

Multimodality in infancy: vocal-motor and speech-gesture coordinations in typical and atypical development

Jana M. IVERSON*

RÉSUMÉ

La multimodalité précoce : coordination du vocal et du moteur, du langage et du geste dans le développement typique et atypique

Dès le début de la vie, le comportement expressif est multimodal ; il se compose de coordinations comportementales précoces qui vont s'affiner et se renforcer avec le temps de façon à être utilisées pour la communication de sens. Parmi ces coordinations communicatives, celles qui impliquent les gestes et le langage sont certainement celles qui ont reçu la plus grande attention empirique, mais on sait peu de chose sur les origines développementales des liens entre gestes et langage. Une des possibilités réside dans le fait que les origines des coordinations gestes-langage reposent sur les relations entre la main et la bouche observées dans les activités sensorimotrices de tous les jours des très jeunes enfants qui n'utilisent pas encore la main ou la bouche pour communiquer du sens. Dans cet article, je passe en revue des preuves qui suggèrent que l'étude des liens entre gestes et langage et les couplages développementalement précoces entre des systèmes vocaux et moteurs dans la petite enfance peuvent fournir un aperçu précieux sur un certain nombre de développements ultérieurs qui reflètent l'interdépendance cognitive entre gestes et langage. Ceci comprend des aspects liés au développement du langage et à des retards dans cette acquisition, les origines infantiles du système gestes/langage que l'on retrouve chez l'adulte et les signes précoces des troubles autistiques. L'implication de ces résultats quant à l'étude du développement de la communication multimodale est envisagée.

MOTS-CLÉS : GESTES, DÉVELOPPEMENT DU LANGAGE, VOCALISATIONS, DÉVELOPPEMENT MOTEUR

*Address for Correspondence: Dept. of Psychology, University of Pittsburgh, 3415 Sennott Square, 210 S. Bouquet St., Pittsburgh, PA 15260, USA ; Email: jiverson@pitt.edu

ABSTRACT

From very early in life, expressive behavior is multimodal, with early behavioral coordinations being refined and strengthened over time as they become used for the communication of meaning. Of these communicative coordinations, those that involve gesture and speech have received perhaps the greatest empirical attention, but little is known about the developmental origins of the gesture-speech link. One possibility is that the origins of speech-gesture coordinations lie in hand-mouth linkages that are observed in the everyday sensorimotor activity of very young infants who do not yet use the hand or mouth to communicate meaning. In this article, I review evidence suggesting that the study of gesture-speech links and developmentally prior couplings between the vocal and motor systems in infancy can provide valuable insight into a number of later developments that reflect the cognitive interdependence of gesture and speech. These include aspects of language development and delay, the infant origins of the adult speech-gesture system, and early signs of autism spectrum disorder. Implications of these findings for studying the development of multimodal communication are considered.

KEY-WORDS: GESTURE, LANGUAGE DEVELOPMENT, VOCALIZATION, MOTOR DEVELOPMENT.

Communication is a multimodal phenomenon. Adult interactions are characterized by a complex, fluid, and rapidly evolving interplay between speech and movement in a variety of forms, including altering facial expression, changing eyebrows or head position, and gestures. Indeed, one of the hallmarks of skilled communication is the production of temporally integrated, semantically cohesive utterances that incorporate communicative behaviors from multiple modalities. While the development of this ability takes place over many years (e.g., Koterba, 2010), even very young infants coordinate expressions from different modalities at greater than chance levels (e.g., vocalizations with facial expressions; Yale *et al.*, 1999, 2003). This suggests that from very early in life, expressive behavior is multimodal, with early behavioral coordinations being refined and strengthened over time as they become used for the communication of meaning.

Of these communicative coordinations, those that involve gesture and speech have received perhaps the greatest empirical attention to date. Gestures are a robust feature of adult communication, tightly linked in time and meaning with co-occurring speech (McNeill, 1992, 2005). This tight linkage, the fact that listeners extract meaning from the gestures that co-occur with a spoken message (e.g., McNeill, Cassell, & McCullough, 1994), and the fact that gestures are even produced by speakers who cannot see gesture, those who are congenitally blind, while talking to an interlocutor also known to be blind (Iverson & Goldin-Meadow, 1998) all suggest that the co-production of gesture and speech reflects their common participation and cognitive interdependence in the communicative act.

While there is evidence of this cognitive interdependence in very young children who are just beginning to produce gestures and speech (e.g., Iverson, Capirci, Volterra, & Goldin-Meadow, 2008), little attention has been devoted to examining the developmental origins of the gesture-speech link. One possibility is that the origins of speech/gesture coordinations lie in hand-mouth linkages that are observed in the everyday sensorimotor activity of very young infants who do not yet use the hand or mouth to communicate meaning. In other words, it is the initial *sensorimotor linkages* of these systems that form the bases for their later cognitive interdependence.

Recent research in my laboratory and in that of others has begun to shed light on this issue. The results of this research strongly suggest that the study of gesture-speech links and developmentally prior couplings between the vocal and motor systems in infancy can provide valuable insight into a number of later developments that reflect the cognitive interdependence of gesture and speech. These include aspects of language development and delay, the infant origins of the adult speech-gesture system, and early signs of autism spectrum disorder (ASD).

