

Expression of motion events in Farsi

Niloofer Akhavan^a, Nazbanou Nozari^{b,c} and Tilbe Gökşun^a

^aDepartment of Psychology, Koç University, Istanbul, Turkey; ^bDepartment of Neurology, School of Medicine, Johns Hopkins University, Baltimore, MD, USA; ^cDepartment of Cognitive Science, Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

The current study examines how gestural representations of motion events arise from linguistic expressions in Farsi, as this language offers many unique characteristics; exhibiting characteristics of both Talmy's satellite- and verb-framed languages. We examined native Farsi speakers' speech and gestures in describing 20 motion events. We focused on two motion event components: path (trajectory of motion like *up*) and manner (how the action is performed like *jumping*). Analyses of syntactic packaging and clause-level correspondence between speech and gesture, as well as parallel ordering of speech and gesture sequences were, for the most part, in support of models that posit a close correspondence between speech-gesture production. However, while Farsi speakers described both path and manner in their speech, gesture was markedly impoverished for manner, suggesting constraints on the one-to-one mapping between linguistic and gestural expressions.

ARTICLE HISTORY

Received 14 November 2015
Accepted 14 December 2016

KEYWORDS

Motion events; gesture;
language and thought;
Persian; Farsi

Expression of motion events in Farsi

The effect of language on thought is considered as either indispensable, such that thought is shaped by one's native language (e.g. Boroditsky, 2001; Whorf, 1956), or as superfluous and existing just when language is recruited to achieve a specific task goal (e.g. Gleitman & Papafragou, 2005, 2013; Papafragou, Hulbert, & Trueswell, 2008; Papafragou & Selimis, 2010a). Another approach highlights the cognitive implications of lexicalisation patterns in languages. As indicated by this *thinking for speaking* approach, thought is provoked by the requirements of a linguistic code. In particular, this theory proposes that the information to be expressed has to be tailored to speaking and must be compatible with the lexical and constructional resources of a given language (Slobin, 1996).


In this paper, we investigate the relation between language and thought by focusing on how Farsi speakers conceptualise motion events in both speech and gesture and the correspondence between these two systems. Languages vary in how they segment and package dynamic motion events and the production of gestures can be susceptible to language-specific aspects, but only during online production of language (e.g. Kita & Özyürek, 2003; Özçalışkan, Lucero, & Goldin-Meadow, 2016a, 2016b). The specific question we pose here is whether Farsi speakers' gestures during describing motion events reflect linguistic characteristics of Farsi

in terms of the components they express and the order in which those components are expressed.

Linguistic representation of motion events

Motion events have four semantic components: figure, ground, path, and manner (Slobin, 1996; Talmy, 1985). *Figure* refers to a particular point in space with respect to another object. *Ground* refers to another physical object, which serves as a reference point with respect to which the figure is located. *Path* refers to the translational motion and *manner* refers to motor pattern of the movement of the figure. Of these four, manner and path are the focus of the current study. Talmy (1985, 1991) categorises most of the world's languages into two major types of *Satellite-framed* (S-framed) and *Verb-framed* (V-framed) languages based on how the path of motion is expressed. S-framed languages such as English (Germanic), Mandarin (Sino-Tibetan), and Russian (Slavic), express manner in the main verb and path with a verb particle or a satellite (e.g. *run down* (the hill): *run* = verb encoding the manner; *down* = satellite encoding the path). Path is occasionally coded in the verb in English (e.g. she *exited* the building), but these cases are infrequent. On the other hand, in V-framed languages such as Spanish (Romance), Turkish (Turkic), and Hebrew (Semitic), manner is rarely expressed in the verb (e.g. *Çocuk aşağıya yuvarlandı: aşağıya =*

CONTACT Tilbe Gökşun  tgoksun@ku.edu.tr

 Supplemental data for this article can be accessed at doi:10.1080/23273798.2016.1276607

© 2017 Informa UK Limited, trading as Taylor & Francis Group

preposition encoding the path; *yuv arlandı* = main verb encoding the manner, “The child rolls down”). Instead, these languages usually express manner in other parts of speech, and sometimes in a subordinated clause (e.g. in Turkish, *Çocuk koşarak* (evden) *ç ıktı*: *koşarak* = subordinate clause encoding the manner; *ç ıktı* = main verb encoding the path, lit: “The child exited the house runningly”). Path of the motion, on the other hand, is expressed either in the main verb (e.g. exit) or in a verb particle (e.g. go out).

Linguistic data from both adults and children across a variety of languages showcases the typological differences in the description of motion events (e.g. Allen, Özyürek, Kita, Brown, Furman, Ishizuka, & Fujii, 2007; Berman & Slobin, 1994; Choi & Bowerman, 1991; Özçalışkan & Slobin, 2003; Papafragou, Massey, & Gleitman, 2002, 2006; Slobin, 1996, 2004). These cross-linguistic differences can be used to examine the influence of language on thought. One approach is to investigate how gestural representations of events during speaking (co-speech gestures) differ based on the properties of a given language. This approach is based on the assumption that gestures reflect how events are mentally represented (Clark, 1973; Landau & Jackendoff, 1993; Miller & Johnson-Laird, 1976; Regier, 1996; Talmy, 1983). If such mental representations, that is, thought, are directly influenced by language, then production of gestures should reflect the properties of the individual’s language. If not, co-speech gesture production may deviate from the linguistic characteristics of the utterance accompanying gestures.

