

A Comparative Study on Representational Gestures in Italian and Japanese Children

Paola Pettenati · Kazuki Sekine · Elena Congestrì · Virginia Volterra

Published online: 12 January 2012
© Springer Science+Business Media, LLC 2012

Abstract This study compares words and gestures produced in a controlled experimental setting by children raised in different linguistic/cultural environments to examine the robustness of gesture use at an early stage of lexical development. Twenty-two Italian and twenty-two Japanese toddlers (age range 25–37 months) performed the same picture-naming task. Italians produced more spoken correct labels than Japanese but a similar amount of representational gestures temporally matched with words. However, Japanese gestures reproduced more closely the action represented in the picture. Results confirm that gestures are linked to motor actions similarly for all children, suggesting a common developmental stage, only minimally influenced by culture.

Keywords Gestures · Picture naming · Cross-cultural/linguistic comparison · Toddlers

Introduction

The present study aims to compare spoken and gestural production in children raised in different linguistic and cultural environments to examine the robustness of gesture use at an early stage of vocabulary development. Cross-cultural observational studies, conducted so far on a restricted number of participants, suggest that all children, regardless of their primary linguistic input, use gestures together with speech during early stages of linguistic development (for a recent review Gullberg et al. 2008). Furthermore, several studies provide clear evidence that gestures do not disappear in children's communication with the development of spoken language and have reported an increase in the use of gestures with

P. Pettenati (✉)

Department of Neuroscience, University of Parma, Via Volturmo 39/E, 43121 Parma, Italy
e-mail: paola.pettenati@nemo.unipr.it

K. Sekine

Japan Society for the Promotion of Science, Tokyo, Japan

E. Congestrì · V. Volterra

Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy

age and linguistic competence growth, especially within spontaneous interaction (Mayberry and Nicoladis 2000) retelling of narratives (Colletta 2004; McNeill 2005), and tasks that require providing explanations or problem-solving (e.g., Goldin-Meadow and Singer 2003; Pine et al. 2004).

A recent study exploring early lexical production during a picture naming task in Italian children between 2 and 7 years (Stefanini et al. 2009), has shown that when children are requested to label simple pictures of objects and/or actions, they are likely to accompany their spoken naming responses with pointing and representational gestures. Furthermore, almost all representational gestures produced represented directly the action shown in the picture or the action usually performed with or by the object presented in the picture (Kendon 2004). This study argued that motor representations produced by children alongside with their early spontaneous naming, contribute towards the creation of an *experiential* dimension and support the linguistic representation expressed by the word. If gesture functions as a motor representation in preschool age children, we could hypothesize that children raised in different cultures may produce gestures despite differences in gesture use within cultures. We shall explore this hypothesis by analyzing the comparative frequency of gesture production as well as speech and gesture timing in children from two different cultural environments.

Previous studies on the early development of gesture were mostly conducted through spontaneous observation in family contexts, as described in the review below. Only very few studies have so far attempted a comparative analysis of gesture development within different cultures relying on a structured experimental setting.

The present study aims to compare gestural production in a controlled experimental setting in two groups of children raised in different linguistic and cultural environments, namely Italian and Japanese children. Italians have traditionally been described as having a rich gesture vocabulary and frequently using gestures in daily communication (De Jorio 1832; Diadori 2003; Efron 1941; Kendon 2004; Munari 1994), a characteristic comparatively less well documented in other cultures. On the other hand, Japanese culture is not considered a ‘gesture-rich’ culture and very few studies document Japanese emblems (Aqui 2004). Given the large gesture repertoire of Italian adults, young Italian children might be expected to produce a larger number of gestures than Japanese children. However if gestures function as motor representations supporting spoken representations in the early stages of language development (as reported in Stefanini et al. 2009), we should expect a similar gestural production in relation to frequency and type in Italian and Japanese children despite the observed cultural differences in adult gesture use. According to this hypothesis we should expect also a similar relationship between the gestural and spoken modalities.

In the remaining sections of this Introduction we will briefly review previous, comparative linguistic studies conducted on the use of gestures in young children from different cultures, to better specify the hypotheses tested in the present study.

Cross Cultural Studies on the Use of Gesture in Toddlers

In a pioneering cross-cultural and cross-linguistic study comparing the gestural and vocal repertoires of 25 Italian and American infants observed between 9 and 13 months of age (Bates et al. 1979), both groups performed schemes of symbolic play (e.g., holding an empty fist to the ear for TELEPHONE¹) and striking similarities were found between early vocal and

¹ All glosses for representational gestures are reported in small capitals following a convention adopted in many studies on children’s gestures.

gestural productions. Another study, based on data from over 50 American infants, Acredolo and Goodwin (1994), highlighted that symbolic gestures (differentially labeled as *referential*, *representational*, or *characterizing gestures*), occurring in a large proportion of their sample and generally preceding their verbal counterparts, were used by infants quite frequently in their daily life and were routinely interpreted by their parents as if they were words. These gestural productions appear at the same age as the first recognizable words and provide a sort of ‘pictographic representation’ similar in meaning and function to early words.

