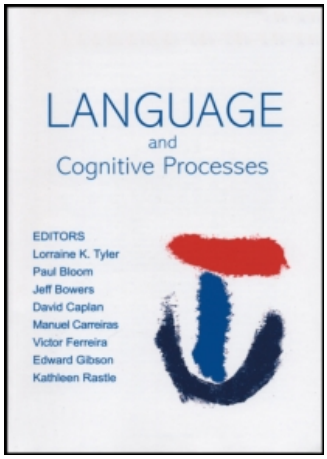
This article was downloaded by:[Max Planck Inst & Research Groups Consortium]
On: 20 December 2007
Access Details: [subscription number 771335669]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production

Sotaro Kita [a]; Asli Özyürek [bc]; Shanley Allen [d]; Amanda Brown [ce]; Reyhan Furman [f]; Tomoko Ishizuka [g]
[a] School of Psychology, University of Birmingham, Birmingham, UK
[b] F.C. Donders Center for Cognitive Neuroimaging, Radboud University, Nijmegen, the Netherlands
[c] Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[d] Department of Literacy and Language, Counseling and Development, Boston University, Boston, MA, USA
[e] Program in Applied Linguistics, Boston University, Boston, MA, USA
[f] Department of Linguistics, Bogaziçi University, Istanbul, Turkey
[g] Department of Linguistics, University of California at Los Angeles, CA, USA

Online Publication Date: 01 December 2007
To cite this Article: Kita, Sotaro, Özyürek, Asli, Allen, Shanley, Brown, Amanda, Furman, Reyhan and Ishizuka, Tomoko (2007) 'Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production', Language and Cognitive Processes, 22:8, 1212 - 1236
To link to this article: DOI: 10.1080/01690960701461426
URL: http://dx.doi.org/10.1080/01690960701461426

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

LANGUAGE AND COGNITIVE PROCESSES
2007, 22 (8), 1212–1236

Psychology Press
Taylor & Francis Group

Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production

Sotaro Kita
School of Psychology, University of Birmingham, Birmingham, UK

Aslı Özyürek
F.C. Donders Center for Cognitive Neuroimaging, Radboud University, Nijmegen, the Netherlands and Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

Shanley Allen
Department of Literacy and Language, Counseling and Development, Boston University, Boston, MA, USA

Amanda Brown
Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands and Program in Applied Linguistics, Boston University, Boston, MA, USA

Reyhan Furman
Department of Linguistics, Bogaziçi University, Istanbul, Turkey

Tomoko Ishizuka
Department of Linguistics, University of California at Los Angeles, CA, USA

Correspondence should be addressed to Sotaro Kita, University of Birmingham, School of Psychology, Birmingham B15 2TT, UK. Email: s.kita@bham.ac.uk

This research was supported by grant BCS-0002117 from the National Science Foundation to the first three authors, and by a Turkish Academy of Sciences (TUBA) Young Scientist Award to Özyürek. Substantial financial and logistical support was provided by the Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, and by Koç University, Istanbul, Turkey. We benefited from feedback at different stages of this project from the members of the Gesture Project at the Max Planck Institute for Psycholinguistics. Harald Baayen and Andrea Krott helped us with the linear mixed effect model analyses used in the mediation analysis, and the anonymous reviewer who is a statistics expert helped us report the mediation analysis clearly. We also greatly benefited from Vic Ferreira and two other anonymous reviewers' comments. We acknowledge with thanks all these contributions to our work.

© 2007 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
http://www.psypress.com/lcp DOI: 10.1080/01690960701461426

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Gestures that accompany speech are known to be tightly coupled with speech production. However little is known about the cognitive processes that underlie this link. Previous cross-linguistic research has provided preliminary evidence for online interaction between the two systems based on the systematic co-variation found between how different languages syntactically package Manner and Path information of a motion event and how gestures represent Manner and Path. Here we elaborate on this finding by testing whether speakers within the same language gesturally express Manner and Path differently according to their online choice of syntactic packaging of Manner and Path, or whether gestural expression is pre-determined by a habitual conceptual schema congruent with the linguistic typology. Typologically congruent and incongruent syntactic structures for expressing Manner and Path (i.e., in a single clause or multiple clauses) were elicited from English speakers. We found that gestural expressions were determined by the online choice of syntactic packaging rather than by a habitual conceptual schema. It is therefore concluded that speech and gesture production processes interface online at the conceptual planning phase. Implications of the findings for models of speech and gesture production are discussed.

When we communicate about action and motion, we not only use speech but also often spontaneously produce gestures that express the spatial aspects of the content of our talk. Gesture production is tightly coupled with speech production in many ways. First of all, gestures are informationally and temporally well-coordinated with the concurrent speech (e.g., Butterworth & Beattie, 1978; Kendon, 1980; Morrel-Samuels & Krauss, 1992; McNeill, 1992). For example, when a speaker produces an utterance, *it was rotating*, he or she may at the same time draw circles in the air with an extended index finger to represent rotation as she says, 'rotating'. The temporal coordination is also evidenced by the fact that in stutterers, gestures are interrupted along with speech (Mayberry & Jaques, 2000). Second, there is a systematic relationship between early language development in children and their gestures. Children already systematically coordinate the contents of their speech and gesture in the one- and two-word stages (e.g., Goldin-Meadow & Butcher, 2003; Volterra, Caselli, Capirci, & Pizzuto, 2005). Co-speech gesture is a resilient feature of human ontogenesis, which develops with minimal or no visual input, as congenitally blind individuals produce such gestures (Iverson & Goldin-Meadow, 1998). Furthermore, there seems to be a strong processing link between speaking and gesturing since we spontaneously produce gestures even when the listener cannot see them (e.g., on the intercom, Cohen, 1977; with a blind listener, Iverson & Goldin-Meadow, 1998). In addition, gesturing (or lack of gesturing) can influence speaking. The execution of a meaningful gesture modifies the sound spectra of a word with the same meaning but not of a meaningless word (Bernardis & Gentullici, 2006). When we are prohibited from gesturing, speech becomes less fluent (Rauscher, Krauss, &

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Chen, 1996). In line with these findings, it has been argued that speech and gesture production processes share a common computational stage (McNeill, 1985).

However, there is considerable theoretical disagreement in the literature regarding the cognitive processes that underlie the link between speech and gesture, especially for so-called *iconic gestures* accompanying speech (McNeill, 1992). Iconic gestures express spatial and motor features of events by means of similarity between the hand movement and the selected aspects of the referent(s) (e.g., a gesture depicting walking with inverted V-shaped wiggling fingers while talking about someone walking).The focus of this paper is to examine the mechanism by which iconic gestures are generated during speaking.

A point of contention in the literature is at which level of computation the production of iconic gestures is linked to the speech production process. One proposal is that iconic gestures are generated from imagery that is formed 'prelinguistically', that is, before linguistic formulation processes (named the Free Imagery Hypothesis). Gestures may be generated either from the spatial imagery in the working memory, activated at the moment of speaking (Krauss, Chen, & Chawla, 1996; Krauss, Chen, & Gottesman, 2000) or from the Conceptualiser (in the sense of Levelt, 1989), which produces a pre-verbal message to be further processed in the linguistic formulation module (de Ruiter, 2000). According to these models, gestures are generated before and without access to linguistic formulation processes. Consequently, both models predict that how gesture expresses a certain idea is *not* influenced by how speech expresses the same idea.