Gestural representation of motion events

Co-speech gestures are bodily motions (mainly hands and arms) that accompany speech. These gestures are classified into four main categories: iconic, metaphoric, deictic, and beat gestures. *Iconic* gestures are used as referential symbols to resemble concrete objects or literal actions (e.g. moving the hand toward left to represent “going left”). *Metaphoric* gestures represent an abstract idea (e.g. moving the hand toward the back of the body to represent past time). *Deictic* gestures are any forms of pointing by any extensible body part (e.g. pointing to an apple with a finger). Finally, *beat* gestures are mere flicks of the hand(s) that adjust to the prosody of the speech without the gesture conveying semantics (McNeill, 1992). Deictic gestures are static (no motion involved), while beat gestures are dynamic (the hand is in motion). Iconic and metaphoric gestures can be in either type. Co-speech gestures are commonly used for communicating information that are visuo-spatial in nature (Alibali, 2005; Kita & Özyürek, 2003), providing a great deal of information about the internal structure of spatial thought.

There has been an unresolved debate about whether speech and gesture form a tightly integrated communication system or whether they originate from the same representational system or two separate but interrelated systems (Alibali, 2005; Alibali, Kita, & Young, 2000; Butterworth & Hadar, 1989; Goldin-Meadow, 2003; Goldin-Meadow & Alibali, 2013; Hostetter & Alibali, 2008; Kita, 2000; Kita & Özyürek, 2003; Krauss, Chen, & Gotfesnum, 2000; McNeill, 1992, 2005; Pouw, de Nooijer, van Gog, Zwaan, & Paas, 2014; De Ruiter, 2007). For example, Kita (2000) proposed that gestures help to organise and package visuo-spatial information into units of language. In an extended version of this claim, Kita & Özyürek (2003) proposed the *interface model*, which claims that gestures follow the language-specific elements of the sentence they accompany.

Evidence for the interface model comes from cross-linguistic studies showing that speakers of different languages produce different gestures for the same concept, and these gestures follow the linguistic structure of the utterances in their language (e.g. Kita, 2000; Kita & Özyürek, 2003; McNeill, 2000; McNeill & Duncan, 2000). Kita and Özyürek (2003) compared English and Turkish speakers’ gestural representations to investigate the language-specific encodings of motion events (see also Kita et al., 2007; Özyürek, Kita, Allen, Furman, & Brown, 2005). They found that in cases where there were differences in the semantic and syntactic encoding of motion event elements (i.e. one- versus multiclausal expressions), gestural representations varied in ways that fit the language specific encoding differences. In particular, English speakers produced one conflated gesture to express both elements of manner and path for concepts expressed in a one-clause (e.g. “running up” was expressed by a gesture of moving the hand upward while simultaneously alternating the index and middle fingers to signal running). In contrast, Turkish speakers produced two separate manner and path gestures for the same concept, which was expressed in a multiclausal in Turkish (e.g. “going up runningly” was expressed by an upward motion of the hand for “go up” and then alternating index and middle fingers for “run” without further vertical movement of the hand) (Kita, 2000; Kita & Özyürek, 2003; Özyürek & Kita, 1999; Özyürek et al., 2005). The developmental trajectory of path-manner packaging is also informative about the language-specific and language-general production of motion event components in speech (Allen et al., 2007) and in co-speech gestures (Özyürek et al., 2008). In a line of research, children speaking S-framed (i.e. English) and V-framed languages (i.e. Japanese, Turkish) were tested using narrations of short animated clips. Allen et al. (2007) investigated the extent to which

universal and language-specific patterns played a role in syntactic packaging of semantic elements of space in early language development. This study distinguished three structural patterns of packaging manner and path information in speech: *tight*, *semi-tight*, and *loose* speech.

Tight speech is defined as a compact unit of speech, involving one verb and one closely associated non-verbal phrase (e.g. “The red guy rolled down.”). *Semi-tight* speech is a unit of speech, involving more than one separate verbal element, one subordinated to the other (e.g. “The red guy went down, rolling.”). Finally, *loose* speech contains more than one sentence with no clausal link (e.g. “The red guy went down the hill. He was rolling at the same time”). Allen et al.’s (2007) results showed that 3-year-old English-speaking children used *tight* speech more often than their Turkish and Japanese counterparts, whereas Turkish and Japanese-speaking children used *semi-tight* packaging more often than English-speaking children, reflecting adult-like patterns of their corresponding languages. However, Turkish- and Japanese-speaking children also used some *tight* constructions to talk about both manner and path. As discussed earlier, these constructs are allowed, but are less frequently used by adult speakers, suggesting that children’s early speech shows both language-specific and language-non-specific preferences for packaging path-manner information. These language-non-specific preferences may reflect universal tendencies (Allen et al., 2007). Özyürek et al. (2008) examined whether co-speech gestures also follow the same developmental trajectory in packaging manner and path information in English and Turkish. They found that at the age of 3, regardless of their language-specific ways of encoding motion in speech, both English- and Turkish-speaking children produced separate gestures to depict manner and path of motion. However, at the age of 5, children’s gestures became more adult-like and showed language-specific patterns (i.e. English speakers used one conflated path + manner gesture whereas Turkish speakers used two separate gestures for manner and path).