The goal of this paper is to review this literature, focusing first on changes in gesture that predate and predict advances in children's language development. This will be followed by a discussion of the developmental origins of the gesture-speech link in infant vocal-motor coordination; and, in conclusion, I will

describe research focused on using patterns of vocal-motor and gesture-speech development to provide information that may be useful in the early identification of children who will eventually receive an ASD diagnosis.

CHANGES IN GESTURE PREDATE AND PREDICT ADVANCES IN LANGUAGE

The longstanding claim that gesture provides a way for very young children to communicate information that they cannot yet express verbally has substantial empirical support (e.g., Acredolo & Goodwyn, 1988; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Capirci, Contaldo, Caselli, & Volterra, 2005; Caselli, 1990). That gesture allows children to communicate meanings that they may have difficulty expressing in words raises the possibility that it may facilitate early language learning. If this is the case, then gesture should not only predate but also predict change in language (see also Ozcaliskan & Goldin-Meadow, 2005). We have explored this hypothesis by examining the role of gesture in two domains of early language: lexical development and syntactic development.

Early lexical development. Children begin to produce their first gestures and first words at around the same point, usually between 8 and 14 months of age (e.g., Bates *et al.*, 1979). Once children begin to acquire words, however, gestures do not disappear; rather, they continue for some time to co-exist with words in children's communicative repertoires. In an initial study, we examined children's early word and representational gesture vocabularies (representational gestures are those that convey relatively fixed content across contexts; e.g., shaking the head "no", flapping the arms for "bird", raising the arms for "big") longitudinally at 16 and 20 months of age, identifying the number of meanings that children conveyed in the two modalities (Iverson, Capirci, & Caselli, 1994).

Perhaps the most striking finding of this study was that at 16 months, there was very little redundancy in the semantic content of items in the word vs. representational gesture repertoires of individual children. Children tended to have either a word or a gesture for a specific meaning, but not both. Thus, for example, one 16-month-old child's word vocabulary consisted of only 3 words ("car", "mommy", "no"), but his repertoire of gestures was much more extensive (12 gestures) and included meanings for which he had no corresponding words (e.g., "bye-bye", "no", "fish", "butterfly"). This pattern was evident across the 12 individual children in the study: fewer than 10% of the gestures and words in children's repertoires at 16 months could be considered equivalent in meaning.

It would appear, therefore, that gestures provide a means for very young children to convey meanings that they cannot yet express verbally, thereby not only expanding their communicative potential but providing a way for new meanings to enter children's communicative repertoires and a means for practicing these new meanings. This in turn may lay the foundation for the eventual appearance of these meanings in speech. To address this possibility, we examined longitudinal data from 10 children observed monthly between the ages

of 10 and 24 months and asked whether children's use of gesture to refer to specific objects was related to the emergence of verbal labels for those objects (Iverson & Goldin-Meadow, 2005).

All instances of object reference made by the child were coded and classified according to whether they occurred in *speech only* (e.g., child only says "ball" during a given session), *gesture only* (e.g., child only points to the ball within the session), or *speech and gesture* (e.g., child says "ball" early in the session but points to the ball later in the session). We then identified meanings that were expressed by these lexical items in multiple sessions (41% of all meanings produced across the observation period) and determined whether they appeared: a. initially in speech and remained in speech; b) appeared initially in gesture and remained in gesture; c) appeared initially in speech and transferred or spread to gesture; or d) appeared initially in gesture and transferred or spread to speech.

Results indicated that children relied heavily on gesture to convey meaning during the initial stages of word acquisition, initially referring to a larger set of objects in gesture (75% of all object references) than in speech (25% of such references). Moreover, a substantial proportion (59%) of lexical items either switched to or spread from one modality to the other in the course of the study, but this differed substantially by modality. Of those that switched or spread from one modality to the other, 85% appeared first in gesture and then switched or spread to speech, while only 15% appeared initially in speech and subsequently in gesture. On average, children produced a gesture for a particular object approximately 3 months before they produced the word for that object. In short, by examining children's earlier use of gesture to refer to objects, we were able to predict a substantial proportion of their later word vocabularies.

Early syntactic development. A number of studies have described the frequent production of gesture + word combinations by one-word speakers prior to the emergence of two-word combinations (Butcher & Goldin-Meadow, 2000; Capirci *et al.*, 2005; Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson, Capirci, Volterra, & Goldin-Meadow, 2008; Morford & Goldin-Meadow, 1992). Initially, children combine single words with single gestures such that the two elements either convey *equivalent* (e.g., shaking the head NO while saying "no") or *complementary* meanings (pointing to mother's coffee cup while saying "cup"). Somewhat later in development, children begin to produce gesture + word combinations in which each component provides a different (but related) piece of information about the referent (e.g., pointing to mother's coffee cup while saying "mommy"). These *supplementary* combinations are of particular interest because their semantic content is identical to that of early two-word combinations (i.e., two meanings; in the previous example, "mommy" + "cup"). Thus, if production of supplementary combinations is indicative of the fact that children are cognitively ready to produce utterances that convey two semantic elements but are not yet able to produce two words in succession, then the emergence of supplementary gesture + word combinations would be expected to predict the onset of two-word combinations.

