Research Article **Open Access**

Kazuki Sekine*, Sotaro Kita

The parallel development of the form and meaning of two-handed gestures and linguistic information packaging within a clause in narrative

DOI 10.1515/opli-2015-0015
Received January 16, 2015; accepted June 12, 2015

Abstract: We examined how two-handed gestures and speech with equivalent contents that are used in narrative develop during childhood. The participants were 40 native speakers of English consisting of four different age groups: 3-, 5-, 9-year-olds, and adults. A set of 10 video clips depicting motion events were used to elicit speech and gesture. There are two findings. First, two types of two-handed gestures showed different developmental changes: those with a single-handed stroke with a simultaneous hold increased with age, while those with a two handed-stroke decreased with age. Second, representational gesture and speech developed in parallel at the discourse level. More specifically, the ways in which information is packaged in a gesture and in a clause are similar for a given age group; that is, gesture and speech develop hand-in-hand.

Keywords: speech, two-handed gestures, hold phase, children

1 Introduction

Much research has investigated the development of co-speech gestures and speech in children. However, few studies have studied the extent to which gesture and speech are similar to one another in the development of form-referent relationships. In the present research, we focused on the use of two-handed gestures in children and adults during narratives. We examined the developmental change among 3-, 5-, 9-year-olds and adults, in terms of the type of two-handed gestures used, and the referents of the two-handed gestures and corresponding developmental changes in speech; specifically, the type of referents expressed in a clause.

Previous research has found that adult speakers use two-handed gestures in systematic ways. Kita, Gijn, and van der Hulst (2014) examined whether the Symmetry Condition (Battison, 1978), which has been attested to exist in all signed languages, was also found in spontaneous co-speech gestures. Two-handed gestures or signs are considered to meet the Symmetry Condition when the two hands are symmetrical or identical in shape, location, and movement. The analysis of co-speech gestures and signs in narrative has shown that co-speech gestures are as symmetrical as signs (Kita et al., 2014).

Article note: Part of a topical issue on sign linguistics and gesture studies, edited by E. Engberg-Pedersen

***Corresponding author: Kazuki Sekine:** University of Warwick, Department of Psychology, Coventry CV4 7AL, United Kingdom, E-mail: kazuki@tkc.att.ne.jp
Sotaro Kita: University of Warwick, Department of Psychology, Coventry CV4 7AL, United Kingdom

Along a similar vein, Enfield (2004) found that a complex two-handed construction used in signed language was also observed in spontaneous co-speech gestures. He documented how two-handed gestures were used in the description of complex artifacts (fish traps) with a symmetrical shape. When describing how such fish traps work, speakers first produced a two-handed symmetrical gesture, to show a shape of a particular part or the whole of the trap. Next, speakers were sometimes noted to use the dominant hand to depict a new referent (e.g., how a fish moves, or the shape of another part), while holding the non-dominant hand in the air to maintain the referent depicted by the first two-handed symmetrical gesture. Enfield suggested that the non-dominant hand (held in the air) provides two types of semantic support for the dominant hand. First, it provides a stable spatial reference point for the referent depicted by the dominant hand. Second, it signals that certain background information continues to be relevant to what is being said, while the dominant hand represents foregrounded/new information.

As Enfield (2004) and Kita et al. (2014) showed, adult speakers can systematically use two-handed co-speech gestures. In the current study, we investigated how the use of these two-handed gestures develops with speech during childhood.

Literature on adult co-speech gestures has suggested two different views about the semantic relationship between co-speech gestures and speech. The first view is associated with the *tradeoff hypothesis*, proposed by de Ruiter (2006) and de Ruiter, Bangerter, and Dings (2012). The authors state, "there is a tradeoff relation between gesture and speech in terms of their communicative load" (de Ruiter et al., 2012, p. 233). Thus, the hypothesis holds that when speech becomes more difficult, gestures compensate for speech by expressing what cannot be expressed in speech. Thus, the content of gestures often does not match that of speech, particularly when it is difficult to express certain concepts in speech. The alternative view is the *hand-in-hand hypothesis*, proposed by So, Kita, and Goldin-Meadow (2009). This hypothesis is based on McNeill's (1992) theory and previous empirical studies (e.g., Kita & Özyürek, 2003; Kita, Özyürek, Allen, Brown, Furman, & Ishizuka, 2007). The hand-in-hand hypothesis proposes that because gesture and speech share a computational stage, gesture represents equivalent content to speech.

Similar contrastive views on the relationship between gesture and speech can be found in the literature on children's gestures. We call them the *gestural supplementation view* (which is similar to the tradeoff hypothesis explained above) and *the gesture-speech parallel development view*. The gestural supplementation view suggests that gestures supplement speech when speech cannot express what children want to express, because their linguistic ability is still limited, or the expressive power of their language is inherently limited.

