may also be familiarity (as a result of production practice) with increasing numbers of distinct sound patterns that leads to a finding reported by Werker, Fennell, Corcoran & Stager (2002): 14-month-olds with vocabularies of over 25 words were found to be able to discriminate a newly learned nonsense word from a minimally different one on which they had not been tested, while children of the same age with smaller reported vocabularies could not (although 17- and 20-month-olds were able to discriminate the minimally different word forms regardless of vocabulary size).

A new systematicity in production appears to accompany the onset of referential word use, possibly supported by growth in representational capacity (Vihman, Velleman & McCune, 1994). Some children develop clearly recognizable word production patterns or ‘templates’, which are used at first for adult words that they closely resemble ([beibi] for baby, [dadi] for daddy, [hai] for hi), but that are then extended to other words that the child ‘assimilates’ to the familiar and ‘pronounceable’ pattern ([ti:ni] for clean, [pâi] for bang: See Vihman, 1996, Ch. 6 and Appendix C). This use
of a well-practised pattern to produce a growing number of words typically results in ‘regression’ in the sense that words may become less accurate, but it is undoubtedly also a factor in the rapid lexical growth commonly reported for this period. The nature of phonological development in this period will be considered in greater detail in Chapter X.

_Twelve to 18 months: The child as both experiencer/communicator and listener/speaker_

**Pragmatic prerequisites established**

Beyond the ability to infer referential intent and to produce broadly communicative signals (point, show, give, as well as stylized facial expressions and body language), seen already by 9 to 12 months, a specific pragmatic prerequisite that emerges only in the second year is the understanding that _conventionalized vocal means_ can be used to communicate. In their study of five infants McCune et al. (1996) traced the evolution of function in grunt vocalizations produced between 9 and 16 months. The communicative grunts (which appear to correspond to the ‘call sounds’ mentioned by Werner & Kaplan, 1984 [1963] as vocal precursors to referential language) were identifiable as being produced more loudly, often repeatedly, and with evidence of communicative intent similar to that found to accompany pointing, showing and giving already in earlier months. A sharp increase in such communicative grunt use was followed closely in each case by referential word production or, in the case of infants who had not yet developed the necessary phonetic resources for identifiable word use, by gestural evidence of referential comprehension. The following observation illustrates the onset of such generalized understanding in a 13-month-old English child M with command of few consonants and little or no word production:

M’s mother points out and names _shoe_ as she reads him a book, then points to the child’s shoe; M responds by pointing to the picture of _shoe_ himself and then to the shoe on _his mother’s_ foot. (T. Keren-Portnoy, p.c., July 2004)

McCune et al. interpreted the developmental sequence in grunt function from physiological effort to attentional effort to communicative expression as reflecting a process identified by Darwin (1965 [1872]): Purely expressive acts derive signal value from earlier uses that are biologically functional. In this case the child may associate the low-intensity vocalization that accompanies moments of focused attention – the grunt or ‘quasi-vowel’ (proposed as a ‘launch-point for speech-like vocalizations’ in one evolutionary scenario: Oller, 2000, p. 355) – with an internal experience of personal meaning, that is, the child’s experience of interest in the object or event that stimulated expression of the attentional grunt. This sets the stage for the transformation of the involuntary vocalization into a voluntarily produced vocal signal. As intentional communication becomes established through gestural means at 9 to 12 months and as the capacity for representing events and sound patterns expands early in the second year, the child comes to use the grunt to communicate his or her attentional focus to others. This understanding of the potential for vocal expression of internal meanings, then, goes beyond the earlier uses of pointing and other nonverbal gestures for communicative purposes and appears to provide a more specific bridge into adult-like language use.
‘Reference functions to select an environmental or mental entity or event for joint attention and conversation’ (McCune & Vihman, 2001, p. 671; cf. Bruner, 1983; Macnamara, 1982; Moore & Dunham, 1995). The notion of ‘reference’ has been defined broadly, to include virtually any conceptual referent (as in Lyons, 1968, where reference is said to be ‘the relationship which holds between words and the things, events, actions and qualities they ‘stand for’: p. 424), as well as in a more restrictive way, to include only nominals: ‘Reference is the contact language makes with the environment; it is the device that enables us to talk about the things we see and touch.’ (Macnamara, 1982, p. viii). This is too narrow a definition for our discussion of the emergence of referential language use, however. ‘Referring expressions’ in adult language – the identified topics of discourse - may typically be nominal, but this is a grammatical convention with no necessary conceptual basis. Any perceptually accessible event – ‘a whiff of perfume’, ‘the collapse of the bridge’, ‘the sound of chimes’ or ‘the flush that spread over her cheeks’ – can be referred to by a nominal expression in English, yet none are prototypical examples of ‘things we see and touch’. We use the term ‘referential language’ here to mean language use that goes beyond routines and narrowly understood words or words produced only as part of the actions they accompany ([Estonian] pai-pai ‘nice-nice’, said while patting the cat [Vihman, 1996], [Italian] bam ‘boom’, said while knocking over a tower of blocks [Bates, 1979]; Barrett, 1995, provides more examples). Referential language is the generalized, flexible deployment of words to pick out a range of different instances of a single meaning-category – whether or not the usage shows a close correspondence with that of adults (e.g., [Estonian] habe ‘beard’ to refer to a picture of a bearded man, the child’s own bearded father, and a bearded playing-card king: Vihman, 1996, p. 138).

