

# DEVELOPMENTAL CHANGE OF DISCOURSE COHESION IN SPEECH AND GESTURES AMONG JAPANESE ELEMENTARY SCHOOL CHILDREN

KAZUKI SEKINE\* · NOBUHIRO FURUYAMA\*\*

*Japan Society for the Promotion of Science, Japan · National Institution of Informatics, Japan*

**ABSTRACT:** *This study investigates the development of bi-modal reference maintenance by focusing on how Japanese elementary school children introduce and track animate referents in their narratives. Sixty elementary school children participated in this study, 10 from each school year (from 7 to 12 years of age). They were instructed to remember a cartoon and retell the story to their parents. We found that although there were no differences in the speech indices among the different ages, the average scores for the gesture indices of the 12-year-olds were higher than those of the other age groups. In particular, the amount of referential gestures radically increased at 12, and these children tended to use referential gestures not only for tracking referents but also for introducing characters. These results indicate that the ability to maintain a reference to create coherent narratives increases at about age 12.*

**KEY WORDS:** *Spontaneous gesture - Speech - Discourse cohesion - Elementary school children - Japan.*

## INTRODUCTION

**A**DULTS often produce gestures accompanying their speech when narrating a story. We already know that young children try to express themselves by using gestures and speech but we do not know much about how they use gestures to narrate a story or an event, or about how the role of gestures changes during their elementary school years. In this study, we tackle these issues by focusing on the relationship between spontaneous gestures and spoken referential expressions in discourse. For the purposes of this study, discourse is defined as a language structure in which several utterance strings are conjoined across the sentence boundary.

As children develop the ability to combine more than two words and produce longer sentences, they begin to make narratives. Many studies have revealed that children begin to recount past events with support of their parents at around 2 years old (Nelson 1989). Nelson observed that a child begins to talk about past experiences and activities they routinely have from their second year of life. However, it seems that temporally or causally combining past events is still difficult for children as young as 3 years old (McCabe & Peterson 1985). In an extensive cross-linguistic study, Berman and Slobin (1994) showed that about 40% of English

- Latest version received June 2010.

\* Corresponding author: Kazuki Sekine, Japan Society for the Promotion of Science/ National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan; sekine@nii.ac.jp.

\*\* Nobuhiro Furuyama, National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan; furuyama@nii.ac.jp.

speaking 3-year-olds do not tend to causally link clauses, but about 70% of 5-year-olds can tie clauses together using conjunctions, such as *and* and *then*. These developmental trends also hold true for non-English speakers, such as Spanish, Turkish, Hebrew, and Japanese (Berman & Slobin 1994; Küntay & Nakamura 2004). For instance, Uchida (1996) reported that by the age of 5 or 6, Japanese children can tell a story based on their daily experiences, fiction or fantasy, and consider the ending of the story in advance. These studies indicate that children acquire the ability to relate and narrate past events during their preschool years. However, the ability to link events into a narrative continues to develop beyond the preschool years (Berman 2004; Bamberg 1987). As Berman's (2004) study showed, even children aged 9 to 10-years-old produced narratives that differed markedly from those of adults in terms of content, morpho-syntax, and lexicon. Thus, we can see that becoming a proficient story teller takes a significant amount of time.

In order to produce a coherent narrative, children need to learn how to tie events to the larger context of a story with coherence. Coherence refers to the way in which the content or the topic of the narrative is held together. Since a narrative has intricately interwoven layers of new and old information, it is crucial for the identities of the referents and topics to be made clear to ensure that the narrative is intelligible. Previous studies have shown that adult speakers create coherence in discourse and identify referents by using a range of linguistic devices. These include 'referential expressions' such as personal pronouns and demonstratives (e.g.: *Peter* said *he* would take *Jessica* to *her* house), 'substitutions' (e.g.: Peter bought a big car. Jessica bought *one* too.), 'ellipses' or 'zero-marking' (e.g.: Peter has a test next month and *ø* is studying a lot), and 'connectives' (e.g.: 'and', 'but', and 'because') (Givón 1983; Haliday & Hasan 1976).

Jisa (2000), who studied the narratives of French children, found that the proportion of clauses devoted to maintaining the subjects across successive clauses increases with age. She also found that although the preferred referential expression used in the maintained clauses for all children was the use of subject pronouns, 7- and 10-year-olds also showed a higher mean use of zero-marking and nonfinite subordination as cohesive expressions than 5-year-olds. These features contribute to a tighter packaging of events through the establishment of a dependency relationship between the two clauses (Berman & Slobin 1994; Jisa 2000).

By asking French and English speakers between the ages of 4 to 9 years to tell the story from a booklet consisting of six pictures, Karmiloff-Smith (1985) investigated the development of cohesive devices. She revealed that by approximately 9 years old, most children can flexibly use referential devices anaphorically. For example, they tended to use demonstratives in the subject slot for the main character rather than only full NPs, and use full NPs in the subject slot for the subsidiary characters. These strategies served to identify the characters and the topic (dis)continuity as the story unfolded. Like these studies, the literature on the development of referential expression, which is mostly in English and French, has so far demonstrated a relatively protracted development of the appropriate use of referential expression beyond the preschool years (Bamberg 1987; Dent, 1984; Hickmann & Hendriks 1999; Levy 2003; Wigglesworth 1990).

Most of this research has focused on language from the Indo-European family because of "the saliency of articles as a distinct grammatical element offering

a formal system of differentiating indefinite references from definite references” (Küntay 2002: p.79). However, there are only a few studies focusing on the development of the marking referent in languages without an obligatory article system, such as Japanese.

For example, Clancy (1992) focused on Japanese discourse, which lacks an article system and in which the use of zero anaphora is extremely common. She investigated the referential expression used in the narrative of children aged from 4 to 7, and that in adults. The study found that NPs were used for introduced and re-introduced referents, and zero-marking was used for the maintained referent for all age groups. As for the age difference, compared to the other age groups, 4-year-olds tended to use zero-marking even when they introduced and re-introduced referents, but there are no significant age differences in the marking of the maintained referents. From these results, she suggested that the ellipsis is the default form of the referential expression for young Japanese children.

Considering that the use of zero anaphora is common in Japanese discourse, to fully understand the development of Japanese children’s discourse, it is important to examine not only the linguistic devices but also the other devices that are used to make referents clearer. Previous studies have shown that narratives appear through multimodal channels, which link words, voice, and body movement (Colletta 2009; O’Neill & Holmes 2002). Although some researchers (Karmiloff-Smith 1981; Wigglesworth 1997) noted the importance of the gestures used in children’s discourse, gestures have not been systematically investigated and little is known about the relationship between paralinguistic expressions and discourse in elementary school children.