Productions of pointing and representational gestures during spontaneous interactions at home between children in the second year of life and their mothers have also been recorded in longitudinal and cross-sectional studies. These studies reported that producing an expression consisting of a gesture and a word was recognized as having a main role in the transition toward two-word speech for Italian as well as American children (Butcher and Goldin-Meadow 2000; Capirci et al. 1996, 2005; Caselli 1994; Iverson and Goldin-Meadow 2005; Iverson et al. 1994; Pizzuto and Capobianco 2005).

One study conducted on three American and three Italian children, followed longitudinally between the ages of 10 and 24 months (Iverson et al. 2008), reported more frequent production of representational gestures by Italian children than by their American peers. In particular, the representational gestures produced by Italian children included several object/action gestures (e.g., EATING) and attributive gestures (e.g., BIG), whereas American children almost exclusively produced conventional gestures (e.g., HI, YES). Despite these differences in gesture vocabulary, in both cultures gesture/speech combinations reliably predicted the onset of two-word combinations (Iverson et al. 2008). These authors concluded that culture and adult input may influence to some extent how the manual modality is used for representational purposes.

Blake et al. (2005) observed the entire bodily gestural repertoire produced by four different infant groups (English Canadian, Italian Canadian, Japanese, and French) between 9 and 15 months during naturalistic interaction with a caregiver. Increases and decreases in gesture categories were remarkably similar across cultures. They found an increase over sessions in comment gestures (i.e., pointing, but not showing), and a decrease in overall request gestures (i.e., reaching). However, some differences appeared in the relative frequency of certain gestures. For example, Japanese infants engaged in a lot of give-and-take with their mothers and produced more frequent object exchanges than other groups at most ages. Italian Canadian infants were highest only in Protest gestures. The authors hypothesize that infants’ gesture repertoire is universal, and that differences between groups, particularly in the use of declarative pointing and give-and-take gestures, are likely to be ascribed to cultural differences in the interaction between child and caregiver.

The Goal of the Present Study

The current article presents the results of a cross-linguistic study to test the hypothesis from Stefanini et al. (2009) that gesture supports early lexical development. A cross-linguistics design, focusing on variables such as frequency and temporal synchrony, allows this study to establish, in an experimental context, a comparative assessment of the role of gesture in lexical development. The aim of the study is to verify if we could find comparable data in Italian and Japanese children using the same task and procedure for data collection. In particular we focus on representational gesture. Representational gestures (e.g., bringing an empty fist to the ear for TELEPHONE; extending the arms for BIG) are defined as pictographic representations of the meaning (or meanings) associated with the represented object or event. This representation can reproduce the action shown in the picture or the action

usually performed with or by the object presented in the picture, but also the size or shape of the object represented or of the object usually associated with the action or the event shown in the picture. The reproduction can be more or less similar to the depicted action or object. A primary goal of this study was to investigate whether Japanese children produce representational gestures, just as Italian children do, using the same naming task.

If motor gestural representation supports spoken naming at a stage of vocabulary expansion (for a more detailed discussion see also Stefanini et al. 2007) we should expect that Italian and Japanese children perform a similar amount and type of representational gestures, having a comparable functional role in speech production. This hypothesis would predict the same temporal relationship between gesture and speech across different cultural groups.

In order to test this hypothesis we have explored in both groups the following variables: (1) The number of correct spoken responses provided, as an index of lexical accuracy in the two spoken languages; (2) The frequency, and the typology (action vs. size and shape) and the relation to the picture (level of reproduction) of representational gestures produced, in order to evaluate cross-cultural similarities and differences; (3) The relationship between use of gestures and word production to determine if gestures are produced to accompany spoken responses (correct or incorrect) or to replace words; (4) The temporal relationship between spoken and gestural modalities in both groups, aiming to explore whether gesture precedes or follows word onset.

Method

Participants

Twenty-two Italian children and twenty-two Japanese children matched for gender (12 female and 10 male) and age (age range 25–37 months; $M = 30$; $SD = 3.6$) participated in this study. Children were distributed evenly across age range with 12 Japanese and 11 Italian children aged 25–29 months, 10 Japanese and 11 Italian children aged 30–37 months. Children exposed to other languages, children with recurrent serious auditory impairment and children with epilepsy or psychopathological disorders were not considered in this study.

Materials and Procedures

A picture naming task, originally developed by Bello et al. (2010) adapted to assess children's level of spoken vocabulary, was used. The version of this task adopted for the present study consists of 46 colored pictures divided into two sets: 24 pictures representing objects/tools (e.g., Comb), animals (e.g., Penguin), food (e.g., Apple) and clothing (e.g., Gloves), and 22 pictures representing actions (e.g., Washing hands), characteristics (e.g., Small) and location adverbs (e.g., Inside-Outside). Examples of pictures are presented in Fig. 1.

Lexical items were selected from the normative data of the Italian version of the MB-CDI (Caselli and Casadio 1995). Only three pictures were substituted in the Japanese version: the picture for “Radiator” with a “Stove” more common in Japanese homes; new versions of the pictures representing the actions of “Crying” and “Laughing” were included, showing a Japanese child performing the same actions.

