Developmental Science 4:1 (2001), pp 1–37

ARTICLE WITH PEER COMMENTARIES AND RESPONSE

The intersensory origins of word comprehension: an ecological–dynamic systems view

Lakshmi J. Gogate,[1] Arlene S. Walker-Andrews[2] and Lorraine E. Bahrick[3]

1. The State University of New York at Brooklyn, USA
2. Rutgers University, USA
3. Florida International University, USA

The learning of language by the child is not simply the associative naming or labeling of impressions from the world. It is also, and more importantly, an expression of distinctions, abstractions, and recognitions that the child is coming to achieve in perceiving.

(J.J. Gibson, 1966, p. 281)

Abstract

How do infants begin to understand spoken words? Recent research suggests that word comprehension develops from the early detection of intersensory relations between conventionally paired auditory speech patterns (words) and visible objects or actions. More importantly, in keeping with dynamic systems principles, the findings suggest that word comprehension develops from a dynamic and complementary relationship between the organism (the infant) and the environment (language addressed to the infant). In addition, parallel findings from speech and non-speech studies of intersensory perception provide evidence for domain general processes in the development of word comprehension. These research findings contrast with the view that a lexical acquisition device with specific lexical principles and innate constraints is required for early word comprehension. Furthermore, they suggest that learning of word–object relations is not merely an associative process. The data support an alternative view of the developmental process that emphasizes the dynamic and reciprocal interactions between general intersensory perception, selective attention and learning in infants, and the specific characteristics of maternal communication.

Introduction

A first step in word learning may involve the discovery that the continuous speech stream contains smaller units or words (Jusczyk, Cutler & Redanz, 1993; Jusczyk & Aslin, 1995; Saffran, Aslin & Newport, 1996). Furthermore, the process of understanding words may entail the realization that unique spoken words (nouns or verbs) are conventionally paired with and refer to perceivable objects or actions in the environment (for a review see Sullivan & Horowitz, 1983). Infants and young children in a language community must recognize that (a) heard speech (words) and seen objects or actions are related, (b) specific words function as symbols for objects or actions, and (c) such words denote specific objects or actions. We propose that the development of the first ability, in particular, has its origins in the interaction between infants' intersensory perception and the specific qualities of adult multimodal communication to infants during the first year (Sullivan & Horowitz, 1983; Zukow-Goldring, 1997; Gogate, Bahrick & Watson, 2000). Focusing primarily on the infant prior to symbolic–lexical development, therefore, we suggest that the development of lexical comprehension is the result of a complex ongoing reciprocal interaction between the organism (the infant) and the environment (the language addressed to the infant). This thesis is consistent with the ecological (J.J. Gibson, 1986) and

Address for correspondence: Lakshmi J. Gogate, Psychiatry, PO Box 1203, SUNY at Brooklyn, 450 Clarkson Avenue, Brooklyn, NY 11203, USA; e-mail: lgogate@netmail.hscbklyn.edu

dynamic systems approaches to development (Thelen & Smith, 1994; Lickliter, in press).

How could the infant's intersensory perception, a general capacity, facilitate the development of lexical comprehension during infancy? The current proposal contrasts sharply with the more popular view of innate language-specific processes (e.g. Chomsky, 1980; Fodor, 1983) and challenges an early statement by Chomsky (1980): 'multi-purpose learning strategies are no more likely to exist than general principles of "growth of organs" that account for the shape, structure, and growth of the kidney' (pp. 138–139). We argue that infants detect invariant, amodal or redundant information in bimodal non-speech and speech events (e.g. the temporally synchronous sight and sound of a bouncing ball, or the mouth movements and vocalizations of a person) very early in life (E.J. Gibson, 1969). According to E.J. Gibson, perception is initially global and becomes increasingly differentiated over time. Infants' general sensitivity to redundant information in bimodal events enables them to detect such relations in arbitrarily linked patterns of audible and visible stimulation, including words and objects or actions. The central role of perception in early lexical comprehension, in particular the perception of temporal contiguity between words and referents, is acknowledged in past research (e.g. Sullivan & Horowitz, 1983; Harris, 1992; Budwig, 1995; Bloom, 1998). However, the studies reported here are the first to show that, around 7 to 8 months, preverbal infants tune in to more precise redundant information (temporal synchrony), to learn the relation between a word and its referent.

We also propose that maternal communication to infants is abundantly endowed with the very information that infants initially detect. Mothers tend to embed words in multimodal events tailored to the perceptual skills of their infants. For example, mothers temporally synchronize their naming and showing of objects to infants when infants most require temporal synchrony to detect arbitrary speech–object relations (Gogate & Bahrick, 1998; Gogate *et al.*, 2000; Gogate, under review; also see Zukow-Goldring, 1997). Perceptual abilities in infancy dovetail with environmental stimulation to promote the development of lexical comprehension. As will be reviewed, experimental research on the development of intersensory perception offers data to explain how infants learn to pair words and objects. Descriptive studies of maternal communication to infants point to ways in which the environment is organized to facilitate infants' word comprehension. The infant's perceptual and environmental (maternal language) systems change with time, and the co-action between these changing systems enables progress in infants' detection of intersensory information and, eventually, comprehension of words in the language. Because the two systems are fluid, the changes in one system are probably driven by a combination of constantly changing dynamics (instabilities) within and external to the system. Furthermore, the systems are self-organizing (i.e. instabilities give rise to temporary points of stability during development). Therefore, the developmental changes of these systems are constrained locally (not by a pre-program) by the properties of each (e.g. developmental factors in the child) and the real-time interactions between them (see Thelen & Smith, 1994, p. 36). The salience of any given parameter over others is a function of its relative weight (e.g. the salience of temporal synchrony may be a function of infants' attention). The environment in a complementary fashion affords different parameters interacting with the infant's developmental abilities. Because of the close fit between the two systems, the changing nature of one fits the changing nature of the other.

In general, the focus of this paper is on the process rather than the end-state of development. Many earlier cognitive accounts of early word learning have suggested that infants simply associate or 'map' words with perceivable objects and events (Ellis & Wells, 1980; Bates & MacWhinney, 1987; Stemmer, 1989). These accounts do not explain exactly how mapping takes place. Other accounts of word learning emphasize the 'association between words and the functions of words in the situation in which the child hears them' (Bloom, 1998, p. 322; Budwig, 1995). These accounts do not consider organismic development (see Bloom, 1998). In contrast, adhering to the ecological and dynamic systems views, we describe and emphasize the interaction between the organism and its environment to show *how* infants learn to relate words and referents. Furthermore, we attribute much more to the process by which infants detect word–object relations than mere association. In our view, infants seek out and attend to the affordances of the environment (the very environmental properties that facilitate the discovery of word–object relations; see J.J. Gibson, 1966). They do not simply 'associate' idiosyncratic patterns of stimulation available to different modalities whenever temporal contiguity or spatial co-location exists. Rather, they detect intermodal invariants, redundant information in patterns of stimulation, especially as their attention becomes educated by both experience and maternal scaffolding. Finally, unlike some modular approaches to language (e.g. Chomsky, 1980; Fodor, 1983), we suggest that language-specific systems are not required to account for the detection of arbitrary relations between words and referents (see Dent, 1990; Bates & Elman, 1996). Infants use the same redundant information, such as temporal synchrony and intensity shifts, to detect arbitrary relations in bimodal speech and non-speech events. Furthermore, there are close parallels in the developmental course of each domain.

Perceptual and lexical development during infancy

Intersensory integration and infants' detection of intersensory redundancy

Prior to the emergence of research findings showing intersensory integration in human infants, J.J. Gibson, Bower and others hypothesized that the senses are integrated or undifferentiated at birth (J.J. Gibson, 1966; E.J. Gibson, 1969; Bower, 1974). According to E.J. Gibson, in particular, perceptual development consists of increasing differentiation and organization. The developmental process is manifested behaviorally in the early detection of redundant information and the later detection of information that is specific to a given modality. Perceptual information such as temporal synchrony, intensity shifts, rhythm and tempo are redundant or 'amodal' because the same information can be perceived across two or more modalities. Recently, the dynamic systems view of development has underscored the importance of intersensory perception for the development of general cognition including language (Thelen & Smith, 1994). In this section, and in the next, we will illustrate how the infant's initial undifferentiated perception, particularly in the auditory and visual domains, enables the perception of arbitrary relations between heard words and seen referents.

A wealth of findings on early intersensory organization across avian and mammalian species provide insights into its phylogenetic and ontogenetic development in humans. Newborns across many avian species detect redundant information across modalities, suggesting a possible innate basis for this ability. Changes in the prenatal environment applied to one sensory modality can influence the postnatal development of a different modality, indicating that the senses are integrated and undifferentiated early in development (Lickliter, 1990a, 1990b; Lickliter & Stoumbos, 1991). These effects of early experience can be seen across tactile, vestibular, auditory and visual modalities in bobwhite quail chicks (Lickliter & Banker, 1994) and in ducklings (Gottlieb, Tomlinson & Radell, 1989; Gottlieb, 1993). Furthermore, the fetal and newborn brains of some animals (e.g. owls) exhibit more plasticity in response to intersensory stimulation than do the brains of adult con-specifics, again presenting evidence for early intersensory integration (Knudsen, 1983; Stein, Meredith & Wallace, 1994).

Similar to their avian counterparts, human neonates show evidence for an integrated multisensory system. For example, human neonates can detect the common shape of a pacifier, and recognize an object's substance across visual and haptic modalities (Meltzoff & Borton, 1979; E.J. Gibson & Walker, 1984; Kaye & Bower, 1994). They can also imitate another person's facial motions (Field *et al*., 1983; Meltzoff & Moore, 1989). With respect to the visual and auditory modes, neonates detect changes in synchrony relations between the sound and sight of objects in motion as early as 4 weeks (Bahrick, 1996). These findings support the view that for humans the senses are integrated early during development and probably lay the foundation for continued intersensory perception across auditory and visual domains, including bimodal speech.

