

## Gestures and words during the transition to two-word speech\*

OLGA CAPIRCI

*Institute of Psychology CNR, Rome.*

JANA M. IVERSON

*The University of Chicago.*

ELENA PIZZUTTO AND VIRGINIA VOLTERRA

*Institute of Psychology CNR, Rome*

(Received 23 February 1995; Revised 18 December 1995)

### ABSTRACT

This study explores the communicative use of the gestural and vocal modalities by normally developing Italian children during the transition from one- to two-word speech. We analysed the spontaneous production of 12 children at 1;4 and at 1;8, focusing on the use of two-element combinations of words and/or gestures. Results indicated that use of gesture and gesture-word combinations during the transition to two-word speech is a robust feature of communicative development across a relatively large number of children in a rich gestural culture, and that the number of gesture-word and two-word combinations increased significantly from 1;4 to 1;8. Number of gestures and gesture-word combinations produced at 1;4 was also predictive of total vocal production

[\*] Portions of this paper were presented at the 17th Annual Boston University Conference on Language Development, 23–25 October 1992, at the Biennial Meetings of the Society for Research in Child Development, New Orleans, 25–28 March 1993, and at the 9th International Conference on Infant Studies, Paris, 2–5 June 1994. The authors would like to thank Emiddia Longobardi and Cristina Parrizi for assistance with data collection, Anna Paola Ercolani for statistical advice, Paolo Desideri for assistance with statistical analyses and tables and Martha Alibali, Elizabeth Bates, John Bonvillian, Maria Cristina Caselli, Susan Goldin-Meadow, Adam Kendon, Marc Marschark, Leslie Rescorla, Donna Thal, Amanda Woodward, Robert Wozniak and two anonymous reviewers for their helpful comments on earlier versions of the manuscript. The research was supported by a Rotary Foundation Scholarship and a National Science Foundation Graduate Fellowship to Iverson and by funding from the CNR Targeted Project 'FATMA: Prevention and Disease Control Factors'. Address for correspondence: Olga Capirci, Istituto di Psicologia CNR, Reparto di Neuropsicologia del Linguaggio e Sordità, Via Nomentana 56, 00161 Rome, Italy.

at 1;8. Findings are discussed in terms of the role of gesture as a transitional device en route to two-word speech.

#### INTRODUCTION

Normally-developing children begin to produce their first recognizable words at around one year of age. For several months, their utterances typically consist of only one word. The ability to combine two linguistic symbols usually develops six or more months later and marks a milestone in the language learning process. From that point, 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 acquisition (see the studies collected in Slobin, 1986).

However, a number of studies conducted in the past twenty years have pointed out that, in the early stages of development, hearing-children's communicative repertoires are not limited solely to the vocal symbols of spoken languages. This body of work has shown that different types of gestures are used productively as independent items or in combinations with other gestures and/or words. Such gestural productions appear to exhibit symbolic and combinatorial properties similar to those attributed to children's early one- and two-word utterances (see the collection of papers in Lock, 1980; Volterra & Erting, 1990). The major indication provided by these studies is that a more accurate description of the role played by different types of gestures and by combinations of vocal-gestural items in early language development may lead to a better understanding of the particular cognitive and symbolic abilities that underlie the task of language acquisition.

The goal of the present study is to contribute to this line of work by examining the structure and content of vocal and gestural elements in the spontaneous production of a group of Italian children observed at 1;4, when vocal production was mostly limited to one-word utterances, and at 1;8, when two- and multi-word utterances were used more consistently. Please note that throughout this paper we will extend the term 'utterance' to characterize both vocal and gestural or vocal-gestural productions. The present work also follows up a recent study conducted by Iverson, Capirci & Caselli (1994), who explored children's lexical development, analysing the interplay between gestures and words, in the same group of subjects. Before turning to the specific questions to be addressed in this study, however, a

brief description of the gestures we focus on and of some relevant findings from previous research is necessary.

