

The role of gesture in the language production of preschool children*

Kazuki Sekine

Japan Society for the Promotion of Science / National Institute of Informatics

The present study investigates the functions of gestures in preschoolers' descriptions of activities. Specifically, utilizing McNeill's growth point theory (1992), I examine how gestures contribute to the creation of contrast from the immediate context in the spoken discourse of children. When preschool children describe an activity consisting of multiple actions, like playing on a slide, they often begin with the central action (e.g., sliding-down) instead of with the beginning of the activity sequence (e.g., climbing-up). This study indicates that, in descriptions of activities, gestures may be among the cues the speaker uses for forming a next idea or for repairing the temporal order of the activities described. Gestures may function for the speaker as visual feedback and contribute to the process of utterance formation and provide an index for assessing language development.

Keywords: spontaneous gesture, context, growth point theory, preschool children

Introduction

Speakers of any age tend to produce gestures along with speech in conversational situations. In this paper, I describe an investigation into the role of spontaneous gestures produced by preschoolers in the language production process. I suggest that gestures not only serve to reinforce or supplement the accompanied speech but also to generate or repair conceptualizations.

The gestures studied in this investigation are those that are characterized by "an obligatory accompaniment of speech, a lack of language-defining properties, idiosyncratic form-meaning pairings, and a precise synchronization of meaning presentations in gesture with co-expressive speech segments" (McNeill, 2000, pp. 22–23). Thus, they are differentiated from other gestures, called *emblems*, that

are form-meaning pairings dictated by social convention and that make sense without accompanying speech. Such gestures are also differentiated from signs as in a sign language, where they are incorporated into a complete linguistic system. Hereinafter, the gestures studied here are referred to as simply “gestures”.

Since McNeill (1992) pointed out that gesture and speech form a single integrated system, many studies have focused on the function of gesture. Kita (2000) surveyed previous studies on the function of gesture and noted that gesture has multiple functions related to delivering substantial information and facilitating speech. He categorized these functions respectively as other-oriented and self-oriented.

The other-oriented functions are those which convey information to the listener and make communication smoother. Research has shown that both adult and child listeners can take a substantial amount of information from gestures and speech, and that even when the information is conveyed only in gestures, the listeners incorporate it into their response (Broaders & Goldin-Meadow, 2010; Cassell, McNeill, & McCullough, 1999; Goldin-Meadow, Wein, & Chang, 1992). Other research has shown that a speaker's gestures not only help regulate turn-taking in conversation (Kita, 2000), but also are shaped interactively by listeners (Bavelas, Chovil, Lawrie, & Wade, 1992; Kendon, 2004; Streeck, 1993). These findings support the view that gestures have other-oriented functions.

It also seems that gestures have self-oriented functions that facilitate language production and verbalization of thought, and reduce cognitive load. For example, studies have shown that speakers produced gestures even when they could not see their listener (Alibali, Heath, & Myers, 2001; Barroso, Freedman, & Grand, 1978; Rimé, 1982), suggesting that gestures may be produced also for the speakers' own benefit. Recent empirical studies have shown that gestures facilitate conceptualization when speaking (Alibali, Kita, & Young, 2000), help retrieval of words (Krauss, Chen, & Gottesman, 2000; Rauscher, Krauss, & Chen, 1996), and reduce cognitive load (Goldin-Meadow, Howard, Spencer, & Wagner, 2001). These findings support the theory that gestures have a self-oriented function.

In this paper, I propose that gesture can be seen as itself creating a context that can be used by the speaker. By subsuming context into the analysis of gestural function, I reveal a new aspect of gestural function that has been overlooked in the studies with an experimental approach.

To clarify this point, it will be useful to outline the growth point (GP) theory proposed by David McNeill (McNeill, 1992, 2005), which I am following in this study. McNeill offered growth point theory in order to explain the relationship between speech, gesture, and thought. A growth point is defined as the smallest idea unit encompassing the unlike semiotic modes of imagery and linguistic encoding. In other words, a growth point is a cognitive unit that includes two

different but inseparable semiotic modes, imagery and linguistic encoding, which eventually become gesture and speech respectively. It is generated when a speaker creates newsworthy or significant information based on contrast with the immediate discourse context (McNeill, 2000). Thus, GP theory sees the contrast or the differentiation from the antecedent context itself as the origin of a growth point.

Embodied gesture and speech have different semiotic values, so even when they convey a single piece of information, they represent that information in different ways. In speech, a concept is analytically segmented into words or phrases and organized in a linear fashion in accordance with grammatical rules. The whole meaning of a linguistic expression (such as a discourse, sentence, or word) is compositionally formed by a number of smaller elements, and each of which have their own independent meanings. In contrast, in gesture, a conceptualization is synthetically represented all at once without being restricted by linguistic code. A gesture is global in that the meaning that part of a gesture (such as a hand shape, orientation, trajectory, or motion pattern) conveys depends on the meaning of the whole gesture (McNeill & Duncan, 2000). Even though gesture and speech have different semiotic values, they emerge from a common origin (a growth point).

McNeill explains the emergence process of gesture and speech by using the phrase “unpacking”. Unpacking is “the process of articulating the implications of the core idea and using these implications as a guide towards a well-formed surface structure” (McNeill, 2000, p. 29). In other words, this is a process in which imagery is abstracted to fit the linguistic category such as a word, a phrase, or a construction, and at the same time, the linguistic categories are embodied along with the imagery. Through this unpacking process, imagery and linguistic categories are formed into gesture and speech on the surface structure. According to McNeill “semiotically a combination of opposites, image and form, the growth point creates a benign instability that fuels thought and speech” (2005, p. 105). Thus, to resolve the instability, imagery and linguistic categories are dialectically activated from the starting point until they are turned into gesture and speech as a stable structure. This unpacking process is always involved with the context. Even in the course of the process from a growth point to a surface form, new meaning could be created. This is because language production is a progressive process in which a speaker’s current expression by gesture and speech creates a context which affects his subsequent (the next or the current unit of) expression.

The assumption of the unpacking process in GP theory contrasts with information processing speech production models based on modular processes, as presented by Levelt (1989), and with the gesture production model built on the speech production model (e.g., de Ruiter, 2000; Krauss et al., 2000). In the information processing approach, the entire process from conceptualizing an idea to formulating speech and gesture progresses linearly, and it is assumed that there is

no interaction between modules and that the context does not affect the modules while information is being processed. This is because contexts or meanings exist outside of and before “language”, and they are treated as input to language. McNeill (2005) states that in the information processing approach, “context can be represented only as a data source, like world knowledge or inputs from the physical environment, viz., it can be handled statically, but cannot be treated dynamically or embodied in the conceptual organization of the utterance, since doing so would render structures unstable and open the module to influences outside the allowed inputs from other modules, and hence undermine its very modularity” (pp. 132–133). Thus, in this approach, gesture and speech are formulated as sequential reactions to inputs from other modules.

