

## “THE EGG AND JERRY”: NARRATION AND GESTURE IN L1 AND L2 BY ITALIAN SCHOOLCHILDREN

CAPUSSOTTI, Giulia<sup>1\*</sup>  
MELUZZI, Chiara<sup>2</sup>

<sup>1,2</sup>University of Pavia

---

**Abstract:** *In this paper, we present sociolinguistic research conducted on Italian schoolchildren learning English as L2. Following on from renowned researchers, we focused on a less studied population, which is school-aged monolingual children. Our participants consisted of 15 students of a fourth-grade class at a primary school in Pavia; all aged around 9 years of age - 7 boys and 8 girls. None of the children present any recorded cognitive problems; they are all Italian L1 speakers with little or no use of other languages at home, and English has been learnt as L2 since the beginning of primary school at age 6. We recorded all the children performing a task of re-narration of a “Tom & Jerry” cartoon, firstly in Italian and then, after one week, in English. The corpus consists of about 2h 45’ of recordings, transcribed and annotated in ELAN. Lexical knowledge in English was also tested through a questionnaire before the recordings. The results were analyzed both qualitatively and, partly, quantitatively. During qualitative analysis, two elements were observed: (1) general tendencies in speakers’ general behavior and (2) differences in the relationship between syntactic-conversational system and gesture system in relation to L1/L2. The quantitative analysis shows a difference in the use of beats gestures and iconic ones between L1 and L2, but also between boys and girls. The main findings of the current research are: (1) frequent use of gesture during L1 recordings, in contrast to a general decrease of gesture during L2 recordings; (2) gesture as linguistic reinforcing or replacement of speech production during L1 recordings, in contrast to gesture as a problem-solving strategy during L2 recordings; (3) generally, offline knowledge in L2 is not reflected in online knowledge.*

**Keywords:** multimodality; speech-gesture; hand gestures; English L2; schoolchildren.

---

\* Corresponding author: [chmeluzzi@gmail.com](mailto:chmeluzzi@gmail.com)

<sup>1</sup>This study has been jointly conceived and written by the two authors, who also jointly made the audio-video recordings of the corpus. However, for academic disclosure purposes, Giulia Capussotti is responsible for sections 2, 3, 4.2 and 5 and 5.1, and of performing the transcription of the data; Chiara Meluzzi is responsible for sections 1,4.1, 5.2, 6 and 7, and for the experimental design.

## **1 Introduction**

The co-occurrence of gesture and speech has been widely studied in connection with a baby's linguistic development, but also in the fields of language impairment and bilingual studies. However, results are not always similar between one language and the other, and there is a substantial lack of knowledge of certain age groups - in particular in relation to the comparison between L1 and a possible L2. For this reason, in this paper we wanted to provide a first insight on narration and gesture in the production of a group of 15 Italian pupils of 9 years of age, video-recorded while performing narrations and short dialogical sequences in both Italian L1 and English L2. A preliminary lexical test was also performed in order to assess the children's knowledge of the main lexical terms, both nouns and verbs, in the cartoon. Data were annotated following a specific annotation protocol on 4 tiers through the ELAN software, in order to identify not only different hand gestures, but also gestures produced with other parts of the body, in particular the head. Data were then analyzed both qualitatively and quantitatively in order to provide a complete picture of the connection between gestures and speech during narrations in L1 and L2. Furthermore, the pupils' sex was also considered, since, as already pointed out in the literature, girls and boys behaved differently, in both L1 and L2 tasks (e.g., Coates, 2015 for a review).

The paper is organized as follows: after a survey of the literature related to gestures and linguistic development and language use in children, our research questions are presented in section 3. Section 4 presents the research protocol for the experiment we performed on 15 Italian schoolchildren, and illustrates the corpus collected and the annotation protocol used for classifying gestures. The analysis is divided between a qualitative and a quantitative approach, with the main results further discussed in the sixth section of the paper, before moving to some preliminary conclusions and further perspectives.

## **2 Theoretical framework**

Co-speech gestures are extensively studied in adults, adolescents and infants. Conversely, much less is known about coordination between speech and gestures in school-aged children, although works by Colletta et al. (2016, cf. also Colletta, 2009) on French children have demonstrated that after the age of nine, children use gestures similarly to adults. However, research has been primarily focused on the initial phases of linguistic development (cf. Capirci et al., 2005), or even on the pre-linguistic phase (cf. Bates et al., 1975), and on linguistic development in children with atypical linguistic development (cf. Bello et al., 2004; Legerstee & Fisher, 2008). Conversely, children between 6 and 10 years old are still poorly investigated, especially interlinguistically, and in relation to foreign language acquisition in the classroom.

During primary school, children's communicative abilities develop and acquire importance in order to determine their social interaction. Indeed, during these years, the interactions with other people (both classmates and adults) increase, and activities become more complex. Therefore, much more participation and attention is required: on the one hand, children receive new input (both linguistic and non-linguistic); on the other hand, they are asked to produce new output. Verbal communication is the first tool that allows teachers to realize what children understand and know. However, children are not always completely aware of what they know, or they are not able to express themselves accurately through speech. In this case, gesture can play a significant role: hand gestures, facial expressions, gaze, and body movements are useful tools to get more information about what children think. So, it is important to go beyond words and look at all communication facets. Pioneers in Gesture Studies, such as McNeill and Kendon, have claimed that gestures can convey concrete messages and provide information about the speaker's mental representations (cf. Kendon, 1980; McNeill, 1985, 1987, 1992). These works have shown that we can assign meaning to those gestures

and that different interlocutors, or observers, tend to assign the same meaning to the same gesture: these observations have emerged in different experimental contexts, such as in narration tasks (McNeill, 1992), and in speech conversations (Kendon, 1980), just to name two.

Concerning children, and specifically schoolchildren, some studies have been conducted on how gesture and speech can co-operate in learning. It has been demonstrated how gestures can change children’s cognitive representations and improve their learning strategies (cf. Church & Goldin-Meadow, 1986; Alibali & Goldin-Meadow, 1993; Alibali et al., 1997; Garber et al., 1998; Garber & Goldin-Meadow, 2002; Church et al., 2004; Wagner Cook & Goldin-Meadow, 2006; Broaders et al., 2007; Ping & Goldin-Meadow, 2008; Wakefield & James, 2015). This is a first insight on the importance of multimodality in improving both teaching and learning strategies. Furthermore, some interesting works have been conducted in the field of second language acquisition and bilingualism: according to scholars, gestures may favor second language memorization of lexical items (Tellier, 2008; Macedonia & von Kriegstein, 2012) and reduce cognitive efforts in bilingual subjects (Church et al., 2004; Pika et al., 2006; Nicoladis et al., 2009).