Studies on gestures have so far revealed that during a narrative, an adult speaker creates coherent discourse not only by using linguistic devices but also by using the *idiosyncratic gestures* (hereafter referred to as ‘gestures’) that co-occur with speech (Gullberg 2006; McNeill 1992, 2005; McNeill & Levy 1993; Yoshioka 2005). Gestures that are performed spontaneously during speech have no standard of well-formedness like sign language, but are created on the fly (McNeill 1992). When introducing characters, themes, or contrastive concepts in a narrative, adult speakers assign them to distinct areas in the space in front of them by using localizing gestures, typically a pointing gesture or an iconic gesture depicting the shape, action, or condition of the referent. The gesture space for the referents is arbitrary and bears no relationship to the actual space if the referents are not present in the discourse context, as seen in the discourse of a signer in Sign Language (Bellugi & Klima 1982). Once such a space is established, it is often maintained throughout the discourse, where speakers refer back to the assigned space by producing gestures within it. Based on adults’ retellings of a cartoon or film, McNeill & Levy (1993) found that the assigned spaces for referents were gestured more frequently when characters were re-introduced than when the narrative maintained focus on one character.

This gestural reference-tracking is attained through the systematic and repeated associations between specific gestural behaviors, which have a recurrence of one or more features, and referential expression in speech (Gullberg 2006). These recurring associations are called a ‘catchment’ (McNeill 2005), and have features that include space, handedness, movement, and orientation. This suggests a common

discourse theme and allows for the establishment of explicit, visual co-reference, and thus, enhances the cohesiveness of the discourse (Gullberg 2006). In Japanese narratives, a topic shift sometimes occurs without any explicit formal means (e.g., full NPs) even when there is more than one possible referent because subject NPs are often elided. However, Furuyama (2001) demonstrated that a reference is often unambiguous because of the recurring gesture features in the accompanying catchments. The current study was designed to understand more about how this unambiguous gesture referencing develops in Japanese children.

Gesture studies on reference maintenance (Gullberg 2006; McNeill 2005; Yoshio-ka 2005) have revealed that reference-tracking is a bimodal phenomenon, and that the two modalities, gesture and speech, are influenced by the grammatical development. However, what is not yet known is the developmental change in children. Few empirical studies on the development of discourse have focused on the use of gestures that work along with speech for maintaining a referent, although McNeill (1992) has argued that gesture and speech develop in parallel in children.

O' Neill & Holmes (2002) examined the ability of English speakers from 3- to 4-years old to re-introduce the main characters in a narrative. They asked them to narrate a picture-book and observed the use of nominal forms and pointing gestures. They found that approximately 50% of re-introductions in narratives could be clearly identified when both the linguistic and paralinguistic factors were taken into account. This proportion is higher than that seen in previous studies, which have not included gestures. However, the study by O' Neill and Holmes was quite limited in that it took into consideration only preschoolers and their pointing gestures, which were directed to a picture-book. We need to know how gestural reference maintenance (including iconic and pointing gestures used anaphorically) changes past the preschool age, along with the ability to produce a coherent narrative.

Moreover, given the fact that adult speakers perform a chain of gestures in the same place to convey continuities and previously established thematic roles (McNeill & Levy 1993), it is also important to examine the change of space where gestures are produced during the elementary school period. Gullberg (2006) reported that referential gestures, which accompany spoken referential expressions from adult speakers are spatially better defined and more differentiated when those gestures can be seen by the addressee than when they cannot. If children consider their gestures as an informational resource used by the addressee, they may produce gestures in a space more easily seen by the addressee, rather than in the peripheral space. Moreover, since young children tend to act out a movement as if they were the character, sometimes by using their whole body (McNeill 1992), we also need to take into account the gesture viewpoint as well as the gesture space.

McNeill (1992) found that a gesture reflects the viewpoint that a speaker takes for the story (s)he is telling, and that younger children tend to produce gestures from the characters' viewpoint. Character viewpoint gestures depict a character's movement by using the speaker's whole body in a larger space, rather than the observer viewpoint in which the speaker depicts the event from outside the scene. Thus, as long as children produce gestures from a character viewpoint, it seems to be difficult for them to produce referential gestures that use the speaker's frontal space in the same way as adults do, because performing character viewpoint gestures tends to require a larger space and more body movement. Conversely,

if children produce a gesture having captured an event from the viewpoint of an observer, it is easier for them to use the frontal space for marking the existence of the referents and maintaining the reference to them.

From these gesture characteristics, we speculated that as children get older, they produce gestures not only to depict the characters' movement in the story, but also to introduce and maintain the referents as topics. At the same time, because gestures are used to show the listener the maintenance of the referents, as children get older, their gestures are produced from the observer point of view in a higher space than the character viewpoint gestures. These gestural performances would change along with development of their linguistic devices. With this prediction, we investigated the change in the use of gestures in the discourse of Japanese elementary school children, especially focusing on: 1) how often children use gestures in a narrative, and 2) how the gestures and linguistic devices for introducing and maintaining animate referents change during the 6 years of elementary school. To demonstrate this, we adopted the same gestural and speech indices used in Yoshio-ka's (2005) study on the narratives of Japanese adults for the ease of comparison.

## METHOD

### *Participants*

Sixty elementary school children and their parents participated in this study. There were 10 children in each grade, from the 1st through the 6th, half boys and half girls. The children's mean age and their age ranges are listed in TABLE 1. In this study, we refer to the different groups of the children by age, rather than by grade, based on the average age of each grade.<sup>1</sup> All the participants were native monolingual Japanese speakers from middle-class families and attending public or private elementary schools in Tokyo, Japan.

TABLE 1. Participants' mean Age and Age Range in Each Age Group.

Age group	Grade	Mean age (year; month)	Age range
7	1	6;11	6;05 - 7;04
8	2	7;11	7;06 - 8;10
9	3	9;04	9;00 - 9;06
10	4	10;00	9;07- 10;05
11	5	11;04	10;09 - 11;11
12	6	12;00	11;05 - 12;10

### *Materials and Apparatus*

Each child watched a 7-minute animated color cartoon of the Tweety and Sylvester series, entitled 'Canary Row' (Warner Brothers, Inc.). This cartoon consists of eight scenes, and in each one, Sylvester the Cat attempts to catch Tweety Bird in a different way. This cartoon was displayed on a 14-inch color computer monitor

<sup>1</sup> Because these data were collected over the course of a year, some children are older or younger than the average age of the next grade.

(Panasonic CF-F8). A mini-DV camcorder (Sony HDR-HC9) was used to record the children's gestures and speech.

