

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/43008>

Please be advised that this information was generated on 2017-12-06 and may be subject to change.

# Gestural introduction of Ground reference in L2 narrative discourse\*

KEIKO YOSHIOKA AND ERIC KELLERMAN

## *Abstract*

*In the field of second language acquisition (SLA) and use, learners' gestures have mainly been regarded as a type of communication strategy produced to replace missing words. However, the results of the analyses conducted here on the way in which Dutch learners of Japanese introduce Ground reference in speech and gesture in narrative show that the two modes of expression are closely related in L1 as well as in L2. First, cross-linguistic variation is observed in both modes of expression, with a tendency for native speakers of Japanese to allocate on-line attention to Ground in both speech and gesture. Second, Dutch learners of Japanese prefer to adopt rhetorical styles more similar to their L1 than the L2 target, and the accompanying gestures are more L1-like. Transfer of L1 to L2 in narrative and the relationship between speech and gesture will be discussed.*

## **1. Introduction**

The present study examines the way in which Dutch learners of Japanese introduce Ground reference in speech and gesture as compared to native speakers of those two languages. Of particular interest here is what learners do when their L1 and the L2 differ typologically. Slobin (1996) suggests that speakers of "satellite-framed" languages and "verb-framed languages" (Talmy 1991) differ in the rhetorical styles they adopt to introduce Ground reference, and that the differences reflect cross-linguistic variation in allocation of on-line attention to particular aspects of events in narrative. Further, given the assumption of a close relationship between speech and gesture proposed in recent studies (e.g., Kendon 1980, 2004; McNeill 1992, 2005), the question arises as to whether such allocation of attention crosses the modal boundary to gesture. Thus, this descriptive study addresses the following questions which concern both cross-linguistic variation and L1-L2 comparison:

*IRAL* 44 (2006), 173–195  
DOI 10.1515/*IRAL*.2006.007

0019042X/2006/044-0173  
©Walter de Gruyter

1. Do speakers of typologically different languages show variation in how they introduce Ground reference in speech and gesture in narrative production?
2. What are the similarities and differences in the introduction of Ground reference in speech and gesture between L2 learners and native speakers?

## 2. Event description in narrative

When producing a narrative, speakers first select and segment the information to be conveyed, and this information is then sequenced and subsequently filtered into the surface form according to the grammar of the language spoken. The decision as to what information is to be conveyed is frequently dependent on the type of narrative. Von Stutterheim and Klein proposed that the organization of the entire narrative text can be viewed as an answer to an underlying question, a *Quaestio*, and that the selection of the relevant information and formulation of a coherent text both at local and global levels is constrained by the demands of this question (von Stutterheim and Klein 1989, 2002).<sup>1</sup> For instance, a description of a layout of a room can be viewed as an answer to a question, “What entities are where in a room?” Alternatively, a question such as “What happened to the principal protagonist?” requires a text which specifies in a chronological order events involving the main character. The present study concerns the latter type of narrative.

When speakers describe events, information about the locations where the events take place is mostly provided in relation to the actions performed by the protagonist(s). The relationship among the elements involved in events can be best captured by the terms proposed by Talmy (e.g., 1991). Talmy states that events are realized by a *Figure*, an entity (animate or not) which is located or moving in relation to another entity, the *Ground*. The course of the movement is termed a *Path*. For instance, in example (1), the “bottle” is the Figure, the “cave” is the Ground and the path of the movement is expressed by the preposition “out (of)”.

- (1) *A bottle floated out of a cave.* (adapted from Talmy 1991)

According to Talmy, languages of the world vary typologically according to how elements of motion events are encoded by lexical items. Talmy’s typology distinguishes satellite-framed (e.g., Chinese and Indo-European languages, excluding Romance) and verb-framed languages (e.g., Romance and Japanese). While the typology is complex and covers various areas, the most relevant to the present article is how speakers of two typologically different languages express Ground and Path of movement in event description.

It has been shown that in satellite-framed languages, Path is encoded by a “satellite” (i.e., adverbial particles such as “out” in example (1) above). On the other hand, in verb-framed languages, Path is generally encoded by verbs. For

instance, in the Spanish example (2), the direction of movement is expressed by the verb, *salió* (exited).

- (2) *La botella salió de la cueva flotando.*  
 the bottle exited from the cave floating  
 ‘The bottle exited the cave floating.’ (adapted from Talmy 1991)

Note that both Examples (1) and (2) have only one Ground element and one verb. However, sometimes movement involves more than one Ground. In such cases, speakers of the two language groups again show variation in how they describe motion events. On the one hand, the availability of various verb particles enables speakers of satellite-framed languages to associate a series of Ground elements with a single verb, as in Example (3).

- (3) *The boy ran out of the house across the field to the cliff.*

In Example (3), there are three Ground elements, “house”, “field”, and “cliff”. They are all compacted in a clause with a single verb “ran”. In contrast, speakers of verb-framed languages cannot compact information about motion in such a way. Ground elements have to be separated into several clauses. Example (4) below is the Japanese equivalent of Example (3):

- (4) *Shonen wa uchi o hashiri-dete, nohara o*  
 boy TOP house ACC run-exit:TE field ACC  
*yokogitte, gake made ikimashita.*  
 cross:TE cliff until go:PAST  
 ‘The boy exited the house running, crossed the field and went to the cliff.’

In Example (4), the three Ground elements are not compacted into one clause with a single verb. Instead, each element is associated with its own verb.

Interesting consequences of this difference in lexicalization between the two typologically distinct groups have been noted by Slobin (1996). Based on the analysis of retellings of a wordless picture story, *Frog, where are you?* (Mayer 1969), and translations by speakers of various satellite-framed and verb-framed languages, Slobin (1996) claims that the effect of lexicalization patterns may be observed not only at sentence level but also in the rhetorical styles preferred by their speakers. According to Slobin, during event descriptions, speakers of verb-framed languages tend to provide more static scene descriptions than speakers of satellite-framed languages, whose descriptions focus on the dynamics of movement. The following examples (from Slobin 1996) show the difference between the two narrative styles of speakers of English (a satellite-framed language) and Spanish (a verb-framed language).