From the standpoint of GP theory, in contrast, currently produced gesture and speech serve as a context for the next unit of the utterance to be produced and subsequent utterance construction may be modified accordingly.

In this study, I investigated the speech facilitative function of gesture in the preschool period by focusing on the context in which gesture and speech occur. There are two reasons for focusing on this period. First, fewer gesture studies have been conducted on preschoolers or school-aged children than on infants or adults (McNeill, 2005). Especially since there has been little study of the self-oriented functions of gesture in children, it is important to investigate how gesture is involved in the language production process in early development. Second, the preschool period is when children acquire the ability to relate and to describe activities. Children begin to recount their activities with the support of their parents at around two years old. By the age of five or six, they can relate multiple activities causally and temporally, and the characteristics that are seen in adult narrative are acquired during this period (Berman & Slobin, 1994; Küntay & Nakamura, 2004). At the same time, as illustrated in the following example 1, the postposing of part of the sentence elements is characteristic of Japanese preschooler’s speech (Uchida, 1996).

Example 1

Toto mo mata kaetta no, kingobachi no tokoro he
 Toto too again come:PAST NE, fishbowl DET place DET
 “Toto (name of the fish) has gone again, to his fishbowl”

The preschool period is also a time when gesture begins to be integrated with language production (Goldin-Meadow & Butcher, 2003). Observations have shown that the frequency of gestures of children becomes similar to that of adults by the last half of the preschool period (McNeill, 1992). Furthermore, studies on the viewpoints that are reflected in gestures revealed that, when children of early preschool

age describe the actions in an activity, they tend to produce enactive gestures from the actor's viewpoint. As they become older, they gradually come to produce gestures that depict the activity from the observer's viewpoint, so their gestures become similar to those of adults during their preschool period (McNeill, 2005)

In this study, I asked children to describe how they slide down a playground slide. This is an activity that even preschoolers can easily describe from their own experience, and when they describe it they use a relatively large number of gestures (Sekine, 2010). More importantly, the sliding activity has clear sequential organization. That is, the child first climbs up the stairs or ladder and then slides down the incline. When speakers talk about an event with a sequence of actions, such as slide play, they can create a contrast of two different actions in a descriptive narrative. This is useful in analyzing the relationship between gesture and context, and in addition, the role that gesture plays in language production when a child describes the activity along a timeline.

Method

Participants

Fifty-two preschool children participated in this study: 15 four-year-olds (9 boys and 6 girls), 17 five-year-olds (8 boys and 9 girls), and 20 six-year-olds (14 boys and 6 girls). The average ages were 4;0 (3;7–4;6), 4;11 (4;7–5;6), and 6;0 (5;7–6;6) respectively. All the participants were native monolingual Japanese speakers from middle-class families who attended a public kindergarten in the Kanto region, Japan.

Procedure

All interviews were conducted individually in a quiet spare room at the school. The child sat in an armless chair, and the experimenter sat in a chair facing the child. First, the experimenter asked the child to tell his name, age, and past kindergarten events such as sports day or a field trip to warm him up and ease his tension. Then, after confirming that the child had experienced playing on a slide, the experimenter asked, "I have never seen a slide, can you tell me what a slide is?" All the children had experienced playing on a slide and could answer the question. To capture their gestures and speech, the interviews were recorded with a camcorder (Sony HDR-HC9), which was positioned at a 45 degree angle from the child's facing to the experimenter.

Speech and gesture coding

All the narratives were transcribed verbatim by the author. For the speech coding, from each transcription, the total number of morphemes (excluding fillers such as “uh” and “ah”, hesitations, and speech errors) were calculated.

I then identified the spontaneous gestures by reviewing the video recordings with a movie player, “A player”.¹ Hand movements were classified as gestures only if they had an identifiable beginning and a clear end and were synchronized with the speech. Then, to determine the viewpoint the children took in the narrative, the gestures were categorized into three types, according to the distinctions made by McNeill (1992). If the speaker reenacted an action that took place during the activity by using the same body part used in the actual activity, the gesture was coded as a “character viewpoint” (C-VPT) gesture. If a speaker depicted the contour, shape of the slide, or action involved with slide from the observer’s perspective, the gesture was coded as an “observer viewpoint” (O-VPT) gesture. In these gestures, the way the hands were used differed from the way they are used in actual play. If the gesture depicted the activity from two different viewpoints simultaneously, such as the speaker stretching out her legs while using her hands to show the slide angle from a lateral view, the gesture was coded as a “dual viewpoint” (D-VPT) gesture.

Reliability

To ensure the reliability of the speech and gesture coding, after all data were coded by the first author, they were re-analyzed independently by a Japanese graduate student who had been instructed in the categories employed in this study. Agreements between the coders were calculated for each index. The two coders agreed more than 94% of the time on all speech indices including the number of morphemes, number of gestures, and gesture viewpoint. The coding disagreements were resolved through discussion.

Analysis

First, the duration of speech and the number of gestures were calculated. Then, if both action components, climbing up and sliding down, were mentioned, the order in which they were mentioned was determined. Next, the relationships among the gesture viewpoint, the component order, and the expressions used for describing the components were analyzed. Finally, in the light of these observations, a proposal is made, in each case, as to how gestures may have contributed to language production.

Results

Expressions used and amount of gesture

One six-year-old boy and two five-year-old boys did not produce any gestures, so their results were excluded from the analyses. The total number of gestures, the number of gestures for each viewpoint, and the number of morphemes produced were calculated (Table 1). An analysis of variance (ANOVA) on each dependent variable with age group as a between-subjects factor showed that there was no main effect of age on any dependent variable. This indicated that the number of gestures, the gestural viewpoint, or the morphemes produced did not change with age in this group of preschool children.

Table 1. Mean number of gestures, gestures for each viewpoint, and morphemes (SD)

	4 years	5 years	6 years	F value
Mean number of all gestures	2.4 (1.5)	2.5 (1.1)	2.7 (1.4)	.19
Observer viewpoint	2.2 (1.5)	1.9 (1.0)	2.2 (1.8)	.18
Character viewpoint	0.1 (0.4)	0.6 (0.9)	0.4 (0.9)	1.37
Dual viewpoint	0.1 (0.3)	0 (0)	0.1 (0.2)	.46
Mean number of morphemes	10.0 (8.7)	14.4 (5.4)	14.5 (8.4)	1.66

Next, to examine how many children mentioned both main action components, climbing up and sliding down, the number of children who mentioned both using speech and/or gestures was counted. Ten of the 15 four-year-olds, 11 of the 15 five-year-olds, and 15 of the 19 six-year-olds mentioned both components. In each age group, therefore, two-thirds of the children mentioned both components regardless of the order of mentioning the activities. A third of the children in each age group mentioned only the sliding-down component. There was no significant relationship between age group and the number of children who mentioned both components (χ^2 (2, $N=49$)=0.65, n.s.). Thus, all children mentioned at least the sliding-down component, which is the more central and important information in describing the sliding activity; there was no age difference in the proportion of the number of components mentioned.