An alternative view is the Interface Hypothesis (Kita & Özyürek, 2003; Özyürek, Kita, Allen, Furman, & Brown, 2005), according to which gestures originate from an interface representation, which is a spatio-motoric representation organised in preparation for speaking. Complex information has to be broken down into chunks with appropriate informational complexity that can be verbalised within a processing unit for speech production (cf. 'thinking for speaking', Slobin, 1996; 'linearisation', Levelt, 1989). It proposes that the spatio-motoric imagery underlying a gesture is shaped *simultaneously* by (1) how information is organised in the linguistic expression that is concise enough to fit within a processing unit for speech production, and (2) the spatio-motoric properties of the referent (which may or may not be verbally expressed).

The results from previous research, as will be illustrated below, have provided support for the Interface Hypothesis, but cannot be explained by the Free Imagery Hypothesis. That is, the representational contents of gestures reflect how the speech production system packages information into

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

units readily encodable within one processing unit.[1] The purpose of the present paper is to further specify how speech and gesture production processes interact with each other. Namely, it investigates whether the influence of speech formulation on iconic gestures is made possible through the *online* interaction between the two systems, as predicted by the Interface Hypothesis, or through the activation of pre-determined language-specific conceptualisation schemas (i.e., Habitual Conceptualisation Hypothesis which will be defined later).

Evidence for the Interface Hypothesis in the literature

The evidence for the Interface Hypothesis (and against the Free Imagery Hypothesis) in the literature is based on cross-linguistic comparisons of iconic gestures depicting motion events. The same motion event can be described differently because languages have different lexical and syntactic resources. These linguistic differences are reflected in differences in how iconic gestures represent motion events (Kita, 1993, 2000; Kita & Özyürek, 2003; Özyürek & Kita, 1999; Özyürek et al. 2005).

For example, a crosslinguistic syntactic difference in the expression of Manner and Path of motion events is reflected in the gestural representations of Manner and Path in different languages. This effect was first demonstrated in a comparison of English, Japanese, and Turkish (Özyürek & Kita, 1999; Kita & Özyürek, 2003). This comparison concerned linguistic descriptions and gestural depictions of an event, in which one of the protagonists rolled down a hill. The linguistic descriptions differed cross-linguistically along the lines discussed by Talmy (1985). English speakers used a verb and a particle or preposition to express Manner (rolling) and Path (descending) of the event within one clause (e.g., *he rolled down the hill*). In contrast, Japanese and Turkish speakers separated Manner and Path expressions over two clauses (e.g., *he descended as he rolled*). Given the assumption that a clause approximates a unit of processing in speech production (Bock & Cutting, 1982; Garrett, 1982; Levelt, 1989), presumably English speakers often processed both Manner and Path within a single processing unit, whereas Japanese and Turkish speakers often needed two processing units. Consequently, Japanese and Turkish speakers should be more likely than English speakers to separate the images of Manner and Path in preparation for speaking so that two pieces of information could be processed in turn. The gesture data

[1] At the same time, iconic gestures are shaped by spatio-motoric properties of the referent. Namely, iconic gestures systematically encode spatio-motoric information that is never linguistically encoded (e.g., directionality of motion as in Kita & Özyürek, 2003 and McCullough, 1993; action, shape, physical dimensions, movement of objects as in Church & Goldin-Meadow, 1986).

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

confirmed this prediction. Namely, Japanese and Turkish speakers were more likely than English speakers to produce gestures that express Manner and Path separately.

Özyürek et al. (2005) demonstrated a tighter link between syntactic and gestural differences in a study of how English and Turkish speakers express Manner and Path in speech and gesture. They found the cross-linguistic gestural difference described above when speakers of English and Turkish used different syntactic means (i.e., one- versus two-clause expressions) to encode Manner and Path, replicating earlier studies (Kita & Özyürek, 2003; Özyürek & Kita, 1999). However, this cross-linguistic difference in gesture was not observed when the syntactic packaging of information was comparable between the two languages, namely, when only Manner or only Path of a given event was expressed in speech. In such cases, the speakers of both languages typically expressed the same information in speech and gesture (e.g., the speakers produced Manner only gestures when they expressed only Manner in speech). Thus, it is not the case that English speakers had a general across-the-board preference for gestural representations in which Manner and Path are expressed simultaneously in one gesture. Rather, the cross-linguistic difference in gesture was specific to the utterances in which syntactic packaging of Manner and Path differed between Turkish and English. This provides further support for the idea in the Interface Hypothesis that gestural representation is shaped in the process of organising information for speaking.

Another demonstration of a tight link between syntactic and gestural packaging of information came from a study of native Turkish speakers' gestures accompanying their second language English (Özyürek, 2002). In this study, speakers at different proficiency levels of English were compared. The most proficient group, which typically produced one-clause expressions to linguistically encode Manner and Path (*he rolled down*), represented Manner and Path simultaneously in one gesture, just like native speakers of English. The less proficient groups, which typically produced two-clause expressions (*he went down as he rolled*), represented Manner and Path in separate gestures, as they would when speaking their native language Turkish. That is, the tight link between syntactic and gestural packaging of information can be demonstrated in a comparison of speakers with different L2 proficiency levels, as well as in a comparison of speakers of different first languages, as predicted by the Interface Hypothesis.

An alternative account for the linguistic effect on iconic gestures

Despite the crosslinguistic evidence for the Interface Hypothesis just reviewed, it still remains unclear whether the linguistic effect on gestural

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

representations found in those studies is really due to online interaction between gestural and linguistic representations during speaking. A possible alternative explanation is that the linguistic effect is due to pre-determined language-specific conceptual schemas that have habitually been formed in line with the typological features of a given language. Let us call this alternative account the Habitual Conceptualisation Hypothesis. According to this hypothesis, gestures are generated from a stable language-specific representation. The rationale behind this hypothesis is as follows. During childhood, speakers acquire typological features of a language such as preferred syntactic patterns for describing events. As the preferred syntactic patterns are repeatedly used, a spatio-motoric representation that is compatible with syntax becomes a conceptual schema, which is habitually activated as a default way of organising information for speaking (Levelt, 1989). Formation of such schemas is desirable as they make speech production more efficient: 'although conceptualising and grammatical encoding *are* interacting for the language-acquiring child, the mature speaker has learned what to encode when preparing a message for expression. He knows by experience whether his language requires a category of medial proximity, number, tense, object shape, or whatever is needed, and he will select the appropriate information in building his preverbal messages. It is no longer necessary for the Conceptualiser to ask the Formulator at each occasion what it likes as input' (Levelt, 1989, p. 105). The idea that such conceptual schemas exist for motion events is made plausible by recent findings that language can influence spatial thinking, in the form of language-specific spatial memory (Lucy, 1992; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Pederson et al., 1998).