### *Procedure*

Our experiments were conducted in a quiet room in the participant's home or a local community center. In some previous studies, the adult participants watched all eight episodes all at once. However, in this study, in consideration of the reduced memory capacity of children, we divided the viewing session into two segments, each with four scenes. At the beginning of the experiment, the children and their parents were told by an experimenter that the purpose of this research was to study a person's memory. The child would have to remember the cartoon story shown on the computer monitor, and retell it to the listener (parent) in as much detail as possible. Then, the child watched the entire cartoon once. Next, the child was instructed to watch the first half of the same cartoon twice in order to more clearly remember the scenes, after which they told the listener what they had seen. While retelling the cartoon, the child was seated in a location different from the one in which (s)he viewed it. The listener was not allowed to ask any questions or provide any prompts such as 'Then?' or 'Next?' unless the child appeared stalled or distracted. They were allowed to respond to the child by head nodding and using back channels during the child's narration. Immediately after the first half of the episode was retold, the child repeated the procedure for the second half of the story. Because a previous study had demonstrated that the seating position can influence the production of gestures (Özyürek 2002), the seating position was counterbalanced in such a way that 3 of the 5 children in each gender/age group sat on a chair to the left of the listener and the other two sat to the right of the listener. The whole session, including the instructions, was recorded using a mini-DV camcorder on a tripod. The narratives of the first and second halves of the story the child told were combined and analyzed.

### *Coding*

#### Speech data

The first author transcribed all the narratives *verbatim*. From the transcriptions, the *mean length of total speaking time* (time spent on the narrative) and the *mean number of clauses* were calculated. A clause was loosely defined as the combination of a noun phrase and a verb phrase. In the following coding clauses, the analytical framework used in Yoshioka's (2005) study was adopted. For each clause, the animate referent(s) was coded according to their identity (e.g. 'Tweety', 'Sylvester', etc.), the linguistic referential expressions used to refer to them (full NP, pronoun, or zero-marking), and their discourse status (e.g. *introduced*, *maintained*, or *re-introduced*). For *maintained* and *re-introduced* referents, only referents in the subject position were analyzed.<sup>1</sup> An *introduced* referent was defined as the first mention of

<sup>1</sup> Restricting the analysis of the referents only to those in subject position is a conscious choice for the purpose of comparing the results of the present study to previous research findings (Gullberg 2006; Yoshioka, 2005). When one of the two referents is mentioned in the immediately succeeding clause, the discourse status of the referent is considered maintained.



the referent in the narrative. A *maintained* referent referred to one that held either the same subject position in the immediately preceding clause, or one that was introduced somewhere in the immediately preceding clause. A referent was coded as a *re-introduced* referent if it was assigned a subject role, had already been introduced prior to the preceding clause, and was thus different from the immediately preceding subject. The following segments are an example of a child's speech. In each segment, the first line is the original Japanese speech, the second line shows the gloss, the third line is the English translation, and the fourth line shows the name of the referent, the type of referential expression, and the discourse status is in brackets. The abbreviations that are used in the interlinear gloss include ACC (accusative), ASP (aspect), DAT (dative), GEN (genitive), INJ (interjection), NEG (negative), NONPAST (non-past), NOM (nominative), PASS (passive suffix), SE (sentence extender), and TOP (topic marker).

- (1) 1 apaato ni neko ga i-te  
 apartment DAT cat NOM exist-and  
 (There is (a) cat in (an) apartment and)  
 <Cat: Full NP, Introduced referent>
- 2 sore wa hantai no apaato ni sumu tori wo mitsukete  
 it TOP opposite GEN apartment DAT live:NONPAST bird ACC find-and  
 (it found a bird living in the apartment opposite and)  
 <Cat: Pronoun, Maintained referent, Bird: Full NP, Introduced referent>
- 3 ø apaato kara dete  
 ø apartment from go-out-and  
 ((it) went out of his apartment and)  
 <Cat: Zero-marking, Maintained referent>
- 4 sorekara sono tori wa heya ni hasittetta  
 then the bird TOP room DAT run-and-go:PAST  
 (then the bird ran into (his) room)  
 <Bird: Full NP, Re-introduced referent>

### Gesture data

First, the spontaneous gestures were identified. Hand movements were classified as gestures only when they had an identifiable beginning and a clear end, and were synchronized with speech. After identifying the gestures, the *total number of gestures* and the *frequency of gestures per second* were calculated. Every child produced gestures. Of the total number of gestures counted, gestures that accompanied spoken referential expressions were coded as *referential gestures*, regardless of the gesture-type classification. Yoshioka (2005) found evidence that when there is an adjective before the nominal, speakers often gesture on the adjective and put their hand in the resting position when the nominals are uttered in speech. However, we excluded these gestures from our analysis, because such gestures were not produced in the present study.

### Gesture space

The gesture space that the children used was examined to investigate where the gestures were produced in the speaker's frontal space. A gesture space is defined as the space assumed to be in front of a speaker where the gestures would be pro-

duced if the speaker was seated. In this study, because the height of the gestures produced in the narrative was the primary focus, the gesture space was divided into the following three sectors based on Kita & Wood's (2006) index: upper, middle, and lower spaces. After coding the heights of the gestures, the production rate of the gestures in each sector was calculated.

### Gesture viewpoint

We categorized the gestures into three types as an index to indicate the relationship between the speaker and the scene in order to examine the viewpoint that children took in a narrative. If a speaker acted out a movement as if he/she had become the character, the gesture was coded as a 'character viewpoint' (CVPT) gesture.<sup>1</sup> If a gesture captured an event from the viewpoint of an observer, it was called an 'observer viewpoint' (OVPT) gesture. If a gesture depicted an event from two different viewpoints simultaneously, it was coded as a 'dual viewpoint' (DVPT) gesture (McNeill 1992).

### Reliability

To ensure the reliability of the speech coding and gesture data, after all the data (except for the total speaking time) were coded by the first author, part of the data was re-analyzed by a trained and independent native Japanese graduate student as a second coder. Four hundreds and sixty clauses (about 10% of the total) were randomly selected and coded by the second coder according to the analytical framework described in the coding 'speech data' section described above. The agreements between the coders were calculated for each index. The two coders agreed more than 94% of the time for all the speech indices including the linguistic referential expressions and the discourse status of the referent for each clause. The clauses that caused disagreement were re-analyzed by both coders until an agreement was reached. As for the gesture data, for practical reasons, 450 gestures (about 12% of the total) were randomly selected for judging by the second coder. The two coders agreed on more than 85% of the narratives for all the gestural indices including the number of gestures (94%), the number of referential gestures (90%), the number of gestures that fell into each gesture space (86%), and the gesture viewpoint (92%). Any coding disagreements were resolved through discussion.

### RESULTS

Our analysis was conducted as follows. First, the length of the narratives and the number of gestures were calculated. Next, the number of clauses, referential gestures, and the relationships between them were investigated. Finally, the spatial characteristics of the gestures, such as the use of space, and gesture viewpoint were examined. Because the speaker's seating position did not affect any of the results, the data were analyzed without regard to the differences in the speaker's position.

<sup>1</sup> When children depict a character's movement, they tend to use not only their hands but also other body parts such as their head or legs, or even their whole body. However, as it is sometimes difficult to judge whether other body parts (excluding hands) are being used as gestures, we focused only on hand movements in this study.