The order in which the components were mentioned

Focusing on only those children who mentioned both components, I investigated the order of mentioning them. Those who mentioned them in chronological order are referred to as the *chronological order group*. Those who mentioned them

in non-sequential order, like mentioning sliding-down first, then moving on to climbing-up, are referred to as the *reverse order group*.

Table 2. Number of children who mentioned components in the chronological and reverse order group

Age	Chronological	Reverse	Total
4 years	4	6	10
5 years	7	4	11
6 years	14	1	15
Total	25	11	36

As shown in Table 2, 11 of the 36 children who described both components with either speech or gesture described them in reverse order. To determine the relationship between the number of children who explained them in reverse order and their age, a Chi-square test was conducted. The result indicated that there is a significant relationship between them ($\chi^2 (2, N=36) = 8.30, p < .05$). To identify the differences among age groups, a residual analysis was used. It showed that a significantly larger number of four-year-olds than six-year-olds described them in reverse order. Evidently, then, four-year-olds tend to begin by explaining the central action of the activity while six-year-olds tend to explain the actions in chronological order.

This result indicates that young children tend to first arrive at a fragmentary spoken and gestured expression of the central action and then add linguistic elements around this image, rather than starting from the beginning or setting up the scene. This phenomenon, in which the discourse focal idea of the moment (which is the most significant action of the sequence, in this case of the slide) first appears in the surface utterance fragment before well-formed utterances are made on the surface structure, can be called a "bare expression of the growth point".

Eight out of the 11 children in the reverse order group (i.e. those mentioning sliding down first) again referred to sliding down after describing climbing up. An example of this is given below in Figure 1. These eight children produced the same hand shape of the gesture at the second mention as was produced at the first. By comparing the first with the second gesture depicting the sliding-down in detail, it may illustrate how the context including gestures and speech affect the growth point.

Let us look at the explanation of the boy, M, in detail (Figure 1 and Example 2). Here, and in subsequent examples, the square brackets in the verbal transcript represent the starting and ending points of the motion of the child's hands, the boldface marks represent the stroke phase of the gesture phrase, which is the phase that conveys substantial meaning, the underlining indicates a hold phase, the "*" represents

self-interruption, and the ‘:’ indicates elongated phonation. The letters in parentheses in the transcription indicating the place where a gesture occurred correspond to the letters labeling the pictures in Figure 1. The abbreviations used in the interlinear gloss include ACC (accusative), ADV (adverbial marker), ASP (aspect), COP (copula), FP (final particle), INJ (interjection), LOC (locative), NL (nominalizer), NOM (nominative), ONOM (onomatopoeic), PST (past), and TOP (topic marker).

Example 2

[*suberidait-te ne anone* (a)*kou yat-te* (b)*SUUNt-te*
 slide-and FP well like.this do-and ONOM-ADV.
 ‘A slide is, well, doing like this and’

[*it-te soshite*] [*soshite* (c)*kaidan nobot-te* (d)*SUUNt-te iku*] *yatsu*
 go-and then then stairs go.up-and ONOM-ADV. go thing
 ‘go like SUUN, then then go up the stairs, and it goes like SUUN’

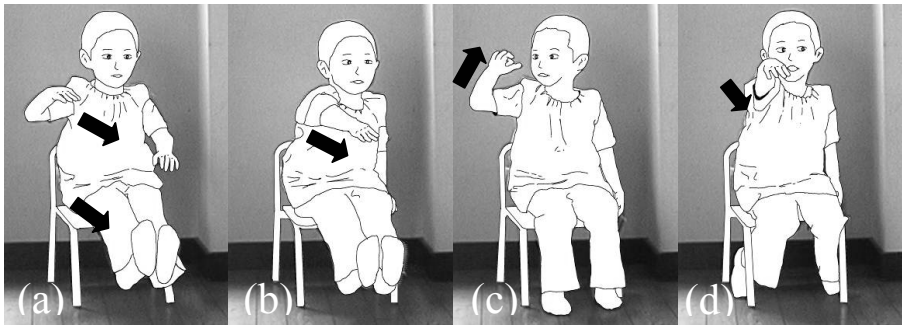


Figure 1. A reverse order explanation of sliding activity given by a four-year-old boy, M

1. M first assumed the posture that he takes when he slides down a slide, by raising both legs up from the floor and stretching them out in front of him. With this posture, he said “*kouyatte*” (“just like this”). At the same time, he lifted his right arm ((a) in Figure 1) as preparation for making a forward movement. Then, while saying “*SUUN*” (onomatopoeia² for going down), he gestured with his right hand to illustrate the contour of sliding ((b) in Figure 1). Thus, he gestured with his right hand to illustrate the contour of sliding, at the same time as he assumes a posture related to the one he would use when actually on the slide. He thus used a dual perspective (D-VPT) gesture: he illustrated the posture of sliding down from the character viewpoint with his whole body, and at the same time the contour of sliding down from the observer’s viewpoint with his right hand.

2. He paused briefly immediately after he said “*SUUNt-te it-te*” (“going like *SUUN*”). While doing this, he held the gesture made immediately before in the air. Then, he used a verb with a conjunctive particle (“*it-te*”) and a conjunction (“*soshite*”)

meaning “then” and moved to the explanation about climbing the stairs. He illustrated climbing the stairs by wiggling his right fingers facing downward ((c) in Figure 1), while saying “going up stairs” (“*kaidan nobotte*”).

3. Lastly, he depicted sliding down by using the same hand shape as used the first time ((b) in Figure 1) along with the onomatopoeia “*SUUN*”. However, this time ((d) in Figure 1), unlike the first time, he incorporated only the observer’s viewpoint.

As described above, he started his explanation from the sliding-down component, and he ended it by again explaining the sliding-down component. In light of the purpose of this study, the gesture used during the second explanation of the sliding-down component is interesting because it implies that a growth point is affected by the context after and/or before the growth point. From this standpoint, it is notable that the gesture used the second time with the elaborated speech had the same specific feature (hand shape) that was used the first time while at same time excluding the body posture element of the gesture used the first time.