All of the children were tested individually in a familiar context: The majority of Italian and Japanese children were tested in their nursery schools, only few children in both groups were tested in their homes. The two pictures sets were presented separately with a



Fig. 1 Examples of pictures used in the picture naming task: clockwise from *top left*, “*comb*”, “*washing hands*”, “*big*”, and “*small*”

break or in two different sessions and the order of picture presentation within each set was fixed. Italian children were tested in two sessions, while Japanese children were tested in one session. This choice was based on schools’ scheduling requirements.

After a brief period of familiarization, the experimenter placed the pictures in front of the child one at a time asking “What is this?” for pictures of body parts, animals, objects/tools, food, and clothing, “What is he/she doing?” for pictures of actions, and “How/where is this?” for pictures depicting characteristics (adjectives or location adverbs). In the case of characteristics, two pictures were put in front of the child: one representing the expected characteristic and another representing the opposite characteristic (e.g., a big ball and a small ball). If the child did not provide the expected label (small) as a first answer, the experimenter said, ‘This is big (pointing to the picture representing the big ball) and how is this?’ (pointing to the picture representing the small ball). A similar procedure with two pictures was used for location adverbs. When the pictures were presented, the experimenter sometimes pointed to the image in order to help the child in focusing her attention on the target but otherwise avoided to produce any other kind of gesture. The mean duration of the task was about 30 min, but short breaks were allowed when needed.

All sessions were videotaped for later transcription. Communicative exchanges occurring between child and experimenter during a time period starting when the picture was initially placed in front of the child and ending when the picture was removed were coded. During these exchanges, children could, in principle, produce multiple spoken utterances and multiple gestures. In particular we examined children’s responses in terms of spoken accuracy, types of gestures produced, and temporal relationship between spoken naming and gesture production, as described below.

Spoken Responses

Answers in the naming task were classified as correct, incorrect, or no response. Responses were coded as correct when the child provided the target word for the picture. In both samples we considered the target word to be the spoken response produced by at least 80% of the participants during the validation study carried out on 20 Italian and 8 Japanese adults (age range 19–33 years). For some pictures, more than one answer was accepted as correct. For example, “Bag” can be called “Sacchetto”, “Busta”, or “Borsa” in Italian, and “Diaper” can be called “Oshime” or “Omutsu” in Japanese. Phonologically-altered forms of correct words (e.g., “lelefono” for the picture of a telephone, intended to elicit the Italian word “Telefono”, “kacha” for the picture of an umbrella, for the Japanese word “kasa”) and onomatopoeia words (e.g., “brum” for “Car” in Italian, “wan wan” for “Dog” in Japanese) were also accepted. Incorrect responses included incorrect labeling of the target items elicited by the pictures (e.g., “scissors” for “suspenders”). When children either stated that they did not know the word corresponding to a picture or did not provide any answer, the item was coded as a no-response. When children gave an incorrect answer or a no-response at their first attempt, a second chance to provide the correct answer was given. A “best answer” criterion was adopted in those cases, such that if the child initially gave an incorrect spoken response and then provided the correct one, s/he was given credit for providing a correct response.

Gesture Production

All visible actions (e.g., posture, body movements, and facial expressions) produced by children interacting with the experimenter were coded as gestures (Kendon 2004). These included gestures produced with and without speech, and those occurring both before and after the spoken response.

Given the specific nature of the task (asking children to name pictures), the criteria for coding an action as a gesture (Pettenati et al. 2010) were as follows: (1) The gesture had to be produced after the adult had made the request to name the picture; (2) The gesture could be performed with empty hand or while holding the picture to be named or by a facial expression and/or a specific posture; (3) The gesture must not be an imitation of the adult’s preceding gesture.

Participants produced various categories of gestures: deictic, representational, conventional, beats, and self-adaptor [for more details on classification of gesture types see Butcher and Goldin-Meadow (2000), Stefanini et al. (2009)]. In the present study we focused only on representational gestures, i.e., gestures depicting pictographic representations of the meaning (or meanings) associated with an object or event.

Regarding the *techniques of representation* used, two types of representational gestures were coded in our study (Stefanini et al. 2009):

- *Action gestures* depicting the action usually performed with the object, by an object, or by a character. In the action gestures (defined by Kendon 2004 as “enactment”; see also Gullberg 1998), body parts engage in a pattern of action that has features in common with the pattern of action that serves as the referent (for example: in front a picture of a comb, the child moves his fingers near his head as if combing his hair)
- *Size and shape gestures* (defined by Kendon 2004 as “modeling” and “depiction”) depict the dimension, form, or other perceptual characteristics of an object or an event. In this case the hand ‘creates’ an object in the air by tracing its shape or direction,

delimiting its size or dimension (for example performing a circle with the index finger extended for “Turning” or moving up the arms to show the length of a pencil for “Long”).

Regarding the *level of reproduction* of the action or event represented by the picture we considered a gesture as a:

- *Complete reproduction*, when it reproduced the object or the action as they appeared in the picture (e.g., child making the gesture of WASHING HANDS reproducing exactly the action shown in the picture);
- *Partial reproduction* when some aspects of the gesture represented the object or the action shown in the picture, but in a different way (e.g., the child reproduces a gesture that represents the action of washing the hands but the position and/or movement of the hands is different from that shown in the picture);
- *Peripheral relation* when the gesture was considered as induced by the picture, while the action was not immediately present in it (e.g., performing the gesture of COMBING in front of the picture of a comb); Peripheral relations between gestures and pictures were found especially in the case of pictures representing objects.
- *Indirect relation* when the gesture represented something related to what was shown in the picture and which clearly “stood for” the picture (e.g., in front of the item “Umbrella” the child makes a gesture that represents the rain).