### **3 Research questions**

The main purpose of our research was to prove the interdependence between speech and gesture in a class of Italian schoolchildren in L1 and L2. In particular, we wanted to ascertain: (1) the co-occurrence of speech and gesture both in L1 and in L2; (2) the role of gesture in narration in both L1 and L2; (3) the type of speech-gesture relation.

In order to answer these queries, we set up a controlled experiment with 15 pupils of a fourth-grade class in a primary school in Pavia. All the children were native speakers of Italian and they share a similar sociolinguistic background; furthermore, they have studied English as part of their school program since the age of 6 (first-grade). As we will explain in the following section, we performed the task first in Italian and then in English in a controlled environment; that is, in a room separated from other classes but still during school hours. Moreover, we used two methods of data collection in order to have a possible stylistic variation between free narration and guided conversation. This choice allowed us to minimize the so-called “observer’s paradox” (Labov, 1972); indeed, during an experiment, subjects are inevitably influenced by the presence of the researcher because they know their behaviour is being observed and analyzed; for this reason, the researcher’s intervention should be limited and the attention oriented towards the content of the task rather than on the form. In this way, we collected our data in a controlled context, but trying to maintain the highest degree of naturalness. Our controlled experiment explicitly wanted to test the following research questions: (1) Do gestures vary both quantitatively and qualitatively in narration in L2 with respect to L1? (2) Do gestures perform a different role in shaping narration in L2 and in L1? (3) Is it possible to highlight a stylistic difference between the speech-gesture relations in free narration when compared to guided conversation?

Based on the previous literature on this topic, we expected an increase in the number of gestures in L2 with respect to L1, especially the so-called “word-finding” gestures, signaling the maintenance of the conversational turn while the speaker is looking for the correct word. In this respect, our expectation was also that gestures would play a major role in L1 in organizing the narrative sequences, whereas in L2 different gestures other than beats would occur. Finally, we also expected that style would play a role in determining the presence of more gestures during the free narration than in the guided conversation, since in the first task the children had to perform a long monologue.

## 4 Methods

### 4.1 The experiment

The subjects of this study consisted of 15 pupils (7 boys and 8 girls) of a fourth-grade class of a primary school in Pavia, a city in the north-west of Italy, near Milan. The pupils were all aged 9 years old at the time of the recordings. The schoolmaster and the teachers were informed of the purposes of the experiment, and it was made clear that no evaluation would be conducted on the data. The teachers also provided the researchers with information about the pupils, in order to verify that they shared a similar sociolinguistic background, and that there were no cases of learning pathologies. A sociolinguistic questionnaire was also shared through the teachers with the pupils and their families, in order to have a more complete sociolinguistic profile of the children. All families were informed of the project, and an ethical statement was shared explaining the procedures of the experiment, also emphasizing that no harm would occur to their children during the data collection. All families reviewed the ethical statement and signed a privacy disclosure module in order to authorize the researchers to collect and share the data in an anonymous form (for more information, see also Capussotti, 2019).

During a first meeting with the class, we explained the experiment we were going to perform, without mentioning the importance or the focus on gestures, in order to avoid a bias in the data. This is a crucial point because the children were not aware of the researcher's focus on speech-gesture cooperation, and so they were not influenced by the goal of the experiment. We explained to the pupils all the phases of the task of the experiment, and what they were supposed to do in front of the camera. It was not anticipated that a phase of guided conversation would follow. We also specified that we would ask them to do the experiment twice over two separate days, and that during the second day they would be asked to tell the story in English. All the children seemed to understand the task, and they all agreed to take part in the experiment.

Data were collected through audio and video recordings by using a Tascam recorder and a Sony  $\alpha$ 6300 camera. Audio signal was recorded in .wav format, at a frequency of 44.1 KHz and a sampling rate of 16-bit. The experiment was performed in the school library during school hours in two different recording sessions, one for narration in Italian L1 and the other for English L2. The school library was not in use during the two days of the recordings; however, it was a completely silent environment and the recordings had to be interrupted during the two breaks imposed by the school timetable.

In the first day of recordings, all children performed the task in Italian. Kids left the room, and watched a 5' video from the "Tom & Jerry" cartoon series. The cartoon is speechless, and tells the story of a woodpecker's egg that falls from the nest and reaches Jerry's home, where he discovers it. The baby woodpecker hatches and believes that Jerry is his mom; since the baby bird starts to destroy all the wooden furniture in the mouse's house, Jerry brings him back to the nest but the bird follows him home again. This short cartoon was selected because of the absence of speech, the presence of only two characters (Jerry, and the egg, and then the baby bird), and simple content counterbalanced by a lot of physical actions performed by the protagonists. No pupil showed signs of not having understood the cartoon or of discomfort during the experiment.

After watching the cartoon twice, the pupil sat in front of the researcher (i.e., the first author of the paper) at a 45-degree angle with respect to the camera; the TASCAM recorder was settled at a convenient distance from both the child and the camera in order to avoid interference. A second researcher (i.e., the second author of the paper), who was also present, was in charge of setting the devices and controlling the quality of the recordings. After a few seconds, all the children immediately stopped paying attention to the second researcher and the camera, thus focusing only on their primary interlocutor, which is the first researcher.

Pupils were explicitly asked to narrate the story on their own with as many details as they remembered, and were also reassured once again that no evaluation would be provided. After the child narrated the cartoon in a monologue form, the first researcher moved directly to the second task, that is a guided interview. The researcher asked questions about more details in the cartoon, to explain the reasons of the protagonists of the cartoon, and/or to explore the feelings showed by the characters. As the pupils were informed during the first meeting, a second recording day took place one week after the first. During these recording sessions, pupils were asked to watch the same cartoon twice, and then to narrate it in English. As for the previous session, the first researcher also conducted a guided interview in English on the same themes involving the cartoon’s characters. The English recordings took place one week apart so as to avoid possible interference between the two activities: if recordings in L1 and L2 had been conducted in the space of just one day, children might have been affected by the lexicon and syntax used during the first interview. Indeed, the two linguistic systems (L1, L2) should not be treated as two independent systems; on the contrary, they should be seen as two interdependent systems (Grosjean, 1989; Grosjean et al., 2013). It is worth specifying that the cartoon was mute, with just background music; in this way no preliminary linguistic input was given to the subjects immediately before the recording took place.