- (5) a. *He (=deer) threw him over a cliff into a pond.*  
 b. *Los tiró a un precipicio donde había harta*  
 them threw at a cliff where there is lots of  
*agua. Entonces se cayeron.*  
 water then they fell  
 ‘(The deer) threw them at a cliff where there was lots of water.  
 Then they fell.’ (adapted from Slobin 1996)

In the English example (5a), the elements that represent the beginning and the end points of a movement are compacted into one clause with a single verb. While information about the location is kept to a minimum, the trajectory of the motion is provided in detail (“over a cliff into a pond”). On the other hand, in the Spanish example (5b), information about the trajectory of the motion is kept to a minimum (“fall”). Instead, the speaker focuses more on a description of the location, (“there was lots of water”).

Given the fact that there is no linguistic reason for why Spanish speakers do not use expressions equivalent to (5a), Slobin (1996) suggests that the implications of the findings go beyond the issue of language per se, and addresses the relationship between language and cognition, i.e., the selection and packaging of information for expressing a motion event reflects language-specific variation. In other words, although the selection and segmentation of information into propositional units is a universal prerequisite for discourse production, language-specific structural constraints may interact with the language production process at the conceptual level of discourse generation.

### 3. Gesture in event description

Despite the fact that gesture has long been considered “non-verbal” activity, the groundbreaking work of Kendon (1980, 2004) and McNeill (1992, 2005) has shown that spontaneous hand movements and the speech they accompany are closely related. In fact, various findings demonstrate that speech and gesture form a single system in conveying meaning (see Kendon 2004; McNeill 2000; Goldin-Meadow 2003 for overview of recent studies). These developments in L1 gesture studies are now beginning to be reflected in gesture studies that focus on second language learners (Gullberg 2003, 2005, 2006; Kellerman and van Hoof 2003; McCafferty 2002, 2004; Negueruela, Lantolf, Rehn Jordan, and Gelabert 2004; Stam 1999; Yoshioka 2005).

One piece of evidence for the interrelationship between speech and gesture comes from studies on gestures accompanying motion event descriptions (in L1 and L2). Underpinned by Talmy’s theoretical frame, these studies show that the difference in lexicalization patterns in motion event description by speakers of satellite-framed and verb-framed languages is reflected in accompanying gesture (Kellerman and van Hoof 2003; Kita and Özyürek 2003; McNeill and

Duncan 2000; Özyürek, Kita, Allen, Furman, and Brown 2005; Stam 1999). For instance, it has been shown that the different ways of conflating manner and path in speech between speakers of typologically different groups (i.e., English vs. Turkish and Japanese) is also reflected in gesture production (Kita and Özyürek 2003; Özyürek et al. 2005). Others show that how speakers differently encode path of motion in speech is reflected in the timing of gestures in L1 Spanish vs. L1 English (Stam 1999) or in L1 Spanish vs. L1 English and L1 Dutch (Kellerman and van Hoof 2003). Furthermore, it has been shown that when manner of motion is not encoded in speech, gesture may nevertheless express the manner (McNeill and Duncan 2000).

Some of the recent studies in L2 also examine gesture production in motion event description (Kellerman and van Hoof 2003; Negueruela et al. 2004; Stam 1999). For instance, Stam (1999) compared the gesture patterns of intermediate and advanced Spanish learners of English with Spanish and English baseline data. The focus of the analysis was the expressions of Path elements in the retelling of a cartoon story. It was shown that the typical Spanish pattern is to place a gesture on the verb which expresses the element of Path in a motion event, a timing which differs from the typical English pattern where the path gesture occurs on the satellite or Ground NP. The question was whether the L1 gestural pattern could be observed in L2 gesture production. The study revealed that this was indeed the case.

Kellerman and van Hoof (2003) also examined the Path gesture patterns of native speakers of English, Dutch and Spanish, and Dutch and Spanish learners of English. As in Stam (1999), the focus was on the synchronization patterns between speech and gesture with respect to information about the Path element in motion events. The results demonstrate that there is cross-linguistic variation in gesture patterning. The gestures produced by Spanish native speakers showed the same pattern as noted in Stam (1999). On the other hand, the speakers of Dutch and English, while belonging to the same typological group, revealed intra-typological differences in their gesture patterns. The patterns found in the Dutch L1 and Spanish L1 speakers were also apparent in L2 production.

Because these studies mostly focus on gesture accompanying manner and path of motion rather than Ground, their results are not directly relevant to the present study. However, the fact that gesture patterning may reflect typological distinctions and that the effect of L1 may be observed in L2 gesture provides the foundation for the present investigation.

#### **4. Gesture in discourse**

Another relevant line of research concerns the relationship of gesture to discourse structure (Gullberg 2003, 2006; Levy and Fowler 2000; Levy and Mc-

Neill 1992; Marslen-Wilson, Levy and Komisarjevsky Tyler 1982). For instance, findings show that mentions of characters at the start of a new episode in a narrative are marked by gesture (Levy and McNeill 1992; Marslen-Wilson et al. 1982). It has also been suggested that gesture production reflects whether information is new or given, with the first mention of a referent being more likely to be marked by gesture than subsequent mentions in the narrative (Gullberg 2003; Levy and Fowler 2000; Levy and McNeill 1992; Marslen-Wilson et al. 1982).