A majority of findings on infants' sensitivity to redundant information in bimodal speech has come from older infants and often cross-sectional approaches. Nonetheless, some developmental patterns can be identified. For example, between $2\frac{1}{2}$ and 4 months, infants attended more to the synchronous than the asynchronous visible lip movements and audible speech patterns of a person reciting a nursery rhyme (Dodd, 1979). Other studies have established that infants detect redundant information in acoustic or phonemic segments, suggesting that infants attend to more specific aspects of bimodal speech. Four-month-old infants matched the visual mouth shape of a person with a corresponding vowel sound across several vowel pairs, in the presence of temporal synchrony between the lip movements and the vowel sounds (Kuhl & Meltzoff, 1982, 1984, 1988). These data suggest that young infants might detect correspondences between the visual mouth movements and vocalizations of adults at a relatively global level, progressing later on to abstract finer correspondences such as specific speech sounds and mouth shapes. In addition, infants detect intensity shifts in bimodal speech across auditory and visual modalities. At 5 months, infants detected the common intensity shift between a gradually opening mouth and an increase in the vowel sound amplitude of spoken disyllables consisting of consonants and vowels. They also detected the common intensity shift between a gradually closing mouth and a decrease in the vowel sound amplitude of these disyllables (MacKain, Studdert-Kennedy, Spieker & Stern, 1983).

Infants perceive redundant temporal synchrony and intensity shifts across auditory–visual non-speech events as well. Around 3 to 4 months, infants selectively attended to a specific object hitting a surface, based on the temporal synchrony between the object's motion and the timing of its sound of impact (Bahrick, 1983, 1987, 1992). Infants also matched approaching and receding objects with an increase and a decrease, respectively, in their sound amplitude (Walker-Andrews & Lennon, 1985; Schiff, Benasich & Bornstein, 1989; Pickens, 1994). Furthermore, at this age infants matched auditory–visual events based on the common (redundant) rhythm and tempo (Spelke, 1979; Bahrick & Lickliter, 2000). These parallel findings across speech and non-speech domains suggest that infants may use general perceptual processes for detecting redundant information in speech and non-speech events.

In summary, the integrated perceptual system is adept at detecting redundant information across different senses by 3 to 5 months of age, if not earlier (Figure 1). How can the infant's general ability to abstract redundant information from auditory–visual events contribute to later word comprehension? The ability to analyze a fleeting array of complex auditory–visual stimulation into coherent signals will be necessary to attach meaning to appropriate words. For example, many similar sounding words in a given language, such as /pin/, /tin/ and /bin/, which differ by a single phoneme only, can refer to different objects and convey very different meanings. For the language learning infant, the detection of redundancies between several speech sounds and their corresponding lip movements or mouth shapes could serve as a handy future tool for distinguishing between spoken words and relating each with its appropriate referent. For example, infants may distinguish the intersensory differences in the vowels of /baba/ referring to 'bottle' and /bobo/ referring to 'the pet dog'. The inability to perform such analyses may be one reason for the sporadic rate of word use by visually impaired compared with sighted infants between 14 and 28 months (Iverson, Tencer & Goldin-Meadow, 1998; see Dunlea, 1989). Thus, intersensory analyses at the phonemic level may serve to initiate sighted infants into the community of speaking humans. This is a plausible hypothesis even though the extent to which various auditory–visual speech segments afford detection of redundant information may vary depending on many factors. For example, the place of articulation makes it easier to hear and see the bilabial stops /b/ and /p/ than the velar stops /g/ and /k/.

Intersensory redundancy and the detection of arbitrary speech–object relations

When observers look at and listen to events, they perceive redundant information across the senses as well as the information specific to each (see Figure 1). Observers may find some of these redundancies to be more informative than others. For example, the sight and sound of a person humming a tune (closed mouth) provides spatial co-location and temporal contiguity between the face and the voice, but not temporal synchrony, or a common rhythm, tempo or intensity shifts. In contrast, the sight and sound of a person singing or speaking with matching mouth movements provides more redundant information, co-location, temporal synchrony, a common rhythm, tempo and intensity shifts. In addition, the types of relations may vary. Given that articulation of a speech sound requires an appropriate mouth shape, the perception of these auditory–visual relations entails abstracting the available identical information across two or more senses. However, perception of the relations between a person's face and his/her specific voice qualities (e.g. fundamental frequency or accent) requires finding similarities between two distinct but co-occurring events, because, taken separately, one event does not convey the same information as the other. Walker-Andrews (1994) asserts that arbitrary relations such as these can be categorized into two types (Figure 1).

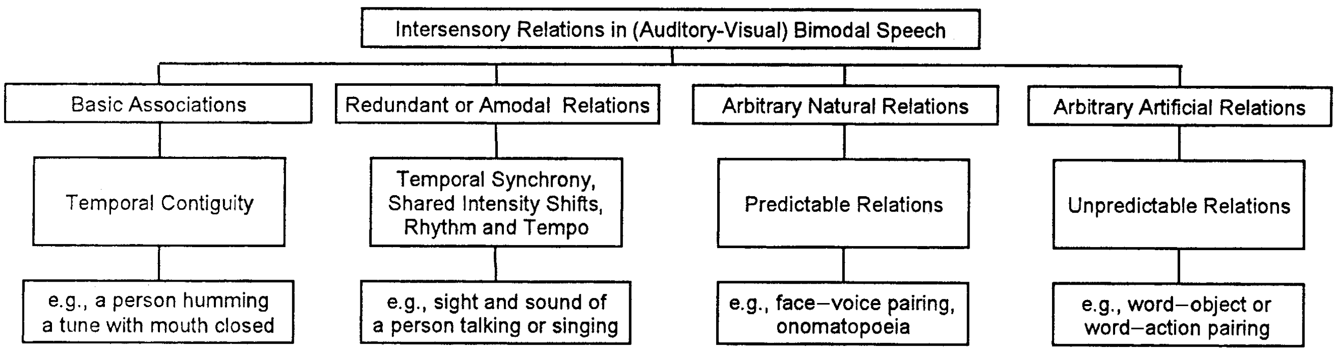

Figure 1 *A taxonomy for intersensory relations in human bimodal communication.*

Some arbitrary relations, such as that between a person's face and the sound qualities of his/her voice, predictably occur together in nature and may be referred to as 'arbitrary natural relations'. The detection of this arbitrary natural relation may be facilitated because a person's face and voice are spatially co-located. 'Arbitrary artificial relations' such as between a spoken word and an object, the color (or shape) of an object and the pitch of its sound on impact, or the color of a container and the taste or smell of the food it contains are idiosyncratic (Reardon & Bushnell, 1988). In the case of word–object relations in particular, detecting the arbitrary relation can be complicated further because the spoken word does not emanate from the object but emanates from a person using the word to refer to the object. There is no spatial co-location. Notwithstanding this complication, the ability to detect these arbitrary relations is necessary for the development of lexical comprehension in the view proposed here.

How do infants begin to detect the arbitrary relations between words and their referents, given their ability to abstract redundant information in auditory–visual events? Some studies of intersensory perception of non-speech events in infancy demonstrate that the detection of redundant relations *precedes* the detection of arbitrary relations within the same context (Bahrick, 1994; Hernandez-Reif & Bahrick, 1995). In these experiments, in addition to redundant information, specific properties of an object for which information is available only to one modality were paired with those available to another, to determine whether infants could perceive redundant as well as arbitrary (natural or artificial) relations. In one set of experiments, infants were habituated to two events in which visible properties such as color and shape were paired with an audible property such as pitch. At test, the arbitrary relation between the two events was mismatched. The 7-month-old infants, but not 3- and 5-month-olds, detected the mismatch in the arbitrary relations between the color and shape of an object and the pitch of its impact sound (Bahrick, 1994). The 7-month-olds detected the changes in arbitrary relations when they were also provided redundant information such as temporal synchrony between the impact sounds and visible motions of the objects. The 3- and 5-month-olds, however, only detected changes in the redundant relations. In a more recent set of experiments, 6-month-olds but not 4-month-olds detected the arbitrary artificial visual–tactual relations set up between the color-pattern and the shape of an object. The 4-month-olds, however, detected the redundant shape information across visual and tactual modalities (Hernandez-Reif & Bahrick, 1995). In these studies, infants detected redundant relations across modalities developmentally prior to detecting arbitrary relations, and when they detected the arbitrary relations it was only in the presence of intersensory redundancy.

Given that young infants detect redundant information in speech and non-speech events, one might expect findings in the domain of speech to parallel those from non-speech research (Bahrick, 1994; Hernandez-Reif & Bahrick, 1995). Specifically, one might expect infants to detect arbitrary relations between speech patterns and objects around 6 to 7 months, but not before. Kuhl, Williams and Meltzoff (1991) reported such findings. Four- to 5-month-old infants did not match two pure tones with the visible mouth shapes of a person articulating the vowels /a/ and /i/. Adults, however, had no difficulty, even though they did not identify the tones as speech sounds. Therefore, the arbitrary relation imposed by the experimenter was the only information uniting the tones and the visible mouth shapes. Given that 3- and 5-month-old infants did not detect the arbitrary relations between the color of an object and the pitch of its impact sound (Bahrick, 1994), it is not surprising that 4- to 5-month-olds (Kuhl *et al.*, 1991) did not detect the arbitrary relations between a mouth shape and a pure tone. Infants at this age did, however, match the mouth shape with a corresponding vowel sound in prior studies (Kuhl & Meltzoff, 1982, 1984, 1988). Findings such as these, from speech and non-speech research, suggest that the sequence in the development of intersensory perception is movement from the detection of redundant relations to the detection of arbitrary relations in both domains (see Bahrick & Pickens, 1994, for a review).