We tested this prediction with the longitudinal data from the 10 children described above, calculating correlations between the ages of onset of complementary combinations, supplementary combinations, and two-word utterances (Iverson & Goldin-Meadow, 2005). The results were striking. While the onset of complementary combinations, which convey only a single semantic element, was only weakly ($r_s = .24$) and non-significantly related to the onset of two-word combinations, the age at which children produced their first *supplementary* combination was highly predictive ($r_s = .91$, $p < .05$) of the age of onset of two-word speech. Children who were among the first to produce supplementary combinations were also among the first to produce two-word utterances. Children who were comparatively slower to produce supplementary combinations were also somewhat slower to produce two-word speech. Thus, it is the ability to combine two different semantic elements within a single communicative act—not simply the ability to produce gesture and speech in combination—that predicts the onset of two-word speech.

The importance of supplementary gesture + word combinations as an index of children's emerging ability to coordinate two distinct meanings within a single, tightly timed communicative message is underscored by findings from children with Down syndrome. Down syndrome (DS) is a genetically based neurodevelopmental disorder characterized, among other things, by general cognitive and expressive language delay (e.g., Chapman, 2003). To assess the emerging ability of children with DS to combine two different meanings within a single message, we analyzed the gesture + word combinations produced by 5 children with DS matched to 5 typically-developing (TD) children on the basis of expressive language ability.

Our matching procedure involved two steps: a) identifying subgroups of TD children whose chronological ages matched the language ages of the children with DS (as determined by administration of the *Primo Vocabolario del Bambino*, the Italian version of the MacArthur Bates Communicative Development Inventory; Caselli & Casadio, 1995); and b) selecting an individually matched TD child for each child with DS on the basis of number of different words produced during a 30-minute play session with a parent (within six words). Although children with DS were, of course, older ($M = 47.6$ months, $SD = 7.95$) than matching TD children ($M = 18.4$ months, $SD = 2.19$), their mental age was generally comparable to (and indeed slightly above) the chronological age of the TD comparison sample ($M = 22.4$ months, $SD = 4.16$). On this basis, one might expect comparable production of two-element combinations. This is not what we found.

First, while three of the TD children each produced a single word + word combination, this structure was not observed in any of the children with DS. Second, while gesture + word combinations were observed in the production of all of the children in both groups and at comparable frequencies ($M_{DS} = 17.0$, $SD = 11.66$; $M_{TD} = 18.4$, $SD = 15.98$), children with DS on average produced approximately twice as many equivalent ($M_{DS} = 11.4$, $SD = 13.09$; $M_{TD} = 5.2$,

$SD = 3.70$) but half as many complementary combinations ($M_{DS} = 5.2$, $SD = 6.8$; $M_{TD} = 10.4$, $SD = 10.36$) as TD comparison children, although neither of these differences was statistically reliable. More importantly, however, supplementary combinations were produced with considerable frequency by 4 of the 5 TD children ($M = 5.6$, $SD = 6.8$) but were almost non-existent among children with DS ($M = 0.6$, $SD = 0.89$).

Taken together, the prevalence of informationally equivalent gesture–word combinations, the relative absence of supplementary gesture–word combinations, and the lack of two-word combinations among the children with DS as compared to TD children at comparable expressive lexical levels and mental age is suggestive of a specific delay in the transition from communication about a single referent to communication about two referents. While delayed language development has been well-documented in children with DS (e.g., Chapman, 2003), analyses of children’s gesture provided a unique source of information regarding this additional pocket of delay that would not have been apparent in an examination of children’s speech alone (see Capone & McGregor, 2004, for additional discussion).

DEVELOPMENTAL ORIGINS OF THE GESTURE-SPEECH LINK

Although a great deal of attention has been devoted to describing the nature and characteristics of the gesture–speech link in adults and children (e.g., see Goldin-Meadow & Iverson, in press, for a review), little research to date has explored its developmental origins. One fundamental characteristic of the adult gesture–speech system is that it is temporally coexpressive. When speakers move their hands to add emphasis to particular words and highlight essential phrases, these gestures are highly synchronous with co-occurring speech, such that the stroke, or active phase, of the gesture is executed just as the related word or phrase is articulated (McNeill, 1992). Gesture and speech, in other words, are closely timed with one another. This relationship is robust even in the face of severe disruptions in the temporal organization of speech (e.g., as in chronic stutterers; Mayberry & Jaques, 2000) and is evident by the time children make the transition to two-word utterances (Butcher & Goldin-Meadow, 2000; Pizzuto, Capobianco, & Devescovi, 2005).

We have recently begun to explore whether vocal-motor linkages in infancy may provide the developmental basis for later speech-gesture co-expression. The rationale for this work is grounded in the fact that connections between the oral/vocal and motor systems are in place at or even prior to birth. The Babkin reflex, for example, can be elicited in newborns by applying pressure to the palm; infants react to this manual stimulation by opening their mouths (Babkin, 1960). Coordination between oral and manual actions is also common in young infants’ spontaneous movements. When newborns bring their hands to the facial area to introduce the fingers for sucking, they open their mouths as the hand is moving toward the facial area, in anticipation of its arrival (Butterworth & Hopkins,

1988; Lew & Butterworth, 1997). Similar movements have also been observed in fetal activity by 12 to 15 weeks gestational age and as frequently as 50-100 times per hour (de Vries, Visser, & Prechtel, 1984). Hand-mouth linkages are also apparent in communicative settings: 9- to 15-week-old infants are especially likely to produce extensions of the index finger with either vocalization or mouthing movements during face-to-face interaction with their mothers (Fogel & Hannan, 1985).

