Gesture and the emergence and development of language

Virginia Volterra, Maria Cristina Caselli, Olga Capirci, Elena Pizzuto [1]
ISTC, CNR

(*to appear in M. Tomasello and D. Slobin, (Eds.) Elizabeth Bates: A Festschrift. Mahwah, N.J.: Lawrence Erlbaum Associates*)

> "Nature is a miser. She clothes her children in hand-me-downs,
> builds new machinery in makeshift fashion
> from sundry old parts…," (Bates, 1979: 1).

Preface (by Virginia Volterra)

During the autumn of 2002, just a few weeks before we would discover about her disease, Liz and I were walking through Rome, talking about work and a dream we had. We wanted to write something together again, something about gesture—a topic in which we had been interested since we started our lifelong collaboration and profound friendship, about thirty years ago. We were aware that many of our old ideas about the role of gesture in children's linguistic development were suddenly becoming extremely "modern" and we were planning to articulate our current perspective (old and new at the same time) by doing a critical review of recent work carried on in different laboratories and countries.

As we had done many times in the past, we started to write the manuscript "a due mani" (two-handed) but despite the attempts we made during the first half of this awful 2003, we did not have enough energy to complete that work.

The present chapter is meant to be a modest, partial attempt to realize that dream: it has been written "a quattro mani" (four-handed) by four people of the "Nomentana Lab" who share a common debt: a debt of immense gratitude to Liz who has forever marked their life with her unique, intense depth and generosity, as a scientist and as a human being.

Introduction

We would like to frame our observations within the context of current discussions of the origins of language, a topic that has been debated, from different perspectives, since antiquity. Early accounts of language origins often contained speculations about the relationship between language and gesture, including the idea that our hominid ancestors communicated through hand signs, which served as the "missing link" in language evolution. Adam Kendon (2002) recently provided a very elegant review of theories about a gestural origin of language showing the relationship between these theories and a deep interest in deaf people and their signed communication. For example the eighteenth-century Neapolitan philosopher, Giambattista Vico, in *La Scienza Nuova* (1744/1953) formulated his theory on the origin of language according to which in the beginning, humans were mute and communicated by gesture, not by speech. Similar ideas on the first forms of language

[1] Institute of Cognitive Sciences and Technologies, Consiglio Nazionale delle Ricerche (CNR)
Via Nomentana 56, 00161 Rome, Italy.
E-mail addresses: v.volterra@istc.cnr.it, c.caselli@istc.cnr.it, capirci@ip.rm.cnr.it, e.pizzuto@istc.cnr.it

being rooted in action or gesture were debated at about the same time in Paris by such thinkers as Condillac and Diderot. For Condillac, language began in the reciprocation of overt actions, making its first form a language of action. This led Condillac to write about the language of gesture, both as this was practiced in the pantomimes of antiquity and as it might be observed among deaf people (Kendon, 2002: 37).

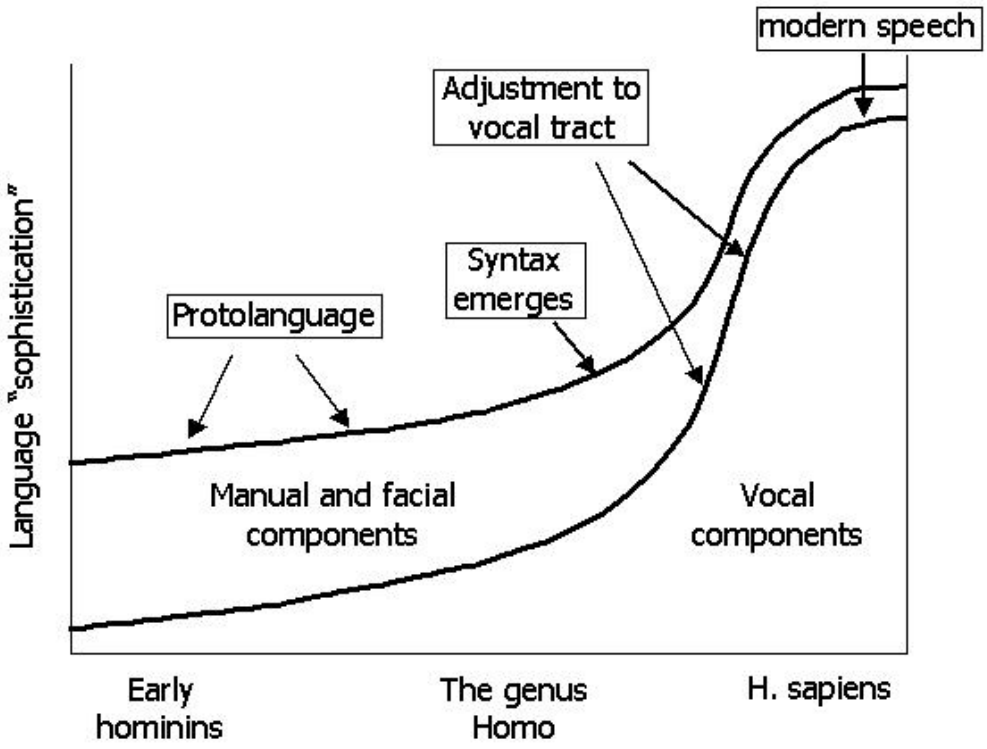
In the nineteenth century comparative linguists like Bopp, Schleicher, Humboldt, and Muller, speculating on a universal language from which all modern languages were supposed to originate, formulated various hypotheses on the use of onomatopoeia and of so-called "acoustic gestures," which may have originally accompanied expressive gestures but then became more sophisticated, and were progressively detached from gestures (Leroy, 1969).

The issue was even too much debated and in 1866 the Société de Linguistique banned papers on language origins, stating that: "The Society will accept no communication concerning either the origin of language or the creation of a universal language." The London Philological Society did the same in 1872. The ban was so effective that the topic of language origins was almost ignored until the second half of the twentieth century when scientists from different disciplines like anthropology, paleontology, primatology, and linguistics came together for a symposium at the 1970 meeting of the American Anthropological Association. Many of the papers presented were collected and published few years later in a volume entitled *Language Origins* (Wescott, Hewes & Stokoe, 1974). In order to get this book into print, Stokoe established a small publishing company, Linstok Press. This book put forward again the theory of a gestural origin of language showing that a great accumulation of ethological, neurological, and paleontological data relevant to the study of language made it possible to develop a scenario for the origins of language.

It is not accidental that around the same years a large body of research was developed in two areas strictly related to the above issue: sign language and language acquisition in deaf signing children. The study of the visual–gestural or signed languages used by deaf people has shown that gestures can, and indeed do develop into full-blown linguistic systems, with functions and properties that are largely comparable to those of vocal languages. This general result highlights the links and continuity that relate gestures to language systems. Due to the gestural substance of these languages, the comparative, crosslinguistic and crossmodal exploration of signed and spoken languages also provides unique insights with respect to the distinctive features of human language, the extent to which these can be influenced by the modality of production, and their evolutionary path (Armstrong, Stokoe & Wilcox, 1995).

The study of language acquisition by children exposed to sign language has highlighted interesting relationships between gestures and signs and between the acquisition of spoken and signed languages. Such comparisons have promoted remarkable advancements in the study of language development in human infants. Significant insights into the organization and evolution of language have been gained through studies aimed at clarifying the interplay between the vocal and the gestural modality in early development, and the more general cognitive roots and developmental precursors of language (Volterra & Erting, 1990).

The most recent formulation of a theory of a gestural origin of language has been provided by Corballis (2002) who, in his recent book *From Hand to Mouth*, has proposed that gesture has existed side by side with vocal communication for most of the last two million years, a hypothesis that has also been put forward by other scholars (Hewes, 1976; Armstrong, Stokoe & Wilcox, 1995; Deacon, 1997).



According to Corballis, something over 30 million years ago great apes differentiated from the Old World monkeys, and by around 16 million years ago larger brains probably heralded an increase in thinking, including enhanced representation of objects in the brain and the capacity of using a form of protolanguage. Around 5 or 6 million years ago, bipedalism was the main characteristic of early hominids that distinguished them from the other great apes and that had freed their hands and arms for more effective gesturing. But the advance from protolanguage to true grammatical language may not have begun until the genus *Homo* emerged, sometime around two million years ago. This branch of hominids was distinguished by an increase in brain size, the invention of stone tools, and the beginnings of multiple migrations out of Africa, and it is likely that language became increasingly sophisticated from then on. For most of this period, language would have been primarily gestural, although increasingly punctuated by vocalizations. An indirect evidence of the gestural origin of language is that articulate speech would have required extensive changes to the vocal tract along with the cortical control of vocalization and breathing. The evidence suggests that these were not completed until relatively late in the evolution of the genus *Homo*.

The adaptations necessary for articulate vocalization may have been selected, not as a replacement for manual gestures, but rather to augment them. Since many species show a left-hemispheric dominance for vocalization (a bias that may go back to the very origins of the vocal cords), as vocalizations were increasingly incorporated into manual gesture, this may have created a left-hemispheric bias in gestural communication as well. *Homo sapiens* discovered that language

could be conveyed more or less autonomously by speech alone, and this invention may have been as recent as 50,000 years ago (Corballis, 2002). Gesture was not simply replaced by speech. Rather, gesture and speech have co-evolved in complex interrelationships throughout their long and changing partnership. If this account is correct, then both modalities should still exhibit evidence of their prolonged co-evolution, reflected in certain universal (or near-universal) interdependencies, as well as predispositions that reflect the comparative recency or antiquity of these abilities (Deacon, 1997).

The tight relationship between language and gesture described above is compatible with recent discoveries regarding the shared neural substrates of language and meaningful actions that, in the work developed by Rizzolatti's laboratory (Galllese, Fadiga, Fogassi & Rizzolatti, 1996; Rizzolatti & Arbib, 1998) have been likened to gestures. Specifically, Rizzolatti and his colleagues have demonstrated that hand and mouth representations overlap in a broad frontal-parietal network called the "mirror neuron system," which is activated during both perception and production of meaningful manual action and mouth movements. The discovery of "mirror neurons" in the monkey brain provided a significant support to the notion of a gestural origin of language. These neurons respond both when the monkey makes a grasping movement and when it observes the same movement made by others. Since the mirror-neuron system is present in both monkeys and humans, it was most likely present in the common ancestor, providing a basis for a form of communication that was voluntary and flexible rather than fixed (Corballis, 2002).

In the present chapter we review a set of studies conducted in our laboratory that bear on the broader issues outlined above. These studies provide evidence on the continuity between prelinguistic and linguistic development, and on the interplay between the gestural and the vocal modalities in both typically developing children and in children with Down and Williams syndromes, whose development proceeds in atypical conditions. Corballis' (2002) evolutionary views on a slow transition from gesture to vocal language appear to be supported by our developmental data, as this transition, and the interdependency between gesture and speech, seem still evident in children's communicative and linguistic development. As observed by Deacon, it is of course unlikely that language development recapitulates "language evolution in most respects (because neither immature brains nor children's partial mapping of adult modern languages are comparable to mature brains and adult languages of any ancestor)" (Deacon, 1997 p. 354), but we can gain useful insights into the organization and evolution of both language and gesture by investigating the interplay between these modalities in the communication and language systems of children with typical and atypical development.