In summary, the evidence reviewed above show speakers’ sensitivity to language-specific properties both in speech and gesture, as predicted by the interface model. Yet, the tight-fit relation between speech and gesture takes time to develop. Young children tend to demonstrate language-non-specific patterns in both speech and gesture. Moreover, these patterns may not correspond well together; young children, irrespective of their native language, are prone to producing *tight* speech but with two separate manner and path gestures, a pattern different from that predicted

by the interface model. These findings leave us with a critical question: Is the lack of correspondence between language and gesture only due to developmental factors or is it possible to observe a similar dissociation in adult speakers? We explore this question by examining the correspondence between language and gesture in Farsi, which has different features than languages studied previously.

One other candidate to address this question is examining the word order people use in different languages. In a motion event all elements are presented simultaneously (figure – path – manner – ground), but the order of using each element in speech is determined by the canonical word order specific to a given language. Goldin-Meadow, So, Özyürek, & Mylander (2008) asked speakers of three subject-verb-object languages (English, Spanish, and Mandarin) and one subject-object-verb (SOV) language (Turkish) to perform two non-verbal tasks. In the first task, participants silently described motion events by gesturing while looking at picture displays of vignettes. In the second task, participants reconstructed the event by putting a set of transparent pictures one by one onto a peg to form a single representation. Note that neither task entailed speech production. They found that in both tasks, speakers of all three languages were strongly inclined to use the same agent-patient-action order, which is similar to the SOV pattern in spoken languages. Similarly, others have proposed the independence of gesture order from the canonical word order in speech and found the preference of SOV order among typologically different languages such as Japanese and Korean for gesturing in reversible events (e.g. the girl kicks the boy) (Gibson et al., 2013). These findings suggest that at least certain aspects of gesture production are universal and independent of speech (see also Özçalışkan et al., 2016a). However, in these studies no speech was involved, thus the interaction of word order in speech with co-speech gestures has not been analysed.

The current study addresses the language and gesture correspondence in a typologically different language: Farsi. As presented above, the question of the sensitivity of gestures based on the structure of the language they accompany has been studied in distinctly categorised S- and V-framed languages such as Turkish and English. Yet, Farsi offers an interesting case to study the link between speech and gesture as it exhibits a mixed typology with characteristics of both S- and V-framed languages, and following a formal SOV but highly flexible word order. In addition, the verbal structure is unique having a small number of manner verbs with a rich set of productive light verb constructs.

Farsi

Farsi (Persian) is the most widely spoken language of the Iranian branch of Indo-Iranian languages, which is itself a branch of the Indo-European languages. Today, it is primarily spoken in Iran, Afghanistan, and a variation of it in Tajikistan, with large communities of speakers in the countries of the Gulf region. Many of the verbs are compounds, created by combining a light verb (e.g. *kardan* “to do”, *shodan* “to become”, *zadan* “to hit”, which may or may not preserve its original meaning in the compound verb) and a non-verb element (e.g. a noun or an adjective) (Folli, Harley & Karimi, 2005). Examples include *sohbat kardan* (lit: talk to do) “to talk”, *penhan shodan* (lit: hidden to become) “to hide”, *ghadam zadan* (lit: step to hit) “to stroll”. The non-verb components vary in how much semantic information they convey. Some, like *harekat* “motion” in *harekat kardan* (“to move”) are broad and underspecified, thus, *harekat kardan* can mean any type of motion. Some, like *ghadam* “step” in *ghadam zadan* (“to stroll”), have more specific semantics, thus, conveying a little more than just the basic action, in this case, walking in a slowly and leisurely fashion. Since many nouns do not carry fine-grained information, details like manner and path are usually left to other parts, such as prepositions and adverbs. These characteristics described above make Farsi a unique case for studying the relationship between language and gesture.

Expression of motion events in Farsi

Using narrations of motion events and following Talmy’s typology, Feiz (2011) claims that Farsi exhibits a mixed typology with characteristics of both S-framed and V-framed languages (see also Verkerk, 2014). The similarity to S-framed languages is apparent in cases where path information is expressed in path satellites and manner in a verb as in English. An example is (1) in which *baala* “up” is a satellite and *daavidan* “to run” is a verb that contains manner information:

(1) [از تپه] بالا دویدن (1)
[az tappe] baala daavidan
lit: [From hill] up run
“to run up [the hill]”

In terms of syntactic packaging, this is equivalent to a *tight* package.

The similarity to V-framed languages is in cases where path information is in the verb, leaving manner information to be expressed in other parts of speech, mostly in adverbial that remain subordinate to the main clause as in Turkish (Example 2a).

(2a) دوان دوان دور [...] چرخیدن
davan davan dor -e- [...] charkhidan
lit: runningly runningly around [...] to circle¹
“running around the [...]”

This construct usually manifests as a *semi-tight* package, where path is encoded by the verb and manner expressed separately through an “adjunct” or an adverb. Another common form of expressing manner and path is through the use of light verb constructs. Recall that the majority of verbs in Farsi are compounds with a light verb combined with a non-verb element, such as a noun. This mode of manner and path expression can take the form of a main clause and a separate adverbial unit, and create a *semi-tight* package (Example 2b):

(2b) لی لی کنان از [...] خارج شدن
ley ley konan az [...] khaarej shodan
lit: Hop Hop doing from [...] exit to become
“To exit the [...] in a hopping manner”

“*ley ley konan*” is an adverb derived from the light verb “*kardan*” (doing).

In addition, manner and path expression can each manifest as two separate clauses, as shown in an example (3). The latter construct is a *loose* package.