Between the two recording sessions, we also conducted a lexical test with the same pupils, in order to ascertain their lexical knowledge of the main English nouns and verbs that would possibly occur in the narration of the cartoon (e.g., nouns like *egg*, *flower*, *mouse*, *bird*, and verbs like *fall*, *jump*, *sit*, etc.). The test was an association-task, where children had to link the Italian words in the first column to the correct English equivalents in the second column. This activity allowed us to evaluate the children’s offline L2 knowledge, thus allowing for a comparison with their ability or inability to use the same knowledge online (that is, during narration and conversation). It is true, however, that this preliminary test conducted in English could be seen as a linguistic input, but it was presented to the children some days before the recordings took place and without specifying that the test was related to the recording session. The results of the test showed that the children had a good offline knowledge: all the words were correctly associated between the two languages, and just minor mistakes were randomly present. From this small test, it could be said that the children knew the words, both nouns and verbs, which were supposed to be used in the narration of the cartoon in English.

#### 4.2 Corpus and annotation protocol

We collected 30 recordings, 2 for each subject, equally balanced between Italian and English. Table 1 quantifies the data collected during the experiment. Over a total of about 2 hours and 45 minutes, the children produced 8,104 words. It should be also noted that recordings in L1 were shorter than in L2, whereas the amount of words pronounced was higher in L1 than in L2, as it was possible to expect (cf. also Table 2 in section 5.1 for the distribution of types and tokens).

**Table 1:** Recordings total duration and number of words classified according to the language and sex variables.

	L1 recordings		L2 recordings	
	Duration	Words	Duration	Words
<b>Males</b>	30’ 03”	2,157	43’ 38”	1,679
<b>Females</b>	38’ 18”	2,297	53’ 05”	1,881
<b>Total</b>	<b>1h 8’ 21”</b>	<b>4,454</b>	<b>1h 36’ 43”</b>	<b>3,560</b>

All video and audio files were time-aligned with the Elan 5.3 software (Wittenburg et al., 2006) in order to create the final corpus. Subsequently, all the recordings, including both the narration and the guided conversation, were orthographically transcribed; no automatic forced-alignment was used, as the children's speech was particularly fragmented, full of pauses and reformulation. Moreover, a manual transcription at this stage of the research was essential for later work on the interconnection between speech and gesture. Transcriptions were performed on two different tiers, one identifying the pupil, and the other the research (for the guided conversation). Narrations and dialogues were then exported automatically as file text in .txt format.

For the annotation of the gestures, three other tiers were created in ELAN, in order to annotate (1) the gestures produced with hands, (2) facial expressions, and (3) gestures produced with other parts of the body (e.g., head). A specific annotation protocol was created (see Capussotti, 2019: 50-53), based on McNeill's (1992) classification for hand gestures within tier 1, and a simplified version of Ekman et al.'s (2002) FACS protocol for facial expression within tier 2, without an indication of the intensity of the gesture, but only the organ involved and the direction of the gesture. Finally, for the third tier, we created a series of labels to indicate the organ performing the gesture and the typology of the gesture, again based on McNeill's (1992) classification. For instance, the tag BEAT\_HEAD corresponds to a beat gesture performed with the head, meaning that the child is using their head to mark the pacing of their narration. It is worth specifying that iconic and metaphoric gestures, in some cases, are realized with the head: indeed, it happens that the subject visually reproduces a referent or an action through the unique use of the head (see also section 5.1). Within our sample, for instance, a subject realizes the tapping produced by the woodpecker's chick, with his beak, through the imitation of the same action with his head. This reflects the definition of iconic gestures, and their role, as McNeill defined them (cf. McNeill, 1992).

The last step was the creation of a fourth tier to annotate gestures in relation to speech, according to five categories:

- ACTION: gestures produced to visually represent an action;
- OBJECT: gestures produced to visually represent an inanimate referent;
- SUBJECT: gestures produced to visually represent an animate referent;
- MANNER: gestures produced to visually represent a behavior or emotional state;
- SPACE: gestures produced to visually represent an imagined space.

For instance, a child could use a deictic gesture in order to visually reproduce an imagined space, as in the example in Figure 1.



**Figure 1:** Example of SPACE: deictic gesture realized with the left hand to represent the imagined space “outside”.

The related transcription extract is as follows (double dots in the transcription indicates an unnatural lengthening of a vowel): “Poi fa trepassi e cosa:?. E va bene. E quindi lo rimette fuori dalla porta e poi lui tata: si fa un buco e rientra” (Eng. Then he takes three steps and what? And it’s ok. And then he puts him outside the door and then he tata: makes a hole and comes back inside). So, as we can see in Fig. 1, the child represents the imagined space “outside” through a deictic gesture realized with his left hand, while uttering the sentence “puts him outside”, to physically represent a space far from the centre of the narration (i.e., Jerry’s house).

Another instance is reported in Fig. 2, representing the category MANNER: the child could employ various parts of his body to visually represent a particular emotional state of one of the characters of the cartoon.



**Figure 2:** Example of MANNER: gesture realized with hands and face to recreate the emotional state of “relief”.

The related transcription extract follows: “Ah sì, il coso piccolo è...triste. Invece Jerry dice nella sua mente "finalmente, un po' di sollievo!"(Eng. Ah yes, the little thing is... sad. Instead Jerry says in his mind “finally, a bit of relief!”). In Fig. 2, we can observe that the subject uses both his arms and facial expression (the image has been darkened for privacy purposes) to convey the idea of “relief”. The gesture starts with the second clause introducing Jerry’s emotional state, and continues until the end of the imagined thought.

A final example for the category of SUBJECT is presented in Fig. 3. In this case, the child realizes an iconic gesture, realized with both her hands, to represent the woodpecker’s chick, while she was presenting the new character by saying “Eh: l'uovo si apre ed esce fuori un cucciolo di picchio eh:” (Eng. And: the egg opens and a woodpecker’s chick pops out).



**Figure 3:** Example of SUBJECT: iconic gesture realized with both hands to recreate the “woodpecker’s chick”.

After the annotation was completed, the number of gestures for each category was inserted in a matrix created in Excel, also reporting the languages in which the task was performed, the codified name of the speakers and the sex of the child. This matrix, together with the transcription and annotation, allows for an integrated quantitative and qualitative analysis of the data at our disposal, as we will see in the following section.