Erting & Volterra (1990) recently compared the major findings of several studies on the role of gesture in the development of spoken and sign language. They noted that the different focus, methodologies, classificatory criteria and terminology employed in such studies often make it extremely difficult to assess the comparability, and thus the generality of their findings. Despite these differences, however, the current body of research on gestural communication has distinguished between two main categories of gestures.

The first category, *DEICTIC GESTURES* (DG's), includes a small set of communicative signals (called 'performatives' in some earlier studies; Bates, Camaioni & Volterra, 1975; Bates, Benigni, Bretherton, Camaioni & Volterra, 1979) that are known to emerge during the prelinguistic period, between the ages of 0;9 and 1;1. This set of DG's includes four different gestures: (1) *REQUEST*, or reaching in the direction of a desired object while opening and closing the palm and looking alternately at the adult's eyes and the object; (2) *GIVE*, or giving an object to the adult, with only the intent to draw the adult's attention to it; (3) *SHOW*, or holding an object in the adult's line of sight, with only the intent to draw the adult's attention to it; and (4) *POINT*, or extending the index finger towards some object/location/event in the environment while alternately looking at the adult and the referent.

We characterize DG's as signals that express the child's communicative intent to request or 'declare': they are used to 'point to', or to draw attention to objects, locations or events in the context of an utterance. The 'semantic content' or, more precisely, the undifferentiated referent of these gestures can be identified only by inspecting the physical context in which communication takes place (e.g. by following the direction of a pointing index finger or by looking at an object that is being 'shown').

The second category, which we call *REPRESENTATIONAL GESTURES* (RG's), includes a larger set of conventional hand or body movements or facial expressions that come to be associated with relatively stable meanings across different contexts of production. While DG's simply 'point' to a given referent, RG's 'stand for' or 'represent' some referent, class of referents, or relations (e.g. bringing a closed fist to the ear to mean *TELEPHONE*; turning the arms, palms up, for *ALL-GONE*; flapping the hands for *BIRD*). Children's first representational gestures and words are initially found as part of routines from which they are progressively detached as they undergo a similar process of decontextualization (Caselli, 1983; Folven & Bonvillian, 1991). In addition, many of the meanings expressed by these gestures are equivalent to those conveyed by first words. The presence in children's repertoires of what we call here representational gestures, and their later appearance relative to DG's have been documented in a number of studies that have described them using a variety of different terms: 'referential' or 'symbolic' gestures

(Caselli, 1983; Caselli & Volterra, 1990) 'characterizing gestures' (Goldin-Meadow & Morford, 1985; 1990), 'signs' or 'symbolic gestures' (Goodwyn & Acredolo, 1993), and 'symbolic play schemes' (Bates *et al.*, 1979).

Most of the current body of research on young children's gestural communication has explored and documented the use of deictic and/or representational gestures in the earliest stages of communicative and linguistic development, namely from the prelinguistic period through the one-word stage (see, among others, Bates *et al.*, 1975; Acredolo & Goodwyn, 1988; Lock, Young, Service & Chandler, 1990; Morford & Goldin-Meadow, 1992; Goodwyn & Acredolo, 1993). Some of the patterns noted in hearing children exposed to spoken languages have also been identified in deaf children acquiring sign languages (Caselli, 1983; Petito, 1990; Bonvillian, Orlansky & Folven, 1990; Caselli & Volterra, 1990; Pizzuto, 1990).

However, very few studies have examined the role that both deictic and representational gestures play in the transition to two-word speech. In our view, these studies have taken two somewhat different approaches to this issue. One approach involves the assessment of the relative complexity of gestural as compared to vocal production by focusing primarily on the TYPES OF GESTURES (i.e. deictic or representational) that children use in single or two-unit utterances and the extent to which early words and gestures possess symbolic features. Caselli (1983), Caselli & Volterra (1990), Caselli, Volterra and Pizzuto (1984) and Volterra (1981) conducted qualitative analyses of longitudinal diary data on the spontaneous vocal and gestural productions of a total of four normally-developing Italian children (age range: ten to thirty months). The 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 (e.g. 'POINT (to chair) - mommy', requesting that mommy sits on the chair), and two-gesture combinations were observed (e.g. EAR - 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 modality was found: combinations of two symbolic, referential words (e.g. 'mommy open') were common, but combinations of two referential gestures were never observed. These data were then compared to similar data on deaf children acquiring sign languages, and the authors reported that clear modality preferences became apparent when children reached the two-word or two-sign stage. Hearing-speaking children combined two referential words (but not two referential gestures) at the same stage of symbolic development when deaf-signing children combined two referential signs. The authors concluded that the capacity to produce symbols can be displayed in both the gestural and the vocal modality in hearing children acquiring spoken languages, but the