### Mean scores of speech and gestural measures

The mean scores of the speech and gestural measures in the age groups were compared using the Kruskal-Wallis nonparametric ANOVA to assess the significant differences among the groups. The significant results were then subjected to a *post hoc* analysis by using the Mann-Whitney *U* test with a Bonferroni correction for multiple comparisons ( $p < .003$  was considered statistically significant).

A Kruskal-Wallis test was conducted on the total speaking time, on the total number of gestures, and on the frequency of the gestures (Table 2). The results revealed a significant difference only in the gesture frequency between the age groups ( $\chi^2(5, N = 60) = 10.84, p = .05$ ). Although 12-year-olds produced gestures more frequently than 7- to 10-year-olds, a Mann-Whitney *U* test showed no significant differences among them. It was clear that the total speaking time did not significantly change during elementary school, and that the total number and the frequency of gestures increased with age. However, the increase in the total number of gestures was not straightforward and seemed to decrease temporarily among 9- and 10-year-olds.

TABLE 2. *Speech and Gestural Performance: Mean Performance across Speech and Gestural Indices for Each Age Group (SD).*

Measures	7 years	8 years	9 years	10 years	11 years	12 years
Total speaking time (second)	403.89 (203.28)	435.80 (191.96)	390.57 (244.76)	397.81 (91.01)	416.35 (88.03)	353.78 (99.45)
Total number of clauses	62.4 (29.93)	69 (20.42)	77.8 (24.68)	77.3 (15.58)	91.9 (20.90)	78.6 (19.47)
Total number of gestures	47.50 (28.88)	57.10 (43.03)	40.20 (33.22)	38.90 (29.56)	80.90 (52.94)	99.80 (73.72)
Frequency of gestures per second	0.12 (0.06)	0.14 (0.09)	0.10 (0.09)	0.09 (0.06)	0.19 (0.13)	0.27 (0.17)
Total number of referential gestures	0.60 (0.84)	1.60 (2.80)	1.30 (2.54)	1.10 (1.29)	4.30 (5.56)	8.10 (4.93)
Proportion of referential gestures in total number of gestures	0.01 (0.02)	0.02 (0.02)	0.01 (0.03)	0.02 (0.02)	0.05 (0.05)	0.10 (0.07)

### Number of Clauses and Referential Gestures

We conducted a Kruskal-Wallis test on the mean number of clauses (TABLE 2). To normalize the percentages, an angular transformation on the mean proportion of clauses that were used was performed for each discourse status and on the mean proportion of linguistic referential expressions used in each discourse status. After this, Kruskal-Wallis tests were conducted on these data. No significant differences were found in all three indices regarding the clauses. Regardless of the age group, about 7% of clauses are used for introduced referents, 67% of clauses are used for maintained referents, and 26% of clauses are used for re-introduced referents. As for the referential expression, when introducing referents, 12-year-olds used only full NPs in the same way adult speakers do, but the other age groups occasion-

ally used zero-marking. The distributions of the referential expressions used for maintained and re-introduced referents were similar between children and adults. Zero-marking is used in more than 80% of maintained referents and in around 40% of re-introduced referents. About 60% of re-introduced referents used NPs in each age group. Thus, zero-marking is most often used for maintained and re-introduced referents in Japanese discourse. These results indicate that the distribution of the discourse status and the referential expressions used in each clause did not change among the age groups of the elementary school children.

To analyze referential gestures, first the number of children who produced referential gestures at least once in their narrative was counted. Out of the 10 children from each year, referential gestures were observed in four 7-year-olds, five 8-year-olds, three 9-year-olds, six 10-year-olds, seven 11-year-olds, and nine 12-year-olds. This showed that more than a half of the children aged 10 or over produced referential gestures. A Kruskal-Wallis test on the mean number of referential gestures was performed. Then, an angular transformation on the proportion of referential gestures for the total number of gestures was conducted before a Kruskal-Wallis test was performed on these data (TABLE 2). The results indicated that there was a significant difference in both the total number of referential gestures ( $\chi^2(5, N=60) = 19.60, p < .001$ ), and in the proportion of referential gestures ( $\chi^2(5, N=60) = 19.26, p = .002$ ).

A Mann-Whitney *U* test showed that the total number of referential gestures produced by 12-year-olds was significantly greater than that of 7- (Mann-Whitney *U*,  $Z = -3.33, p < .001$ ), 8- (Mann-Whitney *U*,  $Z = -2.96, p = .003$ ), 9- (Mann-Whitney *U*,  $Z = -3.01, p = .003$ ), and 10-year-olds (Mann-Whitney *U*,  $Z = -3.14, p = .002$ ), but not significantly different from that of 11-year-olds who showed a 5% usage. It was also found that the mean proportion of referential gestures used by 12-year-olds (10% of the total number of gestures) was significantly greater than that of 7- (Mann-Whitney *U*,  $Z = -3.17, p = .002$ ), 8- (Mann-Whitney *U*,  $Z = -3.10, p = .002$ ), 9- (Mann-Whitney *U*,  $Z = -3.08, p = .002$ ), and 10-year-olds (Mann-Whitney *U*,  $Z = -3.05, p = .002$ ) (less than 3%).

Next, the proportion of referential gestures used for each discourse status was examined. After an angular transformation, a Kruskal-Wallis test was conducted on the mean proportion of clauses that accompanied referential gestures in the total number of clauses used for each discourse status (FIGURE 1). A significant difference was found in the proportion of introduced referents ( $\chi^2(5, N=60) = 15.94, p = .007$ ) and maintained referents ( $\chi^2(5, N=60) = 23.49, p < .001$ ), all of which were accompanied by referential gestures. The proportions of introduced and re-introduced referents showed a similar tendency indicating a gradual increase through the elementary school years, although the proportion of introduced referents showed a slight U-shaped change. A Mann-Whitney *U* test showed that the proportion of 12-year-olds (23%) was greater than that of 10-year-olds (less than 2%) (Mann-Whitney *U*,  $Z = -2.86, p = .004$ ).