Earlier work on gesture and the emergence of language

The first investigation on the role of gesture in the emergence of language conducted at our institution was a longitudinal (for that time pioneering) study by Bates, Camaioni, and Volterra (1975). That study involved three infant girls aged 2, 6, and 12 months, at the beginning of the study, observed (with home visits at two-week intervals) over a period of eight months. At the end of this period the three infants overlapped one another in development.

The study aimed to explore:

- the continuity from precommunicative schemes, to preverbal communication, to verbal interaction;
- cooccurring developments in other domains, such as nonverbal cognition and social relations;
- the kinds of performative intentions (e.g., declaring, ordering, asking) that emerged from the above developments.

The results indicated that the primary cognitive prerequisite for performative intentions was Piaget's sensorimotor stage 5, in particular the ability of tool use (see also Bates, Benigni, Bretherton, Camaioni, & Volterra, 1977). In the same period in which the children we observed used supports or sticks to pull objects (8/10 months of age) they also began:

- to use objects as "tools" to obtain adult attention, while producing communicative behaviors like showing, giving, communicative pointing called *protodeclarative*;
- to invoke adult help in obtaining objects by means of ritualized request or communicative pointing called *protoimperative.*

The first one-word labeling appeared within the same kinds of communicative sequences in a later period, corresponding to Piaget's sensorimotor stage 6, when other abilities like deferred imitation, memory for absent objects or people, and initial form of "pretend" play began to emerge.

The first stage involving intentional communication, but not necessarily speech, was termed the *illocutionary phase* (after Austin, 1962), while the following stage, involving the use of words in the same performative sequences, was termed the *locutionary phase.* The use of terms like phase and stage did not imply sudden and qualitative shift but rather a gradual transition from:

- wordlike sounds in the service of performative functions (e.g., the sound "Mmmm!" used to accompany all requests);
- semi-referential words, in which a relation between sound and referent can be determined only within a ritualized, function-based range (e.g., the word "da!" as an accompaniment to the act of exchanging objects);
- referential words which appear to "stand for" their referents in a range of contexts (e.g., "bau bau" to designate dogs).

In related work (e.g., Bates, 1976; Camaioni, Volterra, & Bates, 1976) the nature of performatives was explored and described in greater detail (for example another behavior noted was "showing off" through the repetition of an arm movement or a facial expression for eliciting adult attention), but the main conclusions reached by the study were confirmed. Of particular relevance for the present review was the finding that the onset of intentional communication between the ages of 9 and 13 months was marked in part by the emergence of a series of gestures — GIVING, SHOWING, POINTING, and RITUALIZED REQUEST — that preceded the appearance of first words.

A subsequent, crosscultural and crosslinguistic study compared the gestural and vocal repertoires of 25 Italian and American infants observed between 9 and 13 months of age (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Volterra, Bates, Benigni, Bretherton, & Camaioni, 1979). Striking parallels between early vocal production and gestural schemes of symbolic play were found. The findings can be summarized as follows (Bates, Bretherton, Snyder, Shore, & Volterra, 1980: 408-409):

- V-symbols (vocal) and G-symbols (gestural) emerge around the same time, and are correlated across the sample in frequency of use, rate of acquisition, and number of different schemes observed by the experimenters and reported by the mothers.
- Patterns of correlation with other measures are quite similar for V- and G-symbols. Both correlate with aspects of tool use and imitation, while neither correlates with spatial relations or object permanence.
- Both kinds of symbols are initially acquired with prototypical objects, in highly stereotyped routines or scripts. At roughly parallel rates, they gradually "decontextualize" or extend out to a wider and more flexible range of objects and events.
- There is tremendous overlap in content for both V- and G-symbols, when "vocabularies" are compiled across the whole sample. There are words for such concerns as eating, dressing, play with vehicles, telephones, games of exchange and peekaboo, sleeping, bathing, and doll

play. The repertoires of conventional [2] gestures involve precisely the same set of concerns: eating, dressing, telephones, exchange games, etc. In short, the two types of symbols refer to, name, or in some sense "mean" the same things."

Already at that time, important differences between vocal and gestural "names" were noted. First, it was clear that children were encouraged by parents to rely much more on vocal symbols for communication. Second, some important differences between the vocal and gestural modalities were to be considered, including differences in short-term memory and a relative propensity for sequential vs. simultaneous presentation of information. Third, unlike the vocal symbols of speech, the symbolic gestures used by 1-year-olds typically involve actions directly on the associated object. Young children are much more likely to name objects vocally while they are manipulating them. However, as observed by Bates et al. (1980: 409) "gestural symbols provide a more unique, defining kind of kinesthetic feedback from their objects. Simply put, you can do more things to a cup while saying "cup" that you can while drinking from it. For this reason, we might expect the gestural symbols of 1-year-olds to be more "closely tied" to their objects than equivalent vocal symbols. By this we mean that gestures may require more perceptual input, and/or that physical contact with the object may more likely to trigger a gestural scheme than a word."

Taken together, these findings highlighted the remarkable similarities between production in the gestural and the vocal modalities during the first stages of language acquisition. They also raised interesting issues with regard to the communicative and linguistic value of early words and gestures. Symbolic actions produced in the gestural modality were often considered to be non-communicative or non-referential despite reports that these gestural schemes can be used productively to communicate about a specific referent in a decontextualized, symbolic manner (Volterra et al., 1979). Consequently, they were often referred to and analyzed separately from verbal production as "symbolic play" regardless of their level of decontextualization. In contrast, words were in general considered to be communicative or referential irrespective of the context or contexts in which they were used. This distinction is highly problematic, however, because it implies that only signals produced in the vocal modality can potentially become referential and be used to name new objects or events in a variety of different contexts.

These issues were addressed by Caselli (1983a) in a longitudinal diary study of one Italian infant from the age of 10 to 20 months. Caselli (1983a) showed that many of the gestures usually set aside as "schemes of symbolic play" (e.g, holding an empty fist to the ear for TELEPHONE, waving the hand for BYE_BYE, or raising the arms, palms up, for ALL_GONE) were in fact frequently used by the child to communicate in a variety of situations and contexts similar to those in which first words were produced. These gestures, characterized as "referential gestures," differed from deictic gestures (such as prototypical POINTING or SHOWING) in that they denoted a precise referent and their basic semantic content remained relatively stable across different situations. The form and meaning of these gestures seemed to be the result of a particular agreement established in the context of child–adult interaction, while their communicative function appeared to develop within routines similar to those which Bruner (1983) has considered fundamental for the emergence of spoken language.

Caselli's (1983a) findings were confirmed by further observations. Caselli, Volterra, and Pizzuto (1984) conducted qualitative analyses of longitudinal diary data on the spontaneous vocal and gestural productions of four typically-developing Italian children (age range: 10-30 months).

[2] It may be useful to clarify that in the studies reviewed here, as in the work of other researchers, the terminology used for different types of gestures observed in children's development is not homogeneous, and it has often considerably changed over the years, even in the work of the same author(s), reflecting parallel changes in methodology and/or perspectives. In this review we have generally chosen not to alter the terminology used in the original studies we refer to. However, it must be noted that gestures that in early work were called "performatives" were subsequently reanalyzed and reclassified as "deictic." More content-loaded gestures that were initially classified as "conventional" or "referential," or also "symbolic play schemes," were subsequently reclassified as "representational" (see the subsequent sections of this chapter and Capirci, Iverson, Pizzuto & Volterra, 1996).

These authors reported that, at the one-word stage, the children's gestural utterances were comparable to their vocal productions. As children moved to the two-word stage, numerous gesture-word combinations were observed (e.g., "POINT (to chair) – mommy," requesting that mommy sit on the chair), and two-gesture combinations (e.g., "EAT - POINT (to food)," requesting to be fed). These types of combinations seemed to precede the first two-word utterances. When two-word utterances appeared, however, a marked difference between the vocal and the gestural modalities was found: combinations of two symbolic, referential words (e.g., "mommy open") were common, but combinations of two referential gestures were never observed.

Since the end of the 70s, research in our laboratory has developed in new directions as we began to explore language development in deaf children and, shortly thereafter, the visual–gestural language used within the Italian (adult) Deaf community, a language that had never been described until we began our investigations, and that is now widely known as Italian Sign language or LIS. The knowledge we began to gain by studying a visual–gestural language like LIS undoubtedly stimulated us not only to compare the acquisition of signed vs. spoken language, but also to refine our methodology in the analysis of the gestures, signs, and words observable in children exposed to a signed or a spoken language.

As noted earlier, comparative research on signed vs. spoken language acquisition can shed new light on the study of language acquisition, most notably in determining which aspects of the acquisition process are dependent on the modality of production and reception, and which ones are unaffected by modality. This is particularly true for studying the role of gesture in the emergence of language. When hearing children acquiring a spoken language make the transition from prelinguistic gestural communication to language, a modality change occurs. In contrast, deaf children acquiring a signed language communicate prelinguistically and linguistically in the same visual–gestural modality. Thus, comparisons between hearing children acquiring a spoken language and deaf children acquiring a signed language may be particularly valuable for clarifying the relationship between prelinguistic communication and language. However, in order to pursue appropriate comparisons it is necessary to use the same criteria, and a uniform terminology, for identifying and distinguishing gestures, signs and words in the communicative productions of both deaf and hearing children.

These methodological concerns were taken into account by Caselli and Volterra, who compared the emergence and development of language in two deaf children of deaf parents and two hearing children of hearing parents (Caselli, 1983b; 1990: Caselli & Volterra, 1990; Volterra, 1981). The children were observed at different ages (one pair during their first year, the other pair during their second year of age), using the same criteria for classifying their gestural and/or linguistic) productions (signs or words. This comparison showed that the same stages and timing characterized the development of communication and language across children, independently of the modality of language reception and expression.

In a first, initial period both deaf and hearing children used only deictic gestures, while referential gestures, signs, and words appeared in a subsequent period. Referential gestures, signs, and words were initially used imitating more or less correctly the model offered by the adult, in response to adult elicitations, in ritualized exchanges that often referred to complex schemes of action not yet analyzed. In a subsequent period, gestures, signs, and words were separated from the action scheme or ritualized exchange, appeared to represent only part of the scheme (for example an object or an action), and were used spontaneously to communicate needs or states, or to name objects, actions and events. Importantly, several referential gestures produced by both the hearing and the deaf children appeared to undergo a similar process of gradual decontextualization, eventually assuming symbolic-like properties.