capacity to PRODUCE and COMBINE symbols in the same modality is manifested only in the modality of the linguistic input to which children are exposed.

The second approach also involves assessment of the complexity of gestures relative to speech but focuses instead on the MEANING AND INFORMATION CONTENT of gestural and/or vocal utterances. This line of work has provided detailed qualitative data on gesture-word and gesture-gesture combinations and their relationship to two-word utterances. Goldin-Meadow and Morford (1985, 1990) conducted a longitudinal study of the spontaneous gestural and vocal productions of three normally-developing American children ranging in age from ten to thirty months. After classifying the children's gestures as either DEICTIC or CHARACTERIZING, the authors attributed semantic meanings to both types, assigning deictic gestures to the categories of 'indicators' or 'cases', and characterizing gestures to the category of 'predicates'.

Goldin-Meadow & Morford found that children produced gesture-gesture and gesture-word combinations before they combined words. Combinations of two distinct gestures, each denoting a separate semantic element, were very infrequent (less than 3% of each child's total gestured production) but had one interesting feature: they disappeared prior to the onset of two-word utterances. All but one such gestural combination consisted of a POINT plus a GIVE gesture. This finding is in agreement with that reported by Caselli & Volterra (1990), who found no combinations of two referential gestures. Goldin-Meadow & Morford (1990: 252 ff.) also observed that children employed two different types of gesture-speech combinations. In the first (and by far most frequent) type of combination observed, called COMPLEMENTARY, the gesture and the word essentially referred to the same referent (though playing different semantic roles depending upon the particular gesture or word; e.g. INDICATOR: POINT at glasses while saying 'glasses'; PREDICATE: GIVE plus the word 'give'). In the second (and much less frequent) type of combination observed, called SUPPLEMENTARY, the gesture denoted a semantic element that was different from that expressed by the word (e.g. POINT at glasses plus the word 'out', to request that glasses be taken out of the case; GIVE plus 'puzzle' to request that the puzzle be given to the child). The authors concluded that children are capable of encoding two distinct semantic elements prior to the development of two-word speech, first via gesture-gesture combinations, and then in a gesture-word form. Gestures thus seem to act as a transitional form en route to two-word speech. These findings have been subsequently expanded and corroborated in more recent studies by Morford & Goldin-Meadow (1992) and Butcher & Goldin-Meadow (1993) who analysed the same set (or subsets) of data in greater detail using different coding and descriptive schemes.

Although these two groups of studies have pioneered the longitudinal

analysis of gesture use in the transition from the one- to two-word speech, there are several questions that remain to be clarified. In the present study, we analyse the spontaneous gestural and vocal production of a group of Italian children prior to and at the onset of two-word speech in order to address three specific questions: (1) When children move to two-word speech, is the use of gesture and gesture-speech combinations a robust developmental phenomenon, exhibiting similar features across different children and cultures, or are there significant individual and cultural differences? (2) Are there differences in the type of units (deictic or representational) that are combined within and across the gestural and the spoken modalities and in the information content that these units can convey at different ages? In particular, are representational elements combined in the same manner in the gestural and the vocal modalities, or are there modality preferences for this type of combination? (3) More generally, is there a developmental relationship between the use of gestures or gesture-word combinations and subsequent language development, i.e. can production of gestures or gesture-word combinations at one age point predict total language output and/or the appearance of two- and multi-word utterances at a later point? The data that are currently available are not sufficient to answer these questions for three reasons.