The gestural introduction (and maintenance) of referents in discourse is not a phenomenon limited to native adult speakers. Research shows that children produce gestures to refer to story characters (O'Neill and Holmes 2002). Adult learners also produce gestures accompanying mentions of referents (Gullberg 2003, 2006). Gullberg examined gestures accompanying the introduction and maintenance of animate referents in narrative, and showed that as in L1, (re-) introduced referents are more likely to be marked by gestures than maintained ones.

Although the aforementioned findings all concern gesture production in relation to the introduction of *animate* referents in discourse, it is plausible that gestures may also be produced when *inanimate* Ground referents are introduced. Thus, based on the previous findings, the present study examines how speakers introduce Ground reference in speech and gesture in L2 narrative with the assumption that speech and gesture are closely related in L2 as they are in L1.

## 5. The present study

### 5.1. Task and participants

The data used in the study are retellings of *Frog, Where are You?* (Mayer 1969), which contains twenty-six separate picture frames. The story is about a little boy and his dog who go out in search of the boy's pet frog, which has escaped. Originally, twenty Dutch learners of Japanese provided their retellings in both L1 and L2. Twenty native speakers of Japanese also participated in the study. However, because some speakers did not gesture and some gestures were performed hidden from the camera (e.g., behind crossed legs), the number of usable narratives was fifteen for Japanese L1, twelve for Dutch L1 and fifteen for L2 Japanese. Speakers were all right-handed.

In order to match the number of hours of instruction and amount of time spent in the country where the target language is spoken, all the non-native narrators were recruited among learners of Japanese who attended the same intensive course at a university in the Netherlands. All learners began the course as absolute beginners, and thus had received the same type and amount of instruction from the same instructors for the same number of hours. At the time

Table 1. Results of the JLPT Level 3 and of the oral examination

Learners	JLPT Level 3 (total 400 points)		oral examination (%)
	Points	%	
L1	340/400	85	73
L2	311/400	77.8	73
L3	260/400	65	68
L4	368/400	92	86
L5	327/400	81.8	83
L6	313/400	78.3	83
L7	306/400	76.5	77
L8	234/400	58.5	53
L9	307/400	76.8	76
L10	279/400	69.8	66
L11	181/400	45.3	75
L12	218/400	54.5	74
L13	296/400	74	81
L14	237/400	59.3	68
L15	257/400	64.3	76

of the data collection, all the learners had had three months of intensive language training followed by seven months' internship in Japan. The proficiency of the learners was measured by the administration of Level 3 of the Japanese Language Proficiency Test and an oral examination.<sup>2</sup> Given that a minimum of 60% is required to pass Level 3, Table 1 shows that most of the learners possess the relevant proficiency (roughly speaking, "low-intermediate").

## 5.2. Procedure

The participants were videotaped individually in a room that had a video camera set up prior to the session. They were given a copy of the story and asked to memorize it as thoroughly as possible. No time constraint was placed on this phase. When the participants decided that they were ready to retell the story, they did so to a native listener. Of the fifteen Dutch participants, seven individuals provided the Dutch narratives and the Japanese narratives with a 10-month interval between the two sets of data collection. Eight participants could not provide the Dutch and Japanese narratives with a sufficient intervening time span, due to practical problems in scheduling. Their Dutch native and Japanese non-native data were thus collected at the same time at the end of the course. For these narratives, the order of the languages used for the retelling task was counterbalanced. Thus, four individuals first provided their L1 Dutch



narrative followed by the L2 Japanese narrative, while four others provided their narratives in the reverse order.

### 5.3. Coding

The video recording was digitised on a computer for analysis. It was considered important that the data were handled digitally for the precise analysis of hand movements and of the association between speech and gesture. Following is the procedure adopted in the present study. First, the speech data (both L1 and L2) were transcribed orthographically (The transcription convention used for Japanese will be found in Appendix A-1). The Dutch data were transcribed by a native assistant who was a student in a linguistics department at a Dutch university. The Japanese transcription was carried out by the first author. The transcribed utterances were then divided into clauses for the purpose of analyzing the introduction of referents. For each clause, the introduction of Ground elements was coded according to their identities (e.g., “tree”, “cliff”, “pond”). By “introduction” we mean the first mention of a particular Ground referent.

For the analyses of rhetorical styles, the newly mentioned referent was also coded for the syntactic role it assumed. For instance, Ground referents are frequently introduced in adpositional phrases in VPs, as in Dutch example (6).

- (6) *en belandt zó op het gewei van een hert dat*  
and lands like this on the antlers of a deer which  
*wegrent en hem in een kleine poel gooit*  
away runs and him in a small pool throws  
‘and (he) lands like this on the antlers of a deer which runs away and  
throws him in a small pool of water’

In Example (6), the new Ground referent, *een kleine poel* (‘a small pool of water’), is introduced in the adpositional phrase with the verb, *gooit* ‘throw’, which describes the action performed by a character in the story. Example (7) is a Japanese introduction of a Ground referent, again in the adpositional phrase in VP.

- (7) *de, otoko no ko to inu wa kawa ni ochichatta*  
and boy GEN child and dog TOP river LOC fall-PAST  
*no*  
SE  
‘and the boy and the dog fell into the river’

On the other hand, Ground referents may be introduced independently of the verb phrase where the action performed by the protagonist(s) is described. For instance, the Ground referents in examples (8) and (9) assume the role of subject or complement, respectively.

- (8) *asai kawa mitaina tokoro ga aru no*  
 shallow river like place NOM exist SE  
 ‘There is a shallow river-like place’
- (9) *sono shita wa kawa ni natterun da*  
 that underneath TOP river COP:ADV become-SE COP  
*kedo*  
 but  
 ‘underneath, there is a river, but’ (lit. underneath has become a river)

In Example (8), the newly introduced Ground referent, *kawa mitaina tokoro* (‘river-like place’), assumes the subject role of the existential construction. Note that the referent is introduced independently of the mention of a verb describing an action. Similarly, in example (9), the Ground, *kawa* (‘river’), is introduced as the complement of the clause which has *~ni natteiru* structure (roughly translated as ‘there is ~’), and its introduction is not associated with a verb that describes an event involving the protagonist(s).