Both deaf and hearing children began to combine two signs or two words in a single utterance at the same age (around 17-18 months), when their observed vocabularies of signs or words comprised about 20-40 distinct items. The combinations the children produced were comparable under one important aspect: hearing–speaking children combined two referential words

at the same stage of symbolic development that deaf–signing children combined two referential signs. Interestingly, the hearing children never combined two referential gestures. Caselli and Volterra thus concluded that in hearing children acquiring spoken languages the capacity to produce symbols can be displayed in both the gestural and the vocal modality, but the specifically linguistic capacity to *produce* and *combine* symbol in the same modality is manifested only in the modality of the linguistic input to which children are exposed. Caselli and Volterra also underscored that, when the vocal and gestural communicative productions of both deaf and hearing children are analyzed according to the same criteria and a uniform terminology, the linguistic advantage reported by some authors (e.g., Bonvillian, Orlansky, & Novack, 1983) in the acquisition of signed languages disappears (see also Volterra & Iverson, 1995).

In the same years in which we were developing our own research, several colleagues around the world had been exploring, from different perspectives, the use of communicative gestures in the first two years of life, and/or early communication in children acquiring signed languages. A collection of papers edited by Volterra and Erting (1990) brought together several studies on these topics, arranged in different sections based on the hearing status of the children examined (hearing and deaf) and the linguistic input they received (spoken or signed). In their concluding remarks, Erting and Volterra (1990) underscored the many points of agreement among the different studies, most notably with respect to the evidence that both hearing and deaf children use gestures to communicate, and that there is a progression in the use of gestures over time. Erting and Volterra also noted relevant discrepancies, especially as concerned the terminology and classificatory criteria used, and the methodological issues these raised. For example, if a child production is labeled as a "sign," the implication is that it is part of a linguistic system and therefore a symbolic act. But, if the same production is labeled as a "gesture," its symbolic status is unclear. Often an author's choice of terms depends upon whether the child is hearing or deaf, or upon the linguistic input to which the child is exposed. Erting and Volterra stressed the need of using a uniform terminology and of defining explicit criteria for deciding upon the status of a gestural production in early infancy: the same criteria should be applied to examine children's vocal, gestural and signed productions in order to determine their communicative, symbolic, and linguistic status.

Further evidence on the relevance of gestures in the emergence of language in typically developing children was provided by the first results of two new research lines we began pursuing at the end of the 80s. One stemmed from work finalized to develop and validate the Italian version of the MacArthur Communicative Development Inventory (MCDI), the now wel- known parental report instrument designed to explore and assess children's early communicative and linguistic development (Dale, Bates, Reznik, & Morisset, 1989; Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick, & Reilly. 1993).[3] Casadio and Caselli (1989) explored children's early word repertoire (receptive and expressive) and production of communicative action-gestures using both a preliminary elaboration of the Italian MCDI (see next section) and a structured interview designed to obtain detailed information from parents on the contexts and the degree of conventionalization / symbolization of children's early word and gesture use. The repertoire of words and action-gestures explored with these parental report tools included:

- a list of 294 words distinguished in different categories related to people, animals, objects, actions, and relations that are commonly encoded in early language (e.g. mommy, dog, water, telephone, feeding-bottle, go, sleep, above),
- a list of 62 action-gestures also distinguished in different categories which, in addition to deictic and referential gestures as defined thus far, comprised real actions (e.g., eating with a spoon), pretend or symbolic play schemes (e.g., pretending to eat with a spoon in the absence of food), and routines (e.g., peekaboo) that are commonly observed in children.

[3] This work, which eventually resulted in the creation of the PVB, the Italian version of the MacArthur CDI (see next section), is another research line inspired by Liz, and it would have never been started nor developed without her constant encouragement, support, and most stimulating collaboration.

Data were provided on a sample of twenty 14-month-old children. The qualitative analyses of the parental reports showed that at this age the number of words children comprehended was markedly larger than the number produced. There were interesting "meaning correspondences" (in a broad sense of the term "meaning") between *words comprehended* and *action-gestures produced* by the children (i.e., in many cases the meaning of words comprehended referred to actions the children produced or which were also encoded by referential gestures noted in children's production). The number of distinct action-gestures produced was larger than that of words produced. The productive action-gestural and vocal systems of the individual children were also found to be highly distinct: in most cases action-gestures and words were related to different types of objects or actions (e.g., if children used a word for 'food' it was unlikely that they had a referential gesture for 'food' or 'eat'). Only a few children had words and action-gestures that somehow conveyed comparable meanings. Information on the contexts of early word and gesture use also indicated that many words were used while performing an action or holding an object related to the word produced. In addition, the parental interview data showed that although parents tended to attribute a communicative value more often to words than gestures, both the words and the referential gestures identified in the children's productive repertoires appeared to undergo a similar process of decontextualization: they were initially produced only in specific and ritualized contexts (nonreferential use), and only later were they used in a symbolic/referential way to anticipate or evoke absent referents.

A second and related investigation aimed at providing new data on children's use of gestures and words during the first two years of life (Caselli, Volterra, Camaioni, & Longobardi, 1993). That study employed a different parental questionnaire, which was originally devised for collecting information from parents of deaf children (Luchenti, Ossella, Tieri, & Volterra, 1988) and subsequently adapted to investigate the communicative and linguistic development of hearing children (Camaioni, Caselli, Longobardi, & Volterra, 1991). In addition to detailed questions on children's vocal and gestural behaviors in different contexts, this questionnaire included:

- a restricted list of early words (N=15) and deictic and referential gestures (N=15), designed to be used with 12- and 16-month-old children (and excluding real as well as symbolic action schemes);
- a much larger vocal vocabulary list comprising 680 items subdivided into 18 different semantic and grammatical categories (e.g., nouns for people, animals, objects; verbs for different actions and states; function words such as articles, prepositions, pronouns).

The items included in these lists were selected from those identified as relevant on the basis of previous studies. The questionnaire was given to the parents of 23 children, with instructions to compile when their children reached the ages of 12, 16, and 20 months. The major results of this study showed that at 12 months the children made extensive use of gestures, at 16 months both gestures and words were used in similar fashion, and only at 20 months did the vocal modality become the predominant mode of communication.

More recent studies

Compared to earlier work, in the more recent studies reviewed below we have extended the investigation of the interplay between gestures and language by examining, with different methodologies:

- larger samples of typically developing children;
- the patterns observable at more advanced stages of communicative–linguistic development;
- input-output relationship in the acquisition of both spoken and signed language;

- the role of gestures in the communicative–linguistic system of children with atypycal learning conditions.

The growth of expressive and receptive vocabulary and its relationship with action and gesture

Evidence on this topic was provided by research conducted using the final version of the "Italian CDI" or "Primo Vocabolario del Bambino" (hereafter: PVB), the parental questionnaire elaborated by Caselli and Casadio (1995) who have collected normative data on a large sample of about 700 children in the age range from 8 to 30 months. The PVB comprises two forms, labeled "Gestures and Words" and "Words and Sentences," designed to collect information on children's early word comprehension and production and action-gesture repertoire, and children's more advanced lexical-grammatical repertoire and sentence structure. In the form that is relevant for the present discussion, Gestures and Words, the repertoire of words potentially comprehended or produced includes 408 items (distinguished in 19 different broad semantic and grammatical categories that range from words for natural sounds, nouns for people, animals, familiar objects, body parts, verbs and adjectives for a variety of actions and states that are commonly encoded in child language, adverbials, pronouns and a subset of function words). The repertoire of action-gestures, listed in a separate subsection, comprises 63 items, distinguished in 7 categories. As in the Casadio and Caselli (1989) study, these categories include not only deictic and referential gestures as defined in several studies, but also real actions and pretend or symbolic play schemes, as well as routines that the children use spontaneously and/or in imitating actions or routines proposed by the adult.

On the basis of data on 315 children (age range: 8–17 months) whose parents compiled the Gestures and Words form of the PVB, Caselli and Casadio (1995) have shown that there is a complex interrelationship between early lexical development in comprehension and production and action-gestures. First, in this age range there are interesting asynchronies between the receptive and the expressive vocabulary, with the first being significantly larger than the second one. Second, in early development the productive repertoire of action-gestures appears to be larger than the vocal repertoire. For example, at 11-13 months the mean number of action-gestures produced is 29, compared to a mean number of 8 words. Third, and what is more interesting, at this early age there is a significant correlation, and also a meaning correspondence, between words comprehended and action-gestures produced. In addition, action-gestures and words appear to develop in parallel through the age of 17 months: at 16-17 months children are reported to produce a mean number of about 40 action-gestures and 32 words. The range of meanings covered by action-gestures and words also appear to be comparable.

These results on a large sample of children support and significantly expand those provided by Casadio and Caselli (1989) and Caselli et al. (1993) on much smaller samples of children. In addition—and although certainly more research is needed to ascertain with more precision the developmental relations between actions, gestures and word—these findings (especially those concerning word comprehension and action-gestures production) suggest that the link between real actions, actions represented via gestures, and children's vocal representational skills may be deeper than it has been ascertained thus far.

Gestures in the transition from the one-word to the two-word stage

As noted above, in our own work preliminary evidence from diary and longitudinal observations (e.g., Caselli et al., 1984; Caselli & Volterra, 1990) suggested that, in typically developing children, gestures are used productively not only in the earliest stages of language development but also when two-word utterance appear. It is well known that the ability to combine two linguistic symbols marks a milestone in the language learning process. From that point on, several major changes in the child's linguistic abilities occur: vocabulary grows at a very fast rate, two- and multi-

word utterances become progressively more frequent and articulated in their meaning and structure, and the acquisition of grammar begins.

Crosslinguistic investigations of a wide variety of languages have shown that the developmental progression from one- to two-symbol utterances takes place in a similar fashion regardless of the particular language and culture to which children are exposed, and can thus be characterized as a universal feature of language learning, in the spoken as in the signed modality (see the studies collected in Slobin, 1985, 1992, 1997). But what is the role of gestures, and how are different types of gestures, as compared to different types of words, distributed in *both* children's repertoire *and* expressive utterances during the transition from the one-word to the two-word stage? Data on these questions were provided by two related studies focused on children's vocal and gestural repertoires (Iverson, Capirci, & Caselli, 1994) and the structure of their vocal, gestural and gestural-vocal or crossmodal utterances (Capirci, Iverson, Pizzuto, & Volterra, 1996). The major results of both studies are summarized and reconsidered here from a unitary perspective that relates the changes in the composition of children's gestural and vocal repertoires to the functions and structure of their vocal and/or gestural utterances (Pizzuto et al., 2000; Capirci, Caselli, Iverson, Pizzuto & Volterra, 2002).

The data used for these studies were videotaped, 45-minute recordings of 12 children observed at home, in different contexts of spontaneous interaction with their mothers (e.g., play with familiar objects, meals or snack time), at two age points: at 16 months, when their vocal utterances consisted for the most of one element, and at 20 months, when two-word utterances appeared in an appreciable number.

