Information packaging in speech shapes information packaging in gesture: The role of speech planning units in the coordination of speech-gesture production

Isabella Fritz[a,*], Sotaro Kita[b], Jeannette Littlemore[a], Andrea Krott[c]

[a] Department of English Language and Applied Linguistics, University of Birmingham, Birmingham B15 2TT, UK
[b] Department of Psychology, University of Warwick, CV4 7AL Coventry, UK
[c] School of Psychology, University of Birmingham, B15 2TT Birmingham, UK

A B S T R A C T

Linguistic encoding influences the gestural manner and path depiction of motion events. Gestures depict manner and path of motion events differently across languages, either conflating or separating manner and path, depending on whether manner and path are linguistically encoded within one clause (e.g., "rolling down") or multiple clauses (e.g., "descends as it rolls") respectively. However, it is unclear whether such gestural differences are affected by how speech packages information into planning units or by the way information is lexicalised (as verb plus particle or as two verbs). In two experiments, we manipulated the linguistic encoding of motion events in either one or two planning units while lexicalisation patterns were kept constant (i.e., verb plus particle). It was found that separating manner (verb) and path (particle) into different planning units also increased gestural manner and path separation. Thus, lexicalisation patterns do not drive gestural depiction of motion events. Rather gestures are shaped online by how speakers package information into planning units in speech production.

Introduction

When we speak, we often spontaneously produce gestures. Gestures are tightly linked to how we encode information linguistically at the temporal and semantic levels (McNeill, 1992, 2005). In terms of synchronisation, gestures co-occur with the element in speech that is closest to the gesture's content, often initiated before their semantic affiliates (McNeill, 1992; Morrel-Samuels & Krauss, 1992; Schegloff, 1984). From a semantic perspective speech-gesture coordination is reflected by linguistic choices. This coordination is evident on both a lexical and structural level. For instance, gestures have been found to adapt to fine-grained differences in verb semantics (Gullberg, 2011; Gullberg & Narasimhan, 2010; Kita & Özyürek, 2003), and it has been suggested that the content of a gesture is linked to how clauses package information in speech (Kita, 2000; Kita & Özyürek, 2003).

Evidence for the influence of linguistic packaging on gestural content stems from cross-linguistic studies on motion event gestures (Kita & Özyürek, 2003; Wessel-Tolvig & Paggio, 2016; Özyürek et al., 2008; Özyürek, Kita, Allen, Furman, & Brown, 2005; Özçalişkan, 2016).

Motion events are linguistically encoded differently across languages. According to Talmy's typology (2000), languages generally fall into two different categories. In so-called satellite-framed languages (e.g., German and English) the manner component of a motion event is usually encoded within the verb, while the path component is encoded in a "satellite" (a particle or an affix). Both components together often form a so-called particle verb (e.g., "to roll down" or "to climb up"). In verb-framed languages (e.g., Spanish and Japanese), path is encoded in the verb, while manner, if encoded in speech at all, is encoded in a phrase, gerund or in a separate clause (e.g., "he descends the hill as he rolls").

Gestures that accompany descriptions of motion depict manner and path in different ways in satellite-framed and verb-framed languages. In satellite-framed languages, where the motion event is linguistically encoded within one clause, manner and path tend to be conflated in a single gesture. In contrast, in verb-framed languages, where manner and path are encoded in two separate clauses, speakers tend to separate manner and path gesturally (Kita & Özyürek, 2003; Kita et al., 2007; Wessel-Tolvig & Paggio, 2016; Özyürek et al., 2008; Özyürek et al.,

2005; Özçalişkan, 2016; Özçalışkan, Lucero, & Goldin-Meadow, 2016). Thus, linguistic encoding in speech influences gestural content; but what is the relevant linguistic level?

These cross-linguistically varying gestural patterns may stem from differences in information packaging during speech production planning (Kita & Özyürek, 2003). We call this view the Planning Unit Account. More specifically, when manner and path are linguistically encoded in a single planning unit for speech production, a single gesture expresses both manner and path. When manner and path are linguistically encoded in two planning units, two gestures express manner and path separately. Clauses are assumed to play a vital role in the coordination of gesture and speech because from a speech production perspective they have been assumed to constitute a good proxy for planning units (Bock, 1982; Levelt, 1989). Thus, clausal packaging of manner and path (one clause in satellite-framed languages, two clauses in verb-framed languages) is related to gestural packaging of manner and path. However, Kita & Özyürek (2003, p. 17) do not claim that gesture-speech coordination is bound to a clausal scope. Rather they argue that gesture-speech coordination is linked to a more general processing unit[1] defined as a unit that "corresponds to what can be processed within one processing cycle for the formulation of speech".