Speech–Gesture Relationship

Modality of Expression

Gesture productions were distinguished between *bimodal* (that is, gestures accompanied with *correct* and *incorrect* spoken answers) and *unimodal* (gestures without speech). *Bimodal* productions in front of the same pictures were also analyzed in terms of temporal relationship using ELAN Software (EUDICO Linguistic Annotator).

Temporal Relationship

Speech and gesture were considered synchronous when the word was produced on the *stroke* of the gesture. Gesture stroke² is defined as the meaningful peak of effort in a gesture. Furthermore the mutual temporal relation between gesture and speech was considered, regardless of synchrony. The analysis included six different possible situations:

- Gesture starts before speech
- Speech starts before gesture
- Speech and gesture start together
- Gesture ends before speech
- Speech ends before gesture
- Speech and gestures end together.

² In a gesture we can distinguish essentially three phases: Preparation (optional): The limb moves away from the rest position into the gesture space where it can begin the stroke. Stroke (obligatory in the sense that absent a stroke, a gesture is not said to occur): The stroke is the gesture phase with meaning; it is also a phase with effort, in the dance notation sense of focused energy. Retraction (optional): The hands return to rest (not always the same position as at the start) (McNeill 2005).

Intercoder Reliability

Reliability between two independent coders was assessed for all spoken and gesture productions. Agreement between coders for the Italian and the Japanese sample was respectively 90 and 95% for spoken answers, and 78 and 83% for gestures. Each disagreement was identified and disagreements were resolved by a third coder, who chose one of the two classifications proposed by the first two coders.

Results

In this section, data for both groups of children, Italian and Japanese are presented. The following aspects are taken into account: spoken production, gesture production (frequency and type, techniques of representation, level of reproduction), speech-gesture relationship (modality of expression, temporal relationship). For each aspect we are considering similarities as well as differences between the two groups. In advance of choosing which statistical procedure to run, tests for normality were carried out at first to examine whether the sampled group is normally distributed. We evaluated the distribution of the quantitative variables (i.e., number of strokes and number of correct answers) by the Kolmogorov–Smirnov test for Gaussian normality. Because the values were not normally distributed, we used non-parametric test Mann–Whitney U test, Spearman's ρ test, and Chi-square test).

Spoken Production

We analyzed the spoken responses provided by the children to determine whether or not they corresponded to the expected word. Correct naming was about 39% in the Japanese and 56% in the Italian sample; incorrect naming was about 46 and 32% while no responses was about 15 and 12% respectively. The Mann–Whitney analysis (Japanese vs. Italians) carried out on the percentages of each type of spoken answers (correct responses, incorrect responses, and no-responses) showed that Italian children produced more correct responses ($U = 109$; $Z = -3.13$, $p < .01$) and fewer incorrect responses ($U = 122$; $Z = -2.82$, $p < .01$) than Japanese children while no significant differences were found in the number of no-responses. Considering more carefully correct spoken responses provided by the two groups, Japanese children interestingly produced a higher percentage of onomatopoeia than Italian children (4% vs. 1.5%; $U = 127.5$; $Z = -2.83$, $p < .01$).

In order to investigate the relationship between age and spoken responses, Spearman correlations were conducted. The result showed that with age the number of correct labels increased significantly for both groups, but the effect was higher for the Japanese group (Spearman $\rho = .65$, $p < .01$) than the Italian group (Spearman $\rho = .42$, $p = .05$).

Gesture Production

Frequency of Gestures

All representational gestures produced (with and without speech) by the 22 Italian and the 22 Japanese children participating in the naming task were analyzed. Forty-one pictures out of 46 elicited at least one representational gesture by one child. All children in the Italian and Japanese samples produced at least one gesture, but a great variability characterized both samples (range 1–18 and 1–28 respectively); the number of gestures produced in each

Table 1 Items eliciting gestures in the Japanese and Italian samples

Objects/animals set				Actions/characteristics set			
Items	JPN	ITA	Total	Items	JPN	ITA	Total
Comb	15	12	27	Washing	16	14	30
Gloves	11	2	13	Crying	9	14	23
Umbrella	2	7	9	Turning	5	15	20
Flags	5	3	8	Phoning	8	10	18
Lion	3	5	8	Heavy	8	9	17
Radiator/stove	2	5	7	Opening	8	6	14
Glass	3	2	5	Swimming	6	8	14
Suspenders	2	3	5	Driving	7	5	12
Table	4	0	4	Pushing	7	4	11
Fork	2	1	3	Long	5	4	9
Picture	2	1	3	Kissing	2	6	8
Beach	1	0	1	Inside	6	2	8
Seal	1	0	1	Laughing	7	1	8
Camion	1	0	1	Eating	4	3	7
Diaper/nappy	1	0	1	Small	3	2	5
Socks	1	0	1	Falling	4	1	5
Bag	1	0	1	In front	2	2	4
Banana	1	0	1	Playing	2	2	4
Dog	0	1	1	Far	1	3	4
Train	1	1	2	Empty	1	2	3
				Clean	1	0	1
Total	59	43	102	Total	112	113	225