For the convenience of distinguishing whether or not Ground reference is introduced in association with a verb describing events (action/location) involving the protagonist(s), in the following, examples such as (6) and (7) are labeled as the “Action/location type” introductions (henceforth “AL types”). Examples such as (8) and (9) are labeled as “Existence type” introductions (henceforth “E types”).

The second stage of coding involved gesture data. It is important to note that since the present study focuses on the gestural introduction of Ground reference, only the gestures accompanying mentions of Grounds were identified. In other words, gestures that occur in the other parts of the clause were ignored. If any phase of gesture accompanied the first mention of a Ground referent, it was coded as one occurrence of a gesture.<sup>3</sup> Qualitative analysis of gestures was undertaken in terms of whether they depicted the form/outline of Ground reference or focused on action or direction.

The coding of gesture was not always straightforward. Given the purpose of the study, it was necessary to exclude disfluency-related gestures, since they had nothing to do with Ground reference introduction (e.g., depicting a missing lexical item in speech). In such cases, these gestures characteristically accompanied filled or unfilled pauses, followed by the native listener’s provision of words that the gestures depicted.

With respect to the reliability of the data coding, part of the data was re-analysed by an independent coder trained by the first author. 92 % inter-rater reliability was established for the speech data. The corresponding figure for the gesture data was also 92 %. Any data that caused disagreement were reanalysed by both coders until agreement was reached.

## 6. Results

In the first half of this section, we will present the results of the analyses of speech and gesture data in L1; the second half will present the results of the analyses of L2 data.

### 6.1. Introduction of Ground reference in L1 speech and gesture

There were a total of 300 instances of speech introductions of Ground reference in the L1 narratives: 108 in L1 Dutch and 192 in L1 Japanese. Figure 1 shows the distribution of the “AL (Action/location) type” and “E (Existence) type” introductions in L1 Dutch and Japanese. The results show that the Dutch native speakers prefer to use “AL type” introduction (97 %, 105/108). Although the Japanese native speakers also show a preference for “AL types” for more than half of the introductions (61 %, 117/192), the difference between the two groups lies in the extent to which speakers use the “E type” in introducing Ground. In L1 Japanese, 39 % (75/192) of the Ground referents are introduced as the clause subjects or complements, while the corresponding figure is 3 % (3/108) for L1 Dutch.

In the analysis of the gesture data, we tallied every gesture accompanying the introduction of Ground reference in the narratives. There were 149 gestures (37 in L1 Dutch and 112 in L1 Japanese). Figure 2 shows the frequency of gestures accompanying the introduction of Ground in speech. 34 % (37/108) of the mention of Ground reference in L1 Dutch is accompanied by gesture, while the figure is 58 % (112/192) for L1 Japanese. It is important to repeat that the present study only focuses on the gestural introduction of Ground reference. Thus, the results in Figure 2 show that these Japanese native speakers are more likely than their Dutch counterparts to produce gesture when Ground reference is mentioned for the first time.

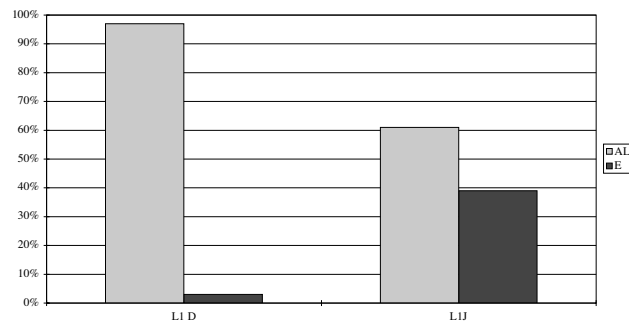


Figure 1. Distribution of “Action/location type” and “Existence type” introductions of Ground in L1 Dutch and L1 Japanese

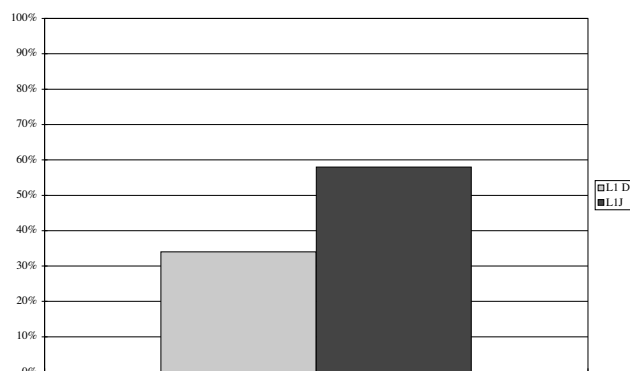


Figure 2. Frequency of gesture accompanying introductions of Ground in L1 Dutch and L1 Japanese

Next, we examined the possible relationship between the type of introduction in speech and its accompaniment by gesture. In L1 Dutch, 34 % (36/105) of the mentions of Ground in the “AL type” introductions were accompanied by gesture. Out of the three cases of “E type” introduction in L1 Dutch, one was accompanied by gesture. In L1 Japanese, 44 % (52/117) of the Ground referents introduced in the “AL type” were accompanied by gesture, while 80 % (60/75) of the “E type” introductions were similarly accompanied.

Next, we will examine gestural form. Following is an example from the L1 Dutch data. In Figure 3, the Dutch speaker introduces the Ground referent, *water* (‘water’), in a prepositional phrase in the VP. The mention of the referent is accompanied by the post-stroke hold of a gesture which depicts a fall. Note that the hands do not depict the physical characteristics of the Ground referent.