A few different lines of studies support the first view, the gestural supplementation view. For example, Evans, Alibali and McNeil (2001) revealed that children with Specific Language Impairment use gesture to compensate for their impaired speech. More specifically, when explaining their answers to Piagetian conservation tasks, children with Specific Language Impairment produced more explanations in which gesture conveyed information that speech did not convey, than typically developing children. Kidd and Holler (2009) asked children to tell stories including homonyms (e.g., *bat* for flying mammal vs. *bat* for sports equipment), and investigated whether the children used gestures to clarify the noun when explaining the homonym. They found that 4-year-olds could use iconic gestures to resolve lexical ambiguity; that is, gesture supplemented information to resolve lexical ambiguity. Church and Goldin-Meadow (1986) found that when children aged 5 to 8 years conveyed information that was difficult for them to verbalize when explaining a novel concept, the children often produced gestures conveying the information that they could not yet explain in speech.

The second view, gesture-speech parallel development, was supported by Nicoladis' (2002) study on gestures produced by 3- to 4-year-old French-English bilingual children. She showed that the rate of iconic gestures was significantly higher when children spoke their dominant language compared to when they spoke their non-dominant language. That is, iconic gestures were produced more often when speaking the better-developed language. Thus, in this view iconic gestures and language are hypothesized to develop hand-in-hand. Additional studies on children's gestures have supported this view. Sekine (2009) examined the relationship between gestural and spoken spatial expression used while providing route directions in children aged 4 to 6 years. He found that as children became able to segment a route into short segments by using landmarks and left-right terms in speech, their gesture sizes became smaller and the frequency

of gestures increased. Gullberg and Narasimhan (2010) focused on the fact that the Dutch language uses two semi-obligatory caused posture verbs ("*leggen*" for lay vs. "*zetten*" for set/stand). These two verbs are selected based on whether the located object is placed horizontally or vertically. However, 5-year-olds tended to overextend the use of *leggen* regardless of object position. This development in speech was mirrored in gesture development. Children who overextended the use of *leggen* produced gestures depicting movement path, but without any information about the moved object. In contrast, children and adults who used the two verbs to distinguish object positions produced gestures depicting the objects (as well as movement path). Finally, Colletta et al. (2015) compared speech and gesture produced in a narrative task in 5- and 10-year-old French-, American-, and Italian-speaking children. The authors found that as the children became able to deliver longer and more detailed accounts of a story, they increased the use of co-speech gestures to represent and track the characters in the story.

The parallel relationship of gesture and speech has been demonstrated at different levels: the frequency of gesture and the complexity of sentences (Nicholadis, 2002), the segmentation of routes in speech and the size and number of gestures (Sekine, 2009), the representation of a moved object in gestures and placement verbs (Gullberg & Narasimhan, 2010), and reference tracking in gesture and speech (Colletta et al., 2015). However, no study has examined the development of two-handed gestures and information packaging in speech.

The two views can be both right depending on the context. For example, when the grammatical or lexical demands are too high, such as when explaining a new concept or complex phenomenon, children's gesture may supplement what is expressed in speech. When the message is easy to express in speech, gesture and speech may simply convey equivalent information. Thus, for a given age, evidence for both views can be found depending on the task. However, in the current study, we are interested in the developmental change in the relationship between two-handed gestures and speech in a narrative task.

Thus, the goal of the current study was to determine which of the two views better characterizes developmental change in semantic coordination between speech and gesture, with a focus on two-handed gestures. More specifically the current study analyzed narratives produced by 3-, 5-, 9-year-olds and adults, and examined the number and type of event components that are packaged in a clause and in a gesture. Event components refer to protagonists and landscape elements that are a part of motion events in this study.

We focused on two-handed gestures because they can express a wider range of referents than single-handed gestures. In two-handed gestures, two hands may jointly express a single concept or each hand may express separate concepts. This makes it interesting to compare how information is packaged in two-handed gestures and in a clause in speech.

The current study used the clause as the unit for assessing semantic packaging in speech. This is because gesture tends to package information in a way that a clause in concurrent speech packages information. This is shown in gestural and linguistic representation of manner and path in motion events (Kita & Özyürek, 2003; Furman, Kuntay, & Özyurek, 2014; Kita et al., 2007; Mol & Kita, 2012; Özyurek, Allen, Kita, Furman, & Brown, 2005). When adult speakers express manner and path in a single clause, they tend to produce a gesture that expresses manner and path together in a single movement. In contrast, when speakers express manner and path in two separate clauses, they tend to express them in two separate gestures (Mol & Kita, 2012). The current study examined linguistic packaging of different aspects of motion events that are different from manner and path, namely, the number and type of event components encoded within a clause.

We elicited narratives using stimuli developed by Özyürek, Kita, and Allen (2001), which consist of 10 animated video clips depicting motion events (see Figure 1 and Method section for more details). Each motion event includes two protagonists and a landscape through which those protagonists are moving. In our study, we firstly examined the use of different types of two-handed gestures. Second, we examined the number and type of event components represented in two-handed gestures and in clauses in speech.

There are two possibilities as to the developmental change in how speech and two-handed gestures are semantically coordinated with each other. According to the gesture-speech parallel development view, a two-handed gesture and a clause should encode similar information at a given age and develop in parallel. According to the gestural supplementation view, a two-handed gesture should express information that the children do not yet express within a clause at a given age.