Thus, in the newborn infant, an initial hand-mouth linkage of the sort just described may provide a foundation for the development of an integrated gesture-speech system. This early linkage has been incorporated as a starting state into a model of vocal-motor development proposed by Iverson & Thelen (1999). This model was developed in an effort to incorporate evidence on vocal-motor linkages in infancy into an integrated view of the infant origins of gesture-speech timing. The model situates the emergence of gesture-speech coordination within the broader context of the development of coordinated movement. Although gesture and speech are produced in order to convey meaning, their co-production requires the ability to produce controlled, voluntary movements in the two effector systems (the vocal tract and the manual system) and to coordinate these movements in time and space. Thus, the development of the gesture-speech system is fundamentally and inextricably tied to the larger problem of the development of motor control.

The overarching goal of the Iverson & Thelen (1999) model is to provide a principled understanding of how the dynamics of change in the strength and stability of early vocal and motor skills can account for the emergence of the ability to link the two modalities in a single, coordinated behavior with common communicative intent. Two of the model's major concepts – oscillation and entrainment – are briefly discussed below.

Oscillation is characteristic of the behavior of developing motor systems. Neuromotor systems under immature or impaired voluntary control have the tendency to oscillate naturally. In the infant motor system, these oscillations take the form of rhythmic, stereotyped behaviors such as shaking, kicking, rocking, and bouncing (Thelen, 1979, 1981a, 1981b). Rhythmic movements of the upper and lower limbs and of the whole body are extremely common throughout the first year, and they appear to be closely associated with moments of transition from no voluntary control over a limb or body segment to adaptive, intentional control (e.g., infants rock on all fours before they crawl and wave their arms before they reach). In the infant vocal system, properties similar to these are apparent in reduplicated babbling. Infants begin to produce reduplicated strings of a single consonant-vowel syllable (e.g., [dadada], [gagaga]) between the ages of 6 and 8 months (e.g., Oller & Eilers, 1988). MacNeilage and Davis (1993) have argued that reduplicated babble begins fundamentally as a mandibular oscillation: the repeated lowering and raising of the mandible results in a perceived contrast between consonants (produced when the vocal tract is closed) and vowels (produced when the vocal tract is in an open configuration). The range of syllabic

patterns produced by infants begins to widen as they gain greater control over the tongue and its position in the vocal tract during phonation.

An implication of this view for the developing gesture-speech system is clear: given the oscillatory properties of the infant motor and vocal systems, and assuming an initial coupling of speech and motor activity based on the research described above, the emergence and production of rhythmic behaviors in one effector system should affect activity in the other. This leads directly to the second key concept of the Iverson and Thelen (1999) model, namely that when oscillators are coupled, each tries to draw the other into its characteristic oscillation pattern. *Entrainment* occurs when one oscillator successfully pulls in the activity of the other, resulting in an ordered patterning of coordinated activity. This cooperativity occurs strictly as a function of the coherence of the parts under certain energetic constraints and occurs without any executive direction. Many such self-organized patterns occur in nature in physical and biological systems, with no cognitive intervention (see Kelso, 1995).

In infancy, as we have seen, a coupling of the vocal and motor systems (particularly the manual system) appears to exist from early in development. This suggests the possibility that development of vocal-motor coordinations may be characterized in terms of entrainment. Given sufficient activation in one component of the system, its influence should extend to the other component. When this happens, we would expect the first system to pull in and entrain the activity of the complementary system. For example, when an infant is engaged in an intense bout of rhythmic limb movement, the level of activation in the motor system may spill over into the vocal system and entrain its activity, resulting in production of a vocalization. Importantly, however, this entrainment is dynamic and flexible, such that activation of one system can have various effects on the other—tight temporal synchrony (e.g., arm swinging accompanied by a string of repeated syllables, each articulating with a movement cycle), or more loosely coupled influence (e.g., arm swinging accompanied by a short vocalization.)

In light of these considerations, we analyzed infants' production of bouts of co-occurring vocal and rhythmic motor behaviors in an effort to: a) address descriptive questions regarding the extent to which such coordinations may share characteristics of adult gesture-speech co-productions; and b) examine hypotheses derived from the Iverson & Thelen (1999) model of vocal-motor coordination. In an initial cross-sectional study, we videotaped groups of 6, 7, 8, and 9-month-old infants at home during naturalistic observation and semistructured play with a primary caregiver (Iverson & Fagan, 2004). Rhythmic limb movements (defined as movements repeated in approximately the same form at least three times at regular intervals of approximately 1 sec or less; Thelen, 1979) and vocalizations (pre-speech sounds excluding laughter, crying, and vegetative sounds) were coded, and vocal-motor coordinations were defined as instances in which vocal and motor behaviors shared some degree of temporal overlap. Coordinations were further classified according to whether

they included manual (fingers, wrist, hand, arm) or nonmanual (leg, foot torso, head) movements. Finally, we gathered data on the onset of reduplicated babble via parent report confirmed by experimenter observation.

From a developmental perspective, if infant vocal-motor coordinations are precursory to the production of mature gesture-speech coordinations, then the frequency of vocal-motor coordinations should increase as infants approach the age of emergence of gestures and meaningful speech. In addition, since adult gestures consist primarily of movements of the hands and arms and the manual movements of gesture are consistently related in time to co-occurring verbal production, either slightly anticipating or occurring in synchrony with speech (McNeill, 1992, 2000), we might also expect these characteristics to be apparent in coordinated bouts of infant vocal and motor behavior.