List of the items of PinG task (objects/animals and actions/characteristics sets) eliciting the production of representational gestures in the Japanese (JPN) and Italian (ITA) samples

sample was not correlated with age (Spearman test: Japanese $\rho = .28$, $p = .21$; Italian $\rho = -.22$, $p = .31$) nor with the number of correct answers given (Japanese $\rho = .31$, $p = .16$; Italian $\rho = -.23$, $p = .30$). We found that the total number of gestures was similar in both groups: 156 in the Italian group and 171 in the Japanese group (Mann–Whitney test: $U = 218.5$, $Z = .55$, $p = .58$), both samples produced significantly more gestures labeling pictures representing actions/object characteristics than labeling pictures representing objects/animals (Wilcoxon test: Japanese $Z = 2.88$, $p < .01$; Italian $Z = 3.28$, $p < .01$) (Table 1).

Techniques of Representation

In both groups the majority of gestures depicted actions, whereas size-shape gestures were less frequent and the difference was significant for both samples (Chi-square test for Japanese: $\chi^2(1) = 98.22$, $p < .001$; for Italian: $\chi^2(1) = 34.41$, $p < .001$). Japanese children produced fewer size-shape gestures than Italian children (Chi-square test: $\chi^2(1) = 7.79$, $p < .01$), while no significant difference in the number of action-gesture between the two samples was found (Fig. 2).

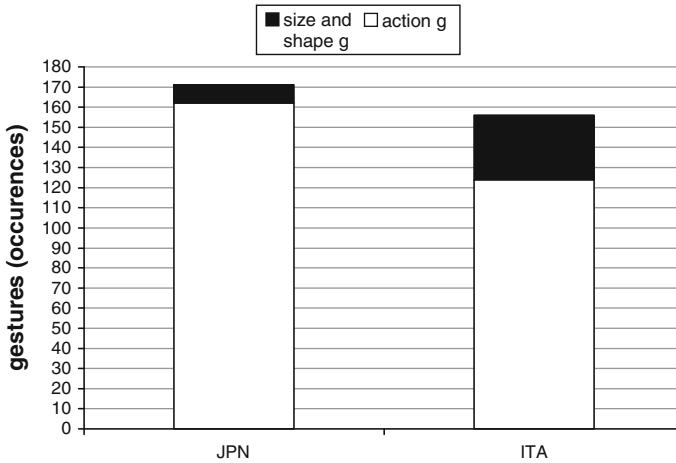


Fig. 2 Number of *action gestures* and *size and shape gestures* produced by Japanese and Italian samples in the two sets of picture naming task (objects/animals and actions/characteristics)

Table 2 Level of reproduction

Categories	JPN	ITA
Complete reproduction	59	16
Partial reproduction	28	50
Peripheral relation	66	81
Indirect relation	18	9

Level of reproduction considering the four categories separately (*complete reproduction*, *partial reproduction*, *peripheral relation* and *indirect relation*) in the Japanese (JPN) and Italian (ITA) samples

Level of Reproduction

Considering the similarity between the gesture performance and the contents depicted in the pictures (subdivided in the four categories described above), Japanese children produced more gestures that were complete reproductions of the picture target than Italian children (Chi-square test: $\chi^2(1) = 15.89$, $p < .001$) (Table 2).

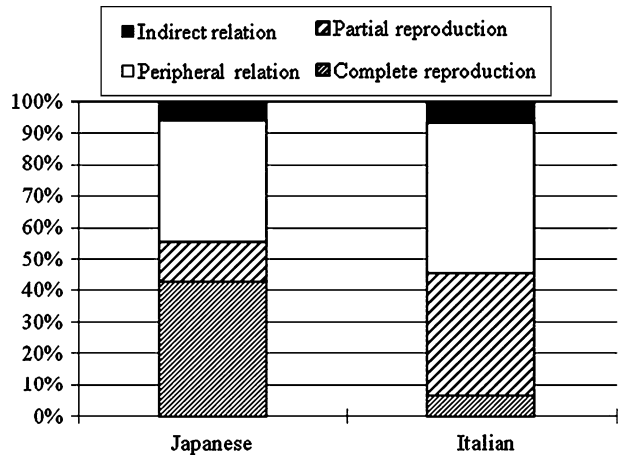
But when the first two categories (complete and partial reproductions) and the other two categories (peripheral relation and indirect relation) were collapsed together, no differences between Italian and Japanese children emerged. This result is shown in Fig. 3 reporting percentages of the four categories produced by Japanese and Italian children.