According to the Habitual Conceptualisation Hypothesis, gestures are generated from these default conceptual schemas that are formed by the habitual use of 'verb-framed' (Japanese and Turkish) and 'satellite-framed' (English) constructions regarding syntactic encoding of Manner and Path expressions. Languages like English typically encode Manner and Path in one clause, whereas languages like Turkish and Japanese typically encode them in two separate clauses (verbs). Thus, Manner and Path are represented together as simultaneous aspects of one event in the default conceptual schemas of English speakers, whereas Manner and Path are represented separately in the schemas of and Japanese and Turkish speakers. These schemas determine the way gestures are shaped and can explain the cross-linguistic differences found in previous studies. Under such a scenario, gestures do not have to have access to the online linguistic formulation process during speaking.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Empirical test for the Interface Hypothesis and the Habitual Conceptualisation Hypothesis

The present study aims to test whether gestural representation is shaped by the online linguistic choice (as predicted by the Interface Hypothesis) or by pre-determined language-specific schemas (as predicted by the Habitual Conceptualisation Hypothesis). As in previous studies, we aim to demonstrate that the linguistic packaging of Manner and Path influences how gesture packages Manner and Path. However, we use data from speakers of only one language (English) in order to test whether the same speakers' gestures will change with their online choice of syntactic construction. As stimuli to elicit narrative, we created animations showing various simultaneous combinations of Manner and Path. These stimuli were designed in such a way that English speakers would verbally express Manner and Path sometimes in a single clause (*he rolled down the hill*) and sometimes in two clauses (*he went down as he spun*, or *he went down and he was spinning*). The former is congruent with the typological classification of English as a satellite-framed language and used habitually to express Manner and Path, and the latter is acceptable but not congruent with the typology.

The design of stimuli was based on Goldberg's (1997) theory of the conditions under which a verb can appear in the change of location construction. According to Goldberg (1997), a verb can appear in this construction when the meaning of the verb is causally related to the meaning of the verb particle of the preposition. For example, a Manner verb, *to crawl*, can appear in the change of location construction with a Path denoting particle (*John crawled in*) or a prepositional phrase (*John crawled into the room*). This is possible because crawling can cause change of location. In contrast, a verb such as *to sing* denotes an activity that may temporally overlap with change of location, but is not causally related to change of location. Thus, the use of *sing* with a Path denoting particle or prepositional phrase in the change of location construction (*John sang in*, or *John sang into the room*) is unacceptable.

Based on the above causality principle, we created two types of stimuli. In half of them, Manner is either driving or facilitating the protagonist's movement forward, that is, Manner is inherent to Path. An example of this type is an event in which a protagonist jumps up a slope (see Figure 1). In the other half of the stimuli, Manner is incidental to change of location. An example of this type is an event in which a protagonist falls from a cliff and rotates as he falls (see Figure 1). We predict that English speakers will produce one-clause expressions (*he jumped up the slope*), which tightly package Manner and Path, more often for Manner-Inherent Events than for Manner-Incidental Events. Conversely, we predict that English speakers will produce two-clause expressions with subordination (*he rotated as he fell*, *he*

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

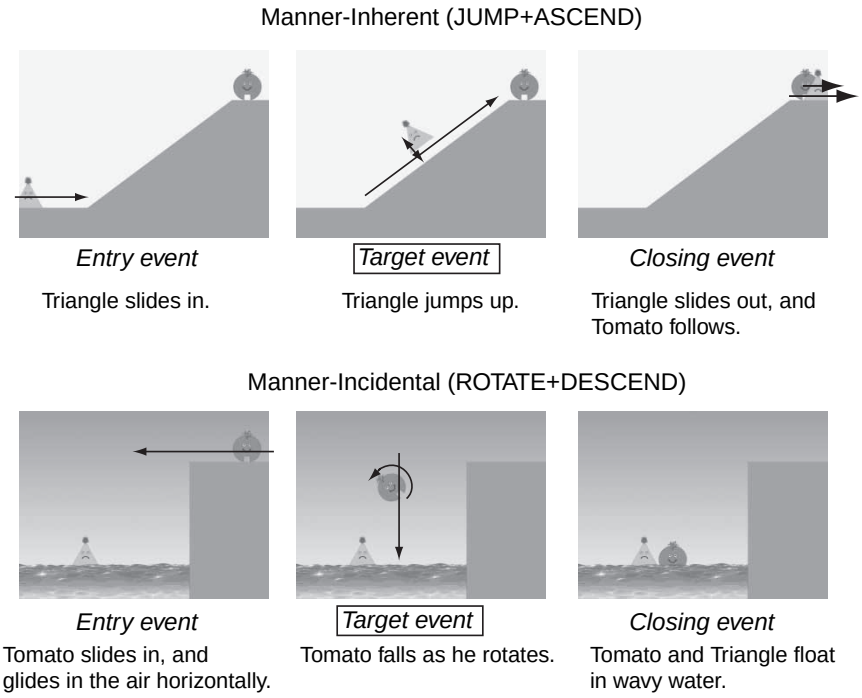


Figure 1. Stimulus examples.

fell rotating) or coordination (*he fell and he was rotating*), which loosely package Manner and Path, more often for Manner-Incidental events than for Manner-Inherent events. However, it is important to note that the association between the two types of events and syntactic encoding types is likely to be a preference, not an absolute rule. Thus, it is also possible for English speakers to use, for example, one-clause expressions to describe Manner-Incidental events or two-clause expressions of Manner-Inherent events, though to a lesser degree.

The predictions for gestures are as follows. According to the Habitual Conceptualisation Hypothesis, English speakers' gestures of Manner and Path are generated from the habitual conceptual schema based on the linguistic typology. Thus, Manner and Path should be gestured simultaneously, as predicted by linguistic typology, and the gestural representation of Manner and Path should not vary as a function of speakers' online linguistic choice. Thus, if the manipulation of event type has an effect on gestural representation, this effect should not be mediated by speakers' online syntactic choice. In contrast, according to the Interface Hypothesis, gestural representation of Manner and Path should vary according to the

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

type of syntactic packaging the speaker actually chooses for a given utterance. Thus, if the manipulation of event type has an effect on gestural representation, this effect should be mediated by the syntactic choice. This is because conceptualisation processes for gesturing and linguistic formulation processes interact online, and representations generated at these two levels tend to converge. Thus, it is predicted that Manner and Path are expressed simultaneously in a single gesture when Manner and Path are packaged 'tightly' in a single clause (henceforth, Tight clauses) in the concurrent speech, and that Manner and Path are expressed in separate gestures when Manner and Path are expressed in two separate clauses (henceforth, Separate clauses) in the concurrent speech.

METHOD

Participants

Participants were 20 native English speakers (age range: 18–40), who were students at Boston University.

Materials

Narratives were elicited by 10 video clips ('Tomato Man movies', Özyürek, Kita, & Allen, 2001) depicting motion events involving simultaneous Manner and Path. Five Manners and three Paths were depicted, yielding the following combinations: JUMP + ASCEND, JUMP + DESCEND, JUMP + GO.AROUND, ROLL + ASCEND, ROLL + DESCEND, ROTATE + ASCEND, ROTATE + DESCEND, SPIN + ASCEND, SPIN + DESCEND, and TUMBLE + DESCEND. The Manner JUMP involves an object moving vertically up and down (always moving along a flat or inclined surface), ROLL involves an object turning on its horizontal axis (always moving along an inclined surface), ROTATE and TUMBLE both involve an object turning on its horizontal axis (always moving vertically through the air), and spin involves an object turning on its vertical axis (always moving along an inclined surface). The five clips depicting the Manners JUMP and ROLL comprised the Manner-Inherent set, while the five clips depicting the Manners ROTATE, SPIN, and TUMBLE comprised the Manner-Incidental set (see Figure 1 for the examples).