The ability to detect redundant information across bimodal events not only *precedes* the detection of arbitrary relations but also *guides* the detection of arbitrary relations within a given context (Sullivan & Horowitz, 1983; Bahrick, 1994). With respect to *arbitrary natural relations*, in a recent set of experiments, infants detected the relations between the faces of two persons and their matching voices given temporal synchrony between the visible mouth movements and audible vocalizations (Bahrick, Hernandez-Reif & Cigales, 1988). Following habituation training to video-films of alternating same-sex faces of two persons reciting a nursery rhyme, each paired with a distinctive synchronous voice, 4- and 6-month-olds detected the arbitrary relations between the faces and voices. When the face–voice pairings were switched (but synchrony between the faces and voices was maintained), infants looked longer to these novel pairings. These results indicate that the presence of redundant information such as temporal synchrony might facilitate the detection of arbitrary natural relations.

With respect to *arbitrary artificial relations*, prior results have shown that 6- to 7-month-old infants also detect these relations in the presence of temporal synchrony that unifies the auditory–visual non-speech events (Spelke, 1979; Lawson, 1980; Bahrick, 1994). However, only recently was the salience of redundant temporal synchrony for the detection of arbitrary artificial relations in speech established by testing infants under synchronous and asynchronous conditions. In one set of experiments, intersensory redundancy (temporal synchrony) resulted in 7-month-old infants learning the arbitrary relations set up between the speech sounds /a/ and /i/ and moving objects (Gogate & Bahrick, 1998). Infants learned the arbitrary relation only when the timing of the vocalizations coincided with that of object motions. Infants failed to learn these relations when temporal contiguity between a vowel sound and a static object or temporal contingency was presented (i.e. when the sounds occurred during the pauses between object motions; Gogate & Bahrick, 1998). Using identical methods, a second set of experiments showed that 8-month-olds (but not 7-month-olds) detected the arbitrary relations between complex monosyllables /tah/ and /gah/ and moving objects, again only in the presence of temporal synchrony (Gogate, under review). Likewise, 8-month-olds did not detect the arbitrary relations between syllables such as /nim/ and /lIf/ and two objects in the absence of temporal synchrony (Werker, Cohen, Lloyd, Casasola & Stager, 1998). These studies underscore the importance of intersensory redundancy for infants' detection of arbitrary word–object relations. They also suggest that intrinsic developmental differences in infants' perception between 7 and 8 months interact with properties of the communicative environment such as syllabic complexity and intersensory redundancy to influence infants' detection of arbitrary relations. One further study, discussed earlier, has shown that the presence of redundant information facilitates the detection of arbitrary relations in other modalities as well. In a study of non-speech events, 6-month-old infants learned the arbitrary relations between the visible shape and color-pattern of objects only if they were given the opportunity to feel (the shape of) the object (Hernandez-Reif & Bahrick, 1995).

The requirement for precise temporal synchrony in concurrent patterns of stimulation across sensory modalities appears to develop gradually during the first 6 months, probably as infants' sensitivity improves (Kellman & Arterberry, 1998). Thus, newborns can detect arbitrary relations in speech events when temporal contiguity (but not temporal synchrony) and spatial co-location are provided, and when the stimuli are quite distinct (Slater, Brown & Badenoch, 1997). In this study, after newborns were familiarized with arbitrary pairings of a red line and the word 'mum' in a male voice, and a green line and the word 'teat' in a female voice, they looked longer at test to alterations in the word–voice–line combinations. In this instance, temporal contiguity was provided by presenting the words only when the infants looked at the visible lines, and spatial co-location was provided when the words were played via a speaker located directly above the visual screen. For newborns, spatial co-location between paired stimuli is necessary for learning sound–object relations (Morrongiello, Fenwick & Chance, 1998). Furthermore, for newborns, temporal contiguity alone may suffice for learning the arbitrary auditory–visual relations (Figure 1). In addition, the task was simplified by using a female voice for one word and a male voice for the other. The newborns could have detected the arbitrary relations at a global level, between the voice quality (male or female) and the red or green line, rather than the specific word and the object. In the studies with 7- and 8-month-old infants (Gogate & Bahrick, 1998; Gogate, under review) intersensory redundancy (temporal synchrony) was provided. Further, all syllables, /a/, /i/, /tah/ and /gah/, were presented by a single female speaker. Thus, detection of the arbitrary relation hinged solely upon learning the pairing between the specific syllable and the object.

Conclusions

In general, these findings support the conclusion that, by the second half of the first year, preverbal infants use intersensory redundancy (temporal synchrony given via visual and auditory experience, or shape given via visual and haptic experience) to detect arbitrary relations. These studies show how intermodal perception might contribute to infants' detection of word–referent pairings. As infants attend to temporal synchrony in speech events, the arbitrary relations are highlighted, allowing infants to discover the links between specific words and referents. In this view, temporal synchrony serves as a bootstrap (and is a control parameter) for infants to tune in to arbitrary word–object relations. The redundant information serves to unify two otherwise arbitrarily co-occurring stimulus patterns. Because temporal synchrony specifies that the vocalization and the object belong together, it probably assists the infant in narrowing down the possibilities to find the most likely referent to pair with the word. An earlier report suggested that symbolic–lexical development may entail knowing that a word 'stands for' a tangible referent rather than knowing that a word 'goes with' a referent, which is indicative of pre-symbolic development (Golinkoff, Mervis & Hirsh-Pasek, 1994). We suggest that the infant's knowledge of how a word and object 'go together' may well be an important developmental precursor to knowing that the word 'stands for' the object. Although we have emphasized the role of temporal synchrony here, other types of redundant information such as intensity shifts, tempo and rhythm between spoken syllables and moving objects were also provided in Gogate's studies (Gogate & Bahrick, 1998; Gogate, under review). These types of information also may facilitate the detection of arbitrary relations. Controlled experiments are needed to examine whether they, too, can facilitate the learning of word–object relations.

The findings suggest that within the auditory–visual domain infants might detect arbitrary natural relations, such as those between faces and voices of the same sex, developmentally prior to detecting arbitrary artificial relations such as those between specific speech patterns and objects. Detecting the arbitrary relations between co-occurring events that are predictable in nature may be easier than detecting the arbitrary relations between less predictably co-occurring events. However, given that this conjecture derives from the results of independent studies (Bahrick *et al*., 1988; Gogate & Bahrick, 1998; Gogate, under review), future controlled experiments are required. Nonetheless, if predictability facilitates detection of word–referent relations, it casts an interesting light on the perception and production of onomatopoeia in infants' early words.[1] For example, infants use terms such as *bow wow* to refer to 'things of action' such as a 'dog' prior to using words such as *doggie* (see Werner & Kaplan, 1963; Nelson, 1973; Fenson *et al*., 1994). In this case, infants may detect the relation between the object and the term easily, because the term resembles the sound that the object makes. Infants may have many opportunities to detect the intersensory relations between the barking animal, its visible opening and closing mouth, and the term *bow wow*. Only later might infants detect the arbitrary artificial relation between the term *doggie* and the same visible animal (Figure 1). Thus, infants' word perception and production may be closely related. In addition, infants' ability to detect more predictable arbitrary relations and their use of onomatopoeia may be reciprocated by adults' frequent use of these forms in their communication to infants. (See Fernald and Morikawa (1993) for mothers' greater use of onomatopoeia to 6-month-olds.) These findings would show, in keeping with the theme of this paper, that maternal language and infants' intersensory perceptual abilities are reciprocally related.

Developmental changes in the prerequisites for detecting spoken label–object relations

Changes in the requisite conditions under which infants learn speech pattern–object relations in experimental settings reflect the dynamic nature of infants' developing perceptual–lexical system. At first, although newborns detect changes in redundant information (Bahrick, 1996), they do not appear to require temporal synchrony to detect arbitrary auditory–visual relations (Slater *et al*., 1997). During this early period, perhaps owing to low visual acuity, temporal contiguity may be sufficient for detecting arbitrary relations between auditory–visual events at a fairly global level (Figure 1). Later, around 7 to 8 months, infants require temporal synchrony to detect speech pattern–object relations (Gogate & Bahrick, 1998; Gogate, under review). Temporal contiguity alone in the absence of temporal synchrony, such as a word uttered in the presence of a static object, is no longer sufficient. Similarly, temporal contingency, when vocalizations occur during the pauses between the motions of an object, does not result in learning. Still later, this requirement seems to diminish with the lexical development that it enables. With advances in perception and memory, coupled with the onset of the implicit knowledge that words stand for objects, and with the growing ability to use words consistently to refer to objects, infants may no longer require temporal synchrony (or other redundant information) to detect arbitrary relations between words and objects. Thus, although 7-month-olds (Gogate & Bahrick, 1998) and 8-month-olds (Werker *et al*., 1998; Gogate, under review) require temporal synchrony, 14-month-olds detected a switch in syllable–object pairs even though vocalizations were presented out-of-synchrony with the moving objects (Werker *et al*., 1998). But, these experiments have also shown that infants require object motion to detect speech–object relations up to the age of 14 months. Neither 7-month-olds (Gogate & Bahrick, 1998) nor 14-month-olds (Werker *et al*., 1998) detected a switch in the speech–object pairs when static objects were presented concurrently with speech patterns. Why did these infants require object motion? Earlier research suggests that seeing objects in motion enables infants to abstract invariant properties of objects across different transformations (see review

[1] For example, infants may detect the arbitrary natural relation between the term *tick-tock* and the object 'clock' by first detecting the redundant relation between the object (the moving hands of the clock) and its temporally synchronous ticking sound (see Nelson, 1973).

in E.J. Gibson, 1991; E.J. Gibson, Owsley, Walker & Megaw-Nyce, 1979). Thus, movement itself highlights the invariant properties of the object and may permit easier detection of arbitrary relations between the objects and vocalizations. However, older infants who have presumably developed a repertoire of receptive and expressive vocabulary can detect word–object relations in the absence of object motion, temporal synchrony or perhaps even temporal contiguity (Baldwin, 1993a).