(3) لی لی کردند و از [...] خارج شدن (3)
ley ley kardan. Az [...] khaarej shodan
lit: Hop Hop do. From [...] exit to become
“To hop. To exit from [...]”

In summary, there are various ways to express manner and path in Farsi, but the most common way involves the use of light verb structures, which entails *semi-tight* or *loose* syntactic packaging.

Word order in Persian

Farsi is a SOV language in formal sentences. However, the sentential constituents have much freedom to move around in a sentence, especially in the spoken language. A secondary goal of the paper is to see whether the flexibility of word order in Farsi is reflected in people’s spontaneous gesture production.

The current study

The current study is the first to investigate how gestural representations of motion events stem from linguistic expressions in Farsi, the unique characteristics of which we reviewed earlier. The paper presents different approaches such as clause-level analysis, syntactic packaging and word order to give a comprehensive account of language and gesture interaction. Farsi speakers are expected to express path of motion with prepositions and manner of motion as verb or adverb together with using light verbs. Our critical prediction

concerns the gesture production. If the clause structure of the language corresponds very closely to gestures, as expected by the interface model (Kita & Özyürek, 2003; Özyürek et al., 2008), we predict that manner and path will be expressed in a single gesture (i.e. manner and path conflated) when manner and path are encoded in a single clause. In contrast, manner and path should be expressed in separate gestures when they are encoded in multiclause. The paper also analyses how gesture production varies according to type of packaging of manner and path within different linguistic units. According to the interface model (Kita & Özyürek, 2003), we predict that: (1) when the speech is *tight*, there would be one conflated gesture representing both manner and path of motion. On the other hand, (2) when the speech is *semi-tight* or *loose* there would be two separate gestures; one referring to the path and the other referring to the manner of motion, as in Turkish. With regard to the pattern of gesture and word order, if gestures are influenced by online processing of language, we would expect the order of manner and path gestures to correspond to the order in which such information is expressed in spoken language.

Method

Participants

Nineteen monolingual native Farsi speakers (9 females) between the ages of 18 and 30 were tested. Participants lived in Iran, were all right-handed, had normal hearing and vision. All participants signed written consent in accordance with the ethical policies of Koç University Institutional Review Board.

Task and stimuli

Participants watched 20 dynamic movie clips, depicting different motion events with combinations of 10 manners (hop, skip, walk, run, cartwheel, crawl, jump, twirl, march, and step) and 9 paths (between, to, out of, under, over, in front of, around, across, and into). Each movie lasted for 3–4 seconds. The clips were previously developed and standardised in English (Göksun, Lehet, Malykhanian, & Chatterjee, 2015). All actions were performed by a woman in an outdoor area (see Figure 1 for sample stimuli and for the full list of events see Appendix A). Critically, all events in the experiment could potentially be expressed using both S-framed and V-framed utterances in Farsi.



Figure 1. Sample stimuli from the experimental task. The pictures are still frames from two motion events: *jump over* (left side) and *walk across* (right side). The yellow arrows indicate the direction of the person's movement.

Procedure

All participants were tested individually in their home environment in a silent room. They were instructed to watch each clip and then describe what they saw. No explicit instruction regarding gesture use was provided. Before the test trials, two practice trials were administered, and participants received feedback on their performance. Test stimuli were displayed on a Dell laptop in three different randomised orders across participants. The testing sessions were audio- and videotaped. The camera was set in a position to capture the hands and the body of the participants but not the heads.

Coding

Speech

The speech was transcribed verbatim by a native Farsi speaker (first author). The transcribed utterances were coded for the use of manner and path of motion. The pattern of speech responses in terms of manner and path was categorised into groups of manner only (only manner information was expressed in the speech), path only, (only path related information was encoded in the speech) and path + manner together, (both manner and path were expressed in the speech). Manner information was further coded into manner as a verb (4a), an adverb (4b), and the noun in a compound verb containing a light verb (4c):

(4a) Verb دویدن <i>Davidan</i> "to run"	(4b) Adverb بدو بدو <i>Bodo bodo</i> lit: runnigly runnigly "in a running fashion"	(4c) Noun + light verb بدو بدو کردن <i>Bodo bodo kardan</i> lit: run run to do "to run"
--	--	---

Path was categorised into path as a preposition (5a), a verb (5b), a verb together with a preposition (5c), a light verb (5d), and a light verb together with a preposition (5e):

(5a) preposition دور <i>dor -e</i> "around"	(5b) Verb چرخیدن <i>charkhidan</i> "to circle"	(5c) Verb + preposition دور چرخیدن <i>dor charkhidan</i> lit: around to circle "to circle around"
(5d) Light verb آمدن <i>aamadan</i> "to come"	(5e) Light verb + preposition از بین رد شدن <i>az bein rad shodan</i> lit: from between pass to become "to pass between"	

For the descriptions that contained both manner and path information, we used two types of coding: clause coding and syntactic packaging. For clause coding, utterances that included both manner and path were classified based on the clause types: (a) *one-clause* expressions (b) *multiclause* expressions. The clause coding was adapted from Allen et al. (2007) and Özyürek et al. (2008), who developed the system to test the predictions of the interface model.

One-clause expressions involve one verb and one closely associated non-verbal phrase. A typical example of this in Farsi includes a manner verb with a postpositional path phrase, but no path verb, as in Example 6a.

(6a) دختر دور [...] میدود
Dokhtar dore [...] midavad
lit: Girl around [...] runs
"The girl is running around the [...]"