2 Method

2.1 Participants

The participants were 40 native speakers of English consisting of four different age groups, with each group consisting of 10 participants: 3-year-olds (mean age: 3;10, range: 3;3 to 4;2, 6 females and 4 males); 5-year-olds (mean age: 5;5, range: 5;5 to 6;0, 6 females and 4 males); 9-year-olds (mean age: 9;4, range: 8;10 to 10;0, 6 females and 4 males), and adults (mean age: 22, range: 19 to 31, 6 females and 4 males).

2.2 Data

The data used in this study were collected as part of a study investigating the cross-linguistic syntactic packaging of 'path' and 'manner' information in motion events (Allen, Özyürek, Kita, Brown, Furman, Ishizuka, & Fujii, 2007; Kita, et al., 2007; Özyürek, et al., 2001; Özyürek, Kita, Allen, Brown, Furman, & Ishizuka, 2008). The current study is a reanalysis of this existing data with a different research purpose—to investigate the way in which two-handed gesture varies with age. In the data, narratives were elicited by a set of 10 animated video clips, which are called "Tomato Man movies" (Özyürek et al., 2001). Each of the clips depicts a motion event consisting of two protagonists (red Tomato man and green Triangle man) and a landscape where those protagonists are moving (Figure 1). Each video lasted about ten seconds.

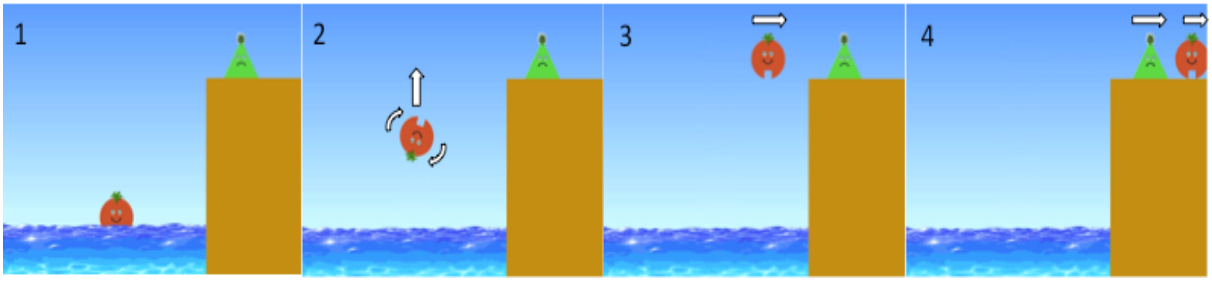

Figure 1. An example of the stimuli shown to participants

2.3 Procedure

Participants were tested individually in a quiet spare room at their school for children and in a laboratory at a university for adults. The experimenter played each of the 10 video clips twice for each participant. Then, the participants were asked to retell what had happened in the clip to a listener who purportedly had not seen it. The experiment lasted for an average of 10 minutes for each participant.

2.4 Coding

All speech produced by each participant for each clip was transcribed and segmented into clauses. For the purpose of this study, clauses included both main and subordinate clauses. Each clause contained one verb (finite or nonfinite) except for the constructions that express the intention of the agent followed by an infinitive verb (e.g., "he tries climbing up", "he wants to go up"), which contained multiple verbs (intention verbs and the infinitive verb expressing the content of the intention). Here are some examples of sentences that were coded as two clauses: "Triangle man spins, while he goes down", "Mr. Tomato is sitting in the water, floating".

Then, we coded the number of protagonists and landscape elements expressed in a clause. For example, in "Mr. Tomato is sitting in the water, floating", there is one protagonist ("Mr. Tomato") and one landscape

element ("the water") in the first clause and there is one protagonist in the second clause ("floating" has a zero pronoun referring to "Mr. Tomato").

As for gestures, we focused our analysis on gestures depicting the protagonists and the landscape (iconic and deictic gestures, respectively, in McNeill's (1992) classification). Each gesture was deconstructed into the following gesture phases, as defined by McNeill (1992): *preparation*, *stroke*, *hold*, and *retraction* phases. The *preparation* is a phase that prepares to execute the stroke phase. The *stroke* phase is the main part of a gesture, as it is the main meaning-bearing part of the gesture, and is carried out with the most effort or tension. The *hold* is a phase where the hand is held in the mid-air in the same position, and may be observed before or after the stroke. The *retraction* is a phase where the hand returns to its rest position. Note that only the stroke phase is obligatory in forming a gesture. Thus, some of the gestures did not have preparation, hold, and/or retraction phases.

Next, we coded the gestures into one of two categories: *single-handed gesture* or *two-handed gesture*. A referent of each representational gesture was determined by the content of the accompanying speech and previous speech context. The single-handed gesture refers to a gesture depicting or indicating one or two event components by a single hand. Figure 2 shows a female adult producing a single-handed stroke, depicting a tomato man that came to the scene and slid up the hill, by extending her index finger and moving it from the left bottom space to the upper right space, while saying "and the tomato just comes, and just **slides up the hill**". The bold face in the speech (in quotation marks) indicates where the stroke of the gesture occurred.