The strong role of multisensory information for detecting novel word–object relations during the first year and its relative loss of importance during the second year is evident from a comparison of Gogate's findings (Gogate & Bahrick, 1998; Gogate, under review) with those of Baldwin and her coworkers (Baldwin *et al.*, 1996). In Gogate's studies, 7- and 8-month-old infants detected the syllable–object relations if an object in a speaker's hand moved synchronously with the speaker's vocalizations, even though her face was not visible (on videotape). In contrast, Baldwin *et al.* (1996) found that older infants required a view of the speaker's face (during live stimulus presentations). For infants of 18 to 20 months, multisensory relations alone between spoken words and objects, in the absence of the speaker's face and referential intent, did not ensure learning of word–object relations. These older infants required *a social criterion* (a visible speaker naming objects while displaying visual interest in them) for establishing stable links between words and objects (Baldwin *et al.*, 1996). Baldwin *et al.*'s finding was substantiated in a recent study where even 13-month-olds detected word–object relations during live stimulus presentations when an object was named while the infant and the speaker attended to it (Woodward & Hoyne, 1999). Thus, during later infancy, joint attention becomes a more critical factor for detecting word–object relations than multisensory information. The increasing importance of joint attention is especially relevant given that even 6-month-olds are capable of following the direction of mother's eye gaze, which is positively correlated with their lexical comprehension at 12 months (Morales, Mundy & Rojas, 1998). Consistent with the dynamic systems view, the pattern of findings suggests that the non-linguistic, contextual requirements for learning word–object relations change with development. In general, a dynamic systems view is unique in that it emphasizes the constant fluctuations in and across parameters rather than a 'genetic blueprint' or a reliance on the most basic associations. Changes over time may look, on the surface, like regressions or U-shaped functions because of the co-action of systems and relative weight of parameters.

Thus far, we have seen that multiple non-linguistic factors interact in a dynamic fashion to facilitate word comprehension during infancy. How do linguistic factors such as syllabic complexity interact with non-linguistic factors? Seven-month-olds (Gogate & Bahrick, 1998) detected the relations between the vowels /a/ and /i/ and objects moving in temporal synchrony but did not detect the relations between more complex syllables, /tah/ and /gah/, and the same objects. Although infants at this age could distinguish the two complex syllables and the two objects from one another, they did not *link* them with the objects. However, 8-month-olds successfully detected these relations under identical conditions (Gogate, under review). In contrast, two prior studies (Stager & Werker, 1997; Werker & Stager, in press) reported that 8-month-olds were unable to 'map sound onto meaning' when two spoken syllables /bIh/ and /dIh/ or /lIf/ and /nim/ were each paired with a moving object. This was the case even though infants perceived the voicing distinction between /bIh/ and /dIh/ and the oral–nasal distinction between /lIf/ and /nim/. However, in these experiments, temporal synchrony was not provided. Such findings show how linguistic and non-linguistic factors can interact and affect infants' detection of word–referent relations.

The perceptual–lexical system: general or language-specific processes?

The findings reviewed above show that young infants are adept perceivers of redundant information across bimodal (auditory–visual) speech and non-speech events, and suggest that the development of general intersensory perceptual abilities may foster lexical comprehension. Specifically, because some of the same types of redundant information guide infants' detection of arbitrary relations in speech and non-speech events, the findings suggest that general or multipurpose perceptual processes are at work. Similar results with older infants (Namy & Waxman, 1998; Woodward & Hoyne, 1999) and young children (Markson & Bloom, 1997) also point to general processes for word comprehension. These data offer an alternative to the view that a unique acquisition device with innate constraints is necessary for the development of early lexical comprehension. Further, these findings suggest that development of intermodal speech does not proceed along paths unique and different from those found for other domains of bimodal perception (cf. Chomsky, 1980; Fodor, 1983).

The current data also call into question the notion that autonomous modules are present at birth for processing language-like and non-language-like patterns of bimodal stimulation. Preverbal infants as young as 7 months detected both arbitrary non-speech (Bahrick, 1994) and speech–object (Gogate & Bahrick, 1998; Gogate, under review) relations following approximately 2–3 min of exposure. We hypothesize, therefore, that infants learn such relations using powerful general perceptual–cognitive processes that may be sensitive to the probability of co-occurrence between bimodal patterns of stimulation. Such a thesis, though speculative, is in keeping with other recent findings. For example, 8-month-olds can detect word boundaries from an artificial language speech stream based solely on the statistical probability of their co-occurrence with adjacent speech sounds (Saffran, Aslin & Newport, 1996; see Bates & Elman, 1996; Seidenberg,[2] 1997). Identical findings have been reported in 8-month-olds' detection of boundaries when tones replaced the syllables in a sequence (Johnson, Saffran, Aslin & Newport, 1998). Thus, infants rely on some process that enables grouping of auditory stimuli based on their transitional probabilities for both syllables and tones. Further studies must establish whether infants access powerful general learning processes and use probabilistic optimization as a means for early word learning and comprehension. Consistent with the dynamic systems view of development, such findings suggest that computation of probabilities is one among many learning strategies available to the infant for word learning (see Kelso, 1997).

The nature of infants' learning

What strategies do young infants use to relate words and their referents? Several recent studies suggest that the detection of word–object relations is the result of associative learning of words and concepts or referents (see discussion in Jusczyk & Hohne, 1997; Stager & Werker, 1997; Werker *et al.*, 1998). In our view, there is much more to the process by which infants detect word–object relations. As discussed earlier, infants seek out and attend to specific affordances of the environment very early in life depending on their perceptual sensitivities (J.J. Gibson, 1966). They are unlikely to simply associate idiosyncratic information available to different modalities whenever temporal contiguity or spatial co-location is provided (see Slater *et al.*, 1997, for data on newborns). Thus, between 3 and 5 months, infants selectively attend to intersensory redundancies only when natural structural correspondences, such as an opening mouth with a gradual increase in vocal amplitude (MacKain *et al.*, 1983) or a single object with a single impact sound (Bahrick, 1988), are also provided. Selective attention to (and detection of) intersensory redundancies facilitates the discovery of arbitrary word–object relations. Thus, in our experiments, 7- and 8-month-old infants did not form idiosyncratic associations between a novel word and a moving hand, or part of the word and part of the object, despite the temporal synchrony between these events. Rather, they selectively attended to the (hand-held) whole object and syllable by tuning in to the intersensory redundancy. They did not learn the arbitrary relations when temporal contiguity between a word and a static object was provided, but required a more precise temporal synchrony between a word and an object's motion. They also did not learn these relations when the object motion consistently preceded the vocalization (delayed conditioning; cf. Markham, Schwartz & Gordon, in press). Therefore, these findings are inconsistent with the association view of early word learning (Bates & MacWhinney, 1987).

Furthermore, the methods we have used to study infants' learning of speech–object relations reflect infants' active participation. Infants set the pace for learning the relations during the habituation sequence (Horowitz, 1974). Although infants learn the arbitrary relations in one context (habituation to two auditory–visual displays, one at a time), they actively search for and select one of the two object displays in a second context (two choice intermodal preference reflecting memory) (Gogate & Bahrick, in press). Schafer and Plunkett (1998) raise a similar point with regard to the difference between their familiarization and testing procedures used to study word–object mapping. These behaviors reflect more than a passive forming of associations that can be induced for two randomly co-occurring (temporally contiguous or contingent) stimuli during conditioning under identical or similar contexts (cf. Rovee-Collier, 1986; see E.J. Gibson, 1969, in press).

Infants might also compute the probability of co-occurrence between words and referents to learn these relations. Some studies of word recognition suggest that the higher probability of occurrence of specific words in the environment may favor infants' recognition of those words. Thus, 4.5-month-olds are familiar with their own name (Mandell, Jusczyk & Pisoni, 1995) and 7.5-month-olds recognize words from a passage or a list after exposure to it (Jusczyk & Aslin, 1995). Similarly, if detection of arbitrary relations between words and objects is influenced by the probability of their co-occurrence, then infants should detect word–object pairings better as the probability of co-occurrence increases. Such findings would explain parsimoniously how infants might learn the most obvious meaning of a word prior to learning secondary meanings of the word. Furthermore, infants should detect arbitrary word–referent relations at a young age if the words and referents occur relatively frequently in their

[2] Seidenberg argues, against modular theories, that language acquisition is based on the frequency of linguistic and non-linguistic sources (see Seidenberg, 1997, footnote 29; and Bates & Elman, 1996).

environment. Tincoff and Jusczyk (1999) conducted experiments that speak to this prediction. Six-month-old infants detected the arbitrary relations between familiar words such as *mommy* and *daddy* and the faces of their own mother and father, but not those of another infant's mother and father. In another study, when a familiar word such as *ball* was presented, older infants (24-month-olds) looked to the named object in a pair even before the word was completely uttered (Fernald, Pinto, Swingley, Weinberg & McRoberts, 1998). Fifteen-month-olds, however, looked at the named object only after the word was completely uttered. Consistent with the dynamic systems view, these findings underscore an interaction between environmental factors such as the probability of co-occurrence and organismic factors such as infants' speed of processing and perceptual acuity in detecting word–referent relations.