In the second explanation of the sliding-down component ((d) in Figure 1), he did not show the movement of his lower body, and he did not stretch his arm less than the first time ((b) in Figure 1). The hand position used in the second explanation was also a little bit higher than in the first. However, the hand movement depicting the direction of movement forward was maintained. This implies that the second onomatopoeia “*SUUN*” had departed from the growth point in which a slightly different contrast was added from the first one. The “*SUUN*” and dual perspective gesture used in the first explanation indicate that he focused on the sliding down itself based on his experience of using a slide. On the other hand, in the second explanation, he focused on simply sliding down in contrast to climbing up because he had already explained the manner of the going-down component and had explained climbing up the stairs right before the second explanation of the sliding-down component. Therefore, it seems that he produced only gestures that emphasized the direction of movement as going down rather than the posture that he takes while sliding. Although the depiction of sliding down was accompanied by the same onomatopoeia in both cases, the emphasis varied depending on the context contrasted. The difference in emphasis can be seen clearly by looking at the gestures the speaker spontaneously produced.

Eight out of the 11 children in the reverse order group explained the sliding-down component again at the end of their narrative, and, during this second explanation, they used a gesture including the same elements (hand shape and movement) as used for the first mention, like M did. Thus, by looking at the child’s gesture accompanying the elaborated speech together with the context, as McNeill (1992) argued, we can speculate about what the growth point in a child’s expression was, and the contrast added to or excluded from it.

Next, I investigated the role the gestures themselves play in the process of correcting speech. That it may do so is illustrated by some observations which will now be described. These suggest that the child may on occasion use his gesture as a resource for checking something he has just said.

The visual feedback function of gestures

The aspect of gesture used as a resource for checking the produced expressions will be referred to as “visual feedback function”. Because a gesture has visibility, it gives the speaker a chance to check the correctness of the expression produced. For example, one five-year-old girl was observed to repair a verbal expression she had produced after glancing at her gesture. She corrected the mismatch between the semantic content in her speech and gesture, as shown in Figure 2 and Example 3.³

Example 3

suberidai wa anone [mazu ne (a) **kaidan kara** ori-te _____ ano
slide TOP well first FP stairs from go.down-and well
‘A slide is, well, first (I/you) go down from stairs, and well’

(b)*kaidan kara ne* (c)*nobot-te kara* suberidai ni (d)*SHUUt-te ne*
stairs from FP going.up-and after slide LOC ONOM-ADV FP
‘After going up the stairs, you go down the slide in the manner of *SHUU*’

naru no
become FP

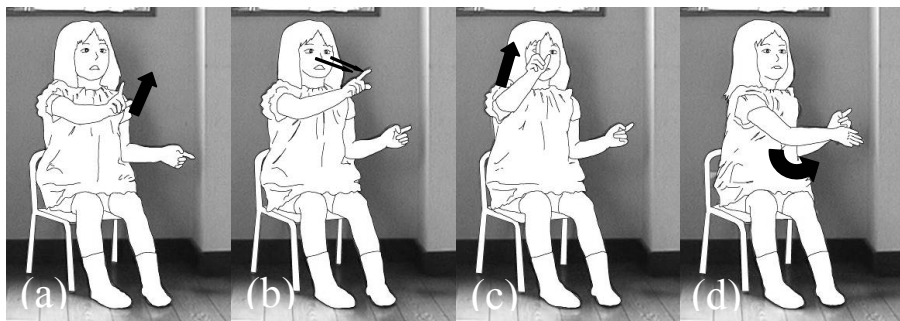


Figure 2. Correction of a mismatch between the semantic content of speech and accompanying gestures by five-year-old girl, O

1. She first tried to explain going up the stairs. This is indicated from her gesture going up to the upper right and the accompanying word “stair (*kaidan*)” ((a) in Figure 2). However, as she moved her hand to the upper right, she introduced the

particle marking a starting point of motion, such as “from (*kara*)” in her speech, and right after that, explained going down the stairs. It seems that she intended to explain going up but mistakenly said the verb phrase “going down”.

2. The selection of the verb “going down (*orite*)” contradicted her intention as understood from her subsequent behavior. First, the utterance “going down (*orite*)” did not accompany a gesture illustrating downward motion, and the hand remained in the air. Second, after the filled pause “well (*ano*)”, she repeated the previous word “from the stairs (*kaidan kara*)” and started repairing her mistake. When she mentioned stairs, she was glancing at her fingers slowly moving to the upper right ((b) in Figure 2, corresponding to speech in italics). It was in the preparation phase of the gesture. After glancing at her fingers, while she was illustrating upwards motion, she started moving her hand toward the upper right again. The hand shape of this gesture was the same as that of the gesture produced with the utterance “from the stairs (*kaidan kara*)”; however, the gesture reached a higher position than the previous one ((c) in Figure 2). These behaviors indicated that she had first intended to explain climbing up rather than sliding down.

3. After explaining the climbing up motion, she finished by using a gesture illustrating the motion of going down, sweeping her right hand down and to the left, while uttering the onomatopoeia “*SHUU*” ((d) in Figure 2).

What is interesting in this case is that she checked her hand shape and gesture movement in the preparation phase by glancing directly at her hands as if she needed to confirm that the combination of the corrected speech and gesture were not semantically mismatched this time. Immediately after glancing at her gesture, she changes it so that it conveyed the same semantic information as the accompanying speech. This implies that gestures can function as a resource for checking generated expressions.

Similar phenomena have also been previously reported. Streeck (1993, 2009) has described how an adult speaker revised her spoken expression after looked at her own gesture. Also, studies that have shown that performance such as recalling Chinese characters (Sasaki, 1987) or counting numbers (Graham, 1999) improved in young children when they looked at their own hand movement (or the shape thereof) may perhaps be considered as further, indirect, evidence supporting the visual feedback function of gestures.

“Verb incomplete expression” of sliding-down motion

Does the explanation of sliding down a slide vary depending on the order of explanation, chronological or reverse? I first examined the linguistic elements that were

used for explaining the sliding-down component in the children's explanation by focusing on the onomatopoeia they used.⁴

As shown in Figure 3, 13 of the children (52%) in the chronological group expressed the sliding-down component as “going down like *SHUU* (*SHUUte oriru*)” by combining a verb (V) and an onomatopoeia (ONOM) for going down. Five of them (20%) expressed it by using only a verb like “sliding (*suberu*)”, four (16%) used a combination of a verb and a noun (N) like “play equipment where (one) goes down (*subette asobumono*)”, two (8%) used expressions consisting of two verbs like “sitting and sliding (*suwatte suberu*)”, and one (4%) used a combination of an adnominal phrase (A) and a verb (“going down like this (*kouyatte suberu*)”).