## 5 Analysis

The analysis has followed two broad approaches, one qualitative and one more quantitative. In the qualitative analysis, we observed the use of gestures during narration/conversation and their interconnection with speech, by emphasizing differences in the amplitude of the gestures. Moreover, we focused on the different role assumed by even the same gestural typology in L1 and L2. Finally, we observed the various typologies of speech-gesture relation. Conversely, the quantitative analysis is narrower, also because of the limits of our corpus, and was entirely focused on hand gestures; indeed, this category presented a quantity of tokens for statistical analysis with the IBM SPSS 20 software. We investigated the quantity of hand gestures in both L1 and L2 recordings, by highlighting whether this difference also assumes statistical significance with respect to variables of gesture and language, and pupils' sex. This last sociolinguistic variable was included after a major difference in gesturing behavior between males and females was detected in the qualitative analysis.

### 5.1 Qualitative analysis

In the study on the use of speech and gesture in narrations, and the speech-gesture interconnection, a first major difference emerged according to the sex of the children. As a general observation, boys tended to talk less and gesture more, with the aid of wide gestures that occupied all the space in front of them and involved both their hands. On the contrary, girls' gestures were more limited, both in quantity and in extension, whereas their speech was richer in both lexical and syntactic terms.

**Table 2:** Tokens, types, subordinates and coordinates in the children's narrations according to speakers' sex, and language of the task. (\* only if expressed in L2)

	L1 - Italian				L2 - English			
	Tokens	Types	Coord.	Subord.	Tokens	Types	Coord.*	Subord.*
<b>Males</b>	2,316	261	113	150	2,025	280	11	5
<b>Females</b>	2,443	330	141	140	2,025	353	28	7

Data in Table 2 reflect this difference in language usage between language boys and girls in L1 and L2. Girls' speech was also more articulated through syntax, whereas the boys opted for a succession of single sentences, especially in L2. Indeed, the data show that girls in L1 used 281 syntactic conjunctions between coordinate and subordinate clauses, whereas boys used 263 syntactic conjunctions, although they tend to use more subordinate clauses, in particular temporal ones. In L2 the difference is more straightforward, since there were 35 conjunctions in English for females (28 coordinative, 7 subordinate), and only 16 for boys (11 coordinative, 5 subordinate). The lexicon also



appeared to be richer for girls than for boys, as testified by the number of types used, in both the languages of the task.

However, this contrast, between boys and girls, was especially evident in L1 recordings, while it became less striking during the L2 recordings. Data in Table 2 for English recordings testify that the difference between boys and girls is not straightforward, neither for the types/tokens difference nor for the use of coordination and subordination. This is probably due to the fact that all the pupils perceived a greater difficulty in performing the task in English rather than in Italian, and this caused a general inhibition. This is also reflected in the generally shorter duration of L2 recordings compared to Italian ones (cf. Tab. 1).

As for gestures, data in Table 3 highlight how boys tended to produce many more gestures than girls did, and that those gestures had a major amplitude. Furthermore, boys often accompanied, or replaced, their speech with gestures. On the contrary, girls generally produced fewer gestures, and, when they did, those gestures were spatially more limited. This difference between boys and girls reduced in the structured dialogue conducted with the researcher, more in L1 than in L2. Moreover, it should be noted how the number of gestures in boys’ narrations was equal in L1 and L2, whereas girls produced 10 more gestures in L2 narrations. The difference is, however, minimal.

**Table 3:** Number of gestures (hands, faces and other body parts) for boys and girls in L1 and L2 narrations and guided interviews.

	L1 - Italian		L2 – English	
	Narration	Dialogue	Narration	Dialogue
<b>Males</b>	168	61	164	30
<b>Females</b>	54	59	64	10
<b>TOTAL</b>	222	120	228	40

As for the relation between gestures and speech, some interesting elements emerged. Firstly, a reduction of the number of gestures in L2 has been observed, but it is specifically limited to the dialogical context. Style played a role in L1 as well. Thus, in general, it is possible to say that the children appeared more spontaneous and talkative when narrating on their own, whereas the dialogue with an adult researcher may have inhibited their communication. Moreover, this dialogical task appeared extremely difficult to be performed in English, thus provoking a decrease in both gesture and speech production.

Secondly, a substantial difference in the role of gestures in L1 and L2 has been observed. On the one hand, gestures were used as intensifiers or substitutes of speech during the narration in the Italian L1. On the other hand, gestures were used as a problem-solving strategy during the narration in the English L2. Fig. 4 and Fig. 5 present two examples of the same child producing the iconic gesture of the ‘egg’ while speaking in Italian and English, respectively.



**Figure 4:** Example of iconic gesture representing the “egg” during L1 narration.

In Figure 4, the child was telling the story in his Italian L1, and he realizes the gesture of the ‘egg’ concomitantly with speech: the sentence produced was “C’è questa mamma che ha un piccolo uovo” (Engl. There’s this mom who has a little egg), and the gesture starts with both hands above the child’s head at the beginning of the nominal phrase introducing the new character (i.e., the egg). It is clear that the child knows the Italian word “uovo” (egg), but he chooses to accompany it with an iconic gesture; thus, he traces an oval with both his hands (up-down oriented) to reinforce his linguistic message. Moreover, the iconic gesture is greatly emphasized: the woodpecker’s egg, that is very small, is reproduced as a gigantic egg. One should also notice the contrast between the speech and the gesture: the child explicitly says that the egg is small, but his gesture represents something extremely big. Therefore, the gesture here should be interpreted not as a mere visual reproduction of the egg, but as a reinforcing strategy to draw the attention of the interlocutor to that specific element of the narration. Pragmatically, we could define it as a focusing strategy: the child wants to highlight the fact that the egg is the real protagonist of his narration, even though it is grammatically introduced on stage as an object in his sentence. So the central role of the egg is emphasized through the use of a wide iconic gesture. This non-verbal information strengthens the simultaneous verbal message.