All communicative gestures and words identified in the children's production were distinguished in two major categories: deictic and representational. Deictic gestures included the REQUEST, SHOW, and POINT gestures that have been extensively described in the literature (e.g,. Bates et al., 1979). Deictic words included demonstrative and locative expressions, personal and possessive pronouns. Representational gestures included both gestures iconically related to actions performed by or with the referent (e.g., wiggling the nose for RABBIT, flapping the arms for BIRD or FLY), and conventional gestures (e.g., shaking the head for NO, turning and raising the palms up for ALL_GONE, culturally-specific gestures proper to the Italian repertoire, such as bringing the index finger to the cheek and rotating it for GOOD or opening-closing four fingers, thumb extended, for CIAO = 'bye-bye'). Representational words included for the most content words that in the adult language are classified as common and proper nouns, verbs, adjectives (e.g., 'mommy', 'flowers', 'Giacomo', 'open', 'good'), affirmative and negative expressions (e.g., 'yes', 'no', 'allgone'), but also conventional interjections and greetings such as 'bravo!', or 'bye bye'.

The notion of utterance was extended to cover not only vocal productions but also gestural and crossmodal (gestural and vocal) productions. In addition, the information conveyed by different combinations of vocal and/or gestural elements was analyzed and three major types of two-element utterances were distinguished: *equivalent*, *complementary*, and *supplementary*[4]. *Equivalent* combinations included only crossmodal productions of two representational units that typically referred to the same referent and conveyed the same meaning (e.g. BIG = *grande* 'big'; BYE_BYE = *ciao* 'bye-bye', where the notation "=" denotes the comparable meaning). *Complementary* combinations, like the equivalent ones, typically referred to a single referent, but had one distinctive feature, denoted by an ampersand (*&*) between the two combined elements: they always included one deictic element (gestural or vocal) which provided non-redundant information, singling out or disambiguating the referent indicated by the accompanying representational element or by another, cooccurring deictic element (e.g., POINT (to flowers) & *fiori* 'flowers'; POINT (to drawing of fish) & FISH; *questa* & *pappa* 'this & food'; POINT (to toy) & *etto* 'this'). *Supplementary* combinations differed from the other two types in that they referred either to the same or to two

[4] Our distinction between complementary and supplementary utterances differs from that proposed by Goldin-Meadow and Morford (1985, 1990). We extended the use of these terms to the classification of both vocal and crossmodal or gestural utterances, and we attribute to them a different meaning as defined above. .

distinct referents, but in all cases each of the combined elements added information to the other one, a feature we notated with a plus sign (+) between the two combined elements. Vocal combinations of two representational words provided the clearest example of this class of utterances (e.g. *piccolo* + *miao miao* 'little + kitty'), but eight other different subtypes were identified (e.g., the crossmodal utterances POINT (to pigeon) + *nanna* 'sleep'; POINT (to game) + *te* 'you', ALL_GONE + *acqua* 'water').

Details on our classificatory/coding procedure and its rationale can be found in Iverson et al. (1994) and Capirci et al. (1996). Two points should be noted, however. First, the label "representational" was applied in these studies (as in subsequent studied discussed below) to those gestures that in much previous work were defined as "referential" (and with a variety of other terms). Second, and even though we are aware of several important problems that still need to be solved, our classification explicitly attempts to provide more accurate information on the relationship between gestures and words by using comparable classificatory criteria (and terminology) in their analysis. This issue has often been neglected in much, even recent research on the topic, where for example deictic words are not distinguished from content-loaded words (our "representational" words), or deictic gestures are attributed more or less complex "representational-like" meanings that to some extent obscure their basic deictic functions (e.g., Butcher & Goldin-Meadow, 2000; Goldin-Meadow & Morford, 1985; 1990. See also relevant discussions in Capirci et al., 1996).

The results concerning the children's gestural and vocal repertoires showed that at both 16 and 20 months gestures constituted a noticeable portion of all children's repertoires (Mean N of types: 9.58 and 10, respectively). At 16 months, six of the twelve children we observed had more or as many gestures as words types. At 20 months, a clear and significant shift toward the vocal modality was observed: ten out of twelve children had more words than gestures.

We also found differences, and developmental changes, in the distribution of deictic and representational elements in the gestural as compared to the vocal repertoire. In fact, while all children had deictic gestures in their repertoire at both age points (with POINT being by far more frequently used compared to REQUEST and SHOW), the same was not true of deictic words: gestural deixis preceded vocal deixis in the repertoire of half the children in our sample, and deictic gestures did not appeared to be supplanted by deictic words, because they continued to be present in the repertoires of all children at 20 months.

Representational gestures were also found, along with representational words, in the repertoires of all children at both ages, and in many children representational gesture types moderately increased from 16 to 20 months. However, and not surprisingly for children exposed to a spoken language, at both ages the repertoire of most children was composed more of representational words (Mean N = 22 and 58, respectively, at the two age points) than representational gestures (Mean N = 6.58 and 7), and a marked increase in the number of representational elements took place in the vocal but not in the gestural modality.

These data demonstrate that, for representational as well as for deictic elements, the clear shift toward the vocal modality observed at 20 months cannot be attributed simply to a contraction of the children's gestural repertoire, but was due to a parallel, and comparatively much greater expansion of the vocal repertoire.

The data on the *different utterance types* produced by the children at 16 and 20 months evidenced distinct developmental patterns for one- as compared to two-element utterances. Within *one-element utterances*, significant developmental changes were noted from 16 to 20 months. At 16 months, most children produced more one-gesture than one-word utterances, and thus showed a clear preference for the gestural modality in the production of one-element utterances. At 20 months, most children shifted to the vocal modality: they produced more one-word than one-gesture utterances, and this increase in the production of one-word utterances was highly significant, while at the same time the number of one-gesture utterances remained essentially the same. It was also

found that most one-word utterances consisted of a representational (not a deictic) word. The opposite was found in one-gesture utterances, composed for the most by deictic gestures.

The data on the composition and the information conveyed by the children's *two-element utterances* provided a more complex and articulate picture of the role that gestures played at both 16 and 20 months of age.

First, we found that at both 16 and 20 months the most frequent type of two-element utterances were crossmodal combinations of a gestural and a vocal element, which also increased significantly from 16 to 20 months (Mean N at the two age points: 15 and 33). These utterances were in fact significantly more frequent not only of two-gesture utterances (almost absent from the children's production), but also of two-word utterances, which began to be consistently produced only at 20 months. Thus, while in one-element utterances the use of gestures declined from 16 to 20 months, in two-element utterances gestures continued to be a constituent structural element. No clear shift towards the exclusive use of the vocal modality was observed, even though—as expected—two-word utterances increased sharply and significantly from 16 to 20 months.

It is of interest to relate this finding to the data on the development of the children's gestural as compared to the vocal repertoire. The significant expansion of the children's vocal repertoire observed at 20 months appeared to have a rather direct effect on the production of one-element utterances (where a shift to the vocal modality was noted at the same age), and was clearly related to the significant increase of two-word utterances. In principle, the growth of the vocal repertoire could have led also to a decrease in the production of crossmodal gesture-word utterances. The fact that these utterances increased instead of decreasing is an additional strong indication of the important role that gestures continue to play even when the children's vocal repertoire and combinatorial abilities had considerably expanded.

Quite differently from crossmodal and two-word utterances, utterances consisting only of gestures remained in a very small number at both 16 and 20 months (Mean N = 1 and 2, respectively), and never included combinations of two representational elements. Taken together with the limited use of one-representational gestural utterances we noted in our children, this result indicates that the exclusive use of representational gestural elements (either alone or in combinatorial structures) is a marginal phenomenon in hearing children who are exposed to a vocal language input. The developmental patterns we noted for the different types of two-element utterances further clarified the role that deictic as compared to representational gestures and words play in children's growing linguistic system. Our results highlighted the special role that deictic gestures (notably POINT) play in development. This role was most evident in two-element utterances, where combinations of a POINT with a representational word were the most productive type of utterance the children used.

The results on the distribution of equivalent, complementary, and supplementary combinations of gestures and/or words provided new information on the structure of crossmodal and vocal communication in the transition from one- to two-word speech. We found that at both age points the most frequent type of two-element utterances produced were crossmodal complementary combinations of a deictic gesture (by far most commonly a POINT) and a representational word (e.g., POINT (to food) & *pappa* 'food'). This type of utterance, which we proposed could be interpreted as a kind of crossmodal "nomination" (somewhat comparable to its vocal-only counterpart made of a deictic word and a name as in *questa* & *pappa* 'this & food'), also increased significantly form 16 to 20 months.

At 16 months, when supplementary utterances of two words had just barely appeared, the second most frequent type of two-element utterances were supplementary crossmodal utterances composed of a deictic gesture and a representational word (e.g. POINT (to balloon) + *grande* 'big'). This type of utterance, which we proposed can be likened to a kind of crossmodal "predication," also increased from 16 to 20 months (Mean N = 5.58 and 12 respectively), but this increase was not significant. In contrast, and quite predictably, the increase of supplementary vocal-only utterances (Mean N = 0.41 and 9.66 at 16 and 20 months respectively) was highly significant.

Finally, equivalent crossmodal combinations of two representational elements (e.g. BIG = *grande* 'big') were present in an appreciable number at both age points (Mean N = 3.75 and 5.5), but were always in smaller proportions compared to the other crossmodal utterances types, and their small increase at 20 months was not significant. In fact, these utterances appeared to be characterizable more as "bimodal one element utterances," where each of the two representational elements somehow "reinforces" or emphasizes the other, than true combinations of distinct elements.

We also performed correlation and two-step regression analyses to evaluate whether single gesture and gesture-word utterances predicted total vocal production. Total vocal production was defined as the total number of all tokens of single- or two- and multiword utterances produced, with or without an accompanying gesture. The correlational and variance patterns we found indicated that both single-gesture and, more significantly, gesture-representational word utterances produced at 16 months were good predictors of total language output at 20 months.

In sum, as noted elsewhere (Capirci et al., 1996), our results suggested that gesture-word utterances serve three different functions for young children as they attempt to communicate. First, the redundancy provided by representational gestures in equivalent combinations may function to reinforce the child's intended message and seems to help the child who is both vocally uncertain and still moderately unintelligible to ensure that her message is understood. Second, the gestural indication contained in complementary utterances provides disambiguating information that helps to locate and identify the single referent in the child's utterances. Third, in supplementary utterances, the gesture and the word refer to two distinct elements and, as a result, the child's intended message is extended.

From a more general perspective, the sheer presence of a consistent number of two-element crossmodal utterances at 16 months, when children's vocal communication is mostly limited to one-word utterances, and the persisting use of such utterances at 20 months, when children begin to produce two-and multiword utterances, suggest that a reappraisal of this developmental period is necessary. The definition of one-word stage appears to be reductive: at this stage, children's utterances are not limited to one-element, but already include two-element crossmodal combinations which appear to convey both nomination and predication structures. On the same grounds, at the two-word stage, a redefinition of the transitional phenomena that lead children to the acquisition of syntax and grammar may be warranted. At this stage a large portion of children's prototypical nomination and predication structures are still expressed via two-element crossmodal utterances. This suggests that in the transition to language proper gestures, most notably POINT gestures, may play an even more crucial role than has been recognized thus far.