However, referential gestures were never produced in the maintained referents of the 7- to 9-year-olds. The proportion of maintained referents that were accompanied by referential gestures was significantly greater for the 12-year-olds (5%) than for the 7-, 8-, and 9-year-olds (respectively, Mann-Whitney *U*,  $Z = -3.25, p = .001$ ,  $Z = -3.43, p < .001$ ,  $Z = -3.43, p < .001$ ). It was mildly significant in 12-year-

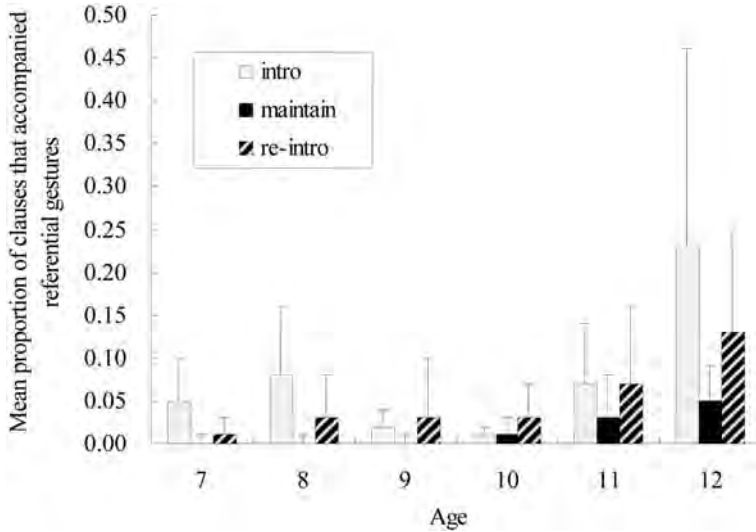


FIG. 1. Mean proportion of clauses that accompanied referential gestures.

olds compared to 10-year-olds (less than 3%) (Mann-Whitney  $U$ ,  $Z = -2.85$ ,  $p = .004$ ). These results indicate that unlike linguistic referential expressions, major changes appear in the gestural referential expressions of children during their elementary school years.

#### *Introduction of Referents in the Narrative*

We also examined whether children produce referential gestures when they first introduce referents into their narrative in speech. There were three main characters in the story; a cat, a canary, and the canary's owner. First, we calculated the number of children who produced referential gestures at least once when they verbally mentioned the characters for the first time. There was one 7-year-old, one 8-year-old, two 9-year-olds, three 10-year-olds, two 11-year-olds, and eight 12-year-olds. There was a significant association between the number of children who produced referential gestures and the age groups (Fisher's exact test,  $p = .005$ ). To find where the significant difference among age groups laid, a residual analysis was used. It indicated that the number of children who produced a referential gesture when introducing a referent at least once was significantly greater in 12-year-olds than for those in the other age groups.

To better understand how 12-year-olds introduce referents in a narrative, the interaction between speech and gestures needs to be investigated in more detail. Let us examine Figure 2 and look at Example 1. Here, the square brackets represent the starting and ending points of the motion of the children's hands, the boldface marks the stroke phrase of the gesture phrase, the underlining indicates a motionless hold phase, an asterisk '\*' represents self-interruption, and a colon ':' in speech indicates an elongated phonation. The numbers used in Figure 2 correspond to the numbers in the transcription in Example 1, which, in turn, indicates the places where gestures occurred.

Figure 2 shows a scene in which a 12-year-old boy is introducing referents into his narrative. First, he introduced the cat by saying ‘a black cat’ while gesturing with his right hand above his lap as if holding an animal and put it in a specific space (1). Then, he mentioned the canary. While saying ‘the yellow chick, bird’, he produced a gesture with the same hand on the left side of the space that he assigned for the cat, as if showing the size of the tiny bird (2). The fact that he set the space for the cat, in a slightly higher position than the one established for the cat, seems to reflect the spatial layout of the characters in the cartoon, in which the canary lives on the third floor of the apartment complex and, in most cases, the cat tries to catch the canary from the ground. This showed that he used gestures to introduce not only the characters in his narrative but also the typical spatial layout in the cartoon. After establishing a contrastive gesture space for the referents, he explained the following scene using gestures to depict the cat going into the apartment (3) and being thrown out of it (4). The listener could understand the causal connection in the scene by looking at the speaker’s gestures produced in the space that had been established for a specific character even if the speaker did not verbally indicate the referents. As shown in this example, most 12-year-olds explicitly introduced referents and established the reference space for the characters using gestures and speech.



FIG. 2. Example of a 6th grade boy introducing referents.

Example 1: 12-year-old boy (id7)

etto [ e: kuro(1)i neko ga mazu yottu warui koto wo shi-te:  
 INJ INJ black cat NOM first four bad things ACC do-and  
 ‘well, first, a black cat did four bad things’

hitotsume wa:] [ e: sono: kiroi (2)hiyoko\* tori ga ie no naka  
 first TOP INJ INJ yellow chick\* bird NOM house GEN inside  
 ‘the first thing is that, well, the yellow chick, bird is inside the house and’

ni i-te so\* sono: manshon ni i-te: sono manshon wa inu  
 DAT exist-and th\* the apartment DAT exist-and the apartment TOP dog  
 ‘is at the apartment and, the apartment does not allow dogs and cats and’

neko okotowari de:] [ saisho wa: doa ni fu (3)tsuuni hait-te  
 cat prohibited and first TOP door DAT normally go straight in and  
 ‘first, (it) normally goes straight in the door’

ittanndakedo (4)oidasarechau]  
 go:PAST-SE-but throw-out:PASS-ASP  
 ‘but is thrown out’

### The Use of Gesture Space

After performing an angular transformation, a Kruskal-Wallis test was conducted on the proportion of total gestures and of the referential gestures produced in each gesture space. Although there was a tendency for 7-year-olds to use the upper space (26%) more than the 12-year-olds (9%) for the proportion of the total number of gestures, there was no significant difference in the use of each gesture space for any age group. However, there was a significant difference in the middle space ( $\chi^2(5, N=60) = 12.93, p = .02$ ) for the proportion of referential gestures (FIGURE 3). A Mann-Whitney *U* test showed that the middle space was used 78% of the time by the 12-year-olds, which was significantly greater than for the 10-year-olds (17%) (Mann-Whitney *U*,  $Z = -3.00, p = .002$ ) and for the 11-year-olds (23%) (Mann-Whitney *U*,  $Z = -2.90, p = .003$ ).

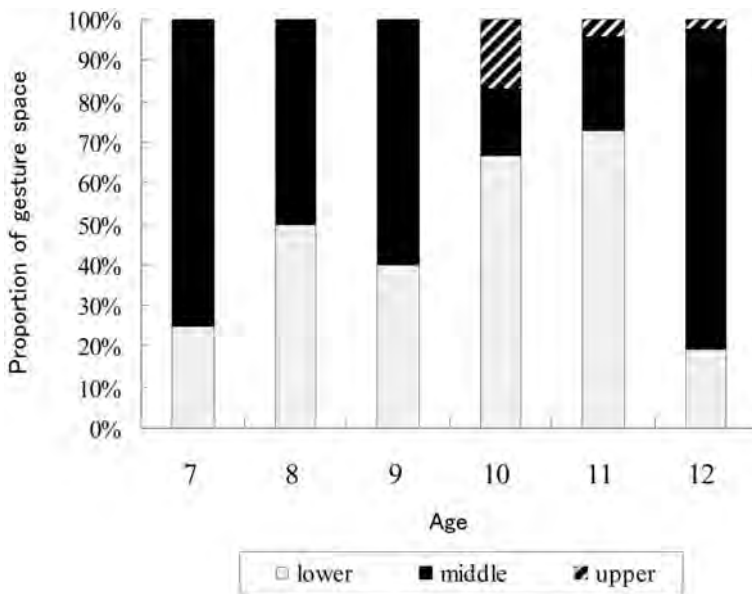


FIG. 3. Proportion of use of gesture space in referential gestures.