Following is another example of gesture accompanying the introduction of Ground reference in L1 Dutch. The referent is introduced in the verb phrase, which describes the location of one of the protagonists. In Figure 4, the mention of the glass jar is accompanied by a gesture which depicts the form of the referent. In L1 Dutch data, 51 % (19/37) of the gestures accompanying the introduction of Ground referents depict action or direction, while 49 % (18/37) of the gestures depict the form of Ground referents.

Following is a Japanese example. The Japanese speaker in Figure 5 describes the same scene as the Dutch speaker in Figure 3. The speaker introduces *gake* (‘cliff’) as the complement of the clause in a *~ni natteiru* construction. The introduction of the referent is accompanied by a gesture depicting the outline of the referent (Figure 5). Three clauses later, the speaker introduces *kawa* (‘river’), again in the same structure. The mention of the Ground element is marked by a gesture which depicts the surface of the water (Figure 5).



*en het ren eh het rendier smijt het jochie [in het water]<sup>4</sup>*  
and the re eh the reindeer throws the boy [in the water]  
'and the ren eh the reindeer throws the boy [in the water]'

Figure 3. Gestural introduction of Ground reference in L1 Dutch



*die [zit in een glazen ^pot]*  
*he [is in an glass ^jar]*  
'he [is in a glass jar]'

Figure 4. Gestural introduction of Ground reference in L1 Dutch

The Japanese speaker in Figure 6 describes the same scene as the Dutch speaker in Figure 4. The speaker introduces the Ground referent, *bin* ('jar'), in the verb phrase which describes the location of the protagonist. The accompanying gesture depicts the outline of the referent.



[*koo gake ni natteru kara*]  
 in this way cliff COP-ADV become-ASP because



[*sono shita wa kawa ni natteru n da kedo*]  
 that underneath TOP river COP-ADV become-ASP SE COP but

*shika wa soko de 'pyu' tto tomaru no tomattara shonen*  
 deer TOP there LOC SSW Q stop: NONPAST SE stop:COND boy  
*wa*  
 TOP

'[Because there is a cliff like this], the deer stops like 'pyu'. When (ø) stopped, the boy, [underneath, there is a river, but']

Figure 5. Gestural introduction of Ground reference in L1 Japanese

In L1 Japanese, 96 % (108/112) of the gestures that accompany the introduction of Ground depict the form or the outline of referents. Gesture depicting action is rarely produced.



*de sono kaeru [ga bin no naka ni haitteru no]*  
 and that frog NOM jar GEN inside LOC enter-ASP:NONPAST SE  
 ‘and (it is that) that [frog is in a jar]’

Figure 6. Gestural introduction of Ground reference in L1 Japanese

### 6.2. Introduction of Ground reference in L2 speech and gesture

There are a total of 83 introductions of Ground reference in speech in L2 Japanese. As in the analysis of speech in L1, these were analysed for syntactic roles and the structures used for the introduction, (“AL type” or “E type”). Recall that the introduction of Ground in an adpositional phrase is labelled “AL (Action/location) type”. Following is an example of this type of introduction from the L2 data.

- (10) *mimi no nagai no doobutsu wa otoko no hito*  
 ear GEN long GEN animal TOP male GEN man  
*o ike ni otoshimasu*  
 ACC pond LOC drop:NONPAST  
 ‘The animal with long ears\* (antlers) drops the man into a pond’

In Example (10), the L2 speaker introduces the Ground referent, *ike* (‘pond’), in the verb phrase in association with the action performed by one of the characters in the story. Example (11) shows an example of “E (Existence) type” in L2 data.

- (11) *ano mori no naka ni ookina ki ga*  
 well forest GEN inside LOC big tree NOM  
*arimasu*  
 exist:NONPAST  
 ‘Well, there is a big tree in a forest’

In Example (11), *ki* (‘tree’) is introduced independently of the mention of the action performed by the protagonist(s).

Figure 7 shows the distribution of the two types of Ground introduction in L2 speech in comparison to L1 Dutch and L1 Japanese. 94% (78/83) of Ground reference introductions occur as “AL types”, while “E types” accounts for 6%

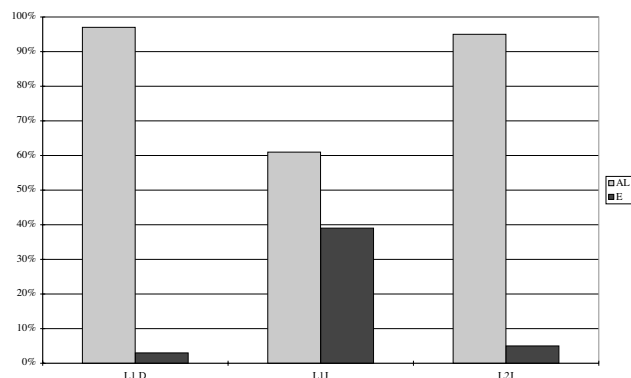


Figure 7. Distribution of “Action/place type” and “Existence type” introductions of Ground in L1 Dutch and L1 Japanese and L2 Japanese

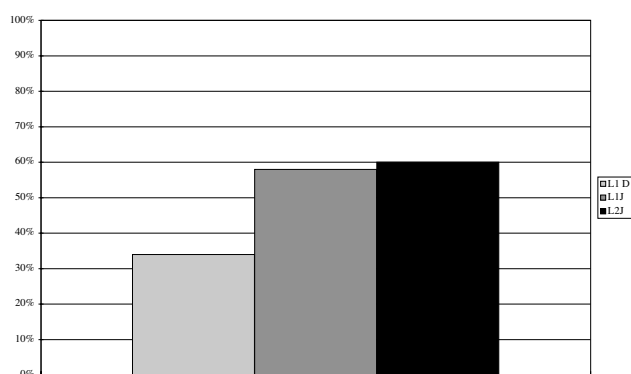


Figure 8. Frequency of gesture accompanying introductions of Ground in L1 Dutch, L1 Japanese and L2 Japanese

(5/83) of all introductions. The figure shows that the learners’ choice of rhetorical styles resembles L1 Dutch more clearly than L1 Japanese.