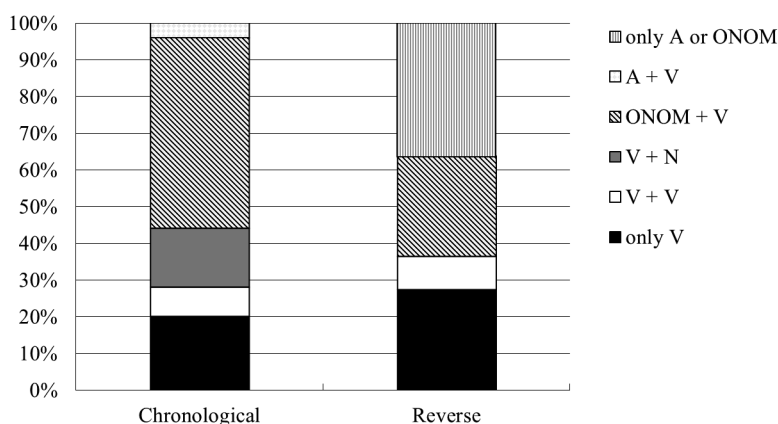


Figure 3. Proportions of linguistic elements used to explain sliding-down motion by group.

The most common expression used by the children in the reverse order group was an adnominal phrase or an onomatopoeia such as “like this” or “*CHUUte*” (four children, 37%). Three (27%) used a combination of an onomatopoeia and a verb, three (27%) used only a verb, and one (9%) used a combination of verbs (Figure 3). The most common expression for the sliding-down component used by this group, linguistic elements without verbs, was not used at all by the chronological order group. Eight of the 11 children in the reverse order group expressed climbing up and sliding down by using both gesture and speech for each action, and four of them ended with an explanation of the sliding-down component without using a verb, even though a sentence in Japanese is supposed to end with a verb. The narratives of these four children are shown in Figure 4 and Example 4 to 7.

Example 4: Four-year-old girl

[(a)*SHU:t-te*] [(b)*kaidan nobot-te* (c)*SU:t-te* ne yaru no]

ONOM-ADV. stairs go.up-and ONOM-ADV. FP do FP
 ‘(A slide) is going up stairs in the manner of *SHUU* and slide down in the manner of *SUU*’



Example 5: Four-year-old boy

[(a)*SURUSURU*-te ne (b)*nobot-te* (c)*SURUSURU*-te ne (d)*suberu* no]
 ONOM-ADV. FP go.up-and ONOM-ADV. FP slide FP
 ‘(I/you) go up like *SURSURU*, and slide like *SURUSURU*’



Example 6: Five-year-old boy

*Chi** [(a)*kou yat-te* (b)*kaidan natte-te*] *soshite nai nai*
*chi** like.this do-and stairs becoming-and then nothing nothing
 ‘*chi** (it) have stairs like this, then nothing nothing’

anone [(c)*te: tsukamaru tokoro ne* (d)*kaidan o nobotte-tte:*
 well hand grab place FP stairs ACC going.up-go
 ‘well, and there is a thing you can grab, and (you) go going up the stairs.’

(e)*tsukama-tte soshite* (f)*SHU:t-te*] *suberu*
 hold-and then ONOM-ADV. slide.down
 ‘while grabbing it, and then slide down like *SHUU*’



Example 7: Six-year-old boy

[(a)*SHU::t-te ne no* (b)*boru desho soshite suberidai o* (c)*suberu no*]
 ONOM ADV FP go.up COP then slide ACC slide.down FP
 ‘(I/you) go up the slide like *SHUU*, you know, then slide down (it)’

*de owat-te mata no** [(d)*nobot-te sorede* (e)*SHU::t-te*] *iku no*
 and finish-and again go* go.up-and then SHU-ADV. go FP
 ‘When it’s done, (you) climb up (it) again and then go (down) like *SHU*’



Figure 4. Examples of verb-incomplete expressions by four children in the reverse order group

The gestures used by these four children to illustrate the sliding-down component were co-produced with onomatopoeia or adnominal phrases. They then moved to the explanation of climbing up before they finished by explaining sliding down with a verb. As mentioned above, Japanese sentences have to end with a verb

phrase. At times, a child moves to the explanation of climbing up after explaining sliding down with an onomatopoeia or an adnominal phrase, without finishing the verb for sliding down. This type of utterance is termed a “verb-incomplete expression”.

The relationship between the number of children who used verb-incomplete expressions and the order of explanation were examined among those who expressed both the climbing-up and sliding-down components using gesture and speech (Table 3). Fisher's exact test showed a significant relationship between the use of a verb incomplete expression and the order of explanation (Fisher's exact test, $p = .003$). In contrast to the chronological order group, in which no child used a verb-incomplete expression, half of the children in the reverse order group used one. In the reverse order group, thus, there were a number of children who expressed two different motions with different linguistic components within a clause, such as “going up like *SHUU* (*SHUU ttene noboru desho*)”. *SHUU* is onomatopoeia for expressing something sliding down.

Table 3. Number of children who produced verb-incomplete expressions

	Verb complete	Verb incomplete	Total
Chronological	20	0	20
Reverse	4	4	8

Why were there children who produced a verb incomplete expression in the reverse order group? It is possible that the children noticed the necessity of explaining going up before explaining going down, so their sentences were incomplete. If so, what was the clue to change the spoken expression? Looking at the explanations of the four children who used a verb incomplete expression, we see that they gestured while first explaining sliding down. In addition, the gesture accompanying the first linguistic element was completed. In other words, the gesturing did not stop until it was completed. Immediately after that, the children moved on to explaining climbing up and produced the repaired gesture and speech expression, which was a verb-completed version of the utterance fragment that preceded the description of climbing up.

Regarding the four children who produced a verb-incomplete expression, the stroke phase of their gestures started at the same time they started speaking or immediately before they started speaking (Figure 4). For the other four children, who did not use a verb-incomplete expression, the beginning of speech preceded the gesture stroke.

These observations suggest that the gesture depicting the motion of sliding down was the clue that caused the children to notice that they did not explain the climbing-up component and the necessity of explaining it first. This means that

not only the production of speech, but also the production of gesture contributes to creating a contrastive context, which in turn leads to the construction of a new meaning. Information expressed in the gesture-speech ensemble (to use a phrase introduced by Kendon (2004)) becomes the background or context, and the new contrastive meaning or information that needs to be conveyed, that is, the growth point, is generated in the speaker's mind from it. Gesturing with speech makes the speaker notice not only what was expressed but also what was not expressed at the time of gesturing, i.e., what should have been expressed. In the sliding case examined here, producing the combination of a sliding-down gesture and an onomatopoeic or adnominal phrase caused the children to notice that they should explain climbing up as well.

This is evident in the results shown in Figure 4. To illustrate the sliding-down component, four children who produced a verb incomplete expression produced two sliding-down gestures in their descriptions. The components of the sliding-down gestures the second time (such as the direction, hand shape, body parts used, and space covered) were included in the sliding-down gestures the first time. Moreover, new elements were added. This means that new contexts were added to the second explanation even if the speaker used the same spoken expression, "going down (*suberu*)". This characteristic is common to the gestures produced by children M and O shown in Figure 1 and 2. This function of gesture, creating the contrastive context itself, will here be called the "context-creating function of gesture". This function was observed not only in the sliding explanation but also in a narrative of a past experience, as we explain next.