The role of gestural input in hearing mother-child interaction

The results summarized above stimulated us quite naturally to extend the investigation to the role and functions of gestures in maternal input. Towards this end, Iverson, Capirci, Longobardi, and Caselli (1999) reexamined the data used for the Iverson et al.'s (1994) and Capirci et al.'s (1996) studies focusing on the gestures produced by the mothers of the 12 children who participated in these earlier studies. Iverson et al.'s study aimed also to assess whether maternal use of gestures changed as children's speech became more complex from the first observation point at which the children were examined (16 months) to the second one (20 months). The mothers' gestures were identified and classified in three major categories: deictic and representational (as defined in the previous section), and emphatic. This last category included gestures that do not have a well identifiable meaning but are often executed during speech in a rhythmic fashion to stress or highlight aspects of discourse structure and/or the content of accompanying speech, essentially comparable to "beats" as described by McNeill (1992) or Magno Caldognetto and Poggi (1995). All maternal utterances containing both speech and gesture were categorized in three major classes

according to the informational role played by gesture with respect to speech: *reinforce* (e.g., nodding YES while saying "Yes, I know that mommy is ugly"), *disambiguate* (e.g., POINT to floor while saying "Put it over there"), and *add* (e.g., POINT to toy telephone while saying "pretty").

Iverson et al. (1999) found that the majority of gestures produced by mothers at both observation points were deictic or representational, while emphatic gestures were rarely observed. Among deictic gestures, POINT was most common. Comparison of maternal gesture patterns at 16 months with those at 20 months revealed no significant differences in the production of any of the gesture types over time. The finding that emphatic gestures were produced so infrequently in this sample is especially interesting given the extensive use of such gestures in Italian culture (Kendon, 1995; Magno Caldognetto & Poggi, 1995).

At both observation points the majority of maternal gestures served to reinforce the message conveyed in speech. Utterances in which gesture disambiguated the verbal message were somewhat less frequent, while utterances in which gesture added information to that conveyed in speech were relatively uncommon. No significant differences were found in the number of utterances in each of the three categories at the two observation points.

Mothers' gestures, in other words, rarely provided information that was not already present in the spoken message. In marked contrast to what is typically reported for adult-adult interactions, in which gesture generally complements or supplements information conveyed in speech (McNeill, 1992), this suggests that Italian mothers are also gesturing less with their children than they would with another adult. This is particularly striking in the light of the fact that, at both observations, the proportion of maternal utterances containing gesture was much lower than that found in the children.

In summary, analyses of maternal production revealed that when mothers gestured, their gestures tended to cooccur with speech, and consisted primarily of deictic gestures that served to indicate referents in the immediate context. In effect, mothers appeared to be using a kind of "gestural motherese" characterized by fewer and more concrete gestures redundant with and reinforcing the message conveyed in speech. Not only were mothers' gestures tightly linked to the immediate linguistic and extralinguistic context, but they appeared to be used with the goal of underscoring, highlighting, and attracting attention to particular words and/or objects. Gestures that that would not fulfill this function, such as the emphatic gestures widely used by Italian adults when speaking to other adults, appeared to be almost completely absent from the communicative repertoire of the mothers examined in this study.

Signed and spoken language input: data from the study of a bilingual child

More detailed evidence on the relationship between maternal input, language modality and gestural-linguistic development was provided by a study on the spontaneous communication of a bilingual hearing child of deaf parents, exposed to sign and spoken language from birth (Capirci, Iverson, Montanari, & Volterra, 2002).

The hearing child of deaf parents who participated in this study (Marco), was observed at monthly intervals between the ages of 10 and 30 months. Both of Marco's parents used Italian Sign Language (LIS) as their primary means of expression, but they frequently used Italian words (voiced or only mouthed) to accompany their signing when interacting with their child. Marco was also exposed to spoken Italian in the nursery he began to attend during the period in which he was observed. Marco was thus exposed from the beginning of his life to LIS and simultaneous signed/spoken communication at home, and to spoken Italian at the nursery.

The analysis of Marco's production focused on his manual gestures, signs and words, and utterance production patterns. To avoid overestimating Marco's sign production, specific, and quite conservative criteria were used to distinguish signs from gestures. Communicative gestural signals were defined as signs only when: a) they resembled adult LIS forms, and b) their form

differed from those that have been identified in the production of typically developing monolingual children. All of Marco's manual signals that failed to meet these criteria were classified as gestures, and further distinguished in the two major classes of deictic and representational gestures according to the criteria proposed in Iverson et al. (1994), and Capirci et al. (1996).

The results of the study showed that Marco's earliest communications consisted primarily of gestures, a finding consistent with all the evidence we have reviewed in the present chapter, indicating that even in a bilingual signing/speaking child the earliest communicative signals are produced in the gestural modality. Acquisition of new words and signs was initially rather slow, and was subsequently followed by a period of rapid growth that occurred first in the word (between the ages of 19 and 22 months) and then in the sign vocabulary (beginning at 25 months). Marco's initial preference for gestural communication was eventually replaced by a preference for verbal and signed communication. By the end of the observation period, Marco's word and sign vocabularies were approximately the same size (82 signs and 93 words), and he used sign and speech to communicate with roughly equal frequency. Two-word utterance first emerged in Marco's production at 16 months, and began to increase markedly in number from 25 months on. Two-sign utterances first emerged at 25 months, outnumbering two-word utterances by 29 months. Interestingly, two-word combinations increased before two-sign combinations, with each increase occurring after rapid growth in word and sign vocabulary size respectively (see also Gregory, 1991).

In order to examine any effect of simultaneous exposure to signed and spoken languages on early communicative development, Marco's gestural and verbal production at 16 and 20 months was compared to that of the group of 12 monolingual children observed at 16 and 20 months in the Iverson et al. (1994) and Capirci et al. (1996) studies previously described. The results of this comparison showed that, aside from an enhanced communicative use of the manual modality, Marco's communication patterns generally followed those observed among children exposed only to speech. Marco's overall vocabulary size and overall verbal/manual productivity fell well within the respective ranges observed in the monolingual children. All of these results are consistent with those of earlier studies on (deaf) signing children (e.g. Caselli, 1983b; Caselli & Volterra, 1990), and further support the view that there is no "sign advantage" in children exposed to a signed language input.

However, an interesting difference was observed when the proportions of deictic and representational gestures produced by Marco were compared to those of the monolingual children. While Marco used proportionately more representational than deictic gestures at both ages —a finding consistent with data also reported by Bogaerde & Mills (1995) —monolingual children produced deictic gestures much more frequently than representational gestures. Although it is possible that this difference may have been influenced by the conservative criteria used for distinguishing signs from gestures, it seems likely that Marco's relatively extensive use of representational gestures was a result of increased facility in the manual modality. Specifically, exposure to sign language may have enhanced the child's appreciation of the representational potential of the manual modality, and this may have been in turn generalized to gesture use.

The study aimed also to clarify whether the signed/spoken input to which Marco was exposed had any significant effect on his production of different utterance patterns, especially with respect to combinations of representational gestures, and of crossmodal combinations of gestural and vocal elements (obviously leaving aside two-sign combinations that are peculiar to language development in the signed modality). Since bilingual signers/speakers have access to linguistic symbols in two modalities, they may in principle be able to produce crossmodal combinatorial structures that are simply not available to monolingual children (e.g., gesture-sign, word-sign). Crossmodal combinations can be used to convey two different pieces of information in a single, integrated utterance, thereby eliminating the problem of coordinating articulatory movements necessary for the production of two words. Crossmodal combinations, in other words, appear to reflect a compromise between "readiness" to produce word combinations and constraints on the

ability to produce two words in succession. These constraints may differ in a bilingual signing/speaking child.

In order to ascertain to what extent Marco's two-element utterances were comparable to those noted in monolingual children, Marco's production at 16 and 20 months was again compared to that of the monolingual children of the study previously described. This comparison revealed that at 16 months Marco's relative distribution of combinations across structure types was quite similar to that for monolingual children. However, at 20 months Marco's overall production of crossmodal combinations was well above the group mean for monolingual children. In fact, Marco produced more crossmodal combinations than any of the monolingual children. Interestingly, though not surprisingly, these combination included not only gesture-word but also sign-word combinations. In addition, at both ages, Marco combined two representational gestures (albeit in a small number of cases), producing structures that were never used by his monolingual peers, and thus appeared to be influenced by his exposure to a signed input.

The large number of crossmodal combinations produced by Marco raised the question of whether these enhanced his communicative potential relative to his monolingual peers. In other words, did Marco make use of sign-word (in addition to gesture-word) combinations to convey two different pieces of information, something that his monolingual non-signing peers can only do using gesture-word?

To address this issue, all of Marco's two-element utterances were classified according to the informational content they conveyed, distinguishing them into the equivalent, complementary, and supplementary type. It was found that at 16 months of age, the overall pattern of production of equivalent, complementary, and supplementary combinations for Marco was roughly similar to that of the monolingual children taken as a group. At 20 months, however, a striking difference emerged. While Marco's production of complementary and supplementary combinations remained similar to that of the monolingual children, he produced many more equivalent combinations than did the monolingual children taken as a group (Marco = 60 vs. group mean = 5.5). One additional relevant differences between Marco and the hearing monolinguals was that Marco's equivalent combinations included not only gesture=word combinations, but also an almost equal number of sign=word combinations (e.g., producing the LIS sign WORK together with the word *lavoro* 'work').

Why was Marco making such great use of equivalent sign=word combinations? It is reasonable to hypothesize that this reflected the nature of the bimodal sign/speech input to which he was exposed: informationally redundant sign=word combinations may be the product of extensive experience with simultaneous communication in everyday interactions (see also van den Bogaerde, 2000). It is of particular interest to note that Marco did not appear to fully exploit the potential of the input to which he was exposed. In principle, considering the large number of representational gestural elements in Marco's had in his repertoire (both gestures and signs), one could have expected that he produced supplementary gesture+word or sign+word combinations. But this was not found, and the types of complementary and supplementary combinations produced by Marco were on the whole comparable to those noted in his monolingual peers. These data provide an additional indication that exposure to a signed input does not affect the informational content of early two-element utterances, regardless of whether they are gestural, vocal, or crossmodal.

Gestures and words in children with atypical development: Down and Williams Syndromes

While we have devoted many years to the study of the relationship between language and gesture in normally-developing children, in deaf children and in children exposed to a signed input, we are only beginning to study the nature and development of gesture in children with atypical patterns of language and cognitive development. Our recent investigations on this topic have been focused on the role of gesture in language development and use in children with two different genetic syndromes: one, Down Syndrome (hereafter: DS), is fairly well known, the other, Williams

Syndrome (hereafter: WS), is a rare genetic condition associated with a microdeletion on chromosome 7q11.23 (Bellugi & St. George, 2001). Children with WS usually present a number of severe medical anomalies, including mental retardation with a specific cognitive profile.