In *multiclause* expressions manner and path were either distributed over separate clauses as path-only or manner-only clauses or one was expressed as an adverb. When path and manner were expressed in separate clauses, manner is described by either a manner verb or manner noun accompanied with a light verb. Path could be constructed by a combination of either a path verb or light path verb with a path preposition. These multiclause expressions are conjoined by discourse markers such as *va* [and] and *baad* [then] in Farsi, as in Example 6b.

(6b) راه میرفت و از روی [...] رفت بالا
"rah miraft va az ruye [...] raft bala"
lit: walk was going and from [...] went up
"[...] was walking and went up the [...]"

When not expressed in separate path and manner clauses, manner could still be expressed separately as an adverb, along with a path preposition and a light

verb, without a manner verb (6c, d) (see Appendix B for more examples). Note that in Farsi, adverbs may or may not be expressed as separate linguistic clauses, but the current coding of adverbial manners as separate clauses allows comparison with previous work testing the interface model.

(6c) دختر بدو بدو کنان به سمت [...] رفت
dokhtar bodo bodo konan be samte [...] raft
lit: Girl run run doing to direction of [...] went
"The girl went toward the [...] while running"

(6d) دختر به حالت بدو بدو به سمت [...] رفت
Dokhtar be halate bodo bodo be samte [...] raft
lit: Girl in manner run to direction of [...] went
"The girl went toward the [...] in a running manner"

For syntactic packaging, we examined manner and path information through the linguistic units they are encoded with. Therefore, we classified only the event descriptions that included both manner and path into one of the three categories: *tight*, *semi-tight*, and *loose*, based on Allen et al.'s (2007) study (see Examples 1–3).

Gesture

For each trial, the number of gestures produced was coded. The gestures were then classified as static or dynamic. Dynamic gestures were further classified into (1) manner only, (2) path only, and (3) path + manner together. Manner-only gestures are those that enact the style of a motion without emphasising the trajectory of the movement (e.g. circular movement of index finger without moving the arm to represent cartwheeling). Path-only gestures show the direction of the movement without representing the manner (e.g. movement of index finger in an arc pattern along the horizontal axis from right to left to represent "across"). Path + manner gestures take two forms: the conflated form contains both components simultaneously (e.g. circular movement of index finger along the horizontal axis from right to left to represent "cartwheeling across"). The separate form still contains both gestures, but they are performed separately and serially. Figure 2 represents these three types of gestures.

Word order and gesture

For the path + manner trials where participants generated separate manner and path gestures, we coded the order in which these gestures were produced.

Reliability

To establish reliability for speech, a second native Farsi speaker independently coded 20% of the data. The agreement between coders was 96.4% ($n = 80$ trials) in assigning manner only, path only, path + manner

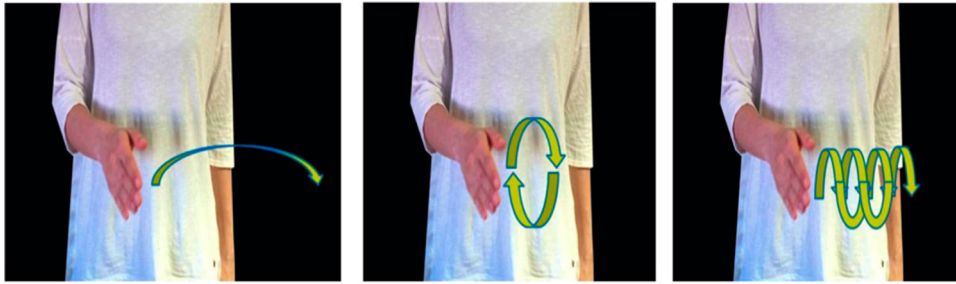


Figure 2. Sample gestures that represent (a) a path-only motion (e.g. across), (b) a manner-only motion (e.g. cartwheeling), and (c) a path + manner (cartwheeling across).

categories to the descriptions and 90.3% ($n = 80$ trials) in segmentation of speech into *tight*, *semi-tight*, and *loose* categories. Furthermore, to establish the reliability of the coding system, a second coder randomly chose and mutely coded 20% participants' gestures. The agreement between coders was 97.4% ($n = 80$ trials) for gesture identification, 90.2% ($n = 80$ trials) for gesture category assignment (dynamic or static iconic gestures), and 87.8% ($n = 80$ trials) for coding gestures that involved manner only, path only, and path + manner.

Results

A total of 377 trials were included in the analyses. Three trials were excluded because the participant did not describe them. One participant's data was excluded as this person's gestures were out of the camera frame.

Speech analyses

Participants expressed both manner ($M = 86\%$, $SD = 8.91$) and path ($M = 89\%$, $SD = 12.31$) information in their speech and a Wilcoxon signed-rank test showed that there was no statistically significant difference between expressing these components in speech ($Z = -.908$, $p = .364$). Even though manner could be expressed in the verbs, participants expressed manner in adverbial form more frequently than in any other forms ($M = 75\%$, $SD = 15.18$), $X^2(2, N = 324) = 137.35$, $p < .001$.² One possible reason for this is the difference between colloquial and formal forms in Farsi. Some manner-heavy verbs (e.g. *khazidan* "crawling") appear more frequently in formal (e.g. in written language) than colloquial language. Avoiding such forms reflects the speaker's choice to use a colloquial style of speaking and remain consistent within that style. However, the same does not hold for all verbs in the set. Some of the verbs that express manner as noun + light verb combinations, for example, "ley ley kardan" (lit = hop-hop doing) are perfectly acceptable in colloquial Farsi. Nevertheless, speakers' preference for expressing manner in adverbs reflects

that even these cases may have been dispreferred in the current experiment. To our knowledge, there are no thorough investigations of the circumstances under which Farsi speakers shift their production from one form to another. Thus, we cannot offer a more conclusive explanation for the observed speech pattern, beyond emphasising that each sentence in the set could have been expressed in a fashion consistent with either S-framed or V-framed languages, which makes this set distinct from a language such as Turkish. Paths were encoded with preposition + light verb more than any other construct, $X^2(4, N = 345) = 554.06$, $p < .001$ (see Figure 3 for the use of manner and path expressions in speech).