Children of all age groups used the lower or middle spaces more frequently overall. As the age increased, the production of gestures in the upper space decreased, while the production of gestures in either the lower or middle space increased. That is, the space used for gestures gradually reduced in size with age. This tendency is different from the use of gesture space for referential gestures. As shown in FIGURE 3, 7- and 8-year-olds never produced referential gestures in the upper space, and although 10- and 11-year-olds produced gestures in the upper space, they mostly used the lower space. About 80% of the referential gestures made by the 12-year-olds, as shown in FIGURE 2, were produced in the middle space.

*Gesture Viewpoint*

After an angular transformation, a Kruskal-Wallis test was carried out on the proportion of each gesture viewpoint in relation to the total number of gestures (TABLE 3).

TABLE 3. Mean Proportions of Gesture Viewpoint (SD)

Viewpoint	7 years	8 years	9 years	10 years	11 years	12 years
CVPT	0.31 (0.18)	0.24 (0.11)	0.23 (0.20)	0.12 (0.09)	0.19 (0.14)	0.08 (0.04)
OVPT	0.68 (0.20)	0.76 (0.12)	0.76 (0.19)	0.88 (0.09)	0.80 (0.15)	0.91 (0.04)
DVPT	0.02 (0.02)	0.01 (0.01)	0.01 (0.02)	0.00 (0.00)	0.01 (0.01)	0.01 (0.01)

A significant difference was seen in the proportion of the CVPT gestures ( $\chi^2(5, N=60) = 15.00, p = .01$ ) and the OVPT gestures ( $\chi^2(5, N=60) = 15.30, p = .009$ ) in relation to the total number of gestures. A Mann-Whitney *U* test showed that the proportion of CVPT gestures was significantly greater for 7-year-olds (31%) than for 12-year-olds (8%) (Mann-Whitney *U*,  $Z = -3.10, p = .002$ ). It also revealed that 12-year-olds produced a greater proportion of OVPT gestures (91%) than 7-year-olds (68%) (Mann-Whitney *U*,  $Z = -2.95$ ). We also counted the proportion of each gesture viewpoint in relation to the total number of referential gestures. However, none of the children produced CVPT and DVPT gestures.

These results indicated that, generally, 7-year-olds tend to depict an event not only from the OVPT but also from the CVPT by enacting the character's motion, while 12-year-olds mostly depict it from the observer's standpoint. When children referred to characters with gestures, all age groups always took the observer viewpoint.

## DISCUSSION

Research on the development of discourse in children has so far mainly focused on the linguistic information produced in narratives. However, given the fact that an adult speaker constructs a coherent discourse by not only using linguistic devices, but also by using gestures, the development of gestures in children needs to be elucidated. To complete research on this point, we examined how Japanese elementary school children introduced and maintained reference in narratives by using gestures and linguistic devices at different age levels.

The results revealed that there is no significant change in the spoken indices, including the discourse status of the referent and linguistic referential expressions, occurred during the elementary school years. As Clancy's (1992) study previously showed, NPs are used for introduced and re-introduced referents, and zero-marking is used for a maintained referent in all the age groups in this study. Because the proportions of referential expressions used in each discourse context in this study were quite similar to those of 6-year-olds or older in Clancy's study, it can be assumed that Japanese children acquire linguistic referential expression at around 6 years of age.

In contrast to the speech indices, significant differences were found between the age groups in most of the gesture indices. The results indicated that children produce many gestures in their narratives, and even use referential gestures at around



7 years of age. Although the total number and frequency of gestures decreased temporarily among 9- and 10-year-olds, after that, it increased dramatically in 12-year-olds. Referential gestures started to be used for introduced and re-introduced referents from the age of 7, but they were only used in maintained referents from the age of 10. Most of the gestures produced by 12-year-olds were observer viewpoint gestures, showing that they can depict a character's action from outside the event. As we predicted, along with the increase in referential gestures, gestures tended to be produced from the observer viewpoint in the middle space, rather than in the lower space.

From these results, the following developmental picture of discourse ability in elementary school children can be outlined. The fact that referential gestures produced with introduced and re-introduced referents appear from 7 years of age suggests that the marking of those referents in a narrative is acquired relatively earlier than those of the maintained referents. Because a character shift generally occurs at the boundary of the event, which is salient for the viewer, even 7- and 8-year-olds seem capable of marking introduced and re-introduced referents using referential gestures. Their gestural characteristics, such as enacting a character's action in a larger gesture space, imply that their discourse is mainly occupied at the narrative level in which the speaker conveys the events of the story in temporal order. Their discourse is not yet coherently organized in that referents are not referred back through the narrative with speech or with gestures. They tend to tell the listener what impressed them most without thinking about any connection to the topic.

At the ages of 9 and 10, the children in this study seemed to begin to be aware of the rules of discourse and communication, such as what information needs to be marked for making a cohesive narrative. They are also gaining an understanding of the way to conduct themselves in their classrooms to suppress fidgeting and hand movements. This is reflected in the production of referential gestures on maintained referents and in the temporal decrease of the total number and frequency of gestures. At around 10 years of age, children not only tend to use referential gestures with introduced and re-introduced referents but also start using them for maintained referents. However, the referential gestures are mostly produced in the lower space, where they are harder to be seen by a listener. We assume from this that 9- and 10-year-olds are in a transitional phase where they begin to be aware of and to internalize, however inflexibly, the rules of discourse and communication.

In this study, when children approached the age of 12, they started to clearly introduce and maintain referents throughout their narrative using gestures. Most referential gestures used by 12-year-olds were produced in the middle space. This seems to be because they are aware that the listener can use their gestures as an informational resource, and because they prefer to use observer viewpoint gestures, which do not require a large space. These results imply that 12-year-olds have acquired the ability to make a narrative coherent with multimodal cohesive devices. This will be discussed in more detail later.

Based on these results, we will discuss: 1) why there were no significant age differences in the use of language devices; 2) why 12-year-olds produce many referential gestures, and 3) why gestural development seems to decrease temporarily among 9- and 10-year-olds.

### Development of Reference Management in Japanese Language

Studies on the development of discourse have revealed that, regardless of age, language learners tend to use more pronouns or zero anaphora than full NPs to mark the maintained referents as they acquire discourse knowledge (Givón 1983; Jisa 2000; Karmiloff & Karmiloff-Smith 2001). However, the present study showed that the linguistic devices used for maintained referents do not change much during the elementary school years. By the age of 12, Japanese children tend to consistently use full NPs for introduced referents, zero-marking for maintained referents, and full NPs and zero-marking for re-introduced referents. These referential patterns are not much different from those of Japanese adult speakers (Yoshioka 2005), but they differ from the results of previous studies where it was found that children gradually use an adequate pronoun for marking a maintained referent during the elementary school years. This seems to stem from differences in linguistic structures. This is conceded as a reason, contrary to our prediction, why this study did not show the parallel development of gestures and linguistic devices.