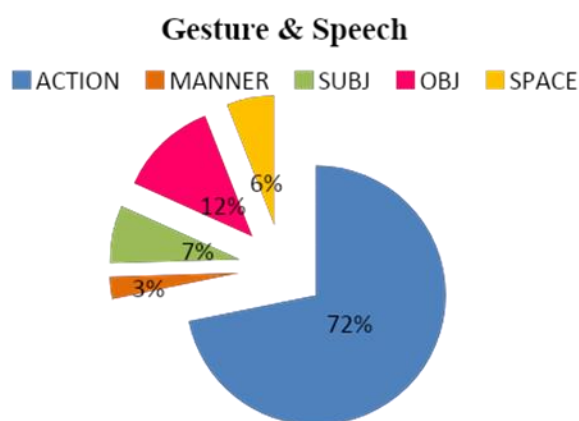


**Figure 5:** Example of iconic gesture representing the object “egg” during L2 narration.

Figure 5 again represents the moment when the egg was firstly introduced in the narration in L2. Again, the child performs an iconic gesture with both hands moving to create an oval shape right in front of his chest. The correspondent transcription extract was as follows, with numbers between brackets indicating pauses: “Ok eh: ok eh: (1.84) / The uovo eh: (4.63) / I get up”. The gestures start when the child says the code-mixed phrase “The uovo”, and also holds the gesture through the long pause (almost 5”) that follows, before releasing it when it switches to the action made by the egg. The example quoted shows three important elements: the abundance of full pauses (eh:); the presence of long silences between one utterance and the other (1.84/4.63); the lack of English lexicon, that is replaced with Italian words (e.g., “uovo”) with the English article “the” (peculiar code-mixing strategy). As in Fig. 4, the same child uses an iconic gesture (Fig. 5) to represent the same object (the egg). However, in this case it is not a reinforcing strategy, but a problem-solving strategy. Indeed, the child cannot remember the English word “egg” and, after a long hesitation, he performs a code-mixing, accompanied by the iconic gesture, to successfully convey his message to the interlocutor. The realization of the gesture, however, differs from the one made with the same referent in the Italian narration; while talking (or trying to talk) in English, the child makes a sort of sphere with both his hands in front of his chest, thus iconically creating a proper small egg. It is worth noting that this time the adjective ‘small’, or an Italian equivalent, does not appear in the narration: this information is, therefore, presented only in gesture.

As has been said before, a lexical test was conducted before the second day of recordings, in order to verify the children’s knowledge of the English lexicon associated with the cartoon. Whilst the results of the test were positive for all the class, at the time of recording the narration in L2, almost all of the children had difficulties in retrieving the adequate lexicon. This proves that offline knowledge does not always correspond to online knowledge.

As for gestural typology in relation to the referent, as annotated in the fourth tier, the pie chart in Fig. 6 sums up the main results. It appears that the majority of gestures visually recreated an action (72% of the total). The remaining gestures were used to reproduced inanimate referents (12%), animate referents (7%), virtual spaces (6%), and in a few cases (3%) to reproduce the emotions or behaviors of the characters in the story.



**Figure 6:** Graphical representation of the relationship between gesture and referent typology calculated on total data sample.

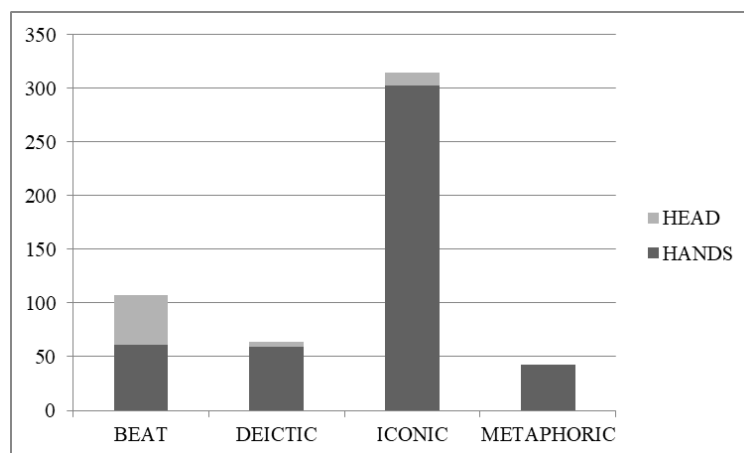
These overall percentages also remained unvaried with respect to the language of the narration and to pupils’ sex. In L2 productions, we recorded an increase of gestures representing animate or inanimate referents, with percentages of 10% and 17%, respectively; in L1 narrations, the same

gestures represented only 5% (animate referents) and 9% (inanimate referents). This could be interpreted as supporting our previous claim that, in an L2 context, gestures were more often used to supply the lack of English lexicon.

To sum up, the qualitative analysis shows a tendency, valid for both male and female, to use gestures differently in L1 and L2, but with an abundance of gestures to denote an action rather than objects. However, iconic gestures denoting animate or inanimate referents increased in L2, probably due to lexical gaps. A difference was also observed between male and female pupils: while girls tended to produce less gestures, and to focus more on a linear syntactic production, boys produced more gestures, and used all the space in front of them for gesturing, but they showed a less elaborate construction of the narration.

## 5.2 Quantitative analysis

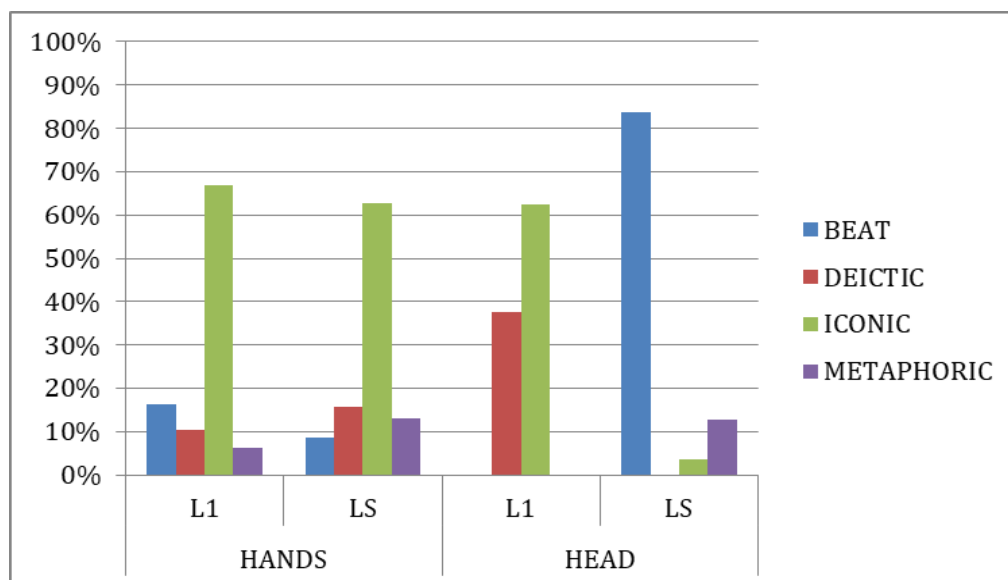
A first descriptive analysis was conducted to observe the distribution of gesture categories within the total sample: Fig. 7 presents the frequencies of the gestures produced with the hands and the head, as divided into the four categories introduced by McNeill (1992). In the following figures, we present the distribution of gestures in percentages, since the absolute numbers of gestures show great quantitative variation between L1 and L2; and without a proportional perspective, the interpretation could be misleading. Furthermore, there are also other gestures in our corpus that are not represented in Fig. 7: these gestures consist in 50 occurrences of total body movement, 2 occurrences of eyes widening, 3 of enhancement of the eyebrows, and one case of a gesture produced with the mouth. These gestures were impossible to classify according to McNeill's categories, which were primarily set for hand gestures, and, therefore, were not further analyzed in this study.