With respect to the analysis of gestural introduction, there are 50 gestures that accompanied the first mention of Ground in L2 narrative. Figure 8 shows the frequency of gestures accompanying the introduction of Ground in L2 with L1 baseline for comparison. 60 % (50/83) of the newly introduced Ground referents in L2 Japanese are accompanied by gesture. Figure 8 indicates that the average frequency of the gestural marking of the newly introduced Ground referents is higher in L2 Japanese than in L1 Dutch but no difference is observed in the frequency between L1 Japanese and L2 Japanese.





otoko no ko wa eto [bin no naka de, ma sono, no  
 male GEN child TOP INJ **jar** GEN inside LOC INJ INJ GEN  
naka <sup>^ni</sup>] chiisai kaeru ga imashita  
 inside ^LOC small frog NOM exist:PAST  
 ‘The boy, well, [inside, uhm, in a jar] had a little frog’ (lit: As for the boy, a little frog  
 existed in a jar)

Figure 9. Gestural introduction of Ground reference in L2

Next, we will examine the form of gestures in L2. As in L1 Dutch and Japanese, L2 speakers also produced gestures which depict the outline or the form of Ground reference. For instance, the speaker in Figure 9 produces a gesture which depicts the outline of a jar as he mentions the referent in speech.

The L2 speaker in Figure 10 introduces the Ground referent, *mizu* (‘water’), in association with the verb, *hairu* (‘enter’). The mention of the referent is accompanied by a gesture which depicts a fall.

In his L1 Dutch narrative, the speaker mentions that the deer throws the boy and the dog into the water. The introduction of *het water* (‘water’) is accompanied by a gesture depicting a throw (shown in Figure 11). In neither L1 or L2 does the speaker manually depict the entities that constitute the scene where the action takes place.

In the L2 Japanese data, 50% (25/50) of the gestures accompanying the introduction of Ground referents depict action or direction, while the other half depict the form of a referent.



*uhm uhm inu to Tanak-kun to issho ni [uh mizu ni]*  
 INJ INJ dog and Tanaka with together COP-ADV INJ **water** LOC  
 / *uh hai-haitte*  
 / *uh en-enter:TE*  
 ‘uhm uhm the dog and Tanaka together uh went [into the uh water]’

Figure 10. Gestural introduction of Ground reference in L2



*en die hert die gooit hem en het hondje [gewoon het water in]*  
 and the deer he throws him and the dog simply the **water** in  
*gekierperd*  
 dumped  
 ‘and the deer he throws him and the dog [simply dumped into the **water**]’

Figure 11. Gestural introduction of Ground reference in L1

## 7. Discussion

The present study examines the way in which Dutch learners of Japanese introduce Ground reference in speech and gesture with L1 data as baseline. As many gesture researchers have demonstrated that speech and gesture are closely related,<sup>5</sup> the present study rests on the assumption that the two modes of expression are also related when one performs in a second language.

Our cross-linguistic analyses of the speech data show similarities and differences in the rhetorical styles the native speakers of Dutch and Japanese adopt in introducing Ground reference. With respect to the similarities, the speakers in both groups preferred to introduce Ground reference in verb phrases as peripheral information associated with actions performed by the protagonist(s) or situations that the protagonist(s) are in (the “AL type” introduction). This general tendency may be due to the type of narrative used in the study. In a story (re)telling, the entire narrative is organized to answer a question, “What happened?” (von Stutterheim and Klein 1989, 2002). Accordingly, in comparison to other types of narrative (such as the description of the layout of a house), speakers, irrespective of the language spoken, may focus more on actions which help push the storyline forward than on the locations where such actions take place.

Nevertheless, cross-linguistic variation is observed with respect to the extent to which the native speakers of Dutch and Japanese introduce Ground reference independently of mentions of actions involving the protagonist(s) (the “E type” introduction). The results reveal that the Dutch speakers overwhelmingly prefer to use “AL types” over “E types”. They rarely introduce Ground reference in existential constructions. In contrast, the Japanese speakers use “E types” in roughly 40% of the Ground introductions. As we have previously noted, Slobin (1996) states that in comparison to speakers of satellite-framed languages, speakers of verb-framed languages allocate more on-line attention to static scene setting in event descriptions. The L1 speech data in the present study match these narrative characteristics of speakers of satellite- and verb-framed languages. The present finding for L1 Dutch is also in keeping with the finding in Hendriks (2005), where native speakers of German (a satellite-framed language) are not elaborate in introducing Ground reference into narrative.

With respect to the gesture data, the results can be summarized as follows. In L1 Dutch, roughly one-third of the introductions of Ground reference are accompanied by gesture. The figure is roughly double in L1 Japanese. Incidentally, this difference in the frequency of gesture production between the two native groups provides evidence against a commonly held view that Asian speakers prefer not to gesture (Chen 1990; von Raffler-Engel 1975, in Neu 1990).

When gestures are produced, roughly half of them in L1 Dutch depict the form or the outline of a referent, and the other half depict either action or direction. In contrast, the majority of the gestures accompanying the introduction of Ground reference in L1 Japanese depict the form or the outline of a referent. The frequency of gestural introduction is higher when Ground reference is introduced independently of the mention of actions involving the protagonist(s) than when the Ground element is introduced in association with events involving the protagonist(s). In the latter case, the figure is roughly half of the former.