The following case comes from a warm up session. Figure 5 shows a scene in which a six-year-old boy, N, is explaining a dance in which he participated at his kindergarten's sports day after being asked about it by an experimenter. In the dance, he turned with other children while holding an arch-shaped stick. In his explanation, he tried to explain how he turned, which is a central action in this dance. However, every time he tried to explain it, he stopped his gesture and speech in the middle of his turning explanation. He seemed to notice that he had to explain the information relevant to the turn motion before finishing the explanation of turning. After several attempts at repair, he eventually completed the explanation of turning that he apparently initially intended to convey. Let us look at his narrative in detail (Figure 5 and Example 8).

Example 8

[(a) *mawa-* *] *da chigau* [(b) *kou yat-te*] [*nanka* (c) *kouiu nanka*]
 Tur- INJ no like.this do-and well like.this something
 '* no no, (it) do like this, well something like this'

[(d) *a:chi mitai no ga at-te sorede minna de* (e) *mawa* mo**
 arch like NL NOM is-and then everyone with tur* ho*
 'there is something like an arch, and then, while everyone was tur* ho*']

(f) *mochi-nagara mawatte tara (hee omoshirokat-ta?) omosihorkat-ta]*
 grab-ASP turn COND INJ fun-PST fun-PST
 'turning with that in your hand, ("oh was (it) fun?") (it) was fun?'

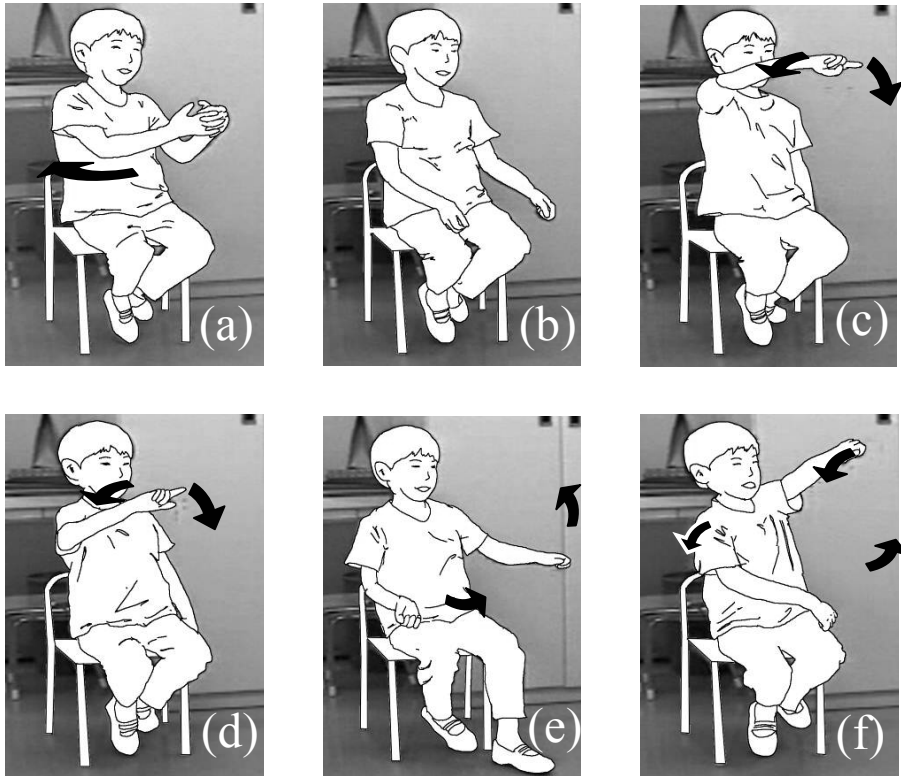


Figure 5. Example of gesture and speech repair by six-year-old boy N

1. First, to explain the turn as a central action in the activity, he tried to enact it by turning his upper body rightward and parallel to the ground, while putting his hands together. However, in the middle of saying "turning", he stopped his gesture ((a) in Figure 5).

2. After negating his own expression by shaking his head and saying "*no (chigau)*", he showed the posture that he took when he turned in the dance while saying "*do like this (koyatte)*" ((b) in Figure 5). However, he interrupted himself without completing a verb phrase. He seemed to notice the necessity of explaining, by

checking the hand shape and movement of his gesture, not only “turning” itself but also turning while “holding” something.

3. To explain the shape of the thing he held in his hands, he explained that what he held was an arched-shape thing by tracing an arch in the air in front of him and saying “something like this (*kouiunanka*)” and “a thing like an arch (*a:chi mitaina*)” ((c) and (d) in Figure 5).

4. Next, exactly like what he did at first, he tried to explain the turning motion. However, again, he stopped the explanation, “*tur (mawa-)*” in the middle. At that moment, he seemed to notice that he had to explain how he turned with the stick ((e) in Figure 5).

5. Finally, he depicted the turning motion by repairing his speech to “turning while holding (*mochinagara mawattetara*)” and rotating both his hands many times rightward while tilting his trunk in the same direction ((f) in Figure 5). Because the experimenter asked him a question in the middle of his speech, his sentence was unfinished, but the verb “turn” was completed.

It is noteworthy that every time he produced the verb “turn” (*mawaru*), the verb was accompanied by a gesture, and that every time he made a repair, the corresponding part of the gesture also changed. Although the first expression “*tur- (mawa-)*” accompanied a horizontal rotation of his upper body including his arms, the movement of his hand in an arch was added to the gesture in the second expression “*tur- (mawa-)*”. Finally, the element of rotation around the vertical axis was added to the gesture. Thus, while the gesture that illustrated “turn” maintained the element of rotation consistently throughout the explanation, other elements such as relevant object, manner, or path of motion were added to or removed from the gesture in the process of explanation.

In this case, the imagery of turning and the verb “turn” (*mawaru*) as the central action of the activity seemed to be generated as a growth point first. Boy N tried to extract it using gesture and the verb “turn”. This is a bare expression of growth point, as mentioned above. Subsequently, each time he made a repair, he added new elements not only to his speech but also to the accompanying gesture, finally completing the explanation of dance.

In the explanation process, what made him notice the necessity of correcting his expression seemed to be the gesture and speech that he himself produced. In other words, the gesture and speech could be the clue that made him notice what he had already explained and at same time what he had still to explain. This observation suggests that gesture has a context-creating function.

Discussion

To investigate the use of self-oriented functions in preschoolers, I examined the gestures and speech they used to explain sliding down a slide. From the standpoint of growth point theory (McNeill, 1992), I examined how gesture-speech ensembles which are created on the basis of the contrast with the context, influence the ongoing language production. From the results, I identified two functions of gesture: visual feedback and context creation. These gestural functions contribute to language production.