**Figure 7:** Number of occurrences of the four gesture typologies (McNeill 1992) produced with the hands or the head.

Later, we investigated the distribution of gesture typologies within the L1/L2 subset, by focusing primarily on gestures performed with the head and the hands. Of a total of 529 gestures, 466 were performed with the hands, and only 63 through head movement. Among these 63 head gestures, 55 were produced during L2 narrations, and only 8 occurred in L1 narrations. Conversely, the hand gestures were more frequent in L1 (268 occurrences) than in L2 (198 occurrences). As shown in Fig. 8, for the hand gestures, in L2 is possible to notice an increase of deictic (from 10.4% in L1 to 15.7% in L2) and metaphoric gestures (from 6.3% in L1 to 13.1% in L2). Conversely, the percentage of iconic gestures slightly decreases from 66.8% in L1 to 62.6% in L2. As for the beats, it is possible to

notice a difference between languages in relation to the organ involved in the gesturing: indeed, beats produced with the head appear only in L2, whereas the beats produced with the hands vary across language, from 16.4% in Italian narrations to 8.6% in English ones. The correlation between hand gestures and language results are statistically significant for  $\chi^2(4) = 13,766$ ,  $p=0,003$ ; Cramer’s  $V=0,172$ .

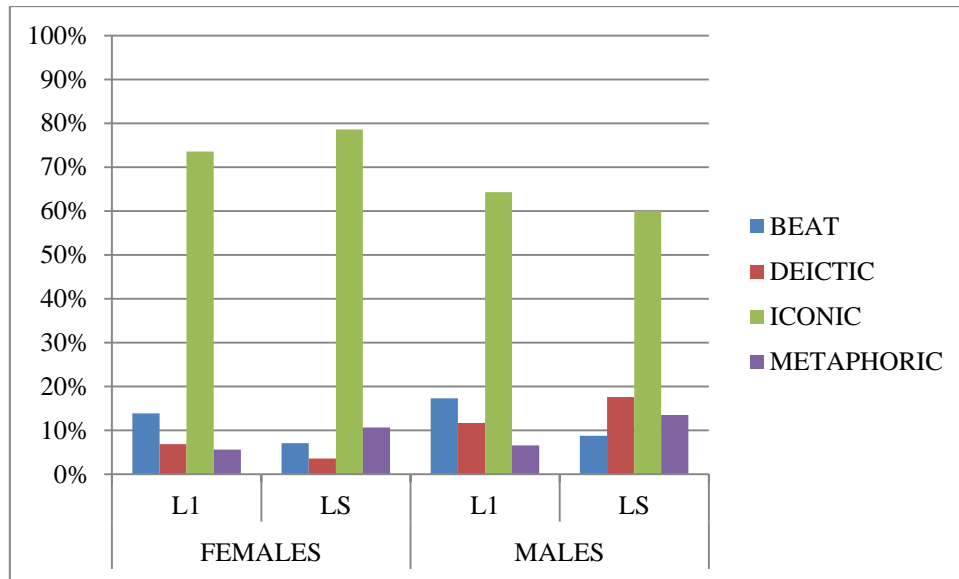


**Figure 8:** Percentages of gestures produced with the head and the hands divided into the four McNeill category, and according to production in L1 or L2.

In general, we can say that gestures realized with the head are more frequent in L2 than in L1, and that they fall especially within the beats category. Indeed, speakers realized only 8 gestures with the head in L1 (5 iconic, 3 deictic gestures), and 55 in L2, of which 46 fall into the beat category, 7 were iconic gestures, and the remaining 2 were deictic head gestures. This could be interpreted as an interactive strategy, designed to maintain the conversation turn, and also the sequential order of the narration. This could strengthen what we noticed in the qualitative analysis: children’s difficulty during their English narration is reflected in their non-verbal language, with an abundant presence of both metaphoric gestures and beats produced with the head.

Focusing on the category of hand gestures, as we can see in Fig. 8, they were produced more in L1 than in L2 (268 vs. 198 hand gestures). The percentage of iconic gestures was similar in both languages (66.8% in Italian, 62.6% in English), while the other categories varied according to the language involved. In particular, in L2 productions, metaphoric and deictic gestures increased, with percentages moving from 6.3% in Italian to 13.1% in English for metaphoric gestures, and from 10.4% in Italian to 15.7% in English. Conversely, beats decreased from 16.4% of occurrences in Italian to 8.6% in English; although it is worth remembering that in English beats gestures occurred often as produced with the head.

During the qualitative analysis, a clear behavioral difference emerged in both the narration and the gestures, as produced by male and female pupils. After restricting the analysis to the hand gestures, the most numerous in our corpus, it was confirmed that males performed 366 gestures, while females showed only 100 gestures. Fig. 9 presents data for the two subsets of males and females divided by hand gesture typology. It should be noted that only the male subset was statistically significant ( $\chi^2(3)=11,809$ ,  $p=0,008$ ; Cramer’s  $V=0,18$ ), while no significance was found for the female subset.



**Figure 9:** Percentages of hand gestures used by female and male pupils divided for gesture typology and language used.

Indeed, the female subset does not present huge differences between the productions in L1 and L2. Conversely, in the male subset it is possible to notice a prevalence of iconic gestures in both languages (64.3% in Italian, 60.0% in English), whereas the presence of deictic and metaphoric gestures increased in L2 with respect to L1. Conversely, beat gestures produced with the hands reduced in L2 (8.8% vs. 17.3% in L1).