In order to validate the grouping of the stimuli into the Manner-Inherent and Manner-Incidental sets, 13 native speakers of English (different participants from the main gesture experiments) viewed the ten video clips, and judged the relationship between Manner and Path. After viewing each stimulus, they rated the degree to which Manner was incidental to the change of location in a five-point scale (1 = very much,

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

5 = not at all). For each rater, the mean incidentality ratings for the items in the Manner-Incidental set and for those in the Manner-Inherent set were calculated. The raters judged Manner to be more incidental (i.e., a smaller value in the scale) for the clips in the Manner-Incidental set ($M = 2.2$, $SD = 0.4$) than in the Manner-Inherent set ($M = 3.9$, $SD = 0.5$), $t(12) = 9.33$, $p < .001$.

Procedure

Each clip was played twice to the participant. Then, the participant retold what happened in the clip to a listener who purportedly had not seen it. This procedure was repeated for each of the ten clips. All interactions were video-recorded for later analysis.

Speech coding

Speech that referred to the stimulus motion event was first segmented into sentences and then into clauses. Clauses differed as to how they syntactically packaged Manner and Path. 'Tight clauses' encoded both Manner and Path tightly within one clause. 'Separate clauses' encoded either Manner or Path (i.e., Separate Manner-Only or Separate Path-Only), but not both within a single clause. See Table 1 for actual examples from the data.

In order to establish reliability of the clause type classification, a second coder judged 20% of the data that had been classified by the original coder. The agreement between coders was 93%.

Gesture coding

We coded all gestures that encoded Manner and Path of the stimulus events using MediaTagger (Brugman & Kita, 1995). The stroke phase of gestures (Kendon, 1980; McNeill, 1992) was isolated by frame-by-frame video analysis, according to the procedure described in Kita, van Gijn, and

TABLE 1
Example of clauses with Tight or Separate syntactic linkage between Manner and Path

Clause type	*Examples*
Tight	(1) *and then tumbles down into the water* (2) *while he's twirling up*
Separate	(1) *he is spinning* (Manner-Only) (2) *doing his spin* (Manner-Only) (3) *and falls into the water* (Path-Only) (4) *as he goes down* (Path-Only)

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

van der Hulst (1998). The stroke phase is the meaning-bearing part of the gesture, and more force is exerted in the execution of the stroke phase than adjacent movement phases (i.e., preparation or retraction phases).

Gestures that encoded Manner and/or Path were classified into three types: *Manner*, *Path*, and *Conflated*. Manner gestures encoded Manner of motion (e.g., a repetitive up and down movement of the hand to represent jumping) without encoding Path. Path gestures expressed change of location without encoding Manner. Conflated gestures expressed both Manner and Path at the same time (e.g., repetitive up and down movements superimposed on a diagonal downward sweep of the hand, representing jumping down the slope). The few gestures that could not be unambiguously classified into the above three types were excluded from further analyses (e.g., two-handed gesture, in which one hand was Manner gesture, the other hand was Conflated gesture). Furthermore, the few gestures that spanned over two types of clauses (a Tight clause and a Separate clause) were also excluded from the analyses because our goal is to compare gestural representations in Tight-clauses vs. Separate clauses.

In order to establish reliability of the gesture type classification, a second coder judged the gesture type (i.e., Manner, Path, Conflated, etc.) for 20% of the relevant gesture strokes that had been identified and segmented by the original coder. The agreement between coders was 89%.

Analysis

As the main aim of the analysis is to assess how the syntactic linkage between Manner and Path affects gestural representation, all of our analyses on speech and gesture focused on descriptions of events that expressed both Manner information and Path information in speech. The 20 participants, on average, produced such an event description in 85% of the 10 stimulus events. Note that Manner and Path information can be expressed with either Tight clauses or Separate clauses. Thus, the following four event descriptions would all be included in the analysis: 'he went up the hill. and he was rolling' (a Separate Path clause and a Separate Manner clause), 'he went up as he rolled' (a Separate Path clause and a Separate Manner clause), 'he rolled up the hill' (a Tight clause), 'he rolled up. he went all the way up the hill' (a Tight clause and a Separate Path clause).'

When the dependent variable was a proportion, it was arcsine transformed before ANOVAs and *t*-tests, as suggested in Howell (2002) and Mosteller and Youtz (1961). When the variables were proportions based on a dichotomous choice (e.g., Tight vs. Separate clauses), they were logit transformed before linear regression in the mediation analysis, as recommended by Gart and Zweifel (1967), Howell (2002), and Zar (1999). Means for the dependent variables reported in the main text, tables, and graphs,

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

however, are all raw proportions prior to the transformation. In all statistical analyses, the alpha level is .05.

RESULTS AND DISCUSSION

The statistical analyses are presented in the following order. First, we analyse the speech and establish that the manipulation of the event types influenced the clause types. Then, we analyse the gestures to establish that the manipulation of the event types influenced the gesture types, and that the clause types influenced the gesture types, independently of the event types. Finally, a mediation analysis, as recommended by Baron and Kenny (1986), is carried out to directly assess whether the effect of event type on gesture type is mediated by clause type.

Analysis of speech

When speakers described both Manner and Path in a stimulus event, the two pieces of information could be expressed within a single clause (Tight clause) or in Separate clauses. The participants produced a total of 140 relevant clauses describing Manner-Inherent events, and 160 relevant clauses describing Manner-Incidental events. Table 2 summarises the distribution of Tight and Separate clauses in the descriptions of the two event types.

Tight clauses were the most frequent choice overall (56% of all clauses, $N = 300$), consistent with Talmy's (1985) typological classification of English. However, the proportion of Tight and Separate clauses varied as a function of event types, as predicted by Goldberg (1997). The Tight clauses were more likely to be used for Manner-Inherent target events than for Manner-Incidental target events, $t(19) = 3.97$, $p = .001$. Conversely, the Separate

TABLE 2
Mean proportions of Tight and Separate clauses in the Manner-Inherent and Manner-Incidental events, with the standard error in parentheses

	Clause type		
	Tight	*Separate*	
Type of events		*Manner-only*	*Path-only*
Manner-Inherent	.71 (.04)	.09 (.02)	.20 (.03)
Manner-Incidental	.42 (.06)	.24 (.04)	.34 (.03)

Note. Only the clauses from event descriptions including both Manner and Path information were analysed.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

clauses expressing either Manner or Path were more likely to be used for Manner-Incidental than Manner-Inherent target events: Separate Manner-Only, $t(19) = 3.679$, $p = .002$; Separate Path-Only, $t(19) = 3.04$, $p = .007$.