Memory for arbitrary speech–object relations in infancy

To understand spoken words on a regular basis, it is not sufficient that infants merely recognize speech patterns and detect the arbitrary relations between the patterns and objects. Infants must also remember the arbitrary relations in order to reference them upon later encounters. The development of lexical comprehension is therefore contingent upon developing memory and can be affected by memory limitations. What factors might facilitate infants' memory for arbitrary relations? Temporal synchrony not only facilitates infants' learning of arbitrary relations but may also enable memory for these relations. In a study discussed earlier, 6-month-olds, but not 4-month-olds, detected a switch in arbitrary face–voice pairings, and remembered them for 10 min following habituation (Bahrick *et al.*, 1998). In another study, 7-month-old infants indicated memory for vowel–object pairings (Gogate & Bahrick, in press). Infants who were habituated to two alternately presented objects whose movements were temporally synchronized with the vocalizations /a/ or /i/ remembered the vowel–object pairs after 10 min or 4 days. Infants who were habituated and tested with temporally asynchronous or static displays of objects did not remember the pairings. The conditions that facilitate young infants' initial learning of arbitrary relations between spoken labels and objects appear to operate during delayed learning or memory for these relations as well. For older infants, other factors such as adult labeling of (static) objects might suffice. Thus, 10- and 14-month-old infants looked longer during a subsequent play period to objects that had been both pointed at and labeled but not to those to which adults had merely pointed (Baldwin & Markman, 1989). Further, at 13 months, infants regulated their own attention to objects, and the act of adult labeling when infants' attention was focused on the objects facilitated infants' memory for the word–object relation (Woodward, Markman & Fitzsimmons, 1994).

Taken together, the findings suggest progression in infants' memory for word–object relations. Seven-month-old infants require redundant information such as temporal synchrony to remember syllable–object relations. Older infants, between 10 and 13 months, remember word–object relations under different conditions, such as when adults label a static object during play episodes, or when adults label an object to which the infant's attention is directed. Future studies might focus on this progression. For example, at what age might infants detect and remember the relation between objects and words that are presented together but asynchronously?

The dynamic nature of infants' perceptual–lexical system: a summary

An integrated multisensory system, present from birth, allows the detection of redundant information in intersensory patterns of stimulation. Thus, very young infants are sensitive to changes in redundant information provided in auditory–visual patterns of stimulation and the ability to detect such changes is well established early in life (Figure 2). For example, infants of 2.5 to 4 months are visually attracted to the temporal synchrony between audible speech sounds and visible lip movements. The ability to detect redundancies, in turn, guides infants to detect arbitrary relations between auditory and visual events by the second half of the first year. The detection of arbitrary relations in temporally synchronous events is possible because temporal synchrony highlights the relation between the bimodal patterns of stimulation. Infants progress thereafter, detecting and remembering first arbitrary relations between predictable events such as talking persons and their specific faces, and later the relations between unpredictable events such as words (spoken syllables) and temporally synchronous moving objects. As discussed earlier, the latter ability is an important precursor to lexical comprehension, because the lexicon consists of words that are arbitrarily assigned to specific objects by convention. Once this ability is well developed, infants no longer require temporal synchrony for the detection of and memory for arbitrary relations. The present findings underscore the dynamic nature of infants' developing perceptual–lexical system, and give rise to questions about what environmental factors might contribute to infants' detection of and memory for word–object relations.

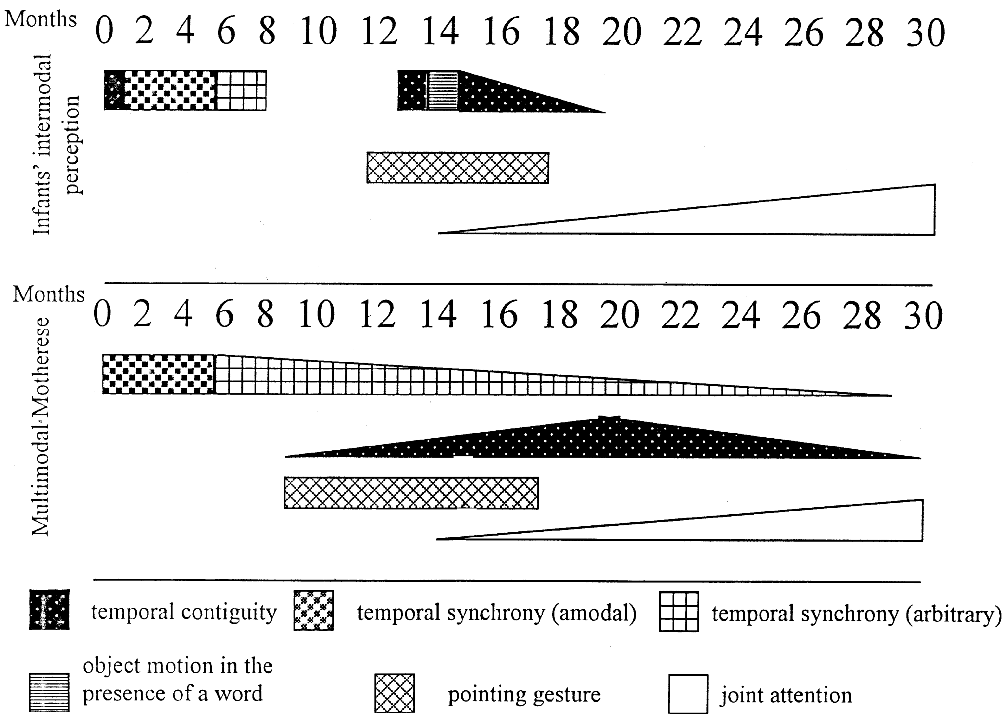

Figure 2 *Reciprocity between infants' perception and mothers' provision of various types of auditory–visual communication.*

Maternal communication and the development of infants' lexical comprehension

In the previous section, we discussed the dynamic characteristics of the infant's developing perceptual system that aid in the detection of and memory for the arbitrary relations between audible speech patterns and visible objects. Given our main hypothesis that lexical comprehension develops from the dynamic interaction between organismic and environmental systems, and given that infants are sensitive to intersensory redundancy in bimodal communication, in this section we discuss specifically how the environment contributes to infants' lexical comprehension. We ask to what extent the information that facilitates detection of and memory for word–referent relations is available to infants in their everyday lexical environment. Next, we ask whether the information provided in the lexical environment is compatible with infants' developing lexical comprehension.

Intersensory redundancy in maternal[3] communication

Soon after birth, maternal interaction with infants is multimodal in nature, providing auditory, visual, tactile and olfactory information about the mother to the infant. Infants show a visual preference and positive affect for simultaneous auditory and visual presentations of both female and male adults' infant-directed speech during the first 6 months (Werker & McLeod, 1989). Of primary importance to the thesis of this paper is that mothers provide a variety of multimodal naming contexts during everyday interactions with their infants. Prior descriptive studies characterize the lexical environment of infants as consisting of numerous instances in which mothers label objects and provide visual stimulation by showing or moving the objects (Messer, 1978; Masur, 1997; see Sullivan & Horowitz, 1983). Mothers also use gestures to highlight meaning by pointing to or touching an object when referring to it (Zukow-Goldring, 1997). They acoustically highlight specific words in the speech stream through the use of exaggerated prosody (greater stress and pitch variation), use shorter sentences and place a word in sentence final position (see review by Fernald, 1992). We suggest that the multimodal aspects of maternal communication may aid in the development of infants' communication.

In particular, mothers provide redundant information such as temporal synchrony when they simultaneously name and show objects or actions to their infants, creating a multimodal event (Zukow-Goldring, 1997). We call this intermodal infant-directed communication

[3] In most research the participant caregiver has been the infant's natural or adopted mother. Therefore, we use the terms 'mother', 'maternal' and 'motherese' throughout.

'multimodal motherese' (Gogate *et al.*, 2000). Temporal synchrony, as noted earlier, facilitates infants' early learning of novel word–object relations (Gogate & Bahrick, 1998; Gogate, under review). Synchronous naming and showing of objects, in particular, occurs in maternal communication to infants as old as 11 months (Messer, 1978). Recently a descriptive study of maternal multimodal communication to 24 American (Caucasian and Hispanic) infants, ranging from the age of 5 months to 30 months, and 12 Asian Indian infants ranging from 5 to 17 months, was conducted. This study showed that mothers exploit temporal synchrony, but not temporal contiguity or contingency, in their communication to infants of all ages. The mothers' utterances of novel words coincided with the objects they moved or actions they performed, and sometimes, even with touch. Second, this study showed that American mothers used multimodal communication styles that were well suited to their infant's age, which is typically correlated with the level of lexical comprehension (Gogate *et al.*, 2000). When naming objects and actions, the mothers of 5- to 8-month-olds used temporally synchronous object motions and words more often than those of older infants (9–17 months and 21–30 months). These results, together with the experimental data discussed earlier (Gogate & Bahrick, 1998, in press; Werker *et al.*, 1998; Gogate, under review), suggest that mothers' use of temporal synchrony actually decreases over time, corresponding with infants' increasing ability to detect word–object relations in the absence of synchrony (Figure 2).