Figure 2. An example of a single-handed gesture.

The two-handed gesture refers to a gesture depicting or indicating one or two event components with both hands. The two-handed gestures were further coded into one of two subcategories: *Gestures with a two-handed stroke* and *gestures with a single-handed stroke and a simultaneous hold*. A gesture with a two-handed stroke refers to gestures in which both hands are involved in the gesture stroke. A gesture with a singled-handed stroke and a simultaneous hold refers to a gesture in which one hand is involved in the gesture stroke, while the other hand is held.

Gestures with two-handed strokes were categorized into one of the following four mutually exclusive categories, based on the referents that the gesture represented: *one landscape element*, *two landscape elements*, *one protagonist*, and *two protagonists*. Figure 3 showed the examples of gesture and the concurrent speech in the four subcategories.

One protagonist: 3 year old girl producing a two-handed stroke depicting a triangle man that jumped into the water by putting her hands together and moving them down to her right side while saying "the green **guy jumped in the water**".

Two protagonists: 5 year old boy producing a two-handed stroke, a tomato man and a triangle man that fell in the water by moving both his hands down while saying "they both **fell in the water** from a cliff".

One landscape element: 9 year old boy producing a two-handed stroke depicting the flat land by horizontally swiping his right hand and the left hand from the midline to his right and to his left respectively while saying "which was **on a low surface**".

Two landscape elements: A male adult producing a both-handed stroke depicting the two different levels of the land by making his hands flat shape at two different levels, while saying "so you got your **two different levels** of elevation".

Figure 3. Examples of gesture and the accompanying speech in the four subcategories of two-handed stroke gestures. The bold face in the speech (in quotation marks) indicates where the stroke of the gesture occurred.

In a gesture with a single-handed stroke and a simultaneous hold, one hand depicts or indicates a referent that is mentioned concurrently in speech, while the other hand is held in the air to indicate the existence of the other protagonist or the landscape. Figure 4 shows a male adult producing a single-handed stroke and a simultaneous hold, depicting a tomato man going up the hill after the triangle man went up the hill. In picture (1) of Figure 4, he introduced the triangle man by his right hand and the tomato man by his left hand. Then, in picture (2), he moved his right hand to the right upper space to describe how the triangle man moved to the top of the hill, while holding his left hand on the desk to depict the tomato man sitting at the bottom. Finally in picture (3), he moved his left hand to the right upper space to depict the tomato man going up to the top of hill, where the triangle man is waiting. In this example, the gestures in pictures (2) and (3) were coded as a single-handed stroke and a simultaneous hold.

"On the ground (1) **the tomato** and **triangle** came up at the same time right next to each other. And the **triangle kept** going. And he **twirled up** the hill until he (2) **got to the top**. And once he got up to the top, **the tomato man**, He (3) **he went up** the hill."

Figure 4. An example of a single-handed stroke and a simultaneous hold and the accompanying speech. In this transcription, the bold face in the speech (in quotation marks) indicates where the stroke of the gesture occurred. A dotted line indicates the hold phase by his left hand (for tomato man), a single underline indicates the hold phase by his right hand (for triangle man), and a double underline indicated the hold phase by both of his hands (for both protagonists). The numbers in the transcription correspond to the numbers given to the pictures in Figure 4.

3 Results

3.1 Mean number of clauses and gestures

For each of the age groups, we calculated the mean number of clauses that were produced in the retelling of the 10 cartoons, the mean number of narrative-related clauses, the mean number of gestures, and the frequency of gestures per clause (Table 1). We defined narrative-related clauses as clauses which encoded information about the protagonists and/or the landscape in the stimulus motion events. Thus, they exclude clauses that encoded only the speaker's comments on stimulus events, or that encoded information irrelevant to what had happened in the clips (e.g., the speaker's personal experience). We focused our analysis on representational gestures (iconic and deictic gestures) that depicted protagonists or landscape of the stimulus stories. We refer to such gestures simply as "representational gestures" henceforth. We conducted analyses of variance (ANOVA) on the four indices with the age groups as a between-subject factor. A main effect of age group was found for the mean number of all clauses, $F(3, 36) = 8.27$, $p < .001$, $\eta^2 = .41$, the mean number of narrative-related clauses, $F(3, 36) = 24.40$, $p < .001$, $\eta^2 = .67$, the mean number of representational gestures, $F(3, 36) = 27.51$, $p < .001$, $\eta^2 = .70$, and the frequency of representational gesture per clause, $F(3, 36) = 11.31$, $p < .001$, $\eta^2 = .49$. Tukey post hoc tests ($p < .05$) showed that the adults produced a significantly greater number of all clauses, narrative-related clauses, and gestures; and produced gestures more frequently than any of the other three age groups. 9-year-olds were found to produce a significantly greater number of narrative-related clauses than 3- and 5-year-olds.