Gesture analyses

Participants produced a total of 527 gestures in 298 out of 377 trials. On average, 68% of these gestures were identified as dynamic iconic, 12% of gestures were static iconic and pointing, and 20% as beat gestures. In this paper, we only focused on dynamic iconic gestures that referred to motions in the clips. Participants produced significantly more path gestures than manner gestures or path + manner gesture together (conflated), $X^2(2, N = 358) = 200.47$, $p < .001$.

For the trial-based analyses, we coded whether participants used only path, only manner, path + manner (separate) or path + manner (conflated) in each trial. As shown in Figure 4, the majority of dynamic gestures were identified as path only ($M = 57\%$, $SD = 15.58$) compared to manner only ($M = 12\%$, $SD = 15.64$), path + manner (separate) ($M = 20\%$, $SD = 16.98$), or path + manner (conflated) confluations ($M = 11\%$, $SD = 13.14$), $X^2(3, N = 268) = 164.01$, $p < .001$. Participants predominantly produced path gestures that indicated the direction of the movement (e.g. toward the building).

Speech – gesture relations

To further explore the information represented for motion event expressions, we analysed whether manner and

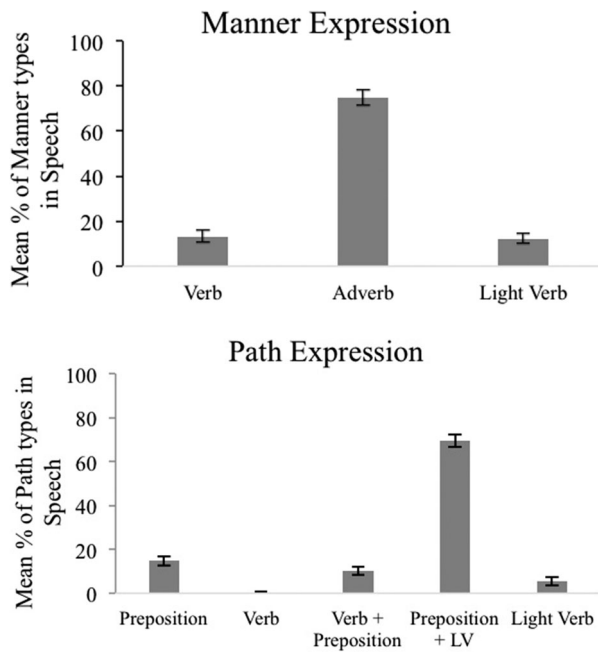


Figure 3. The percentages of manner and path expressions in speech among 377 sentences. The error bars indicate the standard error of mean.

path were conveyed in both speech and gesture or in one of the two. We found that manner and path were expressed differently in speech and gesture. Participants tended to encode path information in both speech and gesture, $X^2(4, N=377)=347.92, p<.001$, whereas manner was mostly produced in speech only, $X^2(4, N=377)=369.98, p<.001$ (Figure 5).

Gestures in different clause types

Of 277 utterances that included path and manner information, 17 were excluded from the analysis because they were not full clauses with verbs and were interrupted

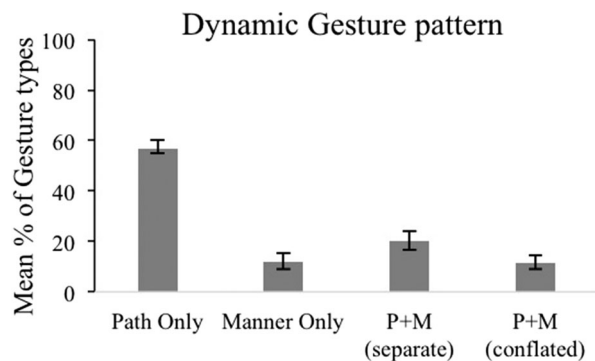


Figure 4. The percentages of path only, manner only, path + manner (separate), and path + manner (conflated) gestures in 268 trials.