The findings, specifically, point to the significant role of integrated multimodal information for promoting the development of lexical comprehension. Bates (1993) raised a similar point. She suggested that infants require integration of information from multiple modalities to comprehend words because they are still in the process of 'cracking the speech code' (p. 237). Even older infants with a working vocabulary predominantly understand words that refer to the immediate context rather than those that refer to previously experienced events. This necessitates that adults use relatively context-bound speech (Nelson, 1973; Messer, 1978; Bates, 1993) in a manner that highlights the arbitrary word–referent relations for young infants. The temporal synchrony provided by mothers when they name objects may facilitate integration of multisensory information perhaps because synchrony highlights the intermodal relations for the infant. The temporal synchrony need only be fairly precise. Infants allow some 'slippage between auditory and visual inputs' (within 350 ms) before they will perceive them as asynchronous (Lewkowicz, 1992, p. 35). Thus, even when mothers loosely synchronize words with moving objects, it may unify the two events and highlight the arbitrary relations, facilitating infants' attention, detection and eventual memory for word–referent relations.[4] Additionally, the loose synchrony may be compensated for by other types of redundant information such as matching intensity shifts, rhythm and tempo between a word and a moving object.

Pointing to static objects in maternal communication

Other types of multimodal communication, including maternal pointing to a static object or touching a static object while naming it, may become useful in directing attention to the arbitrary relations between labels and objects later in infancy, when infants become sensitive to these gestures (Murphy & Messer, 1977). Gogate *et al.* (2000) found that pointing was a relatively minor component of mothers' multimodal communication to young infants. The frequency of pointing was low, perhaps because the mothers named target objects that were in close proximity to themselves and their infant. However, even under these conditions where pointing was sporadic, mothers pointed to objects and labeled them slightly more frequently during interactions with 9- to 17-month-olds than with 5- to 8-month-olds or 21- to 30-month-olds. This finding complements that of studies showing that, between the ages of 9 and 17 months, infants are particularly receptive to the pointing gesture (Butterworth, 1991) and that attention to pointing is a robust predictor of lexical comprehension (Mundy, Kasari, Sigman & Ruskin, 1995). The findings, once again, speak to mothers' sensitivity to infants' capabilities (Figure 2).

Naming of static objects in the presence of joint attention between infant and mother

During interactions with older infants, adults may utter words in the presence of static objects without pointing to the objects or touching the infants (Baldwin *et al.*, 1996). Mothers' and infants' joint attention to objects alone may suffice for the learning of word–object relations, once infants begin to detect these relations on

[4] In two cultures reported thus far, the Kaluli of New Guinea (Schieffelin, 1979) and the Kwara'ae of Malatia in the Solomon Islands (Watson-Gegeo & Gegeo, 1976), mothers may not provide the direct multimodal communication to young infants underscored in this paper. It is our hypothesis that, in these cultures, infants receive redundant multimodal information either indirectly from mothers or directly from other members of the community. These types of secondary input may be important for word learning. Recent studies suggest that infants are just as capable of learning 'object labels when they eavesdrop on others' interactions as when they are directly addressed' (Akhtar, Jipson & Callanan, 1998).

their own and demonstrate a working vocabulary. In accordance with this thesis, Gogate *et al.* (2000) found that mothers named static objects, which they simply held for the infant, significantly more often during interactions with 21- to 30-month-olds relative to the two younger age groups (5–8 months and 9–17 months). In these instances, both mother and infant attended to the object or event being named. Given that the experimental data showed that 14-month-olds did not learn the arbitrary relations between syllables and static objects in the absence of joint attention (Werker *et al.*, 1998), this variation by age is appropriate (Figure 2).

Infant-regulated multimodal communication

In the previous sections, we suggested that mothers regulate the naming context by structuring the environment according to the infants' level of perceptual–lexical competence. Consonant with infants' requirement for integrated auditory and visual information to detect speech pattern–object relations, the descriptive data showed that mothers provide integrated information to younger infants, perhaps highlighting the arbitrary relations between words and referents. Older infants may have other means for gleaning word–referent relations and may not need explicit maternal guidance for learning the relations. Once again, mothers of older infants seem to be attuned to their infants' requirements, and are more likely to simply name objects without also moving them.

During the infants' second year, however, this multimodal communication seems to undergo a qualitative shift, away from predominant adult regulation and towards increased infant regulation. For example, mothers of the oldest group of infants observed by Gogate *et al.* (2000) were likely to attend to and name an object already being explored and manipulated by infants (also see Masur, 1997). Mothers of the younger infants in this study (5–8 months) rarely showed this pattern of behavior (Figure 2). These results support the conclusion that, with experience, older infants regulate the lexical input they receive. This pattern is paralleled by data on infants' word learning. In the second year, infants learned words rapidly when adults named or verbally encouraged infants to act upon the objects on which infants' attention was already focused (Akhtar, Dunham & Dunham, 1991; Woodward *et al.*, 1994). In comparison, their rate of vocabulary learning was slower when mothers redirected infants' attention to objects outside their focus of attention (Akhtar *et al.*, 1991). These patterns speak to the dynamic nature of mother–infant interaction during the development of lexical comprehension. Further studies are required to examine the dynamic nature of maternal communication and to examine the extent of its adaptation to infants' perceptual and communicative requirement.

Conclusion

A diverse set of experimental findings suggests that early lexical comprehension owes much to infants' developing ability to perceive intersensory relations in auditory–visual events. Initially infants attend to and detect global information across the senses, progressing to the detection of more specific relations. In this respect, the changes in the perceptual–linguistic system embody Werner's orthogenetic principle. They show that development proceeds from a state of 'relative globality and lack of differentiation to a state of increasing differentiation, articulation, and hierarchic integration' (Werner, 1957, p. 126; also see Bower, 1974; Walton & Bower, 1993; Hernandez-Reif & Bahrick, 1995; Bahrick, 1996). With respect to bimodal speech, improvement in intersensory perception culminates in infants' detection of and memory for arbitrary relations between different speech patterns and their object or action referents. The development can be directly observed in infants' attention to and selection of salient properties of the environment when learning word–referent relations. Detection of arbitrary word–referent relations begins long prior to an infant's first word production, as illustrated by the precedence of language comprehension skills relative to word production (Benedict, 1979; see Huttenlocher, 1974; Reznick, 1990; Golinkoff & Hirsh-Pasek, 1995). Although infants do not produce words until the second year, in part because of motor constraints, some perceptual abilities necessary for word comprehension are well in place by the end of the first year.

The development of infants' perceptual–lexical system is only part of the story, however. We are not the first to emphasize the maternal communication environment in language acquisition (Garnica, 1977; Cooper & Paccia-Cooper, 1980; Sullivan & Horowitz, 1983; Gleitman, Newport & Gleitman, 1984; Hirsh-Pasek *et al.*, 1987; Kelly, 1992; Lewkowicz, 1996; Bloom, 1998). In addition, however, we wish to call attention to its multimodal nature and the dynamic properties of this particular system (see Tucker & Hirsh-Pasek, 1993). Mothers provide a variety of developmentally compatible multisensory contexts (see Figure 2). When they name objects or actions for infants, they often provide temporal synchrony, they point to objects and actions, or they provide no motion at all. Moreover, they tailor these contextual conditions to the developmental status of their infant. Mothers use temporal synchrony primarily in communication with younger infants, which may scaffold infants' developing ability to detect word–referent relations. They phase out some aspects or provide a different set of conditions for older infants, such as naming an object in motion without temporal synchrony, or naming an object that is held immobile. In addition, mothers may provide socially contingent contexts that facilitate infants' detection of word–object relations. Such a context is provided when infants encounter a speaker who names and visually attends to the object being named, indicating referential intent. As they develop, infants begin to increasingly regulate these contexts. As discussed in the previous section, infants' attention to objects or events in themselves can initiate maternal naming. These findings illustrate that the maternal multimodal communication system ('multimodal motherese') is dynamic and well suited to the infant's developmental status.

We do not suppose that the participants or 'owners' of each system are necessarily cognizant of the dynamic interaction between the two systems. In fact, we would emphasize that mothers and infants are not conscious of the entrainment of the two systems. For example, the infant probably does not view the temporal synchrony provided by the mother as a resource for detecting word–object relations, nor is the infant aware of his/her own ability to detect these relations. In the process of discovering the meanings of words, the infant, motivated to learn, may stumble upon the intersensory redundancy between words and referents. Similarly, the mother is unlikely to provide temporal synchrony consciously nor is she aware necessarily that the infant is relying on such information for detecting a relationship between a spoken word and a moving object. Furthermore, infants' developing lexical competence is not the result of a master plan held by either party, but the 'workings of a myriad of interactions at a local level' (van Geert, 1993, p. 266). The observed interactions between the environmental system (maternal multimodal communication) and the infant system (perceptual–lexical) is the product of acting together of numerous diverse factors, each with multiple degrees of freedom. We have identified some of these factors for developing lexical comprehension – temporal synchrony, motion, and pointing by the mother, and the developing sensitivity to intersensory relations and syllabic complexity on the part of the infant – but there are bound to be many others, each contributing jointly to lexical development. During the ebb and flow of interactions, neither party need be cognizant of these factors, although one provides the environment and the opportunity for the adaptations of the other.

Furthermore, we suggest that the two systems are self-organizing or order seeking (Thelen & Smith, 1994; Reed, 1996; Kelso, 1997). During the course of development, one system provides the affordances encountered by the other. Affordances are properties of the environment that are salient to the perceiver (J.J. Gibson, 1966, 1979; see Reed, 1996). As a result of self-organization, the infant's perceptual–lexical system may regulate itself in response to changes in the affordances of the maternal communication system. Analogously, the maternal communication system may regulate itself in response to systemic pressures exerted by the affordances of the infant's perceptual–lexical system. An outcome of self-organization in the two dynamic systems is that the maternal (adult) communication system strives to adapt itself to the infant as much as the infant's system to maternal communication (see Studdert-Kennedy, 1991).