First, as summarized above, a very small number of subjects have been examined – three American and four Italian children. It is thus not clear whether the developmental phenomena that have been reported are generalizable to larger groups of children. Second, previous studies do not provide sufficient information on the development of deictic words, which have not been distinguished from representational words. The link between gestural and vocal deixis and the role that deictic gestures AND words play with respect to representational gestures AND words in children's early combinations thus remain to be explored. Third, while DG's appear to be fairly equally represented in the 'gestural lexicon' of both the American and the Italian children in the studies reviewed above, inspection of their repertoires of 'symbolic' gestures suggests that, on average and within the same age range, the Italian subjects had a larger number of different referential gestures (from 11 to 19) than the American children (1 or 2 in two subjects; 15 in only one child). Although these discrepancies may be due, at least in part, to different coding procedures, they also suggest that there may be potentially significant cultural differences in the repertoires of symbolic gestures employed by Italian and American children. More detailed information on possible individual and cultural differences may help clarify the question of whether the absence of symbolic combinations in the gestural modality (i.e. of two representational gestures) is simply a feature of few (American or Italian) children's production or a more general indication of

a deeper constraint on the gestural modality in hearing children receiving spoken language input.

New evidence on a sizable number of Italian children observed at 1;4 and 1;8 – two developmental junctures marking the onset of two-word speech – may be particularly relevant to these issues. Italian children are immersed in what Kendon (1995) has recently characterized as a 'gesture-prominent' culture: one in which a fairly large number of conventional gestures are used frequently, along with words, in communication among adults. In addition to De Jorio's (1832) early work on Neapolitan gestures, an index of the significant and perhaps unique relation between gestures and speech in Italian culture is a new edition of a 'dictionary of Italian gestures' by Munari (1994), and a volume by Diadori (1996) designed for learners of Italian as a foreign language. These recent volumes describe more than 100 gestures that are commonly used across the country and specify the linguistic and communicative contexts in which each gesture can be used. Due to the potentially enhanced 'gestural input' that Italian children may receive, an examination of their use of gestures in relation to speech may provide particularly important information about the role of gesture in the early stages of language acquisition and the crosscultural generality of previous findings.

#### METHOD

##### *Subjects*

Subjects were 12 children (six boys and six girls) from upper-middle-class families. All families lived in Rome and were native speakers of Italian. The subjects were selected randomly from a larger group of 23 children observed as part of a study assessing the efficacy of a structured questionnaire in evaluating communicative and linguistic development at 1;0, 1;4 and 1;8 (Camaioni, Caselli, Longobardi & Volterra, 1991). The 12 children selected were videorecorded only at 1;4 and 1;8. Five of the children were first-born, six were second-born, and one was third-born. Nine children had mothers who worked at least part-time outside the home, and three attended day care on a regular basis.

##### *Procedure*

Two videorecorded observations for each child were made: one at 1;4 and the other at 1;8. Two researchers divided the families between them for single-researcher visits. Each session was videotaped in the child's home and lasted approximately 45 minutes, during which mothers were instructed to

TABLE 1. Coding scheme and notational conventions with illustrative examples of the children's vocal and gestural productions<sup>a</sup>