The behavioral phenotypes of these two genetically determined syndromes appear to mirror each other. Children with DS usually exhibit impairments in language acquisition. Problems in morphology and syntax are frequently reported (Chapman, 1995; Vicari, Caselli, & Tonucci, 2000; Fabbretti, Pizzuto, Vicari, &Volterra, 1997). Fowler (1995) has pointed out that the linguistic difficulties of persons with DS may be a consequence of specific difficulties at the phonological level, both in speech perception and in the re-elaboration of acoustic information into a representational form that can be retrieved to serve memory, production, and comprehension. In contrast, children with WS appear to have an unusual command of language: although their comprehension is usually far more limited than their expressive language, this latter tends to be grammatically correct, complex and fluent at least at a superficial level, while under closer inspection it appears verbose and pseudo-mature (Volterra, Capirci, Pezzini, Sabbadini, & Vicari, 1996).

Relatively few studies have examined the relationship between gesture and developing language in children with DS or WS, and such studies have often focused on a limited set of gestures (Franco & Wishart, 1995; Bertrand, Mervis, & Neustat, 1998; Laing et al., 2002). Our own work on the topic is summarized below.

A first study conducted by Caselli, Vicari, Longobardi, et al. (1998) explored the relationship between action-gestures and words in children with DS compared to typically developing (hereafter: TD) children. Caselli et al. (1998) administered the Words and Gestures form of the PVB parental questionnaire (Caselli & Casadio, 1995) to the parents of 40 Italian children with DS (Mean chronological age, hereafter CA: 28.3 months). The scores obtained by the children with DS in the production of action-gestures and words were compared with those of a group of 40 TD children taken from Caselli & Casadio's (1995) normative sample, matched on the basis of receptive vocabulary size. Caselli et al. (1998) found that the children with DS had significantly larger action-gestures repertoires than their TD comparison group. However, this difference only emerged at higher word comprehension levels, i.e., among children with comprehension vocabularies above 100 words. These findings are in agreement with results reported by Singer Harris, Bellugi, Bates, Jones, and Rossen (1997) in a similar (MCDI based) study on American children with DS, and suggest that there may be some sort of "gestural advantage" in children with DS. These children may compensate for poor productive language abilities through greater production of gestures. However, it must be noted that inventories such as the PVB or the MCDI only provide information about whether or not a particular behaviour is in a child's repertoire. The data do not provide any information on the frequency with which children produce gestures when communicating. Furthermore, in the PVB and MCDI actions and gestures are grouped in a single category, and it is thus difficult to assess what is the role of gestures proper in children's developing communicative and language system.

These issues were addressed in a more recent study on the spontaneous production of gestures and words in children with DS conducted by Iverson, Longobardi, and Caselli (2003). In this study five children with DS (three boys and two girls) were examined. The DS children had an average CA of 47.6 months, an average mental age (hereafter: MA) of 22.4 months, and an average language age of 18 months. Language age was assessed on the basis of the PVB expressive vocabulary scores. Each child with DS was matched with a TD child on the basis of gender, language age, and observed expressive vocabulary size. It must be noted that although the children in the two groups were of comparable language age, all the children with DS were still at the one-word stage, whereas all the TD children had already reached the two-word stage. The ten children participating in the study were videotaped for 30 minutes as they interacted spontaneously with their mothers. Their vocal and gestural productions were analyzed

according to the coding scheme proposed by Caselli et al. (1994) and Capirci et al. (1996) described earlier in this chapter.

The results of the study provided evidence for a tight link between gesture and language in children with DS, and revealed interesting similarities as well as differences between the two groups of children examined. Relative to their language-matched TD peers, children with DS produced similar amounts of gesture and words, and combined gestures and words with comparable frequencies. However, relevant difference in the types and distribution of gesture-word combinations were found. When children with DS combined gestures and words, they did so primarily in an informationally redundant fashion. The vast majority of combinations produced by these children were in fact equivalent combinations in which the two representational elements referred to the same referent and conveyed the same meaning (e.g. headshake for NO = 'no'; BYE-BYE = 'bye'). Complementary combinations, in which a gesture is typically used to single out a referent that is being simultaneously labeled in speech were uncommon, and supplementary combinations, in which combined elements add information to one another were virtually non-existent in the production of the children with DS. TD children, on the other hand, made wide use of both complementary and supplementary combinations. Since complementary and especially supplementary combinations can be considered to be cognitively more sophisticated (i.e., convey greater amounts of information) than equivalent combinations, this suggests that children with DS may be somewhat delayed in the production of more advanced types of gesture-word combinations.

In contrast with children with DS, children with WS have been generally described as appearing to prefer the vocal modality. It has been reported that children with WS display a delay in starting to produce gestures (Bertrand et al, 1998), and that they show a limited use of gestures either with a declarative or an instrumental function (Laing et al., 2002).

A study conducted by Capirci, Iverson, Pirchio, Spampinato, and Volterra (2001) aimed to clarify similarities and differences in the use of gestures and words by children with WS, children with DS and TD children. Three preschool children with WS (CA range: 39–51 months; MA range: 26–36 months) were individually matched with three children with DS (CA range: 36–50 months; MA range 26–39 months) and with three TD children matched for mental age (CA range 24–34 months; MA range: 25–37 months). All the nine children examined had already reached the two-word stage. The children were observed at home, in 40-minute free play interactions with their mothers (20 minutes), and with an unfamiliar adult (20 minutes). All interactions were videotaped and all of the children's verbal and gestural communicative productions were fully transcribed and analyzed as described below. The children's utterances were categorized in three major classes: (1) vocal only (utterances consisting only of spoken words), (2) gestural only (utterances consisting only of gestures), (3) mixed, or vocal-gestural utterances (consisting of speech accompanied with gestures). The children's gestures were classified in the following categories: (1) pointing gestures, (2) conventional gestures (i.e., hands and/or body movements that are known to be used within the Italian culture and are associated with stable meaning (e.g,. rotating the index finger on the cheek for GOOD), (3) iconic gestures (i.e., hands and body movement referring to objects, people, places, or events by some idiosyncratic representation of their form or function, e.g,. flapping the hands for BIRDIE, or raising the arms high for TALL), (4) beats (gestures without a clear and stable meaning, that serve to highlight or emphasize aspects of discourse structure and/or the content of accompanying speech, comparable to those that in Iverson & al.'s (1999) study were classified as "emphatic" gestures). Mixed utterances were further coded according to the information conveyed by the gestural elements, and were distinguished into three major types: (1) reinforce (e.g., waving the hand in the gesture meaning 'hello' while saying "hello"), (2) disambiguate (e.g., pointing to a ball while saying "this one is mine"), (3) add (e.g., waving the hand in the gesture meaning 'hello' while saying "mommy").

The results showed that all the children observed produced a greater amount of vocal than gestured utterances. However, the children with DS produced more utterances containing gestures

than did the children of the other two groups. With respect to the type of gestures produced, it was found that almost all the children produced pointing more than other gestures. Children with DS again differed: they produced more iconic gestures than did children with WS and TD children. The information conveyed by gestures in the utterances produced by the two groups of children with genetic syndromes was on the whole comparable to that observed in TD children: gestures were used mainly to reinforce the meaning of verbal utterances, even though gestures with "disambiguate" and "add" functions were also observed.

Children with WS thus appeared to be similar to TD children with respect to the frequency, type, and function of gestures produced, and this result does not support the indications provided by previous studies on a more limited use of gestures by children with WS. In contrast, children with DS produced more and different types of gestures compared to both children with WS and TD children. This result is in agreement with earlier indications on a possible enhancement of gestural communication in children with DS, as provided by Caselli et al.'s (1998) and Singer-Harris's (1997) studies, and as often reported by clinicians. However, this result also differs from the one reported by Iverson et al. (2003), who did not find any "gestural enhancement" in the DS children they studied.

The studies under discussion are all based on relatively small samples of children, and thus their results must be interpreted with caution. While more research is certainly needed, it is also useful to consider the methodological and developmental differences that may explain the discrepancies observed, especially with respect to the role of gestures in children with DS.

From a methodological standpoint, it must be noted that the broader set of action-gestures examined in the Caselli et al.'s (1998) study is not comparable to the much more restricted set of communicative gestures analyzed by both Iverson et al. (2003) and Capirci et al. (2001). The nature of the data used (parental reports vs. observations of children's spontaneous production in different contexts), the different measures used to match children with DS and TD children, the partially different methodologies used for analyzing and coding the children's gestural and vocal productions, all of these factors render very difficult to draw precise comparisons between these studies, and to reach more definite conclusions. In addition, the important role that individual differences may play needs to be further investigated.

Furthermore, it is of interest to recall that the children with DS in the Iverson et al. (2003) study were all at the one-word stage, whereas those of the Capirci et al. (2001) study were at the two-word stage. This developmental difference in language production abilities may be one of the factors involved in determining the somewhat contrasting results of these two observational studies. In other words, one cannot exclude the possibility that children with DS may use less gestures at the one-word as compared to the two-word stage. From this perspective, the results of the studies in question may be less contradictory than it appears at first sight: rather, they may have explored different facets of the developmental process at different developmental stages (Karmiloff-Smith, 1997).

There is ample evidence that the gap between cognition and language skills (especially language production) becomes progressively wider with development among children with DS (Chapman 1995; Fabbretti, et al., 1997). However, with increasing cognitive skills and social experience, and progressively greater difficulty with productive language, children with DS may be able to make use of actions produced in the context of object-related activities and social routines as communicative gestures. Once this happens, they may begin to develop relatively large repertoires of gestures and make enhanced use of gesture to compensate for poor productive language, particularly if they are encouraged to do so through the provision of signed language input (for a review see Abrahamsen, 2000). Thus, while gesture and language may develop in tandem during the early stages of communicative development in children with DS, the nature of the gesture-language link may begin to change as children's cognitive abilities begin to outstrip their productive language skills. A very recent study conducted in our laboratory has focused on the role of gestures in older children with WS (Bello, Capirci, & Volterra, in press). This study investigated