In conclusion, early lexical development can be seen as a process of ongoing reciprocal interactions between the organism and the environment. During the process of adapting to one another, the infant's intersensory perceptual and maternal multimodal communication systems momentarily exhibit stable patterns revealing, at the outset, a consonance between the two systems (preferred states), where predominant factors or control parameters influencing each system are observable. The ongoing changes that give rise to these momentary points of stability may well be driven by a combination of (as yet) unobservable instabilities in one or the other of the two systems (see Kelso, 1997). The instabilities of one system and the self-organization of the second result in changes in the latter that seem consonant with those of the former. In this manner, the changes of one system dovetail with the changes of the other. Environment–organism interactions, therefore, are an important unit for study because everyday social interactions between the mother (or other communicating individuals) and the infant provide rich opportunities for lexical development and ensure the successful transmission of the lexicon from one generation to the next.

Acknowledgements

Writing of this paper was supported by a General University Research Grant to the first author from the University of Delaware. We thank Robert Lickliter and

Roberta Golinkoff for their comments on an earlier version.

References

Akhtar, N., Dunham, F., & Dunham, P. (1991). Directive interactions and early vocabulary development: the role of joint attentional focus. *Journal of Child Language*, **18**, 1–40.

Akhtar, N., Jipson, J., & Callanan, M.A. (1998, April). Learning words by eavesdropping. Poster presented at the International Conference on Infant Studies, Atlanta, GA.

Bahrick, L.E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development*, **6**, 429–451.

Bahrick, L.E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development*, **10**, 387–416.

Bahrick, L.E. (1988). Intermodal learning in infancy: learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development*, **59**, 197–209.

Bahrick, L.E. (1992). Infants' perceptual differentiation of amodal and modality-specific audio-visual relations. *Journal of Experimental Child Psychology*, **53**, 180–199.

Bahrick, L.E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology*, **6** (2), 111–123.

Bahrick, L.E. (1996, April). Increasing specificity in the development of infants' sensitivity to two nested amodal relations. Presented at the International Conference on Infant Studies, Providence, RI.

Bahrick, L.E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, **36**, 190–211.

Bahrick, L.E., & Pickens, J.N. (1994). Amodal relations: the basis for intermodal perception and learning in infancy. In D. Lewkowicz & R. Lickliter (Eds), *The development of intersensory perception: Comparative perspectives* (pp. 205–233). Hillsdale, NJ: Erlbaum.

Bahrick, L.E., Hernandez-Reif, M., & Cigales, M. (1998, June). Personal communication.

Baldwin, D.A. (1993a). Early referential understanding: infants' ability to recognize referential acts for what they are. *Developmental Psychology*, **29**, 832–843.

Baldwin, D.A., & Markman, E.M. (1989). Establishing word–object relations: a first step. *Child Development*, **60**, 381–398.

Baldwin, D.A., Markman, E.M., Bill, B., Desjardins, R., Irwin, J.M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word–object relations. *Child Development*, **67**, 3135–3153.

Bates, E. (1993). Production and comprehension in early language development: a commentary. In E.S. Savage-Rumbaugh, J. Murphy, R.A. Sevcik, K.E. Brakke, S.L. Williams & D. Rumbaugh (Eds), *Monographs of the Society for Research in Child Development*, **58** (3–4, Serial No. 233), 222–242.

Bates, E., & Elman, J. (1996). Learning rediscovered. *Science*, **274**, 13 December, 1849–1850.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Benedict, H. (1979). Early lexical development: comprehension and production. *Journal of Child Language*, **6**, 183–200.

Bloom, L. (1998). Language acquisition in its developmental context. In W. Damon, D. Kuhn & R.S. Siegler (Eds), *Handbook of child psychology* (pp. 309–370). New York: Wiley.

Bower, T.G.R. (1974). *Development in infancy*. San Francisco, CA: Freeman.

Budwig, N. (1995). *A developmental–functionalist approach to child language*. Hillsdale, NJ: Erlbaum.

Butterworth, G.E. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 223–232). Oxford: Blackwell.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Cooper, W.E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.

Dent, C.H. (1990). An ecological approach to language development: an alternative functionalism. *Developmental Psychobiology*, **23** (7), 679–703.

Dodd, B. (1979). Lip reading in infants: attention to speech presented in-and-out of synchrony. *Cognitive Psychology*, **11**, 478–484.

Dunlea, A. (1989). *Vision and the emergence of meaning: Blind and sighted children's early language*. Cambridge: Cambridge University Press.

Ellis, R., & Wells, C.G. (1980). Enabling factors in adult–child discourse. *First Language*, **1**, 46–62.

Fenson, L., Dale, P., Reznick, S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59** (5), 242.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective. In J. Barkow, L. Cosmides & M. Tooby (Eds), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford: Oxford University Press.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, **64**, 637–656.

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G.W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, **9** (3), 72–75.

Field, T.M., Woodson, R., Cohen, D., Greenberg, R., Garcia, R., & Collins, E. (1983). Discrimination and imitation of facial expressions by term and preterm neonates. *Infant Behavior and Development*, **6**, 485–489.

Fodor, J. (1983). *Modularity of the mind*. Cambridge, MA: MIT Press.

Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C.E. Snow & C.A. Ferguson (Eds), *Talking to children: Language input and acquisition* (pp. 63–88). Cambridge: Cambridge University Press.

van Geert, P. (1993). A dynamic systems model of cognitive growth: competition and support under limited resource conditions. In L.B. Smith & E. Thelen (Eds), *A dynamic systems approach to development*: *Applications* (pp. 265–332). Cambridge, MA: MIT Press.

Gibson, E.J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

Gibson, E.J. (1991). *An odyssey in learning and perception*. Cambridge, MA: MIT Press.

Gibson, E.J. (in press). A historical perspective. In *Perceptual development and learning: An ecological view* (Ch. 1).

Gibson, E.J., & Walker, A.S. (1984). Development of the knowledge of visual–tactual affordances of substance. *Child Development*, **55**, 453–460.

Gibson, E.J., Owsley, C.J., Walker, A.S., & Megaw-Nyce, J. (1979). Development of perception of invariants: substance and shape. *Perception*, **8**, 609–619.

Gibson, J.J. (1966). *The senses considered as perceptual systems*. London: George Allen & Unwin.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gleitman, L., Newport, E., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, **11**, 43–79.

Gogate, L.J. (under review). Intersensory redundancy facilitates learning of arbitrary minimal pair-object relations by 8- but not 7-month-old infants: evidence for a dynamic system.

Gogate, L.J., & Bahrick, L.E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel-sounds and objects in 7-month-olds. *Journal of Experimental Child Psychology*, **69** (2), 133–149.

Gogate, L.J., & Bahrick, L.E. (in press). Intersensory redundancy and seven-month-old infants' memory for syllable–object relations. *Infancy*.

Gogate, L.J., Bahrick, L.E., & Watson, J.D. (2000). A study of multimodal motherese: the role of temporal synchrony between verbal labels and gestures. *Child Development*, **71** (4), 878–894.

Golinkoff, R.M., & Hirsh-Pasek, K. (1995). Insights into grammatical development: children's comprehension of sentences. In P. Fletcher & B. MacWhinney (Eds), *Handbook of language acquisition* (pp. 430–461). London: Blackwell.

Golinkoff, R.M., Mervis, C., & Hirsh-Pasek, K. (1994). Early object labels: the case for a developmental lexical principles framework. *Journal of Child Language*, **21**, 125–155.

Gottlieb, G. (1993). Social induction of malleability in ducklings: sensory basis and psychological mechanism. *Animal Behavior*, **45** (4), 707–719.

Gottlieb, G., Tomlinson, W.T., & Radell, P.L. (1989). Developmental intersensory interference: premature visual experience suppresses auditory learning in ducklings. *Infant Behavior and Development*, **12**, 1–12.

Harris, M. (1992). *Language experience and early language development: From input to uptake*. Hillsdale, NJ: Erlbaum.

Hernandez-Reif, M.A., & Bahrick, L.E. (1995, April). Infants' visual–tactual perception of amodal and arbitrary relations. Presented at the Society for Research in Child Development, Indianapolis, IN.

Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P.W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses as perceptual units for young infants. *Cognition*, **26**, 269–286.

Horowitz, F.D. (1974). Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development*, **39** (5–6, Serial No. 158).

Huttenlocher, J. (1974). The origins of language comprehension. In R.L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Hillsdale, NJ: Erlbaum.

Iverson, J.M., Tencer, H.L., & Goldin-Meadow, S. (1998, April). Prelinguistic communication in congenitally blind infants. Poster at the International Conference on Infant Studies, Atlanta, GA.

Johnson, E.K., Saffran, J.R., Aslin, R.N., & Newport, E.L. (1998, April). Statistical learning of non-linguistic stimuli by 8-month-olds. Presented at the International Conference on Infant Studies, Atlanta, GA.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, **29**, 1–23.

Jusczyk, P.W., & Hohne, E. (1997). Infants' memory for spoken words. *Science*, **277**, 26 September, 1984–1986.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, **64**, 675–687.

Kaye, K.L., & Bower, T.G.R. (1994). Learning and intermodal transfer of information in newborns. *Psychological Science*, **5** (5), 286–288.

Kellman, P.J., & Arterberry, M.E. (1998). *The cradle of knowledge: Development of perception in infancy*. Cambridge, MA: MIT Press.

Kelly, M. (1992). Using sound to solve syntactic problems. *Psychological Review*, **99**, 349–364.