Table 1. Mean number and standard deviations (in parentheses) of clauses and representational gestures produced over 10 video clips.

	3 years	5 years	9 years	Adults
Mean number of all clauses	71.8 (30.4)	74.2 (16.2)	89.1 (31.9)	131.5 (38.6)
Mean number of narrative related clauses	43.0 (14.7)	50.4 (8.7)	74.0 (21.2)	106.6 (24.6)
Mean number of representational gestures	26.2 (9.5)	24.1 (11.3)	50.1 (25.0)	112 (40.0)
Representational gesture frequency per clause	0.63 (0.21)	0.48 (0.22)	0.71 (0.32)	1.05 (0.28)

The main results will be presented as follows. First, we will demonstrate age-related differences in forms of gestures. Then, we will discuss age-related change in gesture semantics, that is, what kind of information gestures encode, followed by an analysis of age-related change in speech that parallels the change in gesture. Finally, we will present results on the subtypes of two-handed gestures.

3.2 Gesture form: single-handed *vs.* two-handed

To assess the hand use pattern in representational gesture across age groups, we first calculated the proportion of singled-handed gestures and two-handed gestures, out of all representational gestures for each age group (the first two rows shown in Table 2).

Table 2. The mean proportion and the standard deviations (in parentheses) of each hand use category in representational gesture.

Gesture category	3 years	5 years	9 years	Adults
Single-handed	0.63 (0.34)	0.92 (0.10)	0.79 (0.19)	0.65 (0.25)
Two-handed	0.37 (0.34)	0.08 (0.10)	0.21 (0.19)	0.35 (0.25)
Two-handed stroke	0.36 (0.33)	0.08 (0.10)	0.16 (0.17)	0.14 (0.18)
Single-handed stroke with a simultaneous hold	0.02 (0.04)	0 (0)	0.05 (0.05)	0.21 (0.14)

We conducted an ANOVA on the proportion of single-handed gestures with the age groups as a between-subject factor (note that statistics for the proportion of single-handed gestures is the same as that for the proportion of the two-handed gestures, as the two proportions add up to 1). A main effect of age group was found, $F(3, 36) = 3.32$, $p < .05$, $\eta^2 = .22$. Tukey post hoc tests ($p < .05$) showed that the proportion of single-handed gestures in 5-year-olds was significantly higher than that in 3-year-olds. We also examined whether the proportion of single-handed gestures shows a U-shaped change with age. Polynomial contrasts revealed a significant quadratic trend for the proportion of single-handed gestures as a function of age, $F(1, 84) = 8.30$, $p < .01$, $\eta^2 = .023$, but no linear trend, $F(1, 84)=0.10$, $p = .75$, $\eta^2 =.03$. This result indicates that the proportion of single-handed gestures indeed showed change in a U-shaped curve.

Next, we examined age differences in the use of two types of two-handed representational gestures. We broke down the two-handed gestures into those with *two-handed stroke* vs. *single-handed stroke and a simultaneous hold* (the last two rows shown in Table 2), and calculated the proportion of each type of two-handed gesture of the total number of representational gestures (including single-handed gestures) for each age group. We conducted an ANOVA on the two variables with the age groups as a between-subject factor. A main effect of age group was found for the proportion of gestures with a two-handed stroke, $F(3, 36) = 3.32$, $p < .05$, $\eta^2 = .22$, and for the proportion of gestures with a single-handed stroke and a simultaneous hold, $F(3, 36) = 14.06$, $p < .001$, $\eta^2 = .54$. Tukey post hoc tests ($p < .05$) revealed that the proportion of gestures with a two-handed stroke in 3-year-olds was significantly higher than that in 5-year-olds, and that the proportion of gestures with a single-handed stroke and a simultaneous hold in adults, was significantly higher than the other three age groups. These results indicated that the proportion of two-handed gestures showed a developmental change, as gestures with a two-handed stroke declined with age, and gestures with a single-handed stroke and a simultaneous hold increased with age.

3.3 Gesture semantics: The number of referents represented by two-handed strokes

We categorized all two-handed strokes into one of the following four mutually exclusive categories based on the referents that the concurrent or previous speech represented: *one landscape element*, *two landscape elements*, *one protagonist*, and *two protagonists*. The proportion of each category was calculated by dividing the number of representational gestures in each category by the total number of representational gestures (Table 3). We excluded the participants who did not produce gestures with a two-handed stroke from this analysis. We conducted an ANOVA on the proportion of gestures expressing one landscape element, two landscape elements, one protagonist, and two protagonists, with the age groups as a between-subject factor. A main effect of age group was found for the proportion of gestures expressing one landscape element, $F(3, 36) = 3.13$, $p < .05$, $\eta^2 = .23$. Tukey post hoc tests ($p < .05$) showed that the proportion of gestures expressing one landscape element in adults was significantly higher than that in 3- and 5-year-olds.