	Types of words and gestures
Deictic Words (dw)	qua, là, questo, quello, eccolo, io, tu, mio, tuo, 'here, there, this, that, here-it-is, I, you, mine, yours' POINT, SHOW, REQUEST
GESTURES (DG) Representational Words (rw)	mamma, pappa, apri, bere, 'mommy, food, open, drink' no, sì, più, ciao, 'no, yes, all-gone, bye-bye' FISH, BIRD, GOOD, BIG, OPEN, NO, YES, ALL-GONE, BYE-BYE
GESTURES (RG)	
Two element combinations: modality, components, information	
(a) Modality and components	
Crossmodal DG+rw	(1) POINT (to flowers) fiori, 'flowers' (2) SHOW (a cup) = acqua, 'water' (3) POINT (to pigeon) = nanna, 'nap' (4) BYE-BYE = ciao, 'bye-bye' (5) BIG = grande, 'big'
RG; rw	(6) ALL-GONE = acqua, 'water' (7) POINT (to toy) = eccolo, 'here-it-is' (8) POINT (to game) te, 'you' (9) YES = questo, 'this'
DG-dw	
RG dw	
Unimodal Gestural DG-RG	(10) POINT (to drawing of fish) = FISH (11) POINT (to a closed box) = OPEN (12) POINT (to bottle of water) = POINT (to glass) (13) POINT (to toy) = REQUEST (same toy)
DG-DG	
Vocal	
rw-rw	(14) piccolo = miao_miao, 'little kitty' (15) questa = pappa, 'this = food'
dw-rw	(16) questo = c'entra, 'this = goes-in-here'
dw-dw	(17) eccoli = qua, 'here-they-are = here' (18) qua = mia, 'here = mine'
(b) Information conveyed Equivalent	
RG = rw	(4a) BYE-BYE = ciao, 'bye-bye' (5a) BIG = grande, 'big'
Complementary DG & rw	(1a) POINT (to flowers) & fiori, 'flowers' (2a) SHOW (a cup) & acqua, 'water' (7a) POINT (to toy) & questo, 'this'
DG & dw	(10a) POINT (to drawing of fish) & FISH
DG & RG	(13a) POINT (to toy) & REQUEST (same toy)
DG & rw	(15a) questa & pappa, 'this & food'
dw & dw	(17a) eccoli & qua, 'here-they-are & here'
Supplementary DG + rw	(3a) POINT (to drawing of pigeon) + nanna, 'sleep'
DG + dw	(8a) POINT (to game) + te, 'you'

TABLE 1 (cont.)

	Types of words and gestures
RG + rw	(6a) ALL-GONE + acqua, 'water'
RG + dw	(9a) YES + questo, 'this'
DG + RG	(11a) POINT (to a closed box) + OPEN
DG + DG	(12a) POINT (to bottle of water) + POINT (to plastic glass)
rw + rw	(14a) piccolo + miao_miao, 'little + kitty'
dw + rw	(16a) questo + c'entra, 'this + goes-in-here'
dw + dw	(18a) qua + mia, 'here + mine'

<sup>a</sup> The different categories of gestures, words, and two-element combinations of gestures and/or words are described in the text. Abbreviations used throughout the text for each major type of gestures and words are given in parentheses (i.e. DEICTIC GESTURES = DG). Gestures are shown in small CAPITALS, in English glosses. Italian words are given in lower case letters, followed by their English translation. An underline character indicates single Italian words that in normal orthography are separated by a space (e.g. miao\_miao <kitty>), or single gestures or words which require more than one English word to be glossed (e.g. ALL-GONE for the 'palms up' gesture, 'eccoli' <here-they-are>).

interact and play with their child as they normally would, while the observer carefully avoided interfering with mother-child interaction. The observations were divided equally into three 15-minute segments, so that the children were filmed in three different contexts: play with new examples of familiar objects, play with familiar objects, and during a meal or snacktime. The new objects were a set of toys provided by the experimenter: a toy telephone, a plate, a cup, a toy glass, two animal picture books, a spoon, a teddy bear, two small cars, a ball, a brush and a comb. Familiar objects varied with each child but always included books, toy cars, toy animals and blocks. Differences across contexts were not analysed in the present study, and data will be reported from the observations collapsed across all three contexts.

*Coding and analysis*

All communicative and intelligible gestures and words were transcribed from the videotapes. Gestures and words were considered to be communicative if they were accompanied by eye contact with another person, vocalization or other clear evidence of an effort to direct the attention of another person present in the room (Thal & Tobias, 1992). Only gestures and words meeting these criteria were included in our analysis. Because children's first gestures and words tend to be idiosyncratic forms that may have no conventional meaning or that may have a different meaning than in adult language and culture, we were necessarily broad in our definition of what constituted a gesture and a word. Moreover, the meaning of these early signals is often established in the context of child-caregiver interaction (Bates *et al.* 1979).