Observation of how children describe an activity involving multiple actions revealed that four-year-olds tend to start their explanation from the most important information at the moment of speech or at the central action of the activity. Because playing on a slide in a playground is a circular activity, one might think that the reverse order explanations collected in this study were not due to the children starting their explanation from the central action but simply from the middle of the activity, so that this phenomenon is specific to this activity. However, given that the number of children who explained the actions in the reverse order decreased with age and that young children tend to start from the central action when describing other activities as well, the reverse order explanation can be considered characteristic of narratives produced by children who are just learning to describe an activity in the chronological order. The six-year-olds were much better at explaining the actions sequentially. This hypothesis is supported by findings of a previous study showing that children as young as five and half years old can produce narratives that follow the time course of an activity (Uchida, 1996).

McNeill explains the process of growth point evolution as follows. “Simultaneously representing the same idea unit in opposite modes creates instability, a ‘benevolent instability’ that is resolved by accessing forms on the static dimension — constructions and lexical choices, states of repose par excellence. The GP and the unpacking of it into construction and lexical items is how the two dimensions of language combine — the unstable growth point, itself a combination of imagery and linguistic content, is unpacked into an increasingly well-formed, hence increasingly stable, structure on the static dimension” (McNeill, 2005, p. 18). From this standpoint, the reverse order explanations given by many of the four-year-olds can be interpreted as a phenomenon in which a speaker immediately brings a growth point (or a few linguistic elements) to the surface before the growth point is unpacked into a well-formed stable expression. I call this phenomenon a bare expression of a growth point. Bare expression of growth points may be characteristic of language acts that are frequently seen during early preschool years.

There were children in the reverse order group who used verb-incomplete sentences to explain climbing up before ending with a verb. A noun phrase or

adnominal phrase appearing after a verb phrase in a Japanese sentence has been typically analyzed as a postposition expression and specific to spoken language, and it has been interpreted as a failure of planning (Clancy, 1982) or a demonstration of supplementary information (Kuno, 1978). Although the use of postposition expressions by children has been interpreted as underdeveloped planning, this interpretation cannot explain why they are used. This study suggests that gesture plays an important role as a clue for the production of postposition expressions and that such expressions may be related to a function of gesture.

Two functions of gesture that contribute to the language production process were suggested. One is the visual feedback function. A speaker can ascertain the adequateness or correctness of content expressed by a gesture by looking at it. That is, because of their visibility, gestures can be exploited by a speaker as a resource for checking expressions. The other is the context-creation function. Gestures with speech not only represent the meaning of a referent or event but also form the background or context that contrasts with the meaning. In other words, because gestures represent central actions in the activity, they tell the speaker what information has not been represented and needs to be conveyed, such as the background, peripheral action in the event or grammatically necessary linguistic elements. A speaker can thus generate a new growth point or correct his or her expression by using gestures that make him or her notice that some change in information is needed.

Putting these two functions together, we can think of the role of gesture in speech as follows. A speaker extracts imagery of an activity or object by using a linguistic category to bring a growth point to the surface structure. In this process, imagery and linguistic categories are embodied as gesture and speech, respectively, through dialectical interaction. The gesture produced becomes visible to the conversation participants including the speaker, who can confirm what was expressed if needed. At the same time, the gesture creates a contrastive context indicating the elements that have not been expressed or need to be expressed. It ends with the generation of a new growth point or with the correction of the expression. From this, it seems that gesture can be used in the language production process as a resource to monitor expression and generate new conceptualizations.

McNeill (1992) argued that “a mechanism by which gestures can affect thought is adding, dropping, or changing contrast” (p.251) against the context that the speaker establishes internally. The gestural functions that were suggested in this study support this view. Such gestural functions are more compatible with growth point theory, suggesting that new meaning is generated by taking in the context of speaking in the message production process, rather than with information processing theory, which excludes the effect of the context on the language production. Duncan (2008) found that speech errors of substitution tend to occur when a speaker must alter his or her use of a gesture feature that has heretofore associated

with references to the character in their narrative. This implies the limitation of the information processing model, which has no obvious place for gestural representation that can be the impetus for the errors in lexical access. She concluded that gestures have an impact on moment-to-moment language production. The findings of the present study support her conclusion.

The functions of gesture have been investigated in previous studies, mainly in terms of self- vs. other-oriented functions. However, considering that gesture can be part of a speaker's thinking process and that, at the same time, it can be seen as a potential representation of the speaker's intention for conversation participants, it may be better to think of gesture having multiple functions (Goldin-Meadow, 2003; McNeill, 2005). McNeill (2005) argued that a meaning cannot be made without a carrier for it, and that gesture and speech can be used to construct meaning as the material carrier (Vygotsky, 1962). In light of this view, the production of gesture itself contributes to the expression of thought as well as the production of speech.

How and when is the relationship between speech and gesture established? Previous studies have suggested that gesture and speech become a single integrated system due to the incorporation of gesture into the language production system during the one-word speech period (Goldin-Meadow & Butcher, 2003). However, the synchronization between gesture and speech at the word level does not always suggest a dialectic relationship. Children start narrating past experiences or events with multiple linguistic elements from the end of the one-word speech period. It seems that a dialectic relationship between speech and gesture is established in the process of learning how to extract imagery by using linguistic category during the time when children start narrating with multiple words. Gestures also start to take on the functions of visual feedback and context creation at this time. In traditional developmental psychology, gesture has often been seen as subordinate to or a complement of speech, and it has been thought that, as speech becomes predominant as a communicative means, gesturing becomes less frequent or eventually disappears (cf. Sekine, 2010). However, this study showed that the number of gestures did not decrease even in the later years of preschool. Furthermore, an increase in gesturing was observed in the descriptions of routes by preschoolers (Sekine, 2009) and in the narrations of animated stories by elementary school children (Sekine & Furuyama, 2010). These findings suggest that gestures and speech develop together and that gestures do more than simply convey information.

Few studies have examined changes in gesture and speech after the two-word speech period. An important task for future research is investigating the changes in gestural functions with age and their relationships to speech. McNeill (2005) argued that the principles used to explain mental growth in children, such as differentiation, internalization, dialectic, and reorganization, can be used to explain the online processing of language production in adults, pointing out the

similarity between mental growth processes and language production processes. Future work should also include trying to determine whether the phenomena observed in this study are present in adult narratives as well, and, if so, under what circumstance they occur.

The results of this study showed not only that the production of gesture contributes to the language production process but also that gesture can be a useful index for predicting the development of speech structure. Even if a speaker could express all the elements in the referent, he or she may not be able to place them in the chronological order. The information that a speaker intends to convey is not always fully represented in speech. The results of this study indicate that, to reveal the language production process, the thought that underlies it, and the developmental process, we need to focus on not only speech but also speech and concurrent gestures together.