Table 3. The mean proportion and standard deviations (in parentheses) of referents of gestures with a two-handed stroke

	3 years (N = 9)	5 years (N = 8)	9 years (N = 9)	Adults (N = 10)
One landscape element	0.01 (0.03)	0 (0)	0.08 (0.14)	0.19 (0.25)
Two landscape elements	0.03 (0.06)	0 (0)	0.03 (0.08)	0.01 (0.02)
One protagonist	0.49 (0.38)	0.63 (0.52)	0.63 (0.39)	0.45 (0.37)
Two protagonists	0.46 (0.41)	0.38 (0.52)	0.26 (0.37)	0.35 (0.42)

3.4 Speech: The number of referents expressed in a clause

We examined whether speech showed an age-related change that corresponded to the change in gesture semantics for gestures with a two-handed stroke. To this end, we analyzed what referents were linguistically expressed in a clause. This analysis (and all subsequent analyses of speech) comprised all narrative-related clauses—including the ones without any accompanying gestures—in order to best capture participants' performance in their spoken discourse. We calculated the proportion of clauses having no protagonist (consisting of only landscape elements) (e.g., "there was a cliff on the right", "this is again with the cliff and the water"), the proportion of clauses including only one protagonist expressed by any form, such as a noun phrase, a pronoun, or a zero marking (e.g., "the triangle goes up the hill", "he jumps twice"), and the proportion of clauses including two protagonists by any form (e.g., "then tomato man follows him", "they got out"). These proportions were calculated by dividing the number of clauses in each category by the total number of narrative-related clauses (Table 4). An ANOVA was conducted on the proportion of clauses including no protagonist, one protagonist, and two protagonists, with the age groups as a between-subject factor. A main effect of age group was found for the proportion of clauses including only landscape elements and no protagonist, $F(3, 36) = 17.04$, $p < .001$, $\eta^2 = .59$. Tukey post hoc tests ($p < .05$) showed that the proportion of clauses including only landscape elements and no protagonist was significantly higher in adults than in the other three age groups. This result indicated that the type of referents represented by speech (Table 4) and those represented by gestures with a two-handed stroke (Table 3) showed a similar age-related change. That is, older participants were more likely to represent a landscape element in both gesture and speech. Thus, this result showed that similar types of information were represented in the gesture and speech of our participants.

Table 4. The mean proportion and standard deviations (in parentheses) of clauses including zero, one or two protagonists.

The proportion of clauses including	3 years	5 years	9 years	Adults
Zero protagonist (Only landscape element)	0.03(0.05)	0.01 (0.02)	0.02 (0.03)	0.12 (0.05)
One protagonist	0.37 (0.11)	0.35 (0.09)	0.33 (0.07)	0.30 (0.06)
Two protagonists	0.60 (0.10)	0.64 (0.08)	0.64 (0.02)	0.58 (0.07)

3.5 Gesture form: Number of participants who used gestures with a single-handed stroke with a simultaneous hold

Revisiting the question that we addressed in the analysis summarized in Table 2, we further examined the ways in which the four age groups used gestures with a single-handed stroke and a simultaneous hold. Specifically, we counted the number of participants who produced the stroke-hold combination for at least one out of the 10 stimulus events. The number and proportion of participants who produced the stroke-hold combination in each age group are as follows: two 3-year-olds (20%), zero 5-year-olds (0%), seven 9-year-olds (70%), and eight adults (80%). To investigate whether these proportions differed between age groups, a chi-square test was conducted. Significant differences were found, χ^2 (2, N = 40) = 18.31, $p < .001$, *Cramer's* $V = .68$. A residual analysis indicated that the proportions of both 9-year-olds and adults who produced the stroke-hold combination were significantly higher than for 3- and 5-year-olds. Although 9-year-olds did not use these gestures as frequently as adults (see Table 2), the number of 9-year-olds who produced the stroke-hold combination was almost the same percentage as adults.

This result suggests that 9-year-olds have the ability to produce the stroke-hold combination, but that the frequency is different from adults. In the following section, we will focus on the comparison of the referents of gestures with a single-handed stroke and a simultaneous hold between 9-year-olds and adults.

3.6 Gesture semantics: A subsequently mentioned referent for a hold in gestures with a single-handed stroke and a simultaneous hold.

Analyses conducted so far have revealed that gestures and speech represent referents in similar ways. This parallel relationship may be found at the discourse level as well, that is, in whether or not the referent of gesture/speech was related to the subsequent gestural/spoken context. This section reports analysis on representational gesture and the next section reports on speech. We examined whether the referent of the hold in gestures with a single-handed stroke and a simultaneous hold was represented in any preceding or subsequent gesture in the narrative. We assumed that a speaker uses a hold to indicate a referent maintained from the preceding discourse while expressing another concept in the concurrent stroke, and then re-introduces the referent of the hold in the subsequent discourse. This is because a hold indicates that the referent is sustained in the speaker's mind (McNeill, 1992). We calculated the proportion of gestures with a single-handed stroke and a simultaneous hold in which the referent of the hold was represented in any previous gesture, and which the referent of the hold was represented in any subsequent gesture, by dividing the number of occurrences of each type by total number of gestures with a single-handed stroke and a simultaneous hold (Table 5). The referent of a hold can be either a landscape or a protagonist. We excluded the participants who did not produce gestures with a single-handed stroke and a simultaneous hold from this analysis.