In conclusion, the quantitative analysis confirms what has emerged in the qualitative investigation of the data. Data emphasize a difference between gestures used in L1 and L2, not only with respect to gestural typology but also as it relates to the organ involved. Moreover, a difference between male and female pupils has been confirmed.

## 6 Discussion

Our data show a general trend with a distinction between both the languages of the task and the pupil's sex. On the one hand, boys usually construct their narration/conversation using gestures and other communication strategies, so as to reinforce their linguistic message. On the other hand, girls produce a more limited number of gestures and elaborate their linguistic production more. Some researchers have also observed this tendency since the first phases of communicative development in little girls (see Özçalışkan & Goldin-Meadow, 2010). Those general tendencies can be more or less emphasized in the different participants, based on the psychological and behavioral profile of the subject, and according to the task typology. Indeed, we observed a general decrease of gesture production during L2 recordings: the difficulty of the task caused a higher degree of inhibition during the activity, in particular during the dialogical phase with the researcher.

Gestures also appeared to perform different interactional and pragmatic roles depending on the language of the task (i.e., L1 Italian vs. L2 English). In L1 narrations, gestures are especially used as a reinforcing strategy or in substituting the speech. Conversely, in L2 productions, gestures are mainly used as a problem-solving strategy, following a pattern also observed by other research (e.g., Alibali et al., 2011).

We focused on the semantic-lexical dimensions as connected to the use of gestures. A preponderance of gestures as reproduction of actions, and as representation of inanimate referents, has

emerged. Gesture-action correlation can be easily motivated by the intrinsic dynamic nature of actions; this dynamicity is mirrored in the non-verbal communication, too. However, since children often used gestures to substitute their verbal production, it has been noted how this happened more frequently in the case of actions. This could mean that gestures are probably perceived by children as a more direct instrument of communication: therefore, actions are often reproduced with gestures without linguistic support. Moreover, the use of gestures for inanimate referents representation can be explained with the specific cartoon typology we showed to the children: indeed, the main character of the story is an egg (that later opened and gave birth to a woodpecker’s chick, who continued the story) and its movements cover the majority of the cartoon episode. Therefore, children probably remained so impressed by this character that he took on great importance in their narrations.

As for the organ involved in the gesture activity, we observed that, as was expected, the majority of gestures were produced with the hands, but with a difference based on the pupil’s sex: indeed, boys tended to produce gestures with both hands, whereas girls used only one hand and their gestures were more reduced in amplitude. However, gestures with various parts of the body also occurred in our corpus, as is summarized in Table 4.

**Table 4:** Frequency of gesture typologies within the total sample.

	<b>Gesture typology</b>	<b>Frequency</b>
1°	ICONIC (hands)	303
2°	BEAT (hands)	61
3°	DEICTICS (hands)	59
4°	MOTION (many parts of the body activated)	50
5°	BEAT_HEAD	46
6°	METAPHORIC (hands)	43
7°	ICO_HEAD	12
8°	DEI_HEAD	5
9°	SOP_UP (eyebrows)	3
10°	EYES_OPEN (gaze)	2
11°	MOUTH	1

In the first part of the hierarchy, we observed those gestures realized with hands; the motion category was also quite consistent. Since this last typology is used to represent characters’ movements, the elevated number of occurrences supports what we previously said about the tendency to use gestures to represent the intrinsic dynamicity of actions.

Then, a comment on the beat category has to be made, since it also appears in the first part of the hierarchy: beats realized with the head were abundant in L2 productions (see Fig. 7), while those

realized with hands were more numerous in L1 productions (see Fig. 7). One possible interpretation can be found in the definition of the beat category itself, as it has been firstly provided by McNeill (1992): indeed, beats are a complex category that serve multiple functions, even though the principal role is giving rhythm to the narration. In this case, in their use in the English L2 narrations, beats have the pragmatic function of maintaining the conversational turn whereas, during L1 Italian narrations, beats are used to structure the discourse in terms of syntax and the order of the events. A specific beat role seems to be linked to the activation of a specific body part: the pragmatic function, thus, seems to be associated with head gestures, whereas discourse structure is performed through hands' beats. Finally, the statistical analysis confirms the qualitative analysis, although the strength of this correlation was not very robust, probably because of the relatively small sample collected for this study.

## **7 Conclusions and further perspectives**

This article fits into the scientific fieldwork of Gesture Studies and presents the investigation of speech-gesture relation within a fourth-grade class of 15 pupils. The study focuses on the types of speech-gesture relation in narrations and, partially in dialogical samples, with a primary focus on the difference between L1 and L2 narrations. Furthermore, a difference on the quantity and quality of the gestures has emerged in relation to the pupil's sex.

The qualitative and quantitative analysis presented here allows the answering of our research questions, and partially contradict our previous expectations. Indeed, gestures varied both quantitatively and qualitatively in narration in L2 vs. L1: in L2, there were fewer gestures, in particular iconic ones, and they were produced with less amplitude. Moreover, beat gestures were realized more frequently with the hands in L1, and with the head in L2. This relates to our second research question, as it has emerged that the pragmatic and discursive functions of beats also differ according to language. In particular, in L1, gestures co-occur with speech as part of a reinforcing strategy, or they substitute speech by integrating more details of the narrations. Conversely, due to a lack of active knowledge of English, during L2 narrations gestures assume the pragmatic values of turn-holdings (head-beats), and as part of a problem-solving strategy, verbally signaled by reformulations and long pauses. As for the third research question, there were some differences between the narrative task and the guided (dialogical) conversation with the researcher. However, data was too limited, in particular for L2 dialogues, to allow any further generalizations.

A straightforward difference also emerged with respect to pupils' sex. In particular, girls gesticulate less than boys do, and their gestures have a minor amplitude, both in L1 and in L2. This was true in particular for hand gestures, whereas girls produced many beats with the head also in L2 narrations. This inversely correlates on the linguistic level, with more articulate narrations, in terms of both lexical variety and syntactic structure, produced by girls compared to their male peers.

Although intriguing, these results are obviously limited by the small sample chosen (15 pupils) and by the methodology adopted in the two tasks. In particular, it has been noted how the guided conversation with an adult researcher could have inhibited many pupils, thus resulting in very short dialogues with few or no gestures. It will be necessary, for further research, to widen the sample available, in order to enrich our observations on linguistic and gestural dynamics. Moreover, further studies may consider language-gesture development in L2 in school-children diachronically, by performing the same structured task once a year during different grades, similarly to what has been done by Coletta et al. (2016) in French and by Nicoladis et al. (1999, 2009) with French-English early bilinguals.