Speech–Gesture Relationship

Modality of Expression

Regarding the relationship between gestures and words, both groups of children produced representational gestures with and without spoken responses: the number of bimodal

Fig. 3 Percentages of the four categories of level of reproduction produced by Japanese and Italian children



(gesture + speech) productions was high for both groups (69% for the Italian group and 95% for the Japanese group), but Japanese children produced more bimodal gestures than Italian children (Mann–Whitney test: $U = 128$, $Z = 3.11$, $p < .01$) and Italian children produced significantly more gestures without speech (unimodal gesture production) than Japanese children ($U = 128$, $Z = 3.11$, $p < .01$). The type of spoken responses that was more frequently accompanied by gestures was different for each sample: Japanese children produced a higher number of gestures associated to *incorrect responses* (Wilcoxon test, $Z = 3.14$, $p < .01$), while Italian children exhibited a similar frequency in gestures accompanying *correct and incorrect responses* (Wilcoxon test, $Z = 1.59$, $p = .11$).

Temporal Relationship

Both samples produced a high percentage of gesture in synchrony with speech (82% for Japanese and 77% for Italians in the total number of strokes, Mann–Whitney test: $U = 23$, $Z = .26$, $p = .80$). No significant differences were found between Japanese and Italian children for each index of temporal relationship: in the majority of cases (79% for the Japanese and 83% for the Italian sample, $U = 23$, $Z = .20$, $p = .84$) gestures started before speech (Wilcoxon test for Japanese: $Z = 2.97$, $p < .01$; for Italians: $Z = 3.96$, $p < .01$) and ended after speech (81% for both Japanese and Italian samples, $U = 228$, $Z = .34$, $p = .73$) (Wilcoxon test for Japanese: $Z = 3.52$, $p < .001$; for Italian: $Z = 3.06$, $p < .01$).

Conclusion

The purpose of this study was to verify the hypothesis that gestures function as motor representation at an early stage of lexical acquisition, comparing representational gestures performed by Italian and Japanese children. The present findings showed that a simple picture-naming task, while providing a common ground for data collection, proved to be a favorable structured, experimental setting, as it enabled comparing both the spoken and the gestural production of young children from different languages and cultures. The results

described in the previous section showed that Japanese children produce representational gestures similar for frequency and type to those produced by Italian children confirming that, for both groups of children, gestures function as motor representations supporting the spoken ones during early stages of language development. In the remaining pages of this final section we will discuss more closely the results presented according to each aspect considered: spoken accuracy, frequency and typology of representational gestures, and timing of spoken and gestural production. For each aspect we will also provide explanations of minimal differences found between Italian and Japanese children due to linguistic and cultural diversity.

We found that Italian children produced more *correct* spoken labels than Japanese children. This may be the effect of slight differences in vocabulary acquisition at an early developmental stage, which have been reported also by other researchers as present in younger children. A recent study has shown that the mean age at which Japanese infants acquire the first 50 words is a little higher than the age reported for American infants (Ogura 2007), while previous studies indicate that Italian children acquire the first 50 words at the same age as American children (Caselli et al. 1995). We also found that Japanese children produced more onomatopoeia than Italian children. This finding can be explained by considering the fact that onomatopoeia are lexicalized in the Japanese language, and that this occurs at an early stage of language production (Imai et al. 2008). This is confirmed by the very high number of onomatopoeia in the Japanese CDI. The Japanese language is uniquely rich in relation to this type of expressions, which are frequently used in daily conversations, magazines and newspapers because of their brevity and power to project vivid imagery and represent a peculiarity of Japanese culture, a Japanese way of expressing feelings and/or mental states (Clancy 1990). Fernald and Morikawa (1993) compared Japanese and American mothers' speech to infant at 6, 12, and 19 months and reported that Japanese mothers used onomatopoeia at all considered ages, while American mothers rarely used them at all.

Despite differences in their spoken responses, both Japanese and Italian children produced representational gestures when performing the naming task and with a similar frequency. Both groups performed more gestures when viewing items representing actions or object characteristics rather than when seeing items depicting objects. In addition, the items that elicited the greatest number of gestures were the same for both groups. Moreover, gesture and speech timing was very similar across groups: in both Japanese and Italian children, we observed that in most cases gesture production started before and ended after word production. Our results, showing that gesture stroke *precede* the corresponding spoken word, are consistent with the Lexical Retrieval Hypothesis (Krauss 1998; Pine et al. 2007) which states that gesture use facilitates the retrieval of lexical items from memory and thus plays a direct role in the *speaking* process. No previous study had examined temporal synchrony between the two modalities at this early age.

All these similarities suggest the existence of a common biological basis, which may stand for a shared motor and communicative development in both Japanese and Italian cultures. In young children motor representations appear to support linguistic representations in speech: performing a gestural motor representation may be necessary in order to offer a more experiential dimension and a more precise and concrete image linked to the concept expressed by the word.