The earliest forms of linguistic reference noted by Werner & Kaplan (1963) were learned expressions (words or phrases) that focus on either (a) entities in the environment (object words, or nominals) or (b) relational aspects of events in which objects and persons participate (relational words, e.g., allgone, more, out, uh-oh, up). In their longitudinal study of 20 children recorded at least monthly from 9 to 16 months McCune and Vihman (2001) found that first referential word use could be identified for 13 of the children by the end of the study (between 13 and 16 months for these children); for the remaining children no referential word use was observed by 16 months. Both ‘general nominals’ or common nouns, used to refer to a range of entities relatable by some form of category membership, and relational words, or words belonging to several adult parts of speech but used by the children to refer to dynamic and reversible spatial and temporal events across a variety of contexts (McCune-Nicolich, 1981a; Vihman, 1999; McCune, in press), were considered to be ‘referential’. The nine children who made the transition by 14 months were found to produce twice as many referential words within the two months that followed their first referential use (based on use of at least two distinct referential word types), evidence of a qualitative leap in word learning and use. [refer again to ‘lex spurt’ issue?]

Summary: Precursors and the transition into language

We can now summarize the course of development from birth to the transition to language as we have depicted it. We have posited four distinguishable phases. In the first, the child’s early capacities can be seen to unfold at the same time that rapid neurophysiological changes are taking place. In the second phase, that of the first
**advances**, the child moves from a period of intense social interaction to a new interest in and facility with object manipulation. At the same time, the first adult-like syllable production is heard – the one major speech production landmark of the first year. In the third phase, bringing the strands together, these new vocal abilities (production of adult-like syllables) and new attentional and intentional capacities (combining the child’s interest in objects and in people in the context of episodes of joint attention) are exercised, resulting in the production of identifiable words in context (‘primed word production’) by the more vocally precocious vocal infants. And in the final phase, the transition to language, the now well practiced and thus more familiar and better differentiated vocal abilities are joined to two other emergent kinds of knowledge: (1) dawning awareness of the communicative potential of vocal signals and (2) widening capacities for representation and memory. It is the coming together of these initially unrelated advances in form, meaning, and the link between the two that yields the first flexible and symbolic use of words, or referential language. This shift, then, constitutes a kind of discontinuity within the continuities of both form and meaning that we have traced. Once all the necessary components are in place - requisite neurological support systems and experience with both communication and vocal production, the child seems to actively seek out the word patterns that fit situations of interest to him or her, and to extend the phonological patterns implicit in the words he or she is already producing to new word shapes while at the same time extending the meanings of words beyond particular referent objects and events. Systematic language use - which we might define as the context flexible deployment of arbitrary but conventional, or adult language-based, sound-meaning pairs with voluntary communicative intent - has begun.

We have reviewed the strong biological and social foundations that underlie this development, including the initial perceptual biases favoring the mother’s voice and language, the early receptivity to the affective meanings expressed in speech melodies, and above all the attention to faces and the bonding to caretakers that grows out of the infant’s immediate needs, along with the natural (‘biologically prepared’) responses of the caretakers [REF?]. At least three different kinds of processes can be taken to be responsible for the changes that follow these initial states of child and environment. First, imitation, the ‘common sense’ explanation for language learning and the mainstay of early behaviourist theories, plays a vital role: Where else would the forms of language come from? Secondly, neurophysiological change, the ‘expected’ maturational course of the human infant, underlies some of the behavioural changes we observe, including the shift from the production of vowel-like sounds alone to the first adult-like syllables and also the narrowing in and increasing voluntariness of attention. Neither of these critical maturational changes is wholly internal to the child, however; both depend on a supportive environment, which supplies models for the jaw movements of speech and dynamic ‘attractors’ in the form of social and physical events that hold the infant’s attention and expand representational abilities. Thirdly, and more and more clearly as the scope of the child’s capacities and behavioural choices increases and becomes more evident to the observer, each infant can be seen to integrate vocal and attentional capacities in his or her idiosyncratic way, creatively co-constructing language out of repeated communicative interaction. This is the individual cognitive stamp each child brings to learning; this is one major source of the ‘individual differences’ that are so much in evidence in every aspect of vocal and linguistic development.
Three essential developments fuel the transition to reference. (1) Representational advances make it possible to begin to map early perceptual understandings into the simplest verbal expression of reference to entities and to spatial and temporal events. (2) Emergent vocal motor control, together with the increase in representational capacity, makes possible the organization of the simplest consonantal gestures (vocal motor schemes and precanonical phonetic resources primarily) into one or more stable phonological patterns or templates. (3) Pragmatic insight into the communicative potential of consistent sound-meaning correspondences, which is biologically based yet rooted in social experiences, provides the child with a critical impetus for active engagement in the verbal expression of meaning.