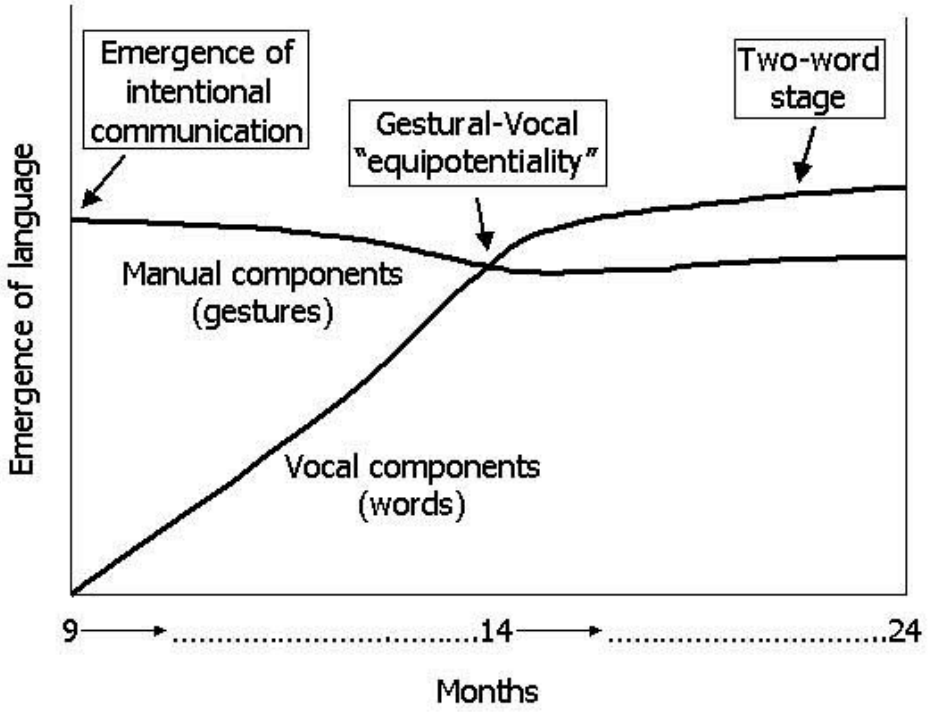
lexical organization and lexical retrieval in children with WS by examining both naming accuracy and the use of accompanying gestures in a picture-naming task such as the Boston Naming Test. This test consists of 60 line drawings representing different objects that children are requested to name. Ten children with WS (Mean CA = 10 yrs;11months; Mean MA = 5 yrs;11 months) were administered the test. These children's performance, and their use of gestures during the task, were compared to those of two distinct groups of TD children: ten TD matched by chronological age (Mean CA = 10 yrs; 8 months) and ten TD matched by mental age (Mean MA = 6 yrs).

It was found that the overall naming accuracy of children with WS was in accord with their mental age. However, compared to both their MA- and CA-matched TD children, the children with WS showed a higher overall rate of gesture production, displayed a richer gestural repertoire than typically developing children, and used a significantly larger number of iconic gestures. The majority of iconic gestures noted in all children (WS and TD alike) appeared to represent the function, rather than the form, of the object depicted (e.g., a child produced a gesture meaning 'brush' moving the extended index finger in the air, as though mimicking the movement of a painting brush). There were few cases of iconic gestures reproducing the form of the represented object (i.e., for 'globe' a child traced a circle in the air with the index finger). In all three groups of children, iconic gestures tended to cooccur with circumlocutions. The analysis of these circumlocutions and of the gestures cooccurring with them indicated that, when the children could not provide the name for an object, they sought the word in the appropriate semantic space, and this appeared to be at least partially expressed and/or codified in the gesture produced. These findings seem to support recent theories on the key role of coverbal gesturing in speech production (e.g., McNeill, 1992), and indicate that the production of iconic gestures, more than other gesture categories, may be triggered by problems in accessing the word for a given object.

Conclusions

In the present chapter we have reviewed studies conducted within our laboratory on the role of gesture in the emergence and development of language. If we consider these studies from a unitary perspective, remarkable changes can be noted over the years. In earlier work, gestures were explored primarily as relevant features of the "prelinguistic" stage, as behaviors that preceded and prepared the emergence of language, which was more or less explicitly identified with speech. At the time, even among scientists, there were very few people who would think of language apart from speech, and it was very difficult to focus on gestures as behaviors that deserved to be fully investigated on their own right, and that continued to be of considerable interest even beyond the stage at which children begin to produce their first recognizable spoken words..

Subsequent studies on deaf signing children who acquire language through their intact visual–gestural modality led us to reconsider the use of gestures also in hearing children. Going back to our "traditional" studies on the acquisition of language under typical conditions, we were able to see that the use of gestures did not stop with the emergence of words but, rather, increased and played an important role in the transition from the one-word to the two-word stage. We would like to try to summarize our present view referring to the developmental scheme in figure 2, a scheme which was inspired by, and indeed mirrors the evolutionary scheme by Corballis presented in the introduction of this chapter.



Before one year of age, children begin to communicate intentionally mainly through gestures, and these gestures are often accompanied by vocalizations. Vocalizations become progressively more sophisticated and similar to the words used in the adult language to which the children are exposed. Around 14 months there is a basic "equipotentiality" between the gestural and vocal channels. In this bimodal period, as aptly summarized by Abrahamsen (2000), "words are not as distinct from gestures and gestures are not as distinct from words as they first appear." As our earlier studies have shown, words and gestures appear to encode similar meanings, and go through a similar decontextualization process. In the following months, the repertoire of spoken words increases dramatically, but gestures are not simply replaced by speech. Rather both the vocal and the gestural modalities are used together, and crossmodal combinations mark the transition to the two-word stage.

The main hypotheses underlying much current work on the interplay between gesture and speech is that there is a continuity between an earlier "preverbal" and a subsequent, somehow functionally "equivalent" linguistic form, and that the use of gesture is a robust developmental phenomenon, exhibiting similar features across different children and cultures. The output systems of speech and gesture may draw on underlying brain mechanisms common to both language and motor functions (Iverson &Thelen, 1999). Within this broad framework, evidence on children with atypical patterns of language and cognitive development may be particularly relevant to assess the resilience of gesture as a developmental phenomenon.

In the studies we have conducted thus far on children with atypical development such as children with Down and Williams Syndromes, we have found that during the early stages of language learning the developmental patterns followed by gesture and speech are on the whole similar to those observed in typically developing children with similar language production abilities. Relevant differences are also observable (e.g., the greater use of redundant, equivalent

combinations of gestures and words noted in children with DS, or the different use of iconic gestures found in children with WS compared to typically developing children). Further research is clearly needed to explore the ways in which the role of gesture in relation to speech changes developmentally, and the extent to which this role may vary among individuals with language difficulties.

The studies reviewed in the present chapter, based on children who varied in cultural experience (Italian vs. North American), input language (spoken vs. signed, English vs. Italian) and cognitive profiles (typical development, Down Syndrome, Williams Syndrome), strongly support the view that there is a remarkable continuity between prelinguistic and linguistic development, and that symbolic skills that are most evident in vocal linguistic productions are inextricably linked to, and co-evolve with more general cognitive and representational abilities, as is most apparent in the tight relationship between gestures and words, which continues through adulthood (McNeill, 1992, 2000).

This view appears to be particularly plausible in the light of the neurophysiological studies on "mirror neurons" we mentioned earlier (Gallese et al., 1996; Rizzolatti & Arbib, 1998) that have discovered powerful links between motoric and representational abilities in both monkey and human brains, with strong implications for a clearer understanding of the relationship between structured action, gestures and vocal language in humans. A central question that has always been hotly debated since ancient times, and which is still much discussed at present, is whether language development in human infants is driven primarily by specialized (and species-specific) innate structures that are for the most unrelated to those underlying more general cognitive and symbolic abilities, a sort of "language instinct" as characterized by one of the proponents of this view (Pinker, 1994), or whether on the contrary it is intricately but solidly linked to more general cognitive and neuro-sensory-motor structures that language shares with other domains (e.g., memory, sensory-motor coordination), and that are put in the service of language in a unique way, as proposed by other leading scholars (Deacon, 1997; Elman et al., 1996; Tomasello, 1999). Proponents of the first view underscore the dissociation of language from other cognitive domains, and the discontinuity between prelinguistic and linguistic development. Proponents of the second view highlight the interrelation between language and other cognitive domains, and the continuity between prelinguistic and linguistic development. As observed by Tomasello (1999), language did not come out of nowhere nor did it arise as some bizarre genetic mutation unrelated to other aspects of human cognition and social life. Natural language is a symbolically embodied social institution that arose historically from previously existing social-communicative activities. Long before children learn to speak, they are able to communicate, meaningfully and intentionally, with their caretakers. In learning a language, children are acquiring a more effective and elaborate means of doing something that they can already do in a more primitive fashion. As suggested by Bates' earlier and more recent work: "*...Language is a new machine built out of old parts*" (Bates & Goodman, 1997), "*emerging from a nexus of skills in attention, perception, imitation, and symbolic processing that transcend the boundaries of 'language proper'"* (Dick & Bates, in press).

Liz Bates has been for a long time a strong advocate of this second view, a view which we fully share not on "a priori" bases, but on the grounds of our research. Much of this research has been inspired by Liz's work, and often developed together during Liz's frequent visits at "her" Nomentana Lab in Rome. Liz focused from the start on key questions that only later were to be "rediscovered" and recognized as being of central relevance for an appropriate understanding of human symbolic skills. Many of Liz's "old ideas" appear today surprisingly "new" and are supported by the most recent findings in both behavioral and neurophysiological studies.

Acknowledgements

The research reported in the present chapter is supported by grant from EUROCORES program "The Origin of Man, Language and Languages" (OMLL)

References

Abrahamsen, A. (2000). Explorations of enhanced gestural input to children in the bimodal period. In K. Emmorey & H. Lane (Eds), *The Signs of Language Revisited: An antology to honor Ursula Bellugi and Edward Klima*. Mahwak: Erlbaum, pp. 357-399.

Armstrong, D.F., Stokoe, W.C., & Wilcox, S. E. (1995). *Gesture and the Nature of Language*. Cambridge University Press

Austin, J.L. (1962). *How to do things with words.* New York: Oxford University Press.

Bates, E: (1976). *Language and Context. The acquisition of pragmatics*. New York: Academic Press.

Bates, E. & Benigni, L., Bretherton, I., Camaioni, L., Volterra, V. (1979). *The emergence of symbols: cognition and communication in infancy.* New York: Academic Press.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1977). From gesture to the first word: on cognitive and social prerequisites. In L. Lewis and L. Rosenblum (Eds*.), Interaction, conversation and communication in infancy* . New York: John Wiley, pp. 247-307.

Bates, E., Bretherton, I., Snyder, L., Shore, C., & Volterra, V. (1980). Vocal and gestural symbols at 13 months. *Merrill Palmer Quarterly, 26,* (4), 407-423.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merril Palmer Quarterly, 21*, (3), 205-226.

Bates, E. & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, 12(5-6), 507-584.

Bello, A., Capirci, O. & Volterra, V. (in press). Lexical production in children with Williams syndrome: spontaneous use of gesture in a naming task. *Neuropsychologia*.

Bellugi, U. & St. George, M. (Eds.). (2001) *Journey from cognition to brain to gene. Perspectives from Williams syndrome*. Cambridge MA: MIT Press.

Bertrand, J., Mervis, C.B., & Neustat, I. (1998). Communicative gesture use by preschoolers with Williams syndrome: a longitudinal study. Paper presented at the *International Conference of Infant Studies*. Atlanta, GA.

Bogaerde, B. van den (2000). *Input and Interaction in Deaf Families* (Doctoral disseration). Utrecht, NL: LOT.