Analysis of gestures

The analysis focused on gestures produced in the verbal descriptions that referred to both Manner and Path of the stimulus event. In these descriptions, the participants produced a total of 78 Manner gestures, 146 Path gestures, and 176 Conflated gestures, concurrently with clauses expressing Manner and/or Path.

As a first step in the analysis, we examined whether the manipulation of the event types had an impact on the types of gestures produced. In the second analysis, which is crucial for our hypothesis, we investigated whether syntactic packaging had its own effect on gestures, independently from the event types.

The first analysis in assessing the effect of event type concerns the likelihood of producing Manner gestures as a function of event type. This analysis focused on gestures that co-occurred with clauses that expressed Manner information (i.e., Separate Manner-Only clauses and Tight clauses). The dependent variable was the proportion of Manner gestures among all the relevant gestures (comprising Manner gestures, Path gestures, and Manner-Path Conflated gestures). The proportion of Manner gestures was significantly higher in Manner-Incidental events ($M = .29$, $SD = .18$) than in Manner-Inherent events ($M = .16$, $SD = .15$), $t(19) = 2.88$, $p = .01$. The second analysis concerns the likelihood of producing Path gestures. This analysis focused on gestures that co-occurred with clauses that expressed Path information (i.e., Separate Path-Only clauses and Tight clauses). The proportion of Path gestures among all relevant gestures was significantly higher in Manner-Incidental events ($M = .45$, $SD = .19$) than in Manner-Inherent events ($M = .29$, $SD = .18$), $t(19) = 3.78$, $p = .001$. The third analysis focuses on the likelihood of producing Conflated gestures. This analysis focused on gestures that co-occurred with clauses that expressed Manner and/or Path (i.e., Separate Path-Only clauses, Separate Manner-Only clauses, and Tight clauses). The proportion of Conflated gestures among all relevant gestures was significantly lower in Manner-Incidental events ($M = .36$, $SD = .16$) than in Manner-Inherent events ($M = .57$, $SD = .22$), $t(19) = 4.31$, $p < .001$. Thus, Manner-Incidental events elicited more separated representations of Manner and Path, and Manner-Inherent events elicited more conflated representations of Manner and Path in gesture. Because the manipulation of event type also affected the syntactic packaging of Manner and Path in speech, as seen in the previous subsection, a question arises as to the extent to which this difference in gestural representations in the two conditions is due to the type of syntactic packaging in concurrent speech.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

The crucial prediction of the Interface Hypothesis is that gestural representation is shaped by the speaker's online choice of linguistic packaging of information. Thus, it predicts that syntactic packaging of Manner and Path has a direct influence on gestural representation of Manner and Path, distinct from that of event type. In order to assess this prediction, we repeated the above analyses of gesture with an additional independent variable, namely, the type of clause used in the expression of Manner and/or Path in the concurrent speech: Tight clauses and Separate clauses. Thus, we classified all gestures into the ones that temporally overlapped with Tight clauses and the ones that temporally overlapped with Separate clauses.

First, we investigated the likelihood of Manner gestures. The proportion of Manner gestures among all the relevant gestures was entered into a 2×2 repeated measures ANOVA with clause type (Tight vs. Separate) and event type (Manner-Incidental and Manner-Inherent) as independent variables. We analysed the data from the seven participants who produced relevant gestures (i.e., Manner, Path, and Conflated gestures co-occurring with clauses that expressed Manner information in Separate Manner-Only Clauses and Tight Clauses) in all four conditions in the 2×2 design. The mean proportions of Manner gestures in the four conditions are shown in Figure 2(a). The proportion of Manner gestures was significantly higher in Separate Manner-Only clauses ($M = .51$) than in Tight clauses ($M = .19$), $F(1, 6) = 11.7$, $MSE = 0.0545$, $p = .014$, $\eta_p = .661$. The main effect of event type was not significant, $F(1, 6) = 0.174$, $MSE = 0.0720$, nor was the interaction between clause type and event type, $F(1, 6) = 0.0298$, $MSE = 0.127$.

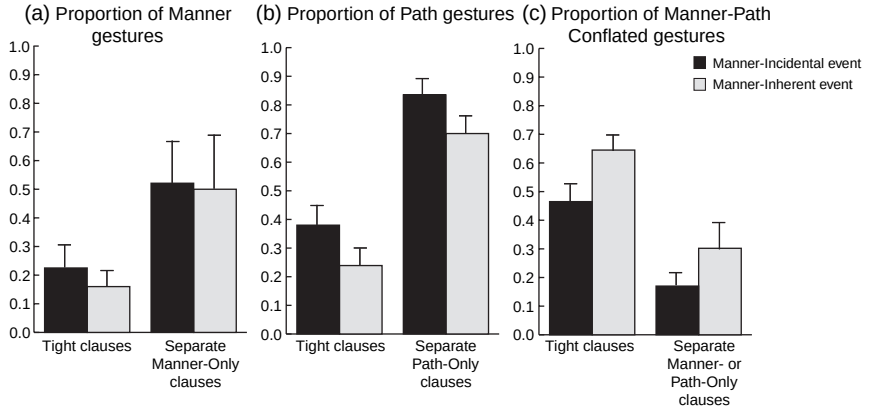


Figure 2. The mean proportions of Manner, Path, or Manner-Path Conflated gestures used with Tight clauses and Separate clauses, in the description of two types of events (Manner-Inherent and Manner-Incidental). The error bars represent the standard errors of means.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Second, we investigated the likelihood of Path gestures. The proportion of Path gestures among all the relevant gestures was entered into the same 2×2 repeated measures ANOVA as before. We analysed the data from the 13 participants who produced relevant gestures (i.e., Manner, Path, and Conflated gestures co-occurring with clauses that expressed Path information in Separate Path-Only Clauses and Tight Clauses) in all four conditions. The mean proportions of Path gestures in the four conditions are shown in Figure 2(b). The proportion of Path gestures was significantly higher in Separate clauses ($M = .77$) than in Tight clauses ($M = .31$), $F(1, 12) = 88.8$, $MSE = 0.0271$, $p < .001$, $\eta_p = .881$. The proportion was also higher in Manner-Incidental events ($M = .61$) than in Manner-Inherent events ($M = .47$), $F(1, 12) = 6.19$, MSE = 0.0425, $p = .029$, $\eta_p = .340$. There was no significant interaction between clause type and event type, $F(1, 12) = 0.030$, $MSE = 0.0350$.

Finally, we investigated whether the likelihood of Conflated gestures differed as a function of the syntactic packaging in the concurrent clause. The proportion of Conflated gestures among all the relevant gestures was entered into the same 2×2 repeated measures ANOVA as before. We analysed the data from the 15 participants who produced relevant gestures (i.e., Manner, Path, and Conflated gestures co-occurring with clauses that expressed Manner and/or Path information) in all four conditions. Their mean proportions of Conflated gestures in the four conditions are shown in Figure 2(c). The proportion of Conflated gestures was significantly higher in Tight clauses ($M = .55$) than in Separate clauses ($M = .24$), $F(1, 14) = 49.2$, $MSE = 0.0294$, $p < .001$, $\eta_p = .779$. The proportion was also higher in Manner-Inherent events ($M = .47$) than in the Manner-Incidental events ($M = .32$), $F(1, 14) = 4.71$, $MSE = 0.0660$, $p = .048$, $\eta_p = .252$. The interaction between clause type and event type was not significant, $F(1, 14) = 0.196$, $MSE = 0.0401$.