Given that gesture expresses what is relevant, salient or important to the speaker (McNeill 1992: 125), differences in the form of gesture between the two L1 groups appear to reflect the following fact: the Japanese speakers make a bigger effort than their Dutch counterparts to gesturally specify the static scene in event descriptions. Thus, the tendency observed in the speech data is mirrored in the gesture data. These parallel findings in the two sets of data offer support for the view that speech and gesture form a single system where meaning is conveyed in a multimodal manner (e.g., McNeill 1992, 2000, 2005; Kendon 1980, 2004).

While previous studies have shown that gestures accompanying descriptions of manner and path of motion events reflect typological variation in the language spoken (Kellerman and van Hoof 2003; Kita and Özyürek 2003; McNeill 2000; McNeill and Duncan 2000; Negueruela et al. 2004; Özyürek et al. 2005; Stam 1999), the present results show that the gestural introduction of Ground reference may also reflect the typological characteristics of the language spoken.

The analyses of the L2 data show that the Dutch learners of Japanese introduce Ground reference mostly in VPs as information complementary to the action performed by the protagonist(s), and rarely independently of mentions of events involving the protagonist(s). The choice of rhetorical styles in L2 resembles L1 more than it does the L2 target. Interestingly, this result contradicts a previous finding which shows that Danish learners of Spanish provide static descriptions of a particular scene from the *Frog story* in a manner more similar to the L2 target than to their L1 (Cadierno 2004). One possible reason for the discrepancy in the findings may be the difference in tasks employed in the two studies. The learners in Cadierno's study wrote the story with a provided bilingual list of key nouns (but not verbs) which contained "necessary vocabulary to describe each picture" (Cadierno 2004: 22). This provision of key words and the request for written data may have resulted in Cadierno's learners producing picture *descriptions* rather than narratives, with an attendant increase in attention to Ground elements; the learners in the present study, on the other hand, retold the story from memory without any external aids, and thus will have produced narratives rather than descriptions.

The tendency for the Dutch learners in the present study not to allocate much attention to Ground in speech is mirrored in the gesture data. Half of the gestures accompanying the introduction of Ground reference in L2 depict action or direction. These gestures, which are observed in L1 Dutch but rarely in L1 Japanese, appear to suggest transfer. Previous studies also identify a role for transfer in L2 gesture production (Kellerman and van Hoof 2003; Stam 1999). The present study shows that the phenomenon may occur at discourse level in L2 production.

As for the frequency of gesture in L2, at the moment we do not have a clear explanation as to why learners gesturally mark the introduction of Ground reference more frequently in L2 than in their L1. Since L2 speech is characteristically associated with disfluency, the possibility that some gestures in the L2 data may have been compensatory in nature cannot be ruled out. However, this will remain speculation until a finely tuned investigation is conducted focusing on the relationship between disfluency and gesture in L2.

In conclusion, the present study reveals that the introduction of Ground reference in L2 speech and gesture reflects the fact that in the process of producing narrative, learners draw upon their L1 as the source of knowledge. The way in which speech and gesture relate in L2 is a question that is beyond the scope of the current study. However, the data presented in this article at least support the view that speech and gesture are closely related, an assumption that in fact predicts such behaviour in L2 speech and gesture.

*Leiden University*  
<k.yoshioka@let.leidenuniv.nl>  
*Radboud University Nijmegen*  
<e.kellerman@let.ru.nl>

## Appendix A

### *Transcription conventions*

#### *Transcription conventions for Japanese utterances (based on Iwasaki 2002)*

(A-1)

ACC	accusative	NOM	nominative
ADV	adverbial form	NONPAST	non-past
ASP	aspect	PASS	passive suffix
COP	copula	PAST	past
COND	conditional	Q	question marker
DAT	dative	QT	quotative
GEN	genitive	SE	sentence extende
INJ	interjection	SSW	sound-symbolic word
LOC	locative	TE	-te (conjunctive)
NEG	negative	TOP	topic marker

*Transcription conventions for speech and gesture (A-2)*

[ ]	segment where gesture occurs
[bold]	illustrated gesture
[underline]	gesture hold
^	superimposed beat (short vertical movements)
/	pause (not measured)
NL	native listener

**Notes**

- \* We would like to thank Marianne Gullberg (Max Planck Institute for Psycholinguistics, Nijmegen) and two anonymous *IRAL* reviewers for their comments and suggestions for revision. All remaining errors are ours.
- 1. The *Quaestio* model is an analytical framework used to investigate the early stage of language acquisition by non-guided second language learners in a large cross-linguistic project financed by the European Science Foundation (Klein and Perdue 1992; Perdue 1993).
- 2. The JLPT is an international examination administered by the Association of International Education, Japan, and copies of previous tests are available. The tests consist of 3 parts, vocabulary (including kanji (character) recognition), listening and reading/grammar. The oral examination was administered by a certified language instructor.
- 3. Each gesture may have the following phases: *preparation*, *pre-stroke hold*, *stroke*, *post-stroke hold* and *retraction* (or *recovery*) (see McNeill 1992 for the definition of gesture phases). Kendon (2004) defines a combination of the *stroke*, any *preparation* and *hold* phases as the gesture *phrase*, where the meaning of a gesture is most clearly expressed.
- 4. The transcription convention used for gesture is in Appendix A-2.
- 5. There are a few theories about how speech and gesture are coordinated in L1 production (see de Ruiter in press; Furuyama 2003 for review). However, given the focus of the present study (L2 discourse), our assumption does not go beyond the fact that speech and gesture are closely related in L2.

**References**

- Cadierno, Teresa (2004). Expressing motion events in a second language: a cognitive typological perspective. In *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*, Michel Achard and Susanne Niemeier (eds.), 13–49. Berlin: Mouton de Gruyter.
- Chen, Si-Qing (1990). A study of communication strategies in interlanguage production by Chinese EFL learners. *Language Learning* 40 (2): 155–187.
- de Ruiter, Jan Peter (in press). Postcards from the mind: the relationship between speech, imagistic gesture and thought. *Gesture*.
- Furuyama, Nobuhiro (2003). A review of theories of speech-gesture coordination. *National Institute of Informatics Journal* 5: 29–63.
- Goldin-Meadow, Susan (2003). *Hearing Gesture: How our Hands Help us Think*. Cambridge, MA: Harvard University Press.