Despite these similarities, some differences were also noted that could be explained by referring to cultural differences. First, Italian children produced more gestures without speech, while Japanese children produced more gestures with incorrect spoken responses, confirming that both groups resort to gestures when the spoken label is unavailable or

difficult to retrieve. It could be possible that Italian children are influenced by an environment where adults often use gestures as emblems without resorting to speech (Kendon 2004). Second, as for the representational techniques and the level of reproduction adopted, Japanese children produced fewer size-shape gestures than Italian children, offering gestures that reproduced more closely the action represented in the picture. The tendency of Japanese children to reproduce a model more precisely may be related to a learning style typical in Japan. Literature on Japanese culture suggests that knowledge and skills are often transmitted without verbal explanation, as shown in the art of Japanese sushi making (De Waal 2001; Matsuzawa 2001) where the apprentice learns the art of sushi by observing what the master is doing. It seems that learning by observing is more common in the Japanese culture, while in the Italian culture active teaching based on verbal and gesture modalities is more common. Compared to Indo-European culture and languages, skills in Japan may tend to be conveyed through observation or imitation. The basis of the Japanese learning style seems to be a set of cultural values that emphasize *omoiyari* (empathy). Feeling of *omoiyari* is so widely shared that overt verbal communication is often not required (Clancy 1990; Rothbaum et al. 2000). As reported by Azuma (1994) empathy is fostered in young Japanese children because it is the cornerstone of the child's willingness to imitate and to please the parent. Studies on early mother–child interaction have revealed patterns emphasizing nonverbal communication at an extremely early stage. A study by Fernald and Morikawa (1993) comparing Japanese and American mothers' speech to infants found that mothers' speech in both cultures shared common characteristics, such as linguistic simplification and frequent repetition, and mothers made similar adjustments in their speech to infants of different ages. However, American mothers labeled objects more frequently and consistently than Japanese mothers, while Japanese mothers used objects to engage infants in social routines more often than American mothers. Further studies on the communicative interaction between mothers and very young children are needed in order to investigate if parental attitude toward gesture use may impact gesture production.

To conclude, our study shows the robustness of gesture use in a naming task by children at an early stage of lexical development. Our findings suggest that when 2-year-old children label pictures depicting objects or actions, occasionally they still need to perform an 'action' in the form of a 'gesture' (Stefanini et al. 2009). The function of these gestures may be to recreate a direct link with the object or the action to be labeled. This suggests that words may not yet be fully de-contextualized, and the production of a gesture may recreate the context in which the word was initially acquired (Capone 2007). There were also interesting similarities and differences between Italian and Japanese children in the way in which a depicted item was represented. Motor representations may be needed to support linguistic representations in speech, irrespective of the cultural environment in which the child is raised, but the way gestures are produced may be influenced by culture even from an early developmental stage. The connection between a body representation and speech representation needs to be examined too. So far, research has tried to reveal this in younger children also by using a correlational analysis or comparing the mean numbers of action and gesture (A/G), speech comprehension and speech production based on parental reports (Caselli et al. 2012): These analyses and this methodological frame appear to confirm that A/G and speech are tightly related. Recent observational studies have shown also how caregivers guide infants, with verbal and nonverbal messages, to direct their attention to relevant affordances of objects and effectivities (i.e., bodily abilities) through an "assisted imitation" strategy (Zukow-Goldring 2006; Zukow-Goldring and Arbib 2007). Mother–child interaction and assisted imitation contribute to expanding and enriching the representational properties of the motor system. What must still be

understood is the full impact of a child's culture and language versus his/her natural predisposition to resort to motor representations to support verbal development at different ages and for different communicative purposes. Future research may examine whether similar findings could be reported for other cultures and for other age groups. Recent studies (Gullberg 2009; Özyürek et al. 2008) have already shown that the use of gesture to describe motion events is associated with the structure of the spoken language. Improvement of such investigations might greatly contribute to our understanding of how and when culture and language influence the development of gesture and speech.

Acknowledgments The work reported in this paper was partially supported by the Fondazione Monte di Parma (Research Group for Study on Children's Motor and Language Development, University of Parma), by a national research grant PRIN 2008—2008J2WEEK (“Gestures and language in children with atypical and at risk developmental profiles: relationships among competences, mother–child interaction modalities and proposals of intervention”) and the “Executive programme of cooperation in the field of Science and Technology between Italy and Japan” of the Ministero degli Affari Esteri, Italy. We are very grateful to Marianne Gullberg for her insightful suggestions on an early presentation of the study and to Aliyah Morgerstern, Paula Marentette and Laura Sparaci for helpful comments on previous versions of this manuscript. We especially thank the children who participated in the study, their parents and schools.

References

- Acredolo, L. P., & Goodwin, S. W. (1994). Sign language among hearing infants: The spontaneous development of symbolic gestures. In V. Volterra & C. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 68–78). Washington, DC: Gallaudet University Press.
- Aqui, H. (2004). *70 Japanese gestures: No language communication* (A. Chang, Trans.). Berkeley: Stone Bridge Press.
- Azuma, H. (1994). Two modes of cognitive socialization in Japan and the United States. In P. Greenfield & R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 275–284). Hillsdale, NJ: Erlbaum.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.
- Bello, A., Caselli, M. C., Pettenati, P., & Stefanini, S. (2010). *Parole in gioco: Una prova di comprensione e produzione lessicale*. [Picture naming game: A lexical comprehension and production task]. Firenze: Giunti, Organizzazioni Speciali.
- Blake, J., Vitale, G., Osborne, P., & Olshansky, E. (2005). A cross-cultural comparison of communicative gestures in human infants during the transition to language. *Gesture*, 5, 201–217.
- Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one to two word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–258). Cambridge: Cambridge University Press.
- Capirci, O., Contaldo, A., Caselli, M. C., & Volterra, V. (2005). From action to language through gesture: A longitudinal perspective. *Gesture*, 5(1/2), 155–177.
- Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gesture and words during the transition to two-word speech. *Journal of Child Language*, 23, 645–673.
- Capone, N. C. (2007). Tapping toddlers' evolving semantic representation via gesture. *Journal of Speech, Language, and Hearing Research*, 50(3), 732–745.
- Caselli, M. C. (1994). Communicative gestures and first words. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 56–67). Washington, DC: Gallaudet University Press.
- Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., et al. (1995). A crosslinguistic study of early lexical development. *Cognitive Development*, 10, 159–201.
- 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*. [The first child's vocabulary. Guide to using the MacArthur questionnaire for the assessment of communication and language in the first year of life]. Milano: Franco Angeli.