Vocal-motor coordinations were frequently observed and were produced by 41 of the 42 infants in the study, suggesting that they are a robust feature of spontaneous infant behavior (see also Ejiri & Masataka, 2001). Consistent with our expectations, the rate of production of vocal-motor coordinations increased almost tenfold across the 6- to 9-month age range. In addition, a majority of infant coordinations consisted of manual (generally involving one arm) rather than nonmanual movements. Finally, with regard to timing, the vast majority (83%) of infants' vocal-motor coordination bouts were either movement-initiated or synchronous. These latter two findings parallel those for the adult system described above.

With regard to predictions derived from the Iverson & Thelen (1999) model of gesture-speech development, we explored a pair of hypotheses regarding entrainment as a possible mechanism underlying the coordination of vocal and motor activity. The first had to do with the effect of entrainment on coordinated behaviors. Because entrainment results in fundamental alteration of features of the component behaviors, the rhythmic organization of limb movements should be reflected in co-occurring vocalizations. In the vocal system, rhythmic organization is typified in consonant-vowel structure; and thus vocalizations coordinated with rhythmic limb movement should be especially likely to contain consonant-vowel (CV) repetitions. Data were consistent with this prediction: among vocalizations containing a CV repetition, a significantly greater proportion ($M = 24\%$) were produced in coordination with rhythmic movement than in isolation ($M = 17\%$).

The second hypothesis had to do with the impact of reduplicated babble onset on the likelihood of vocal-manual entrainment. When infants begin to babble, babbled vocalizations reflect rhythmic organization in much the way that developmentally-prior limb movements involve rhythmic organization. Thus, the onset of reduplicated babble marks the emergence of opportunity for *mutual* entrainment of the vocal and manual systems: rhythmic organization can now spill over from either of the component systems to the other, thereby increasing the overall likelihood of vocal-manual entrainment. This leads to the expectation that among infants who have begun to babble, the proportion of rhythmic

manual behaviors coordinated with vocalization should be higher than that for infants who have not yet begun to babble.

To control for potential effects of age on frequency of coordination, we tested this prediction by examining data taken exclusively from 6-month-old infants. Of these infants, 10 had begun to babble, and 6 were prebabblers. Results revealed that the median proportion of rhythmic manual behaviors coordinated with vocalization was slightly more than twice as high for babblers ($Mdn = .18$) than prebabblers ($Mdn = .08$). Furthermore, the respective distributions of these proportions across individual infants in the two groups were relatively nonoverlapping.

To summarize, our work to date on infant-vocal motor coordination suggests that vocal-motor coordinations are commonly observed in infant behavior, that infants coordinate vocalizations with limb movements (particularly movements of the hands and arms) prior to the emergence of gesture and speech, and that these coordinations become increasingly frequent as infants approach the age range in which gesture and speech typically begin to emerge. Moreover, bouts of vocal-motor coordination share some features of adult gesture-speech co-productions. They more often involve the hands and arms than other limbs, and, as is the case for adult gesture-speech co-productions, they are exceedingly likely to be initiated by movement or to be synchronous. These data, which are generally consistent with Iverson and Thelen's (1999) proposal that entrainment is a candidate mechanism underlying coordination of vocal and motor activity, suggest that performance of bouts of coordinated behavior provide infants with opportunities to practice integrating activity across the two modalities, a skill that is required for the synchronous production of gesture and speech.

VOCAL-MOTOR DEVELOPMENT, SPEECH-GESTURE DEVELOPMENT, AND THE EARLY IDENTIFICATION OF AUTISM SPECTRUM DISORDERS

Autism spectrum disorder (ASD) is a neurodevelopmental disorder involving primary deficits in social interaction, language, and communication (American Psychiatric Association, 1994). Although not included in the diagnostic criteria, motor difficulties (e.g., Fournier *et al.*, 2010) and vocal atypicalities (e.g., Sheinkopf, Mundy, Oller, & Steffens, 2000; Wetherby, Cain, Yonclas, & Walker, 1988) are common in children with ASD; and these difficulties appear to be closely correlated. Thus, for example, in a longitudinal study of 35 children with ASD, Stone and Yoder (2001) found that after controlling for expressive language level at 2 years of age, motor imitation was a significant predictor of language abilities at age four. In addition, production of early oral and manual-motor behaviors in children with ASD is related to language outcomes in early childhood, with nonverbal children reported to produce significantly fewer of both types of behaviors than children who eventually acquire fluent speech (Gernsbacher *et al.*, 2008).

Despite the fact that many parents of children with ASD report having been concerned about the child's development prior to 12 months (e.g., Coonrod & Stone, 2004), ASD is very difficult to diagnose reliably before age two (e.g., see Rogers, 2001). Because an ASD diagnosis involves delayed development in a variety of domains, such as pointing, language, and symbolic play, clinicians must wait until well after children have reached the typical age of emergence for such behaviors before considering such a diagnosis. This has led to a surge of interest in the identification of early behavioral markers of risk for a later ASD diagnosis.