- Gullberg, Marianne (2003). Gestures, referents, and anaphoric linkage in learner varieties. In *Information Structure, Linguistic Structure and the Dynamics of Language Acquisition*, Christine Dimroth and Marianne Starren (eds.), 311–328. Amsterdam: John Benjamins.
- (2005). To mean what you say: Gestures, placement events, and advanced second language acquisition. Paper presented at the 2nd Conference of the International Society for Gesture Studies (ISGS) 'Interacting bodies', Lyon, France, June 15–18.
- (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, 56 (1): 155–196.
- Hendriks, Henriëtte (2005). Structuring space in discourse: A comparison of Chinese, English, French and German L1 and English, French and German L2 acquisition. In *The Structure of Learner Varieties*, Henriëtte Hendriks (ed.), 111–156. Berlin: Mouton de Gruyter.
- Iwasaki, Shoichi (2002) *Japanese*. Amsterdam: John Benjamins.
- Kellerman, Eric and Annemarie van Hoof (2003). Manual accents. *IRAL* 41: 251–269.
- Kendon, Adam (1980). Gesticulation in speech: Two aspects of the process of utterance, In *The Relationship of Verbal and Nonverbal Communication*, Mary R. Key (ed.), 207–227. The Hague: Mouton Publishers.
- (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kita, Sotaro and Asli Özyürek (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.
- Klein, Wolfgang and Perdue, Clive (1992). *Utterance Structure: Developing Grammars Again*. Amsterdam: John Benjamins.
- Levy, Elena T. and Carol A. Fowler (2000). The role of gestures and other graded language forms in the grounding of reference in perception. In *Language and Gesture*, David McNeill (ed.), 215–234. Cambridge: Cambridge University Press.
- Levy, Elena T. and McNeill, David (1992). Speech, gesture and discourse. *Discourse Processes* 15(3): 277–301.
- Marslen-Wilson, William, Elena Levy, and Lorraine Komisarjevsky Tyler (1982). Producing interpretable discourse: The establishment and maintenance of reference. In *Speech, Place, and Action: Studies in Deixis and Related Topics*, Robert J. Jarvella and Wolfgang Klein (eds.), 271–295. Chichester: John Wiley.
- Mayer, Mercer (1969). *Frog, Where Are You?* New York: Dial Press.
- McCafferty, Steven (2002). Gesture and creating zones of proximal development for second language learning. *The Modern Language Journal* 86: 192–203.
- (2004) Space for cognition: gesture and second language learning. *International Journal of Applied Linguistics* 14 (1): 148–165.
- McNeill, David (1992). *Hand and Mind*. Chicago: University of Chicago Press.
- (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- McNeill, David (ed.) (2000). *Language and Gesture*. Cambridge: Cambridge University Press.
- McNeill, David and Susan D. Duncan (2000). Growth points in thinking-for-speaking. In *Language and Gesture*, David McNeill (ed.), 141–161. Cambridge: Cambridge University Press.
- Neguera, Eduardo, James P. Lantolf, Stefanie Rehn Jordan, and Jaime Gelabert (2004). The “Private Function” of gesture in second language speaking activity: A study of motion verbs and gesturing in English and Spanish. *International Journal of Applied Linguistics* 14 (1): 113–147.
- Neu, Joyce (1990). Assessing the role of nonverbal communication in the acquisition of communicative competence. In *Developing Communicative Competence in a Second Language*, Robin C. Scarcella, Elaine S. Andersen, and Stephen D. Krashen (eds.), 121–138. New York: Newbury House.
- O'Neill, Daniela K. and Amanda C. Holmes (2002). Young preschoolers' ability to reference story characters: The contribution of gestures and character speech. *First Language* 22: 73–103.

- Özyürek Asli, Sotaro Kita, Shanley Allen, Rayhan Furman, and Amanda Brown (2005). How does linguistic framing of events influence co-speech gestures? Insights from crosslinguistic variations and similarities. *Gesture* 5 (1/2): 219–240.
- Perdue, Clive (ed.) (1993). *Adult Language Acquisition: Cross-linguistic Perspectives*. Volume 2. The results. Cambridge: Cambridge University Press.
- Slobin, Dan (1996). Two ways to travel: Verbs of motion in English and Spanish. In *Grammatical Constructions: Their Form and Meaning*, Masayoshi Sibatani and Sandra. A. Thompson (eds.), 195–219. Oxford: Oxford University Press.
- Stam, Gale (1999). Speech and gesture: What changes first in L2 acquisition? Paper presented at the Second Language Research Forum, Minneapolis, September.
- Talmy, Leonard (1991). Path to realization: A typology of event conflation. In *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, Laurel A. Sutton, Christopher Johnson and Ruth Shields (eds.), 480–519. Berkeley, CA: Berkeley Linguistics Society.
- von Stutterheim, Christiane and Wolfgang Klein (1989). Referential movement in descriptive and narrative discourse. In *Language Processing in Social Context*, Rainer Dietrich and Carl F. Graumann (eds.), 39–76. Amsterdam: North-Holland.
- (2002). Quaestio and L-perspectivation. In *Perspective and Perspectivation in Discourse*, Carl F. Graumann and Werner Kallmeyer (eds.), 59–88. Amsterdam: Benjamins.
- Yoshioka, Keiko (2005). *Linguistic and Gestural Introduction and Tracking of Referents in L1 and L2 Discourse*. Groningen dissertation in linguistics 55. Groningen: Groningen University.