Kelso, S. (1997). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Knudsen, E.I. (1983). Early auditory experience aligns the auditory map of space in the optic tectum of the barn owl. *Science*, **222**, 25 November, 939–942.

Kuhl, P.K., & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science*, **218**, 10 December, 1138–1141.

Kuhl, P.K., & Meltzoff, A.N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development*, **7**, 361–381.

Kuhl, P.K., & Meltzoff, A.N. (1988). Speech as an intermodal object of perception. In A. Yonas (Ed.), *Perceptual development in infancy: The Minnesota Symposia on Child Psychology* (Vol. 20, pp. 235–266). Hillsdale, NJ: Erlbaum.

Kuhl, P.K., Williams, K.A., & Meltzoff, A.N. (1991). Cross-modal speech perception in adults and infants using non-speech auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, **17** (3), 829–840.

Lawson, K.R. (1980). Spatial and temporal contiguity and auditory–visual integration in infants. *Developmental Psychology*, **16**, 185–192.

Lewkowicz, D.J. (1992). The development of temporally-based intersensory perception in human infants. In F. Macar, V. Pouthas & W.J. Freidman (Eds), *Time, action and cognition: Towards bridging the gap* (pp. 33–44). Dordrecht: Kluwer.

Lewkowicz, D.J. (1996). Infants' response to the audible and visible properties of the human face: 1. Role of lexical–syntactic content, temporal synchrony, gender, and manner of speech. *Developmental Psychology*, **32** (2), 347–366.

Lickliter, R. (1990a). Premature visual stimulation accelerates intersensory functioning in bobwhite quail neonates. *Developmental Psychobiology*, **23**, 15–17.

Lickliter, R. (1990b). Premature visual experience facilitates visual responsiveness in bobwhite quail neonates. *Infant Behavior and Development*, **13**, 487–496.

Lickliter, R. (in press). Dynamical systems and psychological science. In B.D. Midgley & E.K. Morris (Eds), *Modern perspectives on J.R. Kantor and interbehaviorism*. Connecticut, NH: Greenwood Press.

Lickliter, R., & Banker, H. (1994). Prenatal components of intersensory development in precocial birds. In D.J. Lewkowicz & R. Lickliter (Eds), *The development of intersensory perception: Comparative perspectives* (pp. 59–80). Hillsdale, NJ: Erlbaum.

Lickliter, R., & Lewkowicz, D.J. (1995). Intersensory experience and early perceptual development: attenuated prenatal sensory stimulation affects postnatal auditory and visual responsiveness in bobwhite quail chicks. *Developmental Psychology*, **31**, 609–618.

Lickliter, R., & Stoumbos, J. (1991). Enhanced prenatal auditory experience facilitates postnatal visual responsiveness in bobwhite quail chicks. *Journal of Comparative Psychology*, **105**, 89–94.

MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left-hemisphere function. *Science*, **219**, 18 March, 1347–1349.

Mandell, D.R., Jusczyk, P.W., & Pisoni, D. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, **6** (5), 314–317.

Markham, M.R., Schwartz, B., & Gordon, W.C. (in press). *Learning and memory*. Pacific Grove, CA: Brooks-Cole.

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, **385**, 27 February, 813–815.

Masur, E.F. (1997). Maternal labelling of novel and familiar objects: implications for children's development of lexical constraints. *Journal of Child Language*, **24**, 427–439.

Meltzoff, A.N., & Borton, R.W. (1979). Intermodal matching by human neonates. *Nature*, **282**, 22 November, 403–404.

Meltzoff, A.N., & Moore, M.K. (1989). Imitation in newborn infants: exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology*, **25**, 954–962.

Messer, D. (1978). The integration of mothers' referential speech with joint play. *Child Development*, **49**, 781–787.

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, **21** (2), 373–377.

Morrongiello, B.A., Fenwick, K.D., & Chance, G. (1998). Crossmodal learning in newborn infants: inferences about properties of auditory–visual events. *Infant Behavior and Development*, **21**, 543–554.

Mundy, P., Kasari, C., Sigman, M., & Ruskin, E. (1995). Nonverbal communication and language development in children with Down syndrome and children with normal development. *Journal of Speech and Hearing Research*, **38**, 157–167.

Murphy, C.M., & Messer, D. (1977). Mothers, infants and pointing: a study of gesture. In H.R. Schaffer (Ed.), *Studies in mother–infant interaction* (pp. 325–350). London: Academic Press.

Namy, L., & Waxman, S.R. (1998). Words and gestures: infants' interpretation of different forms of symbolic reference. *Child Development*, **69** (2), 295–308.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, **38** (1–2, Serial No. 149).

Pickens, J.N. (1994). Perception of bimodal distance relations by 5-month-old infants. *Developmental Psychology*, **30** (4), 537–544.

Reardon, P., & Bushnell, E.W. (1988). Infants' sensitivity to arbitrary pairings of color and taste. *Infant Behavior and Development*, **11**, 245–250.

Reed, E.S. (1996). *Encountering the world: Toward an ecological psychology*. Oxford: Oxford University Press.

Reznick, S.J. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, **11**, 145–166.

Rovee-Collier, C. (1986). The rise and fall of classical conditioning research: its promise for the study of infant development. In L.P. Lipsitt & C. Rovee-Collier (Eds), *Advances in infancy research* (Vol. 4, pp. 139–159). Norwood, NJ: Ablex.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning in 8-month-old infants. *Science*, **274**, 13 December, 1926–1928.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, **69** (2), 309–320.

Schieffelin, B. (1979). Getting it together: an ethnographic approach to the study of the development of communicative competence. In E. Ochs & B. Schieffelin (Eds), *Developmental pragmatics*. New York: Academic Press.

Schiff, W., Benasich, A.A., & Bornstein, M.H. (1989). Infant sensitivity to audiovisually coherent events. *Psychological Research*, **51**, 102–106.

Seidenberg, M.S. (1997). Language acquisition and use: learning and applying probabilistic constraints. *Science*, **275**, 14 March, 1599–1603.

Slater, A., Brown, E., & Badenoch, M. (1997). Intermodal perception at birth: newborn infants' memory for arbitrary auditory–visual pairings. *Early Development and Parenting*, **6**, 99–104.

Spelke, E.S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, **15**, 626–636.

Stager, C.L., & Werker, J.F. (1997, July). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, **388**, 381–382.

Stein, B., Meredith, M.A., & Wallace, M.T. (1994). Development and neural basis of multisensory integration. In D.J. Lewkowicz & R. Lickliter (Eds), *The development of intersensory perception: Comparative perspectives* (pp. 81–105). Hillsdale, NJ: Erlbaum.

Stemmer, R. (1989). Empiricist versus prototype theories of language acquisition. *Mind and Language*, **4**, 201–221.

Studdert-Kennedy, M. (1991). Language development from an evolutionary perspective. In N. Krasnegor, D.M. Rumbaugh, R. Scheifelbush & M. Studdert-Kennedy (Eds), *The biological and behavioral determinants of language development* (pp. 5–28). Hillsdale, NJ: Erlbaum.

Sullivan, J., & Horowitz, F.D. (1983). Infant intermodal perception and maternal multimodal stimulation: implications for language development. In C. Rovee-Collier & L.P. Lipsitt (Eds), *Advances in infancy research* (Vol. 2, pp. 183–239). Norwood, NJ: Ablex.

Thelen, E., & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Tincoff, R., & Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, **10** (2), 172–175.

Tucker, M.L., & Hirsh-Pasek, K. (1993). Systems and language: implications for acquisition. In L.B. Smith & E. Thelen (Eds), *Dynamic systems and development: Applications* (pp. 359–384). Cambridge, MA: MIT Press.

Walker-Andrews, A.S. (1994). Taxonomy for intermodal relations. In D.J. Lewkowicz & R. Lickliter (Eds), *The development of intersensory perception: Comparative perspectives* (pp. 39–56). Hillsdale, NJ: Erlbaum.

Walker-Andrews, A.S., & Lennon, E. (1985). Auditory–visual perception of changing distance by human infants. *Child Development*, **22**, 373–377.

Walton, G.E., & Bower, T.G.R. (1993). Amodal representation of speech in infants. *Infant Behavior and Development*, **16**, 233–243.

Watson-Gegeo, K., & Gegeo, D. (1976). Calling-out and repeating routines in Kwara'ae children's language socialization. In B.B. Schieffelin & E. Ochs (Eds), *Language socialization across cultures*. Melbourne, Australia: Cambridge University Press.

Werker, J.F., & McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responses. *Canadian Journal of Psychology*, **43** (2), 230–246.

Werker, J.F., & Stager, C.L. (in press). Developmental changes in infant speech perception and early word learning: is there a link? In J. Pierrehumbert & M. Broe (Eds), *Papers in laboratory phonology, V*.

Werker, J.F., Cohen, L.B., Lloyd, V.L., Casasola, M., & Stager, C.L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, **34** (6), 1289–1309.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D.B. Harris (Ed.), *The concept of development* (pp. 125–148). Minneapolis, MN: University of Minnesota Press.

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought*. New York: Wiley.

Woodward, A., & Hoyne, K.L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, **70** (1), 65–77.

Woodward, A., Markman, E., & Fitzsimmons, M. (1994). Rapid word learning in 13- to18-month-olds. *Developmental Psychology*, **30**, 553–566.

Zukow-Goldring, P. (1997). A social ecological realist approach to the emergence of the lexicon: educating attention to amodal invariants in gesture and speech. In C. Dent-Read & P. Zukow-Goldring (Eds), *Evolving explanations of development: Ecological approaches to organism–environment systems* (pp. 199–252). Washington, DC: American Psychological Association.

Received: 28 January 1999
Accepted: 9 December 1999