In summary, the type of clause in the concurrent speech makes a unique contribution to gestural representation, independently from the effect of event type. More specifically, Manner gestures and Path gestures were more likely in Separate clauses than in Tight clauses. Conversely, Conflated gestures were more likely in Tight clauses than in Separate clauses. As for the effect of event type, Path gestures were more likely in Manner-Incidental events than in Manner-Inherent events, and Conflated gestures showed the opposite pattern. There was no significant event type effect for Manner gestures.

Mediation analysis of speech and gesture

One of the strengths of the above analyses of gestures was that clause-level synchronisation with speech was taken into account. This is appropriate given that a clause is the theoretically relevant unit of speech in the Interface

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Hypotheses. However, the above analyses have some limitations as well. First, the analyses included only a subset of participants who produced the relevant type of gestures in all four cells of the 2×2 design. One analysis included as few as seven participants (the Manner gesture analysis). Second, the gesture analyses could not be directly linked to the speech analyses even though conceptually clause type is a variable that mediates the effect of event type on gesture type. To overcome these difficulties, we conducted an additional mediation analysis, following Baron and Kenny (1986).

Mediation analyses use a set of regressions to test a causal model in which the independent variable (in our case, event type) influences the outcome variable (in our case, gesture type) both indirectly via a mediating variable (in our case, clause type) and directly (Baron & Kenny, 1986). The Interface Hypothesis predicts the existence of a significant indirect path, in which the independent variable influences the mediating variable, which in turn influences the outcome variable.

In our analysis, the mediating variable was the proportion of Tight clauses as opposed to Separate clauses in each event description, that is the number of Tight clauses divided by the sum of the numbers of Tight and Separate clauses. The outcome variable was the proportion of Conflated gestures as opposed to Single-Information gestures (Manner-Only gestures and Path-Only gestures) for each event description, that is the number of the Conflated gestures divided by the sum of the numbers of Conflated and Single-Information gestures. Note that these proportions were calculated from a dichotomous choice (e.g., Tight vs. Separate clauses). In order to make such proportions compatible with linear regression, we applied the following logit transformation (Gart & Zweifel, 1967; Zar, 1999): the logit transformed mediating variable = Log ((the number of Tight clauses + 0.5)/(the number of Separate clauses + 0.5)), the logit transformed outcome variable = Log (the number of Conflated gestures + 0.5)/(the number of Single Information gestures + 0.5). Unlike in the typical logit transformation, 0.5 is added to the numerator and the denominator in order to reduce the bias of estimating the log-odds based on small samples (Gart & Zweifel, 1967). As in the case of the speech and gesture analyses in preceding sections, we restricted our analysis to event descriptions that linguistically expressed both Manner information and Path information in one way or another. Based on the results from the preceding sections, our predictions for the mediation analysis were as follows. Event type has a significant effect on the proportion of Tight clauses (the path (1) in Figure 3) and on the proportion of Conflated gestures (the paths (1), (2), and (3) combined in Figure 3). Finally, event type and clause type each have their own influence on gesture type (the path (3) and path (2) in Figure 3).

As the regression method, we used a linear mixed-effect analysis of covariance with participant and stimulus item as crossed random effects

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

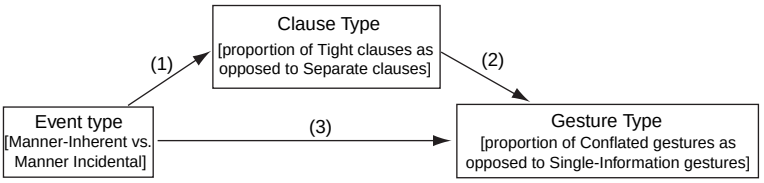


Figure 3. The summary of the mediation analysis of the effect of event type manipulation on clause type and gesture type. (The variables used in the regressions for the mediation analysis are described in the square brackets.)

(Bates & Sarkar, 2005; Baayen, 2007; see de Vaan, Shreuder, & Baayen, 2007 for the use of this technique in psycholinguistics). The data from all 20 participants were used in all of the regressions in the mediation analyses.

The first step in the mediation analysis tested whether event type affected the mediating variable (the significance of the path (1) in Figure 3). We entered event type (Manner-Inherent = 1, Manner-Incidental = 0) as the only variable (along with participant and stimulus item as crossed random effects) into a model that predicts the (logit-transformed) proportion of Tight clauses. The proportion was significantly higher in the Manner-Inherent events ($M = .79$, $SD = .31$; the M and SD reported here and henceforth are untransformed proportions) than in the Manner-Incidental events ($M = .48$, $SD = .44$), $B = 0.87$, $SE\ B = 0.21$, $t(167) = 4.17$, $p < .001$. B is the regression coefficient of the event type in the model and SE of B is used to assess the significance of the factor in the model. (The intercept $= -0.24$, $SE = 0.16$.) This is consistent with the result from the speech analysis in the preceding section, and with the idea that Manner-Inherent events were likely to elicit Tight clauses (as opposed to Separate clauses) and Manner-Incidental events were likely to elicit Separate clauses. We complete the specification of the model with the estimated standard deviations for the random effects, which are inferred from the data: 0.24 for participant, 0.20 for stimulus item, and 1.06 for the residual error.

The second step in the mediation analysis tested whether the event type affected the outcome variable (the significance of the paths (1), (2), and (3) combined in Figure 3). We entered event type (Manner-Inherent = 1, Manner-Incidental = 0) as the only variable (along with participant and stimulus item as crossed random effects) into a model that predicts the (logit-transformed) proportion of Conflated gestures. The proportion was significantly higher in the Manner-Inherent events ($M = .66$, $SD = .37$) than in the Manner-Incidental events ($M = .41$, $SD = .32$), $B = 0.69$, $SE\ B = 0.20$, $t(167) = 3.42$, $p < .001$. (The intercept $= -0.37$, $SE = 0.16$.) This is consistent with the result from the gesture analysis in the preceding section, and with the idea that Manner-Inherent events were likely to elicit Conflated gestures (as opposed to Single-Information gestures) and Manner-Incidental events were likely to

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

elicit Single-Information gestures. We complete the specification of the random effect model with the estimated standard deviations for the random effects: 0.31 for participant, 0.22 for stimulus item, and 0.95 for the residual error.