Table 5. The proportion (and the standard deviation) of the gestures with a single-handed stroke and a simultaneous hold in which the referent of the hold was represented in any previous or subsequent representational gesture.

Referent of hold is expressed in the....	9 years (N=7)	Adults (N=8)
previous representational gesture	0.37 (0.41)	0.98 (0.02)
subsequent representational gesture	0.28 (0.20)	0.84 (0.09)

Independent t-tests showed that the gesture-hold combination in which the referent of the hold was represented in previous context, was significantly higher in adults than in 9-year-olds for previous gesture, t (13) = 4.21, $p < 0.01$, $r = 0.76$, and for subsequent gesture, t (13) = 7.21, $p < 0.01$, $r = 0.90$.

This result indicated that adults tended to maintain a referent, which had already been introduced in the story, by using a hold gesture, and to use the referent in the subsequent context, but 9-year-olds did not.

3.7 Speech: clauses that re-introduce previously mentioned referents

We examined the management of referents in speech that were parallel to the maintained referent in a hold in gestures, which was discussed in the preceding section. To this end, we calculated the proportion of clauses in which a protagonist or a clause was re-introduced (Table 6). A clause was coded as a *re-introduced clause* if the referent(s) in the clause had already been introduced in any grammatical position prior to the immediately preceding clause, but not mentioned in the immediately preceding clause. When a referent is a protagonist(s), we focused on whether the referent(s) in the subject position had already been introduced in any grammatical position prior to the immediately preceding clause (following the definition of re-introduction in Yoshioka (2005)). When a referent is a landscape element, it could appear in any grammatical position.

An independent t-test showed that the proportion of clauses that have a re-introduced referent in adults was significantly higher than that in 9-year-olds, t (18) = 6.00, $p < .001$, $r = .82$. Taking account of the finding from the previous section, both gesture and speech tended to be related to the subsequent context in adults, but not in 9-year-olds. Again, this suggests that gesture and speech develop hand in hand at discourse level too.

Table 6. The proportion of clauses in which a referent was reintroduced

	9 years (N=10)	Adults (N=10)
The proportion of clauses including a re-introduced referent	0.13 (0.04)	0.26 (0.05)

4 Discussion

In order to examine how two-handed gestures change during childhood, we investigated the relationship between two-handed gesture and speech in terms of form and referents in 3-, 5-, 9-year-olds and adults.

There are two main findings from our study. First, two types of two-handed gestures (*gestures with a two-handed stroke* and *gestures with a single-handed stroke and a simultaneous hold*) showed different developmental changes. This pattern appears to stem from the fact that one type of two-handed gesture was found to decrease with age, and the other type to increase. The two-handed gestures in 3-year-olds were predominantly gestures with a two-handed stroke, and this type of two-handed gesture was less frequent in the 5- and 9-year-olds' gestures. However, gestures with a single-handed stroke and a simultaneous hold appeared to become more frequent in children of the latter age groups.

In terms of our second finding, two pieces of evidence appear to support the view that speech and gesture develop hand-in-hand. First, the adults encoded landscape elements more often in clauses and in two-handed gestures than children. Second, the adults were more likely to use both gesture and speech to track a referent by relating it to the previous and the subsequent context, but 9-year-olds were not. Again, this suggests that gesture and speech develop hand in hand at the discourse level too. These findings indicate that the ways in which information is packaged in gesture and in a clause are similar for a given age and develop together both at the sentence and discourse levels.

Our second finding is in line with the gesture-speech parallel development view (Colletta et al., 2015; Gullberg & Narasimhan, 2010; Nicoladis, 2002). This is in contrast with the tradeoff view (Acredolo & Goodwyn, 1988; Church & Goldin-Meadow, 1986; Iverson & Goldin-Meadow, 2005; Kidd & Holler, 2009). Thus, we propose that gesture and speech develop hand-in-hand in relation to two-handed gestures and

clausal packaging of event components in speech. Our findings also fit the broader literature suggesting that semantic coordination between gesture and speech changes with age at the clause level (Özyürek, et al., 2008) and at the discourse level (McNeill, 1992).

To summarize, the current results indicate that the speech-gesture parallel development view best characterizes how semantic coordination between speech and gesture in narrative develop after three years of age. However, this is not to deny that under certain contexts, gestures may supplement information not expressed in speech for a given age group.

In the development of gesture, as with other developmental domains, the same forms may have more than one function in later stages of development (Karmiloff-Smith, 1986). This phenomenon was found to relate to two of our study findings. First, the two-handed gesture appears to have only one predominant function at 3 years of age; to depict one protagonist. In contrast, the same form of gesture was noted to have different functions in 9-year-olds and adults; to depict one protagonist and one landscape in 9-year-olds, and to depict two different protagonists in adults. Second, 9-year-olds predominantly used the stroke-hold combination to depict a landscape element and a protagonist, but adults were noted to use this combination to depict not only a protagonist and a landscape, but also two protagonists. Thus, even when the form of gesture appears similar, adults used the two-handed gesture in functionally different ways. In other words, hand usage for gesture appears to become more multifaceted with age.