Bogaerde, B. van den, & Mills, A. (1995). *Propositional content in different modes: An analysis of the language production of deaf and hearing children of deaf parents.* Paper presented at the Child Language Seminar, Bristol, April 7-9.

Bonvillian, J.D:, Orlansky, M.D. & Novack, L.L. (1983). Developmental milestones: Sign language acquisition and motor development. *Child Development*, *54*, 1435-1445.

Bruner, J. (1983). Child's talk: learning to use language. New York: Norton.

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*. Cambridge: Cambridge University Press: 235-257.

Camaioni, L., Caselli, M.C., Longobardi, E., & Volterra, V. (1991). A parent report instrument for early language assessment. *First Language, 11*, 345-359.

Camaioni, L., Volterra, V., & Bates, E. (1976) *La comunicazione nel primo anno di vita*. Torino: Boringheri

Capirci, O., Iverson, J., Montanari, S. & Volterra,V. (2002). Gestural, signed and spoken modalities in early language development: The role of linguistic input. *Bilingualism Language and Cognition*, 5 (1)*,* 25-37.

Capirci, O., Iverson, J.M., Pirchio, S., Spampinato, K. & Volterra, V. (2001). Speech and gesture in discourse produced by children with Williams syndrome and with Down syndrome. Paper presented at the *Xth European Conference on Developmental Psychology*, 22-26 Agosto, Uppsala: Sweden.

Capirci, O., Iverson, J.M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child language, 23*, 645- 673.

Capirci, O., Caselli, M.C., Iverson, J.M., Pizzuto, E., & Volterra, V. (2002). Gesture and the Nature of Language in infancy: the role of gesture as transitional device enroute to two-word speech. In D. Armstrong, M. Karchmer, & J. Van Cleeve (Eds.) *The Study of Sign Languages - Essays in Honor of William C. Stokoe* Washington, D.C.: Gallaudet University Press, pp. 213-246.

Casadio, P., & Caselli, M.C. (1989). Il primo vocabolario del bambino. Gesti e parole a 14 mesi. *Età Evolutiva, 33*, 32-42.

Caselli, M.C. (1983a). Gesti comunicativi e prime parole. *Età Evolutiva, 16,* 36-51.

Caselli, M.C. (1983b). Communication to language: deaf children's and hearing children's development compared. *Sign Language Studies, 39,* 113-144.

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*. New York: Springer-Verlag, pp. 56-67.

Caselli, M.C. & Casadio, P. (1995). *Il primo vocabolario del bambino. Guida all'uso del questionario MacArthur per la valutazione della comunicazione e del linguaggio nei primi anni di vita.* Milano: Franco Angeli. Caselli, M.C., Vicari, S., Longobardi, E., Lami, L., Pizzoli, C. & Stella, G. (1998). Gestures and words in early development of children with Down syndrome. *Journal of Speech, Language and Hearing Research, 41,* 1125-1135.

Caselli, M.C. & Volterra, V. (1990). From Communication to language in hearing and deaf children . In V. Volterra & C. Erting (Eds.) *From gesture to language in hearing and deaf children*. Berlino: Springer Verlag, pp. 263-277 (1994 - 2nd Edition Washington, D.C.: Gallaudet University Press).

Caselli, M.C., Volterra, V., Camaioni, L., & Longobardi, E. (1993). Sviluppo gestuale e vocale nei primi due anni di vita. *Psicologia Italiana*, IV, 62-67.

Caselli, M.C., Volterra, V., & Pizzuto, E. (1984, April). The relationship between vocal and gestural communication from the one-word to the two-word stage. Paper presented at the *International Conference on Infant Studies*, New York, NY.

Chapman, R.S. (1995). Language Development in Children and Adolescents with Down Syndrome, In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language*. Oxford: Blackwell, pp. 641-663.

Corballis, M.C. (2002), *From hand to mouth - The origins of language*. Princeton, NJ: Princeton University Press.

Dale, P.S., Bates, E., Reznick, J.S. & Morisset, C. (1989), The validity of a parental report instrument of child language at twenty months. *Journal of Child Language*, *16*, 239-429.

Deacon, T. (1997). *The symbolic species. The coevolution of language and the human brain.* The Penguin Press

Dick & Bates, E. (2002). Language, gesture and the developing brain. *Developmental Psychobiology*, 40, 293-310

Elman, J.L., Bates, E., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996). *Rethinking Innateness*. Cambridge, Mass: MIT press.

Erting, C., Volterra, V: (1990). Conclusion. In V. Volterra, C. Erting (eds), *From gesture to language in hearing and deaf children*. Berlino: Springer Verlag, pp. 299-303. (1994 - 2nd Edition Washington, D.C.: Gallaudet University Press).

Fabbretti, D., Pizzuto, E., Vicari, S., & Volterra, V. (1997). A story description task with Down's syndrome: lexical and morphosyntactic abilities. *Journal of Intellectual Disability Research, 41,* 165-179.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S. & Reilly, J.S. (1993), *Communicative Development Inventories*, San Diego, CA: MacArthur

Fowler, A.E. (1995). Language variability in persons with Down syndrome. In L. Nadel & D. Rosenthal (Eds.), *Down syndrome: Living and learning in the communit*y. New York: Wiley-Liss.

Franco, F., Wishart, J.G. (1995). The use of pointing and other gestures by young children with Down syndrome. *American Journal of Mental Retardation, 2,* 160-182.

Gallese, V., Fadiga, L, Fogassi L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593-609.

Goldin-Meadow, S. & Morford, M. (1985). Gesture in early child language: Studies of hearing and deaf children. *Merrill-Palmer Quarterly, 31,* 145-176.

Goldin-Meadow, S. & Morford, M. (1990). Gesture in early child language. In V. Volterra & C.J. Erting (Eds.), *From gesture to language in hearing and deaf children.* New York: Springer-Verlag, pp. 249-262.

Gregory, S. (1991). The first signs and words: Language development in a bilingual environment. In J. Kyle (Ed.), *Growing up in sign and word*. Bristol, UK: Centre for Deaf Studies, pp..

Hewes, G.W., (1976). The current status of the gestural theory of language origin, *Annals of the New York Academy of Sciences*, vol. 280, 482-604.

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., Longobardi, E., Caselli, M.C., (1999). Gesturing in Mother-child Interaction. *Cognitive Development,* 14, 57-75

Iverson, J., Longobardi E. & Caselli M.C. (2003). The Relationship between Gestures and Words in Children with Down Syndrome and Typically-Developing Children in the Early Stages of Communicative Development. *Intenational Journal of Language and Communication Disorders, 38,* 179-197.

Iverson, J.M., Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies, 6,* 19-40.

Karmiloff-Smith, A. (1997). Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology*, *13(4),* 513-524.

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics, 23,* 1-33.

Kendon, A. (2002) Historical observations on the relationship between research on sign languages and language origins theory. In D. Armstrong, M. Karchmer, J. Van Cleeve (Eds.) *The Study of Sign Languages - Essays in Honor of William C. Stokoe.* Washington, D.C.: Gallaudet University Press, pp. 35-52.

Laing, E., Butterworth, G., Ansari, D., Gsodl, M., Longhi, E., Paterson, S. & Karmiloff-Smith, A. (2002) Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science 5 (2),* 233-246.

Leroy, M. (1969) *Profilo storico della linguistica moderna*. Bari: Laterza.

Luchenti, S., Ossella, T., Tieri, L., & Volterra, V. (1988). Una scheda di osservazione per il bambino sordo alla prima diagnosi audiologica. *Giornale di Neuropsichiatria dell'Età Evolutiva, 8*, 4, 313-328.

Magno Caldognetto, E., & Poggi, I. (1995). Conoscenza e uso di gesti simbolici: Differenze di sesso e di età [Awareness and use of symbolic gestures: Sex and age differences*]*. *Atti del*

Convegno Internazionale di Studi "Dialettologia al Femminile". Padova, Italy: CLEUP, pp. 399-412.
McNeill, D. (1992) *Hand and Mind - What gestures reveal about thought*. Chicago: University of Chicago Press.
McNeill, D. (Ed.) (2000) *Language and Gesture*. Cambridge: Cambridge University Press.
Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: William Morrow.
Pizzuto, E., Capirci, O., Caselli, M..C., Iverson, J.M., & Volterra, V. (2000). Children's transition to two-word speech: Content, structure and functions of gestural and vocal productions. Paper presented at the *7th International Pragmatics Conference*, Budapest, July 9-14 2000.
Rizzolatti, G. & Arbib, M.A. (1998). Language within our grasp. *TINS*, *21*, 188-194.
Singer Harris, N., Bellugi, U., Bates, E., Jones, W., & Rossen, M. (1997). Contrasting profiles of language development in children with Williams and Down Syndromes. *Developmental Neuropsychology, 13(3),* 345-370.
Slobin, D. I. (Ed.) (1985). *The crosslinguistic study of language acquisition. Vol. 1: The data, Vol. 2: Theoretical issues*. Mahwah, NJ: Lawrence Erlbaum Associates.
Slobin, D. I. (Ed.) (1992). *The crosslinguistic study of language acquisition. Vol. 3*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Slobin, D. I. (Ed.) (1997). *The crosslinguistic study of language acquisition. Vol. 4*. Mahwah, NJ: Lawrence Erlbaum Associates.
Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
Vicari, S.,. Caselli, M.C. & Tonucci, F. (2000). Asynchrony of lexical and morphosyntactic development in children with Down Syndrome. *Neuropsychologia, 38,*634-644.
Vico, G. (1744). *La Scienza Nuova*. Naples: Stamperia Muziana [1954 edition by F. Nicolini, *Giambattista Vico - Opere*. Verona: Stamperia Valdonega., pp. 531-550]
Volterra, V. (1981) Gestures signs and words at two years: When does communication become language? *Sign Language Studies, 33*, 351-362.
Volterra, V., Bates, E., Benigni, L., Bretherton, I., Camaioni, L., (1979) First words in language and action: A qualitative look. In E. Bates, *The emergence of symbols: cognition and communication in infancy*. New York: Academic Press, pp. 141-222.
Volterra, V., Capirci, O., Pezzini, G., Sabbadini, L., & Vicari, S. (1996). Linguistic Abilities in Italian Children with Williams Syndrome.*Cortex, 32,* 663-677.
Volterra, V., Erting, C. (Eds) (1990). *From gesture to language in hearing and deaf children*. Berlino: Springer Verlag. (1994 - 2nd Edition Washington, D.C.: Gallaudet University Press).
Volterra, V., Iverson, J.M. (1995). When do modality factors affect the course of language acquisition? In K. Emmorey, J. Reilly (Eds.), *Language, Gesture, and Space*. Hillsdale, NJ: Erlbaum, pp. 371-390.
Wescott, R, Hewes, G. & Stokoe, W.C. (Eds.) (1974). *Language origins*. Silver Spring. MD: Linstok Press.