The final step of the mediation analysis tested whether the mediating variable and event type affected the outcome variable (the significance of the path (2) and the path (3) in Figure 3). Of particular theoretical importance is whether the mediating variable affected the outcome variable when the effect of event type is controlled (the path (2) in Figure 3). We entered event type, the (logit-transformed) proportion of Tight clauses, and interaction between the two variables (along with participant and stimulus item as crossed random effects) into a model that predicts the (logit-transformed) proportion of Conflated gestures. Both the mediating variable and event type had a significant effect on the dependent variable. More specifically, the proportion of Conflated gestures was higher when the proportion of Tight clauses was higher, $B = 0.17$, $SE\ B = 0.083$, $t(165) = 2.05$, $p = .042$, and also the proportion of Conflated gesture was higher in Manner-Inherent Event than in Manner-Incidental Events, $B = 0.41$, $SE\ B = 0.19$, $t(165) = 2.19$, $p = .030$. The interaction between the two explanatory variables was not significant, $B = 0.19$, $SE\ B = 0.14$, $t(165) = 1.33$. (The intercept $= -0.33$, $SE = 0.14$). Thus, both the mediating variable and event type had unique and independent influence on the outcome variable. We complete the specification of the random effect model with the estimated standard deviations for the random effects: 0.28 for participant, 0.15 for stimulus item, and 0.93 for the residual error.

In summary, the mediation analysis revealed that the effect of event type on gesture type is partially mediated by clause type. Namely, the paths (2) and (3) in Figure 3 were both significant (as well as the path (1)). This conclusion is consistent with the results from the gesture analysis from the preceding subsection. Most importantly for the purpose of this paper, it is also consistent with the idea that Tight clauses were likely to lead to Conflated gestures (as opposed to Single-Information gestures), and Separate clauses were likely to lead to Single-Information gestures. There is also evidence that event type had direct (non-mediated) influence on gesture type.

Causal interpretations of the mediation analysis results

The predictions for the mediation model were all supported. The results from this mediation analysis and those from the speech and gesture analyses in preceding sections provide a converging picture that the clause types influence the gesture types in the way predicted by the Interface Hypothesis.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

It needs to be acknowledged, however, that there are a couple of potential alternatives to this interpretation. First, it is possible that gesture influences speech, rather than the other way around. Though this study cannot rule out such a possibility, the findings from crosslinguistic studies (Kita & Özyürek, 2003; Özyürek et al., 2005) are not compatible with such an interpretation. That is, it is implausible that Turkish and English have different syntax for expressing Manner and Path *because* they gesturally express Manner and Path differently. The most parsimonious account for the current study and the previous crosslinguistic studies is that speech (as well as the type of events) influences gesture. Second, the manipulation of the event types might have caused an unknown mediating variable to vary, which in turn separately caused the change in the clause types and the change in the gesture types. Though the current study cannot rule out such a possibility, we prefer a more parsimonious model in which the clause types influence the gesture types. Nevertheless, we acknowledge that the particular causal interpretation we propose for the results from the current study needs to be taken with caution.

GENERAL DISCUSSION

The current study tested two possible underlying mechanisms for the semantic coordination between speech and gesture. More specifically, it tested whether the linguistic effect on gestural representations is due to online interaction between linguistic conceptualisation and gestural representations (the Interface Hypothesis) or due to language-specific conceptual schemas that reflect typological differences between languages but do not vary as a function of speakers' online syntactic choice (the Habitual Conceptualisation Hypothesis). In order to provide support for the Interface Hypothesis, we examined how speakers of the same language (English) gesturally express Manner and Path when speech syntactically packages these pieces of information in two different constructions: (1) a verb plus a preposition or verb particle (i.e., the construction congruent with the typological classification of English as a 'satellite-framed language'; Talmy, 1985) and (2) two separate clauses for Manner and Path. The two types of syntactic packaging were elicited by manipulating whether Manner is causally linked to change of location (Path) or not in the stimulus events that were used to elicit narratives and spontaneous gestures.

The main finding for speech was that when Manner and Path in the stimulus event were not causally linked in a clear way (Manner-Incidental events), the speakers were more likely to deviate from the typology and express Manner and Path in two separate clauses, as predicted by Goldberg (1997). More importantly, the participants produced both single-clause and

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

multi-clause expressions of Manner and Path in speech, which set the basis for the testing of the above two hypotheses.

Next, we tested the influence of event type and syntactic packaging on the gesture types. The findings showed that both event type and syntactic packaging type influenced gestural representation independently. Thus, the most important finding with regard to testing the two hypotheses was that syntactic packaging had an effect on type of gestural representation, regardless of event type. Namely, when the concurrent speech syntactically linked Manner and Path expressions tightly within a single clause, gesture tended to express Manner and Path simultaneously. When the concurrent speech syntactically linked Manner and Path expressions in a less tight way in two separate clauses, gestures tended to separate Manner and Path information. Furthermore, no interaction was obtained between the event type (Manner-Inherent vs. Manner-Incidental) and the clause type. We confirmed these outcomes in a separate mediation analysis. It was concluded that the speaker's online choice of the clause type had a unique and independent influence on gestural packaging of Manner and Path. This provides support for the Interface Hypothesis, and they are at odds with the Habitual Conceptualisation Hypothesis.

The above conclusion dovetails with earlier cross-linguistic findings that Turkish and Japanese speakers were more likely than English speakers to separate Manner and Path in their gestures, mirroring the difference in linguistic expressions (Özyürek & Kita, 1999; Kita & Özyürek, 2003; Özyürek et al., 2005). It is also compatible with Özyürek's (2002) finding that when native speakers of Turkish narrated in English, the gestural representations of Manner and Path were English-speaker-like only for those speakers whose English was proficient enough to produce one-clause expressions of Manner and Path. All these findings suggest that representational gestures are generated from the interface representation, arising in the process of packaging spatial and motoric information into chunks that are readily verbalisable within one processing unit for speech production. This is made possible through online interaction between gesture and speech production processes (Kita & Özyürek, 2003).

The findings of this study also have theoretical implications for speech production models per se. Gesture's online sensitivity to syntactic packaging of information suggests that conceptual message representations and syntactic representations are generated interactively during speaking (Kita & Özyürek, 2003; Vigliocco & Kita, 2006). This is at odds with a strong modular view on formulation processes in speaking, which provide no online feedback to conceptualisation processes (de Ruiter, 2000; Levelt, 1989). A more recent model on word production by Levelt and his colleagues (Levelt, Roelofs, & Meyer, 1999) relaxed this assumption to some degree, and proposed that there is online interaction between the activation of lexical

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

concepts and the morpho-syntactic representation of lexical items. The current findings suggest that such online interaction may also be present in the planning of conceptual message representation and that of phrasal and sentential syntax.

Kita and Özyürek's (2003) model of speech and gesture production suggests further that the message generation process for speech (Conceptualiser in Levelt, 1989) interacts online with the process that determine the content of gestures ('Action Generator'). The Action Generator takes into account both the information in spatio-motoric working memory and the message representation for speech in the Conceptualiser. The series of online interactions between the speech formulation process, the message generation processes, and the gestural content generation processes account for the alignment of the information organisation in speech and gesture. This interface mechanism between spatio-motoric thinking and speaking allows the speaker to prepare spatio-motoric information into units that are readily verbalisable.