Acknowledgements

I am sincerely grateful to David McNeill, Susan Duncan, Mika Ishino, and Adam Kendon for their insightful comments on earlier versions of this manuscript. I am also grateful to the children and their parents for their generous cooperation.

Notes

* This article is a revised version of the earlier version published in the *Journal of Japanese Qualitative Research in Psychology*.

1. <http://www.vector.co.jp/soft/win95/art/se182771.html>.
2. Onomatopoeias are conventionalized mimetic expressions of natural sounds and are lexicalized in the Japanese language. They are widely used regardless of the age and the communicative context (Shibatani, 1990).
3. This case is interesting in terms of the speech error itself. In growth point theory, a speech error is interpreted as “false” in order to segment a group of concepts or imagery representations as a unit into time components (McNeill, 1987). Looking at O’s speech error in light of growth point theory, although she conceptualized the sequence of sliding actions as a unit, she failed to segment it into linguistic elements. In other words, she apparently failed to extract the imagery. As a result, there was a mismatch between her expression (“going down”) and the accompanying gesture (illustrating “climbing up”).
4. The linguistic elements used to describe the sliding-down motion by the children who mentioned only sliding down were distributed similarly to those used by the children in the chronological order group.

References

- Alibali, Martha W., Heath, Dana C., & Myers, Heather J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169–188.
- Alibali, Martha W., Kita Sotaro & Young, Amanda J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, 15, 593–613.
- Bavelas, Janet B., Chovil Nicole, Lawrie, Douglas A., & Wade Allen (1992). Interactive gestures. *Discourse Processes*, 15 (4), 469–489.
- Broaders, Sara C. & Goldin-Meadow, Susan (2010). Truth is at hand: How gesture adds information during investigative interviews. *Psychological Science*, 21 (5), 623–628.
- Baooros Felix, Freedman Norbert, Grand Stanley, & Van Meel, Jacques (1978). Evocation of two types of hand movements in information processing. *Journal of Experimental Psychology*, 4, 321–329.
- Berman, Ruth A. & Slobin, Dan I. (Eds.) (1994). *Development of linguistic forms: English. Relating events in narrative: a crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cassell, Justine, McNeill David, & McCullough Karl-Erik (1999). Speech–gesture mismatch: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics and Cognition*, 7, 1–33.
- Clancy, Patricia M. (1982). Written and spoken style in Japanese narratives. In D. Tannen (Ed.), *Spoken and written language: exploring orality and literacy* (pp. 55–76). Norwood, NJ: Ablex.
- de Ruiter, Jan Peter (2000). The production of gesture and speech. In David McNeill (Ed.), *Language and gesture* (pp. 284–311). Cambridge: Cambridge University Press.
- Duncan, Susan (2008). Gestural imagery and cohesion in normal and impaired discourse. In I. Wachsmuth, M. Lenzen, G. Knoblich (Eds.), *Embodied communication in humans and machines* (pp. 305–328). Oxford: Oxford University Press.
- Goldin-Meadow, Susan., Nusbaum Howard, Kelly Spencer D., & Wagner Susan (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, 12 (6), 516–522.
- Goldin-Meadow, Susan (2003). *Hearing gestures: how our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, Susan & Butcher, Cynthia (2003). Pointing toward two-word speech in young children. In Sotaro Kita (Ed.), *Pointing: where language, culture, and cognition meet* (pp. 84–107). Mahwah, NJ: Lawrence Erlbaum Associates.
- Goldin-Meadow, Susan, Wein, Debra, & Chang, Cecilia (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction*, 9, 201–219.
- Graham, Theresa A. (1999). The role of gesture in children's learning to count. *Journal of Experimental Child Psychology*, 74, 333–355.
- Kendon, Adam (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.
- Kita, Sotaro (2000). How representational gestures help speaking. In David McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.
- Krauss, Robert M., Yihsiu Chen, & Gottesman, Rebecca F. (2000). Lexical gestures and lexical access: A process model. In David McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge: Cambridge University Press.
- Kuno, Susumu (1978). *The grammar in discourse [Danwa no bunpo]*. Tokyo: Taishūkan.

- Küntay, Aylin C. & Nakamura, Keiko (2004). Linguistic strategies serving evaluative functions: A comparison between Japanese and Turkish narratives. In S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative*. Vol. 2: *Typological and contextual perspective* (pp. 329–358). Mahwah, NJ: Lawrence Erlbaum Associates.
- Levelt, Willem J. M. (1989). *Speaking*. Cambridge, MA: MIT Press.
- McNeill, David (1987). *Psycholinguistics: a new approach*. New York: Harper & Row.
- McNeill, David (1992). *Hand and mind*. Chicago: University of Chicago Press.
- McNeill, David (2000). Growth points, catchments, and contexts. *Cognitive Studies*, 7, 22–36.
- McNeill, David (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- McNeill, David & Susan Duncan (2000). Growth points in thinking-for-speaking. In David McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge: Cambridge University Press.
- Rauscher, Frances H., Robert M. Krauss, & Yihsiu Chen (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, 7, 226–231.
- Rimé, Bernard (1982). The elimination of visible behavior from social interaction: Effects on verbal, non-verbal and interpersonal variables. *European Journal of Social Psychology*, 12, 113–129.
- Sasaki, Masato (1987). Why do Japanese write characters in space? *International Journal of Behavioral Development*, 10 (2), 135–149.
- Sekine, Kazuki (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.
- Sekine, Kazuki (2010). *Jihatsuteki-miburi kara miru youjiki no shiten henka* [The developmental change of perspective-taking in preschool age in terms of spontaneous gestures]. Tokyo: Kazama Shobou.
- Sekine, Kazuki & Furuyama, Nobuhiro (2010). Developmental change of discourse cohesion in speech and gestures among Japanese elementary school children. *Rivista di psicolinguistica applicata*, 10 (3), 97–116.
- Shibatani, Masayoshi (1990). *The languages of Japan*. Cambridge: Cambridge University Press.
- Streeck, Jürgen (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs*, 60 (4), 275–299.
- Streeck, Jürgen (2009). *Gesturecraft: The manu-facture of meaning*. Amsterdam: John Benjamins
- Uchida, Nobuko (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.
- Vygotsky, Lev S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Author's address

Kazuki Sekine
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
Tokyo 101-8430
JAPAN
sekine@nii.ac.jp

About the author

Kazuki Sekine is a research fellow at the Japan Society for the Promotion of Science, and is working at the National Institute of Informatics, JAPAN. His research interest is the development of gestures in children. He is investigating when idiosyncratic gestures emerge and how culture influences the production of gestures. He is exploring these issues in terms of the relationship between narrative development, spatial cognition, and culture.