The presence of difficulties and delays in the vocal and motor systems of children with ASD raises an intriguing possibility with regard to potential early markers of risk for an ASD diagnosis. If, as previously argued, the origin of gesture-speech coordination lies in infant vocal-motor coordination, and in light of the large literature documenting the existence of disorders and delays in language, gesture, and motor abilities in older children with ASD, then atypical vocal-motor and speech-gesture development and coordination in infants could serve as an early diagnostic marker for ASD.

Since ASD is a relatively low incidence disorder, with a prevalence of approximately 9.0 per 1000 children, prospective identification of infants in the general population who will ultimately receive an ASD diagnosis would require longitudinal assessment of an impossibly large sample (e.g., recruiting over 2000 children in order to ensure the presence of a subgroup of approximately 20 children with an eventual ASD diagnosis). This has led a number of researchers to adopt a strategy in which infants at especially high risk for ASD (HR infants) are targeted for study. One such high-risk group involves infants who have an older sibling already diagnosed with the disorder. Inasmuch as the recurrence risk for ASD in later-born siblings of children with autism is approximately 18% (e.g., Yirmiya *et al.*, 2007; Zwaigenbaum *et al.*, 2005), more than 200 times that in the general population (Ritvo *et al.*, 1988), focusing on these infants is likely to yield a subset of infants who will go on to receive an ASD diagnosis.

In a recently completed longitudinal study, we followed a group of 21 HR infants and a comparison group consisted of 18 infants with a typically-developing older sibling and no family history of ASD (Low Risk infants; LR) between the ages of 5 and 36 months. All infants were videotaped at home monthly for approximately 45 minutes between the ages of 5 and 14 months with a follow-up session at 18 months; sessions included naturalistic observation and semistructured play segments. At 36 months, the Autism Diagnostic Observation Schedule (ADOS; Lord *et al.*, 2000) was administered to all HR infants for purposes of diagnostic outcome classification; 3 HR infants received an autism diagnosis.

One of the goals of this study was to explore whether vocal-motor and gesture-speech links in HR infants later diagnosed with autism differed from those in LR infants and in HR infants with no eventual ASD diagnosis. With regard to vocal-motor links, studies of TD infants have reported that the onset of reduplicated babbling is accompanied by an increase in frequency of rhythmic

arm activity, followed by a subsequent decline (Ejiri, 1998; Thelen 1979). We therefore analyzed the frequency of rhythmic arm movements at three sessions: that coinciding with the onset of reduplicated babble (identified by parent report of the production of reduplicated syllables and confirmed via experimenter observation) and the sessions one month prior to and one month after onset (Iverson & Wozniak, 2007).

Data from the LR group replicated the pattern of change reported in previous studies of TD infants, with a clear increase in arm movement from the pre-babble to the babble onset and a subsequent decline. An increase from pre-babble to babble onset was also apparent among HR infants, but it was somewhat attenuated: the relative difference in rate of rhythmic arm movement between the babble onset session and the pre- and post-babble sessions combined was substantially lower for HR relative to LR infants. With regard to the infants who eventually received an autism diagnosis, one failed to produce reduplicated babble at any session, a second babbled very late (onset at 18 months), and the third (who babbled at 7 months) exhibited attenuated change in rhythmic arm movement across the three sessions comparable to that observed in other HR infants with no subsequent ASD diagnosis. Thus, although there is considerable variability in this regard, these findings suggest that delays or atypicalities in the nature and strength of early vocal-motor links may be characteristic of some HR infants, perhaps especially HR infants who go on to receive an autism diagnosis (see Iverson & Wozniak, 2007, for further discussion). Delays or atypicalities of this sort might be expected to influence the emergence of later gesture-speech links.

To examine the emergence and development of the gesture-speech link, we coded communicative gestures, vocalizations, words, and gesture + speech combinations (i.e., gesture + vocalization, gesture + word) produced spontaneously at the sessions when children were 13 and 18 months of age respectively (Iverson, Paradé, Winder, & Wozniak, 2009). The first finding of interest was that the overall rate of spontaneous communication in the two sessions (i.e., collapsing across gestures, words, vocalizations, and combinations) was significantly lower for HR relative to LR infants at both ages, and was lowest for the three infants who eventually received an autism diagnosis. On average, HR infants produced approximately 7 and 14 spontaneous communications per 10 minutes at 13 and 18 months respectively. By contrast, the relevant LR group means were 15 and 20 communications per 10 minutes. The infants later diagnosed with autism were at the bottom of the distribution at both ages, producing fewer than 5 spontaneous communications per 10 minutes even by 18 months.

With regard to gesture-speech coordination, the mean proportions of gestures produced in combination with speech (either a vocalization or a word) were similar for the HR and LR infants; roughly 40-50% of children's gestures were coordinated with a verbalization at both ages. This was not the case, however, for the three infants eventually receiving an autism diagnosis. All three exhibited extreme delays in production of gesture-speech combinations and were once

again at the bottom of the distributions at both ages. At 13 months, none of the three infants produced a single two-element combination. By contrast, all of the LR children and all but two of the HR children combined gestures with speech at this age. By 18 months, all children except those later diagnosed with autism produced gesture-speech combinations with some frequency. Two of the children diagnosed with autism, however, produced only a single gesture + vocalization combination during the session; and combinations continued to be absent from the production of the third child.