In addition to the effect of the clause types, the event types had a unique contribution to how gestures represent Manner and Path. The proportion of Path gestures was higher in Manner-Incidental events than in Manner-Inherent events. Conversely, the proportion of Conflated gestures was higher in Manner-Inherent events than in Manner-Incidental events. One possible interpretation of these effects is that the strength of causal linkage between Manner and Path may be positively correlated with the likelihood of Manner and Path being expressed simultaneously in gestures (as a Conflated gesture). In other words, an iconicity principle which maps causal linkage to temporal linkage (i.e., simultaneity) may be at play. Though statistically not significant, Manner gestures also show a tendency compatible with this interpretation (see Figure 2a). This interpretation is also in line with the Interface Hypothesis according to which a gesture is shaped *both* by how information is organised (1) by the linguistic expression that is concise enough to fit within a processing unit for speech production, and (2) the spatio-motoric properties of the referent which may or may not be expressed in speech. In previous research we have found other spatio-temporal properties of the referent such as the direction of the movement also to be encoded in gestures independent of the linguistic coding of the event across three different languages (Kita & Özyürek, 2003). If the causal linkage between Manner and Path can be seen as a relevant 'spatio-motoric property' of an event, it should also shape gestural representation of the event.

Thus, the gesture types are determined by multiple different factors including both clausal packaging of Manner and Path and the strength of causal linkage between Manner and Path. This raises a question as to what other factors might contribute to the choice of gesture types. There are a few

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

different possibilities, for example, at the social-communicative and discourse levels. First, there might be a tendency for gestures to reflect the structure of the event as accurately as possible to the interlocutor (Kita & Özyürek, 2003). In the stimulus events, Manner and Path were always simultaneous. This may lead to a tendency to prefer Conflated gestures over Manner-Only and Path-Only gestures for communicative reasons. This may explain some of the Conflated gestures produced with Separate clauses. Second, there might also be a tendency to gesturally express information that is important in discourse (Kita & Özyürek, 2003; McNeill, 1992). In the stimulus events, the change of location (Path) is crucial for the story development as it enables the closing event of the story (see Figure 1), which makes Path information more important than Manner information. This may lead to a tendency to prefer Path gestures to Manner gestures. This tendency is visible in Figure 2a and 2b. Path gestures tend to be more frequent than Manner gestures, overall. Finally, gestures in discourse might also influence each other. For example, if one produces a Path gesture first in a description of an event, and if one wants to express Manner information in the next gesture, one might prefer a Manner gesture to a Conflated gesture. This is because the latter would contain redundant Path information that has already been expressed in the first gesture. Further exploration of various factors that influence gesture types at these levels would be a very interesting topic for future research.

This paper argued for the Interface Hypothesis as an explanation for information coordination of speech and gesture during speaking by reviewing evidence in the literature and by providing new evidence that further specifies the computation behind this coordination. We showed that gestures are generated from the online interface between spatio-motoric thinking and speaking in which spatial imagery is packaged into units that are suitable for verbalisation, rather than from pre-determined language-specific spatial conceptual schemas. This online adjustment of spatial representations for speaking is reflected in linguistic effects on gestural representations. More specifically, how gestures represent events is influenced by how concurrent speech syntactically packages information about the events into clauses (which approximate the units for speech production). In summary, we showed that during speaking, speakers coordinate linguistic and gestural representations in an online and interactive fashion.

REFERENCES

Baron, R. M., & Kenny, D. A. (1986). The moderator mediator variable distinction in social psychological-research–conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bates, D. M., & Sarkar, D. (2005). *The lme4 library.* http://lib.stat.cmu.edu/R/CRAN/.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Baayen, R. H. (2007). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, *44*, 178–190.

Bock, K., & Cutting, J. C. (1982). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Brugman, H., & Kita, S. (1995). Impact of digital video technology on transcription: A case of spontaneous gesture transcription. *KODIKAS/CODE: Ars Semeiotica, An international journal of semiotics*, *18*, 95–112.

Butterworth, B., & Beattie, G. (1978). Gesture and silence as indicators of planning in speech. In R.N Campbell & P.T. Smith (Eds.), *Recent advances in the psychology of language: Formal and experimental approaches* (pp. 347–360). New York: Plenum.

Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communication*, *27*(4), 54–63.

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

de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284–311). Cambridge: Cambridge University Press.

de Vaan, L., Shreuder, R., & Baayen, R. H. (2007). Neologisms and lexical memory: Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *Mental Lexicon*, *2*, 1–23.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

Gart, J. J., & Zweifel, J. R. (1967). On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika*, *54*, 181–187.

Goldberg, A. E. (1997). The relationships between verbs and constructions. In M. Verspoor, K. D. Lee, & E. Sweetser (Eds.), *Lexical and syntactical constructions and the construction of meaning* (pp. 383–398). Amsterdam: John Benjamins.

Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 85–107). Mahwah, NJ: Lawrence Erlbaum.

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, *396*, 228.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Kay (Ed.), *The relation between verbal and nonverbal communication* (pp. 207–227). The Hague: Mouton.

Kita, S. (1993). *Language and thought interface: A study of spontaneous gestures and Japanese mimetics*. Unpublished doctoral dissertation. University of Chicago.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.

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

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement Phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction, International Gesture Workshop Bielefeld, Germany, September 17–19, 1997, Proceedings*. Lecture Notes in Artificial Intelligence, Volume 1371 (pp. 23–35). Berlin: Springer-Verlag.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Krauss, R.M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. Zanna (Ed.), *Advances in experimental social psychology 28* (pp. 389–450). Tampa, FL: Academic Press.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge: Cambridge University Press.

Levelt, W. J. M. (1989). *Speaking*. Boston, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Lucy, J. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case of space. *TRENDS in Cognitive Sciences*, *8*, 108–114.

Mayberry, R., & Jaques, J. (2000). Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In D. McNeill (Ed.), *Language and gesture* (pp. 199–214). Cambridge, UK: Cambridge University Press.

McCullough, K.E. (1993). *Spatial information and cohesion in the gesticulation of English and Chinese speakers*. A paper presented at the annual meeting of American Psychological Society. Chicago.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 615–622.

McNeill, D. (1985). So you think gestures are non-verbal? *Psychological Review*, *92*, 350–371.

McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago Press.

Mosteller, F., & Youtz, C. (1961). Tables of the Freeman-Tukey transformations for the binomial and Poisson distributions. *Biometrika*, *48*(3/4), 433–440.

Özyürek, A. (2002). Speech-gesture synchrony in typologically different languages and second language acquisition. In B. Skarabela, S. Fish, & A. H. J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 500–509). Somerville, MA: Cascadilla Press.

Özyürek, A., & Kita, S. (1999). Expressing manner and path in English and Turkish: Differences in speech, gesture, and conceptualization. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society* (pp. 507–512). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

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

Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from cross-linguistic variations and similarities. *Gesture*, *5*(1), 215–237.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S. C., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language*, *74*, 557–589.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, *7*, 226–230.

Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, UK: Cambridge University Press.

Talmy, L. (1985). Semantics and syntax of motion. In T. Shopen (Ed.), *Language typology and syntactic description: Vol. 3. Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge University Press.

Downloaded By: [Max Planck Inst & Research Groups Consortium] At: 10:43 20 December 20

Volterra, V., Caselli, M. C., Capirci, O., & Pizzuto, E. (2005). Gesture and the emergence and development of language. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Vigliocco, G., & Kita, S (2006). Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes*, *21*(7–8), 790–816.

Zar, J. H. (1999). *Biostatistical analysis* (4th ed.). London: Prentice-Hall International.