- Caselli, M. C., Rinaldi, P., Stefanini, S., & Volterra, V. (2012). Early action/gesture “vocabulary” and its relationship with word comprehension and production. *Child Development*. doi:10.1111/j.1467-8624.2011.01727.x.
- Clancy, P. (1990). Acquiring communicative style in Japanese. In R. Scarcella, E. S. Andersen, & S. D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 27–33). Rowley, MA: Newbury House.
- Colletta, J. M. (2004). *Le développement de la parole chez l'enfant âgé de 6 à 11 ans. Corps, langage et cognition*. [The speech development in children aged 6 to 11 years. Body language and cognition]. Hayen: Mardaga.
- De Jorio, A. (1832). *La mimica degli antichi investigata nel gestire napoletano*. [The mimicry of ancient peoples investigated through Neapolitan gestures]. Napoli: Fibreno.
- De Waal, F. (2001). *The ape and the sushi master: Cultural reflections of a primatologist*. New York: Basic Books.
- Diadori, P. (2003) *Senza parole. 100 gesti degli italiani*. [Without words. 100 gestures of the Italians] (4th ed.). Roma: Bonacci.
- Efron, D. (1941). *Gesture and environment*. New York: King Crown Press.
- Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Phonetica*, 57, 242–254.
- Goldin-Meadow, S., & Singer, M. A. (2003). From children's hands to adults' ears: Gesture's role in the learning process. *Developmental Psychology*, 39(3), 509–520.
- Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Sweden: Lund University Press.
- Gullberg, M. (2009). Gestures and the development of semantic representations in first and second language acquisition. *Acquisition et Interaction en Langue Etrangère*, 1, 117–139.
- Gullberg, M., De Boot, K., & Volterra, V. (2008). Gestures and some key issues in the study of language development. *Gesture*, 8(2), 149–179.
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sounds symbolism between a word and an action facilitates early verb learning. *Cognition*, 109, 54–65.
- Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development*, 1, 23–43.
- Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication of Italian vs. American children. *First Language*, 28(2), 164–181.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *American Psychological Society*, 16, 367–371.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Krauss, R. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science*, 7(2), 54–60.
- Matsuzawa, T. (2001). *Primate origins of human cognition and behavior*. Tokyo: Springer.
- Mayberry, R., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Psychological Science*, 9(6), 192–196.
- McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- Munari, B. (1994). *Il dizionario dei gesti italiani* [The dictionary of Italian gestures]. Roma: Adnkronos Libri.
- Ogura, T. (2007). Early lexical development in Japanese children. *Gengo Kenkyu*, 132, 29–53.
- Özyürek, A., Kita, S., Allen, S., Furman, R., Brown, A., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, 44(4), 1040–1050.
- Pettenati, P., Stefanini, S., & Volterra, V. (2010). Motoric characteristics of representational gestures produced by young children in a naming task. *Journal of Child Language*, 37(4), 887–911.
- Pine, K. J., Lufkin, N., Kirk, E., & Messer, D. (2007). A microgenetic analysis of the relationship between speech and gesture in children: Evidence for semantic and temporal asynchrony. *Language and Cognitive Processes*, 22(2), 234–246.
- Pine, K. J., Lufkin, N., & Messer, D. J. (2004). More gestures than answers: Children learning about balance. *Developmental Psychology*, 40(6), 1059–1067.
- Pizzuto, E., & Capobianco, M. (2005). The link and differences between deixis and symbols in children's early gestural-vocal system. *Gesture*, 5, 179–199.
- Rothbaum, F., Pott, M. L., Azuma, H., Miyake, K., & Weisz, J. (2000). The development of close relationships in Japan and the U.S.: Paths of symbiotic harmony and generative tension. *Child Development*, 71, 1121–1142.

- Stefanini, S., Bello, A., Caselli, M. C., Iverson, J., & Volterra, V. (2009). Spoken and gestural lexicon in a naming task: Developmental data. *Language and Cognitive Processes, 24*(2), 168–189.
- Stefanini, S., Caselli, M. C., & Volterra, V. (2007). Spoken and gestural production in a naming task by young children with Down syndrome. *Brain and Language, 101*(3), 208–221.
- Zukow-Goldring, P. (2006). Assisted imitation: Affordances, effectivities, and the mirror system in early language development. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 469–500). Cambridge: Cambridge University Press.
- Zukow-Goldring, P., & Arbib, M. A. (2007). Affordances, effectivities, and assisted imitation: Caregivers and the directing of attention. *Neurocomputing, 70*(13–15), 2181–2193.