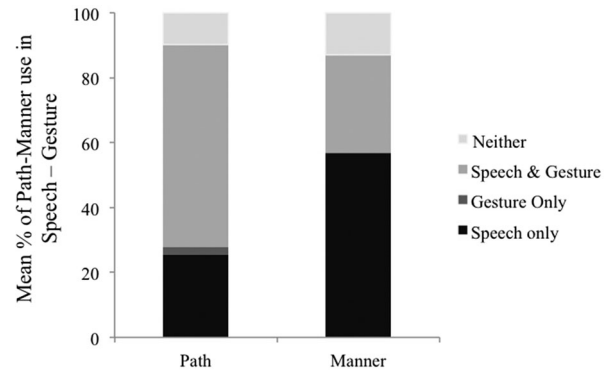


Figure 5. The distribution of the manner and path expressions in speech, gesture, and both.

before completion. On average, 43% of these descriptions were encoded by one-clause and the rest (56%) by multiclausal expressions. Moreover, we investigated what type of gestures people produced depending on the clause type, out of all the event descriptions that included both manner and path. The results are summarised in Table 1. In this analysis, we included only data from clauses that were accompanied by dynamic gestures. The results showed that people predominantly produced path-only gestures with their one-clause expressions, $X^2(3, N=89)=81.29, p<.001$. Such cases are depicted in the example below:

“دختر دور درخت میدود”
Dokhtar dore derakht midavad
 lit: Girl around tree runs
 “The girl is running around the tree”
Gesture: an index finger moving spherically to represent “around”

Moreover, according to previous studies, we would expect participants’ gestures to overlap mostly with typologically congruent expressions, that is, more manner and path conflated gestures to overlap with one-clause expressions (Kita & Özyürek, 2003; Özyürek et al., 2005). However, the results did not support this. There was no reliable difference between the use of separate and conflated gestures in one-clause expressions, $X^2(1, N=18)=0.22, p=.63$ (see the example below).

“دختر بپر بپر به سمت در رفت”
Dokhtar bepar bepar be samte dar raft
 lit: Girl hop hop to direction of door went
 “The girl hopped toward the door”
Gesture: In the case of separate gestures, both hands moved up and down repetitively to represent “hopping”, followed by one of the hand moving forward on a horizontal axis to represent “toward”. In the case of a conflated gesture, one of the hands moved up and down repetitively while also moving forward.

Table 1. The distribution of different gestures (path only, manner only, path + manner (separate), path + manner (conflated) and other trials among clause structure (One-clause and Multiclausal) types in all of all the event descriptions that included both manner and path ($N = 260$).

	Path only	Manner only	P + M (separate)	P + M (conflated)	Other
One-clause ($n = 158$)	59	12	8	10	69
Multiclausal ($n = 202$)	90	11	51	19	31

In multiclausal expressions too, people predominantly used path-only gestures (52%), $X^2(3, N = 171) = 90, p < .001$. An example is given below.

دختر دويد، دور درخت چرخيد

Dokhtar davido, dore derakht charkhid

lit: Girl ran, around tree circled

"The girl ran and circled around the tree"

Gesture: the index finger moved spherically to represent "around"

The results also revealed that the distribution of separate versus conflated gestures differed only in the condition of multiclausal expressions, such that there were more separate gestures with this type of expression, $X^2(3, N = 88) = 54.1, p < .001$ (see the example below).

دختر چرخ فلک زد، رفت به سمت خانه

dokhtare charkho falak zado, raft be samte khane

lit: Girl cartwheel hit, went to direction of home

"The girl cartwheeled and went toward home"

Gesture: a circular movement of index finger followed by the index finger moving forward on horizontal axis

Syntactic packaging and gesture type

Another way of looking at the correspondence between speech and gesture is through the linguistic units that manner and path information are encoded by. To investigate whether the syntactic packaging in Farsi affects gestural expressions, we examined the type and pattern of gesture production when both manner and path were expressed in speech. First, in speech, participants used *semi-tight* packaging (58%) significantly more than other types: *tight* (28%) and *loose* (14%) syntactic packaging, $X^2(2, N = 260) = 76.95, p < .001$. Second, analysis of gestures showed that path + manner (separate) gestures were produced more with *semi-tight* packaging compared to *tight* and *loose* packages, $X^2(2, N = 56) = 23.56, p < .001$. Yet, in both *semi-tight* and *tight* packages, people mostly used path-only gestures, $X^2(3, N = 123) = 77.43, p < .001$ and $X^2(3, N = 45) = 29.60, p < .001$, respectively. Having only

Table 2. The distribution of different gestures (path only, manner only, path + manner (separate), path + manner (conflated) and other trials among 3 syntactic packaging (*tight*, *semi-tight*, and *loose*) types in 260 trials.

	Path only	Manner only	P + M (separate)	P + M (conflated)	Other
<i>Tight</i> Package ($n = 74$)	27	5	6	7	29
<i>Semi-tight</i> Package ($n = 150$)	69	9	35	10	27
<i>Loose</i> Package ($n = 36$)	13	0	15	6	2

34 gestures in *loose* packaged sentences, no significant difference was obtained for gesture type (see Table 2).

Word order and gesture sequence

For this analysis, we included the trials that contained gestures for both manner and path ($N = 58$). Four trials were excluded from the analyses because the speech either lacked path or manner information or the sentence was incomplete. The order of manner and path expressions in gesture and speech (i.e. which one came first) was coded. Results showed that speakers often used gestures for manner information before path information, $X^2(1, N = 54) = 46.29, p < .001$. Similar to the gesture patterns, in 98% of the cases, we observed the same word order sequence in people's utterances.

Discussion

To our knowledge, this is the first study on motion event conceptualisation in speech and gesture in Farsi. We investigated how motion events are expressed in speech and gestures of the native speakers of Farsi, a language that has characteristics of both Talmy's S- and V-framed languages (Feiz, 2011; Verkerk, 2014). Critically, we used this language to test if linguistic variables such as clause structure, syntactic packaging of manner and path information and canonical word order determine the type and order of the produced gestures, as expected by the interface model (Kita & Özyürek, 2003).