Overall, these data are consistent with the notion that delays and/or disruptions in early infant vocal-motor coordination may be related to subsequent delays in the emergence of the ability to combine speech with gesture in a single, well-timed utterance. Among HR infants who did not receive an ASD diagnosis, an attenuated pattern of change in arm rhythmicity was apparent at babble onset. And although gestures were subsequently combined with speech in proportions similar to those for LR children, HR infants, particularly at 13 months of age, exhibited a more restricted repertoire of gesture-speech combinations suggestive of reduced flexibility in the ability to coordinate communicative behaviors across the two modalities. Finally, the three infants who eventually received an autism diagnosis exhibited extreme delays in production of gesture-speech combinations, delays that were developmentally preceded by atypical patterns of vocal and motor development. While these findings are preliminary and based on a very small number of children, they point to the potential utility of vocal-motor coordination in general and gesture-speech combinations in particular as a marker of future ASD diagnosis.

CONCLUSIONS

The research described in this paper indicates that changes in gesture predate and predict advances in children's language development, that the developmental origins of the adult gesture-speech link appear to lie in infant vocal-motor coordination, and that patterns of vocal-motor and gesture-speech development may be useful in the early identification of ASD. Two major conclusions can be drawn from this research.

First, the data reported here support the general notion that long before infants make use of the hand or mouth for intentional communication, the sensorimotor linkages in these systems provide the basis for their later cognitive interdependence. Multimodal behavioral coordinations, in other words, exist from very early in life; and as development proceeds, these coordinations are refined and strengthened so that when the child becomes capable of the intentional communication of meaning, that communicative expression is, from the outset, multimodal.

Second, these data also reinforce the value of studying multimodal behavioral coordinations in infancy. Although research in our laboratory has focused primarily on vocal-motor and speech-gesture coordinations as they provide insight into later developments that reflect the cognitive interdependence of

gesture and speech, it seems likely that broadening this approach to include facial expression and other communicative movements as well as gesture will provide critical developmental information that is not provided by a study of speech alone.

Author Note

The research reported in this article was supported by grants from the National Institute of Health (R01 HD41677 and R01 HD54979) and Autism Speaks. I am grateful to members of the Infant Communication Lab, University of Pittsburgh, for discussion of many of the ideas presented here and to Robert H. Wozniak for extensive and insightful comments on previous versions of the manuscript.

REFERENCES

- Acredolo, L. P. & Goodwyn, S. W. (1988). Symbolic gesturing in normal infants. *Child Development, 59*, 450-466.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th Edition). Washington, D.C.: American Psychiatric Association.
- Babkin, P.S. (1960). [The establishment of reflex activity in postnatal life.] In *Central nervous system and behavior (Trans. from Fizjologii, Zburnal [USSR], 1953, 922-927)*, Russian Scientific Translation Program. Bethesda, MD: National Institutes of Health.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L. & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.
- Butcher, C. & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill, (Ed.), *Language and gesture* (pp. 235-257). Cambridge: Cambridge University Press.
- Butterworth, G. & Hopkins, B. (1988). Hand-mouth coordination in the new-born baby. *British Journal of Developmental Psychology, 6*, 303-313.
- Capirci, O., Contaldo, A., Caselli, M.C. & Volterra, V. (2005). From action to language through gesture: A longitudinal perspective. *Gesture, 5*, 155-177.
- Capirci, O., Iverson, J.M., Pizzuto, E. & Volterra, V. (1996). Communicative gestures during the transition to two-word speech. *Journal of Child Language, 23*, 645-673.
- Capone, N.C. & McGregor, K.K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech Language and Hearing Research, 47*, 173-186.
- Caselli, M.C. (1990). Communicative gestures and first words. In : V. Volterra, & C.J. Erting, (Eds.), *From gesture to language in hearing and deaf children* (pp. 56-67). New York: Springer-Verlag.
- Caselli, M.C. & Casadio, P. (1995). *Il Primo Vocabolario del Bambino*. Milan: Franco Angeli.
- Chapman, R.S. (2003). Language and communication in individuals with Down syndrome. In : L. Abbeduto, (Ed.), *International Review of Research in Mental Retardation: Language and Communication, vol. 27* (pp. 1-34). New York: Academic Press.