Yoshioka (2005) argued that the use of zero anaphora is extremely common for maintained referents in Japanese narrative discourse, because the so-called 'dummy subject' is absent from the Japanese language and the use of third-person pronouns such as *she*, *he*, and *they* is very limited. These linguistic characteristics are different from those of the Indo-European languages, such as English, German, and French, which require a speaker to use such pronouns to maintain referents. Japanese children do not necessarily need to learn how to systematically use pronouns to mark the maintained referents. In addition, it seems that the use of zero-marking is a default device for marking referents (Clancy 1992), considering that 7- through 11-year-olds sometimes use zero-marking even when introducing referents. These Japanese linguistic characteristics might influence the results of this study, for example, the fact that there are no significant differences in the proportion of referential expressions in each discourse status used by each age group.

Küntay (2002) suggested that different languages offer a varying range of linguistic devices to introduce and maintain referents. For example, unlike languages using formal article systems, Mandarin Chinese uses word order devices to mark (in)definiteness, Warlpiri uses case assignment, and Turkish and Finnish utilize definite multifunctional devices such as optional lexical items, case assignment, and word order. As for Japanese, post-nominal particles could be regarded as linguistic elements that are sensitive to the discourse status of the referents (Nakamura 1993). Examining the relationship between the change of gestures and the language-specific referential expression will be our next task.

### Development of Reference Management in Gestures

In contrast to the linguistic devices of speech, there were clear developmental differences in the gestures used in Japanese narratives. The total number of gestures and referential gestures used for maintaining referents drastically increased at age 12. In addition, 12-year-olds explicitly marked referents when they first introduced them with both full NPs and gestures. So, we needed to answer the question of why gestures abruptly increase at this age.

From the findings in this study, it can be inferred that by age 12, children have acquired at least two types of discourse-related meta-communicative knowledge, which are the underlying causes of the sudden increase in gesture usage by children at this age. One type consists of recognizing what information they should give to help the listener build a coherent narrative in their mind. With this knowledge, children might notice the necessity of telling the listener about the event that the characters brought about and of maintaining the referent(s) as a topic. The other type of knowledge enables children to recognize what means of communication effectively convey the intended information. With this knowledge, children find that gestures can also be effectively used to convey a character's action and to refer to the referent as a topic. The fact that 12-year-olds tend to produce more referential gestures in the middle space, where they are most easily seen by the listener, and that they apparently do so to increase the effectiveness of their conveyance of necessary information, seems to suggest that two types of meta-communicative knowledge are acquired around the age of 12.

These speculations are consistent with previous studies. Cassell (1991) found that the proportion of meta-narrative clauses, which are comments on the representation of the story, increase at around the age of 11, and she suggested that children of late elementary school age appear to start noticing the structure of a narrative and managing how to convey it to a listener. Karmiloff-Smith (1986) examined the development of the metalinguistic awareness of a system of nominal markers used for referential and descriptive expressions between the ages of 4 and 12. She found that about 80% of metalinguistic comments spontaneously made by children between 11 and 12 years of age were about the intra-linguistic system, and such comments were less than 50% in younger age groups. These findings seem to partly support the speculations proposed here. Being a proficient speaker of a given language might involve an expert and flexible interplay between augmented linguistic knowledge on the one hand, and greater experience with language and gesture usage in varied communicative contexts, on the other (Berman 2004).

### Temporal Decline of Gestures

A temporal decline in the total number of gestures was found in Japanese middle-grade elementary school children. We then answered the question of why this happens to children aged 9 and 10. A part of this phenomenon appears to stem from their learning environment. In Japanese classrooms, pupils are sometimes both implicitly and explicitly warned by their teacher to avoid fidgeting or moving their hands when the teacher or another child is speaking or sometimes even when they themselves speak. A teacher often conveys this message by simply telling a child that they should not move their hands or by indirectly indicating that they should put their hands on their thighs.

Another aspect of this temporal decline seems attributable to its location within the phases of a child's development. The ages of 9 and 10 might be a period in which children are very sensitive to the rules taught in a classroom, and in which they try to internalize them as much as they can. Karmiloff-Smith (1986) revealed that a substantial number of children explicitly made spontaneous metalinguistic comments from 9 years old onwards. Thus, children around the middle grades seem to begin noticing that there are some underlying rules in the communica-

tion or language system. This cognitive change seems to influence how children produce or constrain gestures.

At the same time, meta-communicative knowledge and attentional resources seem to be limited at this age, which makes it difficult for them to explain and simultaneously represent what they think through gestures. Similar phenomena have been reported in other studies on the discourse development of Japanese children. For example, studies found that 10-year-olds show poorer performance in usage of the postpositional particle *WA* than the other age groups (Ito & Tahara 1985; Nakamura 1993). In addition, Küntay and Nakamura (2004) revealed that Japanese children around 9 years old used the fewest evaluative devices, compared with preschoolers, 7-year-olds, and adults. As Ito and Tahara suggested, because children in this period are just beginning to notice and attempt to grasp the multifunctional nature of language devices, their performance seems to temporarily decline.

With these results presenting a qualitative change in Japanese discourse skills, the ages of 9 and 10 can be seen as the transitional period in which Japanese children begin noticing discourse and communicative rules and try to embody them. Through this period, children gradually acquire discourse skills and the two types of meta-communicative knowledge previously mentioned. By the age of 12 they can make a coherent discourse, similar to that of adults, by using referential gestures.

Further investigations are needed to clarify the factors that influence the temporal decline in 9- and 10-year-olds. To reveal these factors, it will be important to observe how children actually learn about discourse and communicative knowledge in their classrooms. At the same time, comparing our data with that of language learners in other cultures is also necessary in order to see whether the temporal decrease is specific to Japanese children. In addition, given the fact that previous studies have mainly focused on early childhood and adulthood, it would be helpful to observe how discourse skills, including gestures, change beyond elementary school. This study did not show significant developmental change in the linguistic devices used for narratives in terms of the criteria that have been used in most previous research on the discourse in Indo-European languages. It would be worth investigating how linguistic elements in Japanese language develop in parallel with the change of gestures to make a narrative cohesive.

#### CONCLUSION

By focusing on the gestures that are produced spontaneously during the narration of a story, we demonstrated that children already use many gestures in their narratives from as early as 7 years old. It was also revealed that although the proportion of linguistic referential expressions does not change during the elementary school years, a dramatic difference was found in the referential gesture use between 11- and 12-year-olds. This difference was that, by the age of 12, children use their hands not only for depicting a character's action but also for maintaining the referent as a topic. Through elementary school, Japanese children acquire the use of gestures, which visually shows topic continuity and discontinuity and gradually can help them to produce coherent narratives, seemingly based on learning two types of meta-communicative knowledge.