We see the transition to reference as constituting a bridge from prereferential to linguistic structure. In the earlier period, a relatively unsystematic or ‘piecemeal’ accumulation of knowledge or skills can be identified both in the conceptual (or semantic) and the vocal motor (or phonetic) domains. Both semantic and phonetic structure are already emergent in that period, however, as evidenced by the fact that ambient language influence is detectible in the earliest uses of words (Choi & Bowerman, 1991, Gopnik & Choi, 1995) and phonetic categories (Boysson-Bardies & Vihman, 1991, Vihman & Boysson-Bardies, 1994).

The shift to reference involves continuing interaction with a particular ambient language, which allows the child to further develop lexicosemantic categories while at the same time effecting the transformation of perceptual into conceptual meanings (Piaget, 1962; Chapter 8). In the domain of verbal expression or form, interaction with the adult language allows the child to begin the systematization of phonetic patterns or structures that constitutes the first step toward phonological organization. Once incipient phonological and semantic systems have been constructed the child can begin to attend to or be influenced by the more complex linguistic organization implicit in the verbal input with which he or she is engaged.
References


### EARLY CAPACITIES

<table>
<thead>
<tr>
<th>FUNCTION or MEANING</th>
<th>LINKED FORM and FUNCTION</th>
<th>VOCAL FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>From birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHILD AS EXPERIENCER</td>
<td>Attention is caught by salient perceptual events: Moving faces and voices</td>
<td>CHILD AS LISTENER</td>
</tr>
<tr>
<td>CHILD AS COMMUNICATOR</td>
<td>Expresses physiological needs and global feeling tone</td>
<td>CHILD AS 'SPEAKER'</td>
</tr>
<tr>
<td><strong>From two months</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHILD AS EXPERIENCER</td>
<td>Begins to selectively attend to objects</td>
<td>CHILD AS LISTENER</td>
</tr>
<tr>
<td>CHILD AS COMMUNICATOR</td>
<td>Smiles, frowns: Social responses ('Primary intersubjectivity')</td>
<td>CHILD AS 'SPEAKER'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>From four months</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHILD AS EXPERIENCER</td>
<td>Recognizes objects (based mainly on motion)</td>
<td>CHILD AS LISTENER</td>
</tr>
<tr>
<td>CHILD AS ACTOR</td>
<td>Explores objects (alternately looks, mouthes, touches)</td>
<td>CHILD AS 'SPEAKER'</td>
</tr>
<tr>
<td><strong>FIRST ADVANCES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>From six months</strong></td>
<td>CHILD AS EXPERIENCER</td>
<td>CHILD AS LISTENER</td>
</tr>
<tr>
<td></td>
<td>Attention becomes more voluntary and flexible</td>
<td>CHILD AS 'SPEAKER'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BRINGING THE STRANDS TOGETHER</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

31
### Between nine and 12 months

**CHILD AS EXPERIENCER**
- Participates in episodes of joint attention (passively at first)

**CHILD AS COMMUNICATOR**
- Communicates intentionally (‘secondary intersubjectivity’): ‘Shows’, points, ‘gives’
- Engages in single pretend play acts

**CHILD AS LISTENER**
- Begins to show word comprehension (in context only at first)
- No longer discriminates consonantal contrasts not in ambient language
- Responds to familiar word forms (even out of context)

**CHILD AS ‘SPEAKER’**
- Grunts with effort of focal attention
- Produces words in priming context
- Imitates word forms

**“Phase shift”: representation, intentional communication, phonetic capacity**

### TRANSITION TO LANGUAGE

**Between 12 and 18 months**

**CHILD AS EXPERIENCER/COMMUNICATOR and LISTENER/SPEAKER**

**REPRESENTATIONAL PREREQUISITES**
- Engages in combinatory pretend play acts
- Uses ‘substitute’ objects (symbols)

**PRAGMATIC PREREQUISITES**
- Uses grunts to communicate (influence other’s focus of attention)

**PHONETIC PREREQUISITES**
- Achieves stable production control of two or more VMS
- Develops consistent word forms (“templates”)

**REFERENTIAL LANGUAGE USE**
- Understands that words refer to categories
- Uses words as symbols