- Coonrod, E.E. & Stone, W.L. (2004). Early concerns of parents of children with autistic and nonautistic disorders. *Infants and Young Children*, 17, 258-268.
- de Vries, J.I.P., Visser, G.H.A. & H.F.R. Prechtl, (1984). Fetal motility in the first half of pregnancy. In H.F.R. Prechtl, (Ed.), *Continuity of neural functions from prenatal to postnatal life* (pp. 46-64). Oxford: Spastics International Medical Publications.
- Ejiri, K. (1998). Relationship between rhythmic behavior and canonical babbling in infant vocal development. *Phonetica*, 55, 226-237.
- Ejiri, K. & Masataka, N. (2001). Co-occurrence of preverbal vocal behavior and motor action in early infancy. *Developmental Science*, 4, 40-48.
- Fogel, A. & Hannan, T.E. (1985). Manual actions of nine to fifteen week-old human infants during face-to-face interactions with their mothers. *Child Development*, 56, 1271-1279.
- Fournier, K.A., Hass, C.J., Naik, S.K., Lodha, N. & Cauraugh, J.H. (2010). Motor coordination in Autism Spectrum Disorders: A synthesis and meta-analysis. *Journal of Autism and Developmental Disorders* (DOI 10.1007/s10803-010-0981-3, published online 2 March 2010).
- Gernsbacher, M.A., Sauer, E.A., Geye, H.M., Schweigert, E.K. & Goldsmith, H.H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry*, 49, 43-50.
- Goldin-Meadow, S. & Iverson, J.M. (in press). Gesturing over the lifespan. In R. Lerner, and W. Overton, (Eds.), *Handbook of lifespan development. Volume 1: Methods, biology, neuroscience, and cognitive development*. New York: John Wiley.
- Iverson, J.M., Capirci, O. & Caselli, M.C. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23-43.
- Iverson, J.M., Capirci, O., Volterra, V. & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication of Italian vs. American children. *First Language*, 28, 164-181.
- Iverson, J.M. & Fagan, M.K. (2004). Infant vocal-motor coordination: Precursor to the gesture-speech system? *Child Development*, 75, 1053-1066.
- Iverson, J.M. & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, 396, 228.
- Iverson, J.M. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 367-371.
- Iverson, J.M. & Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies*, 6, 19-40.
- Iverson, J.M., Paradé, M.V., Winder, B.M. & Wozniak, R.H. (2009, April). "Early development of the gesture-speech system in infants at risk for ASD." *Paper presented in the symposium "Hand and Mind in Autism: Co-Speech Gestures in Autism Spectrum Disorders" (I-M. Eigsti, Chair) at the Biennial Meetings of the Society for Research in Child Development, Denver, CO.*
- Iverson, J.M. & Wozniak, R.H. (2007). Variation in vocal-motor development in infant siblings of children with autism. *Journal of Autism and Developmental Disorders*, 37, 158-170.
- Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

- Koterba, E.A. (2010). *Conversations between friends: Age and context differences in the development of nonverbal communication in preadolescence*. Unpublished doctoral dissertation, University of Pittsburgh.
- Lew, A.R. & Butterworth, G. (1997). The development of hand-mouth coordination in 2- to 5-month-old infants: Similarities with reaching and grasping. *Infant Behavior and Development*, 20, 59-69.
- Lord, C., Risi, S., Lambrecht, L., Cook Jr., E.H., Leventhal, B.L., DiLavore, P. C., Pickles, A. & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205-223.
- Mayberry, R.I. & Jaques, J. (2000). Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In: McNeill, D. (Ed.), *Language and gesture* (pp. 199-214). Cambridge: Cambridge University Press.
- Morford, M. & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language*, 19, 559-580.
- MacNeilage, P.F. & Davis, B.L. (1993). Motor explanations of babbling and early speech patterns. In : B.de Boysson-Bardies, S.de Schoenen, P. Jusczyk, P.F. MacNeilage, & J. Morton, (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 341-352). Dordrecht: Kluwer.
- McNeill, D. (1992). *Hand and mind: What gesture reveals about thought*. Chicago: University of Chicago Press.
- McNeill, D. (Ed.) (2000). *Language and gesture*. Cambridge: Cambridge University Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- McNeill, D., Cassell, J. & McCullough, K-E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction*, 27, 223-237.
- Oller, D.K. & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development*, 59, 441-466.
- Ozcaliskan, S. & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of language development. *Cognition*, 96, B101-B113.
- Pizzuto, E., Capobianco, M. & Devescovi, A. (2005). Gestural-vocal deixis and representational skills in early language development. *Interaction Studies*, 6, 223-252.
- Ritvo, E.R., Jorde, L.B., Mason-Brothers, A., Freeman, B.J., Pingree, C., Jones, M.B., McMahon, W.M. & Petersen, P.B. (1989). The UCLA-University of Utah epidemiologic survey of autism: Recurrence risk estimates and genetic counseling. *American Journal of Psychiatry*, 146, 1032-1036.
- Rogers, S.J. (2001). Diagnosis of autism before the age of 3. In L.G. Masters, (Ed.), *International review of research in mental retardation* (vol. 23, pp. 1-31). New York: Academic Press.
- Sheinkopf, S.J., Mundy, P., Oller, D.K. & Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders*, 30, 345-354.
- Stone, W.L. & Yoder, P.J. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism*, 5, 341-361.

- Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behaviour*, 27, 699-715.
- Thelen, E. (1981a). Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology*, 17, 237-257.
- Thelen, E. (1981b). Kicking, rocking, and waving: Contextual analyses of rhythmical stereotypies in normal human infants. *Animal Behaviour*, 29, 3-11.
- Wetherby, A., Cain, D., Yonclas, D. & Walker, V. (1988). Analysis of intentional communication of normal children from the prelinguistic to the multiword stage. *Journal of Speech and Hearing Research*, 31, 240-252.
- Yale, M. E., Messinger, D. S., Cobo-Lewis, A. B., Oller, D. K., & Eilers, R. E. (1999). An event based analysis of the coordination of early infant vocalizations and facial actions. *Developmental Psychology*, 35 (2), 505-513.
- Yale, M. E., Messinger, D. S., Cobo-Lewis, A. B. & Delgado, C. F. (2003). Facial expressions of emotion: A temporal organizer of early infant communication. *Developmental Psychology*, 39, 815-824.
- Yirmiya, N., Gamliel, I., Shaked, M. & Sigman, M. (2007). Cognitive and verbal abilities of 24- to 36-month-old siblings of children with autism. *Journal of Autism and Developmental Disorders*, 37, 218-229.
- Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J. & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, 23, 143-152.