These results suggest that gestures can be seen as a useful index to predict discourse development, and at the same time that referential gestures help language learners to construct cohesive discourse by providing spatial scaffolding for cohesion. This study revealed that gesturing is a crucial component in the construction of discourse. Discourse development is not just a linguistic matter, but rather a phenomenon that is related to language, body, and culture.

#### ACKNOWLEDGEMENTS

This work was supported by a grant-in-aid for scientific research provided by the Japan Society for the Promotion of Science. We would like to thank David McNeill, Elena Levy, Gale Stam, and Keiko Yoshioka for their insightful comments on our earlier versions of this manuscript. We would also like to thank the children and the parents for their cooperation.

#### REFERENCES

- BAMBERG, M. (1987), *The acquisition of narrative: Learning to use language*, Berlin, de Gruyter.
- BELLUGI, U., KLIMA, E. (1982), *From gesture to sign: Deixis in visual-gestural language*, in *Speech, place and action: Studies in deixis*, Eds. R. J. Jarvella, W. Klein, Chicester, John Wiley & Sons, pp. 297-313.
- BERMAN, R. A. (2004), *Between emergence and mastery: The long developmental route of language acquisition*, in *Language development across childhood and adolescence*, Ed. R. A. Berman, Amsterdam, Benjamins, pp. 9-34.
- BERMAN, R., SLOBIN, D. (Eds.), (1994), *Development of linguistic forms: English. Relating events in narrative: A crosslinguistic developmental study*, Hillsdale, NJ, Lawrence Erlbaum Associates.
- CASSELL, J. (1991), *The development of time and event in narrative: Evidence from speech and gesture*, Unpublished doctoral dissertation, University of Chicago.
- COLLETTA, J. M. (2009), *Comparative analysis of children's narratives at different ages: A multimodal approach*, «*Gesture*», 9(1), pp. 61-96.
- CLANSY (1992), *Referential strategies in the narratives of Japanese children*, «*Discourse Processes*», 15(4), pp. 441-467.
- DENT, C. H. (1984), *Development of discourse rules: Children's use of indexical reference and cohesion*, «*Developmental Psychology*», 20(2), pp. 229-234.
- GIVÓN, T. (1983), *Topic continuity in discourse: Quantitative cross-language studies*, Amsterdam, Benjamins.
- GULLBERG, M. (2006), *Handling discourse: Gestures, reference, tracking, and communication Strategies in early L2*, «*Language Learning*», 56(1), pp. 155-196.
- FURUYAMA, N. (2001), *De-syntactizing the theories of reference maintenance from the viewpoint poetic function of language and gesture: A case of Japanese discourse*, Unpublished doctoral dissertation, University of Chicago.
- HALLIDAY, M. A. K., HASAN, R. (1976), *Cohesion in English*, London, Longman.
- HICKMANN, M., HENDRIKS, H. (1999), *Cohesion and anaphora in children's narrative: A comparison of English, French, German, and Mandarin Chinese*, «*Journal of Child Language*», 26, pp. 419-452.
- ITO, T., TAHARA, S. (1985), *A psycholinguistic approach to the acquisition of multifunctionality in Japanese particles wa and ga*, «*Descriptive and applied Linguistics*», 18, pp. 121-131.
- JISA, H. (2000), *Increasing cohesion in narratives: A developmental study of maintaining and reintroducing subjects in French*, «*Linguistics*», 38(3), pp. 591-620.

- KARMILOFF, K., KARMILOFF-SMITH, A. (2001), *Pathways to Language: From Fetus to Adolescent*, Cambridge, MA, Harvard University Press.
- KARMILOFF-SMITH, A. (1981), *The grammatical marking of thematic structure in the development of language production*, in *The child's construction of language*, Ed. W. Deutsch, London, Academic Press, pp. 121-147.
- KARMILOFF-SMITH, A. (1985), *Language and cognitive processes from a developmental perspective*, «Language and Cognitive Processes», 1(1), pp. 61-85.
- KARMILOFF-SMITH, A. (1986), *From meta-processes to conscious access: Evidence from children's metalinguistic and repair data*, «Cognition», 23, pp. 95-147.
- KITA, S., WOOD, C. (2006, June), *Semiotic characteristics of speech-accompanying gestures in children and adults' narrative*. Paper presented at the Current Developments in Gesture Research, Hatfield, UK.
- KÜNTAY, A. C., NAKAMURA, K. (2004), *Linguistic strategies serving evaluative functions: A comparison between Japanese and Turkish narratives*, in *Relating events in narrative*, Volume 2: *Typological and contextual perspective*, Eds. S. Strömquist, L. Verhoeven, Mahwah Lawrence Erlbaum Associates, pp. 329-358.
- KÜNTAY, A. C. (2002), *Development of the expression of indefiniteness: Presenting new referents in Turkish picture-series stories*, «Discourse Processes», 33(1), pp. 77-101.
- LEVY, E. (2003), *The Roots of Coherence in Discourse*, «Human Development», 46, pp. 169-188.
- MCCABE, A., PETERSON, C. (1985), *Developing narrative structure*, Hillsdale, NJ, LEA.
- MCNEILL, D. (1992), *Hand and mind*, Chicago, University of Chicago Press.
- MCNEILL, D. (2005), *Gesture and thought*, Chicago, University of Chicago Press.
- McNeill, D., LEVY, E. T. (1993), *Cohesion and gesture*, «Discourse Processes», 16(4), pp. 363-386.
- NAKAMURA, K. (1993), *Referential structure in Japanese children's narratives: The acquisition of wa and ga*, «Japanese / Korean linguistics», 3, pp. 84-99.
- NELSON, K. (Ed.), (1989), *Monologue as representation of real-life experience*, «Narratives from the crib», Cambridge, MA, Harvard University Press.
- O'NEILL, D. K., HOLMES, A. C. (2002), *Young preschoolers' ability to reference story characters: The contribution of gestures and character speech*, «First Language», 22, pp. 73-103.
- ÖZYÜREK, A. (2002), *Do speakers design their co-speech gestures for their addressees? The effects of addressee location on representational gestures*, «Journal of Memory and Language», 46, pp. 688-704.
- UCHIDA, N. (1996), *Kodomo no discourse no hattatsu: Monogatari sanshutsu no kiso katei* [The development of children's discourse: A basic process of narrative production], Tokyo, Kazama Shobou
- WIGGLESWORTH, G. (1990), *Children's narrative acquisition: A study of some aspects of reference and anaphora*, «First Language», 10, pp. 105-125.
- WIGGLESWORTH, G. (1997), *Children's individual approaches to the organization of narrative*, «Journal of Child Language», 24, pp. 279-309.
- YOSHIOKA, K. (2005), *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse*, Nijmegen, Radboud University.