A final point we want to highlight, as possible future research, is the study of gesture as a didactic tool. As previously noted by McCafferty (2004), the results presented here also highlight how gestures could play a positive role in improving children's speech and performance in particular in a



foreign language. In other words, gesture could be studied in terms of language teaching and language learning, from both the teacher’s and the student’s point of view.

## REFERENCES

1. Alibali MW, Goldin-Meadow S. *Gesture-Speech Mismatch and Mechanisms of Learning: What the Hands Reveal about a Child’s State of Mind*. *Cognitive Psychology*, 25: 468–523, 1993.
2. Alibali MW, Flevares LM, Goldin-Meadow S. *Assessing Knowledge Conveyed in Gesture: Do Teachers Have the Upper Hand?* *Journal of Educational Psychology*, 89(1): 183–193, 1997.
3. Alibali MW, Spencer RC, Knox L, Kita S. *Spontaneous gestures influence strategy choices in problem solving*. *Psychological Science* 22(9): 1138–1144, 2011.
4. Broaders SC, Cook S, Wagner MZ, Goldin-Meadow S. *Making children gesture brings out implicit knowledge and leads to learning*. *Journal of Experimental Psychology: General*, 136: 539–550, 2007.
5. Bates E, Camaioni L, Volterra V. *The acquisition of performatives prior to speech*. *Merrill-Palmer Quarterly of Behavior and Development* 21(3): 205–226, 1975.
6. Bello A, Capirci O, Volterra V. *Lexical production in children with Williams syndrome: Spontaneous use of gesture in a naming task*. *Neuropsychologia*, 42: 201–213, 2004.
7. Broaders S C, Wagner Cook S, Mitchell Z, Goldin-Meadow S. *Making Children Gesture Brings Out Implicit Knowledge and Leads to Learning*. *Journal of Experimental Psychology: General*, 136(4): 539–550, 2007.
8. Capirci O, Contaldo A, Caselli M C, Volterra V. *From action to language through gesture: A longitudinal perspective*. *Gesture* 5(1-2): 155–177, 2005.
9. Capussotti G. *La multimodalità in italiano e inglese L2: relazione tra gesto e discorso nei bambini di una scuola primaria*. Unpublished Master’s Degree Thesis, University of Pavia, Italy, 2019.
10. Church RB, Goldin-Meadow S. *The mismatch between gesture and speech as an index of transitional knowledge*. *Cognition*, 23: 43–71, 1986.
11. Church RB, Ayman-Nolley S, Mahootian S. *The role of gesture in bilingual education: Does gesture enhance learning?* *Bilingual Education and Bilingualism*, 7: 303–319, 2004.
12. Colletta JM. *Comparative analysis of children’s narratives at different ages: a multimodal approach*. *Gesture*, 9(1): 61–97, 2009.
13. Colletta JM, Pellenq C, Guidetti M. *Age-related changes in co-speech gesture and narrative: Evidence from French children and adults*. *Speech Communication*, 52(6): 565–575, 2016.
14. Coates J. *Women, men and language: A sociolinguistic account of gender differences in language*. London: Routledge, 2015.
15. Ekman P, Friesen W V, Hager J C. *Facial Action Coding System*. Salt Lake: Research Nexus, 2002.
16. Garber P, Alibali M W, Goldin-Meadow S. *Knowledge Conveyed in Gesture Is Not Tied to the Hands*. *Child Development*, 69(1): 75–84, 1998.
17. Garber P, Goldin-Meadow S. *Gesture offers insight into problem-solving in adults and children*. *Cognitive Science* 26: 817–831, 2002.
18. Grosjean F. *Neurolinguists, beware! The bilingual is not two monolinguals in one person*. *Brain and Language*, 36: 3–15, 1989.
19. Grosjean F, Li P. *The Psycholinguistics of Bilingualism*. Oxford: Wiley-Blackwell, 2013.
20. Kendon A. *Gesture and speech: two aspects of the process of utterance*. In Mary R. Key (eds.), *Nonverbal Communication and Language*, The Hague: Mouton, 207–227, 1980.
21. Labov W. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania, 1972.
22. Legerstee M, Fisher T. *Coordinated attention, declarative and imperative pointing in infants with and without Down syndrome: Sharing experiences with adults and peers*. *First Language*, 28: 281–312, 2008.
23. Macedonia M, von Kriegstein K. *Gestures enhance foreign language learning*. *Biolinguistics*, 6(3-4): 393–416, 2012.

24. McCafferty S G. *Space for cognition: gesture and second language learning*. International Journal of Applied Linguistics, 14(1): 148–165, 2004.
25. McNeill D. *So You Think Gestures are Nonverbal?* Psychological Review 92(3), 350–371, 1985.
26. McNeill D. *Psycholinguistics: A new approach*. New York: Harper & Row, 1987.
27. McNeill D. *Hand and Mind. What Gestures Reveal About Thought*. Chicago-London: University of Chicago Press, 1992.
28. Nicoladis E, Mayberry RI, Genesee F. *Gesture and early bilingual development*. Developmental Psychology, 35(2): 514–526, 1999.
29. Nicoladis E, Pika S, Marentette P. *Do French-English bilingual children gesture more than monolingual children?* Journal of Psycholinguistic Research, 38: 573–585, 2009.
30. Özçalışkan S, Goldin-Meadow S. *Sex differences in language first appear in gesture*. Developmental Science 13: 752–760, 2010.
31. Ping R, Goldin-Meadow S. *Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity*. Developmental Psychology, 44(5): 1277–1287, 2008.
32. Pika S, Nicoladi E, Marentette P. *A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer?* Bilingualism, 9: 319–327, 2006.
33. Tellier M. *The effect of gestures on second language memorisation by young children*. Gesture, 8: 219–235, 2008.
34. Wagner Cook S, Goldin-Meadow S. *The Role of Gesture in Learning: Do Children Use Their Hands to Change Their Minds?* Journal Of Cognition And Development, 7(2): 211–232, 2006.
35. Wakefield E M, James K H. *Effects of learning with gesture on children's understanding of a new language concept*. Developmental Psychology, 5: 1105–1114, 2015.
36. Wittenburg P, Brugman H, Russel A, Klassmann A, Sloetjes H. *ELAN: a Professional Framework for Multimodality Research*. In: Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation, 2006.