Alternatively, the cross-linguistically varying gestural patterns might stem from differences in motion event conceptualisation based on lexicalisation of motion concepts and its implication for clausal structure. We call this view the Lexicalisation Account. This hypothesis is based on the argument that clauses are not just important units for speech production but also "conceptual units" (Pawley, 1987, 2010). This view can also account for Kita and Özyürek's (2003) cross-linguistic finding. According to this account, satellite-framed languages, which encode manner and path in a single clause, represent manner and path within a single conceptual unit. In contrast, verb-framed languages represent manner and path in two separate conceptual units. What is expressed in a single conceptual unit is expressed as a single gesture. The influence of lexicalisation patterns on motion event conceptualisation is also in line with Slobin's (2003, 2006) thinking-for-speaking hypothesis, which states that during speech production we have to filter our thoughts through linguistic encoding possibilities. Hence, lexicalisation patterns guide the speaker's attention to different aspects of the motion event (Slobin, 2000). Slobin (2000, 2003) argues that due to the obligatory encoding of manner in satellite-framed languages (within the main verb) in combination with the path component which is governed by the verb (i.e., particle verbs), speakers of these languages tend to perceive motion events as a "single conceptual event". Since it is not obligatory to encode manner linguistically in verb-framed languages, speakers of these languages do not perceive manner as inherent to the motion event. Manner is rather perceived as an activity that accompanies the path element of the motion event which is encoded in the main verb (e.g., exit, enter). These differences in conceptualisation would also explain the prominence of manner and path conflated gestures in satellite-framed languages and manner and path separated gestures in verb-framed languages.

Importantly, Kita et al. (2007) found that linguistic encoding only has an online effect on motion event conceptualisation and that gestural content is not bound to a habitual way of thinking based on how a language predominantly encodes motion events (satellite-framed versus verb-framed construction). In their study, Kita et al. (2007) compared gestures accompanying the two types of constructions within English: a satellite-framed construction (one verb framing, e.g., "he rolled down the hill") or a verb-framed construction (two verb framing, e.g., "he went down as he rolled"). They hypothesised that if habitual (dominant) language-specific event conceptualisation shapes motion event gestures, conflated manner and path gestures would be expected regardless of the construction type. However, in Kita et al.'s study (2007) the participants' gestures differed between satellite-framed and verb-framed constructions. When participants used satellite-framed constructions they accompanied speech with the expected conflated manner and path gesture. But when participants used verb-framed constructions manner and path were not only linguistically but also gesturally separated. Furthermore, essentially the same effect of construction types on gesture was also found in Dutch, another satellite-framed language (Mol & Kita, 2012). Hence, these studies suggest that gestures are shaped during speech production based on online linguistic choices and not on habitual language-specific event conceptualisations. Thus, conceptual events, which the Lexicalization Account associates with gestural information packaging, must be generated online at the moment of speaking.

One shortcoming of studies on motion events so far is that they have not been able to establish whether differences in gestural patterns stem from differences in how speech is packaged into planning units or differences in lexicalisation patterns (what information is encoded in a clause). Hence, different accounts could explain the gestural differences between verb-framed and satellite-framed constructions. Thus, the main aim of the present study is to provide unambiguous evidence for the Planning Unit Account. To this end, we manipulated the linguistic distance between manner and path components while keeping the lexicalisation pattern constant.

In Experiment 1 we tested whether increasing the linguistic distance between manner and path elements within the same clause of a satellite-framed construction can break up the planning unit into a manner and a path component, and consequently separate manner and path into two different gestures. Crucially, we asked German speakers to insert a sub-clause ("as seen in the video") between manner and path elements, which should make speakers plan manner and path in two different planning units. This assumption is based on previous research suggesting that clauses are a good proxy for planning units (Bock, 1982; Levelt, 1989) and thus inserting a sub-clause (i.e., another planning unit) between manner and path should lead the participants to process manner and path in separated planning units. Moreover, Wagner, Jescheniak, & Schriefers (2010) found evidence that the scope of planning changes when processing load increases. For the sentence structure in our experiment, we assume that the insertion of the sub-clause within a main clause will increase processing load which then results in the production of manner verb and path particle in different planning units.

It is possible in German to insert extra linguistic elements between manner and path elements. German particle verbs (e.g. "hinunterrollen" – "to down-roll") can be linguistically combined into one word or split up into two (potentially) distant words, depending, among other factors, on the clause type (main clause versus subordinate clause). German main clauses have an S-V-O structure where the verb always has to be placed in the second position of the clause and the particle comes in the final position. As seen in (1) and (2), the verb (e.g., *klettert*, "climbs") and the particle (e.g., *hinauf*, "up") can be separated by inserting elements such as prepositional phrases, direct objects (e.g., *einen Regenbogen*, "a rainbow") or even whole clauses (e.g., *wie im Video gesehen*, "as seen in the video").

(1) Der Elefant **klettert** einen Regenbogen **hinauf**.
     "The elephant **climbs** a rainbow **up**."
(2) Der Elefant **klettert**, wie im Video gesehen, einen Regenbogen **hinauf**.
     "The elephant **climbs**, as seen in the video, a rainbow **up**."

In German subordinate clauses, the verb and the particle are in reverse order compared to main clauses. Importantly, these two elements are contracted (e.g., *hinaufklettert*, "up-climbs") in the final position of the clause (3).

(3) Ich sehe, dass der Elefant einen Regenbogen **hinaufklettert**.
     "I see that the elephant a rainbow **up-climbs**."

---

[1] In this paper we will use the term "planning unit" synonymously to Kita and Özyürek's (2003) "processing unit".