We focused on the development of two-handed gestures in this study, but did not investigate the different types of two-handed gestures. For example, one may symmetrically or asymmetrically use the two hands for a gesture (Enfield, 2004). Thus, future research could examine whether the symmetric or asymmetric two-handed gestures change with age, and if so, which aspect of language development is related to the two types of two-handed gestures.

The current study found that two types of two-handed gesture showed different developmental patterns, and that the semantic packaging in two-handed gestures and in clauses developed hand-in-hand. In future research, we aim to examine the types of cognitive and linguistic abilities that underlie such changes.

Acknowledgments: The data collection in this research was financially supported by Grant BCS-0002117 from the National Science Foundation. This work was supported by a postdoctoral fellowship awarded to the first author by the University of Warwick. We would like to thank Priya Silverstein for her help with our coding. We would also like to thank the Tomato Man project team (Asli Özyürek, Shanley Allen, Reyhan Furman, and Amanda Brown) for the opportunity to analyze their data.

References

Acredolo, L. P., & Goodwyn, S. W. 1988. Symbolic gesturing in normal infants. *Child Development*, 59, 450-466.

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii M. 2007. How language specific is early syntactic packaging of Manner and Path? A comparison of English, Turkish, and Japanese. *Cognition*, 102, 16-48.

Battison, R. 1978. *Lexical borrowing in American Sign Language*. Silver Spring, MD: Linstok Press.

Boyatzis, C. J., & Watson, M. W. 1993. Preschool children's symbolic representation of objects through gestures. *Child Development*, 64, 729-735.

Church, R. B., & Goldin-Meadow, S. 1986. The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, 23, 43-71.

Colletta, J-M., Guidetti, M., Capirci, O., Cristilli, C., Demir, Ö. E., Kunene-Nicolas, R. N., & Levine, S. 2015. Effects of age and language on co-speech gesture production: An investigation of French, American, and Italian children's narratives. *Journal of Child Language*, 42, 122-145.

Enfield, N. J. 2004. On linear segmentation and combinatorics in co-speech gesture: A symmetry-dominance construction in Lao fish trap descriptions. *Semiotica*, 149, 57-123.

Field, A. 2005. *Discovering statistics using SPSS* (2nd ed.). London: SAGE Publications.

Furman, R., Kuntay, A., & Ozyurek, A. 2014. Early language-specificity of children's event encoding in speech and gesture: Evidence from caused motion in Turkish. *Language, Cognition and Neuroscience*, 29, 620-634.

Gullberg, M., & Narasimhan, B. 2010. What gestures reveal about how semantic distinctions develop in Dutch children's placement verbs. *Cognitive Linguistics*, 21, 239-262.

Iverson, J., & Goldin-Meadow, S. 2005. Gesture paves the way for language development. *Psychological Science*, 16, 367-371.

Karmiloff-Smith, A. 1986. From meta-processes to conscious access: Evidence from children's metalinguistic and repair data, *Cognition*, 23, 95-147.

Kidd, E., & Holler, J. 2009. Children's use of gesture to resolve lexical ambiguity. *Developmental Science*, 12, 903-913.

Kita, S., Gijn, I., & van der Hulst, H. 2014. The non-linguistic status of the Symmetry Condition in signed languages: Evidence from a comparison of signs and speech-accompanying representational gestures. *Sign Language & Linguistics*, 17, 209-232.

Kita, S., & Özyürek, A. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16–32.

Kita, S., Ozyurek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. 2007. Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, 22, 1212-1236.

McNeill, D. 1992. *Hand and mind*. Chicago & London: The University of Chicago Press.

Mol, L., & Kita, S. 2012. Gesture structure affects syntactic structure in speech. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 761-766). Austin, TX: Cognitive Science Society.

Nicoladis, E. 2002. Some gestures develop in conjunction with spoken language development and others don't: Evidence from bilingual preschoolers. *Journal of Nonverbal Behavior*, 26, 241-266.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. 2008. Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, 44, 1040-1054.

Özyürek, A., Kita, S., & Allen, S. 2001. Tomato Man movies: Stimulus kit designed to elicit manner, path and causal constructions in motion events with regard to speech and gestures. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics, Language and Cognition group.

de Ruiter, J. P. 2006. Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech-Language Pathology*, 8, 124–127.

de Ruiter, J.P., Bangerter, A., & Dings, P. 2012. The Interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science*, 4, 232-248.

Sekine, K. 2009. Changes in frame of reference use across the preschool years: A longitudinal study of the gestures and speech produced during route descriptions. *Language and Cognitive Processes*, 24, 218-238.

So, W.C., Kita, S., & Goldin-Meadow, S. 2009. Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, 2009, 33, 115-125.

Yoshioka, K. 2005. *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse*. Nijmegen, Radboud University.