Motion event expressions in speech and gesture

Farsi speakers expressed manner and path components of motion events with similar frequency. The manner information was mostly expressed through adverbs, whereas prepositions and light verbs were used more to describe path of the events. In contrast to speech, participants' dynamic gestures mainly reflected the path, not the manner of motions. These findings are in line

with other results from our lab, which indicate the predominant use of path gestures in contrast to manner gestures among both English and Turkish speakers (Karaduman, Çatak, Bahtiyar, & Göksun, 2015). Using similar stimuli, we found that both English and Turkish speakers preferred to produce path gestures in tasks involving spontaneous gesture use or gesture-only expressions, regardless of the accompanying speech.

The results of our clause-level and syntactic packaging analyses, which point out the relationship between the processing units of language and gestural representations, provided additional evidence for the relation between gesture and language. Regardless of the clause type expressions (one-clause and multiclausal) or syntactic characteristics (mainly for *tight* and *semi-tight* packaging) used in the speech, there was a bias for path gesture. Previous research showed that gestural representations were sensitive to linguistic packaging such that English speakers used conflation of manner and path gestures as a result of their *tight* packaging of path + manner in one-clause. In contrast, Turkish speakers used a combination of manner and path since their speech has characteristics of *semi-tight* packaging and multiclausal structure (Allen et al., 2007; Kita & Özyürek, 2003; Özçalışkan et al., 2016a; Özyürek et al., 2005; Özyürek et al., 2008). Although the expressions in Farsi could take any of the clausal and packaging forms, the predominance of path gestures regardless of linguistic expression of manner and path is not predicted by the interface model.

Other aspects of the data, however, can be taken as evidence in support of the interface model. First, in the examination of how path + manner gestures overlapped with the two types of clauses, we found that the distribution of separate versus conflated gestures differed in the condition of multiclausal expressions. As expected by the interface model, there were more separate gestures produced with the multiclausal expressions. Second, the findings from the syntactic packaging are in line with the clause analysis; despite the prevalence of path gestures in all types of syntactic packaging, the path + manner (separate) gestures were produced more often in trials where manner and path were separated in two linguistic units (28.5%) (*semi-tight* packaging) than trials with *tight* expression of manner and path information (13.5%).

Why do Persian, Turkish, and English-speaking participants prefer path gestures to manner gestures? We cannot answer this question with certainty, but we will discuss several possibilities. Dominance of path gestures might be the salience of path compared to manner for event descriptions (Talmy, 2000). According to Slobin (2004) “people are led to focus on and elaborate

manner if they use a language with high codability in this domain” (p. 237). Consequently, Slobin attributes saliency differences in encoding motion event components to the lexicalisation patterns in languages. S-framed languages tend to include more information about path than V-framed languages. As another categorisation system, Ibarretxe-Antuñano (2004a, 2004b) argues that languages regardless of their typological differences can be classified on the basis of their path saliency. Ibarretxe-Antuñano (2009) presented six factors that contribute to the degree to which a given language is high-path-salient: (1) the preponderance of path particles, (2) the final position of verb in the canonical word order, (3) the tolerance for verb omissions, (4) the existence of dummy verbs, (5) cultural systems in which space and motion play important role, and (6) free word order in the oral language. Farsi possesses three of the above factors. First, it stands out by its preponderance of path particles. Farsi speakers extensively rely on combinations of light verbs and rich set of locative marker prepositions in their expressions of paths. Second, Farsi is a verb-final language. Third, although in Farsi adverbs usually come before the verbs, the other elements of a sentence can move freely around, especially in the spoken language. Thus, Farsi allows expressing path elements in several places in the sentence. On the other hand, manner verbs are few in Farsi, and it has been argued that manner is less salient in languages with limited manner verbs (Slobin, 2004, Verkerk, 2013). In short, the predominance of path gestures in Farsi may be in part due to the path-salient nature of the language. However, this explanation does not apply to Turkish and English, in which we found a similar pattern of path predominance in gestures.

The path-bias could also be due to the nature of the videos, in which path information could be more salient than manner information. This would not be the case. People talked about path and manner information in similar frequency. If the stimuli were path prominent, we would obtain more path than manner information in both modalities.

Finally, the predominance of path gestures may stem directly from how people represent events. In event representations, some components are more noticeable than others, which is essential to language production in guiding the mappings between conceptual structures and linguistic systems (Fisher, 1996; Grimshaw, 1981; Pinker, 1989). One example of prominence in conceptual structure is evident in asymmetries between source and goal of motion. In the case of source – goal relationship, children and adults show a “goal bias” as they are more likely to mark the information about goals of motion events accurately than

the sources of motion events (Fisher, Hall, Rakowitz, & Gleitman, 1994; Lakusta & Landau, 2005; Lakusta, Wagner, O'Hearn, & Landau, 2007; Papafragou, 2010). Although some researchers might only interpret this bias as an asymmetry between goal and source of information, this bias can also be regarded as attention to the *path* of an intended goal. For example, Gergely, Nadasdy, Csibra, & Biro (1995) showed that 12-month-old infants looked longer when an agent did not take the most direct path to its intended goal, suggesting the infants were surprised by the failure to directly follow intentions. Furthermore, infant studies also suggest that extracting path information within motion events develops earlier than extracting manner information in non-linguistic dynamic events (Pruden, Göksun, Roseberry, Hirsh-Pasek, & Golinkoff, 2012; Pruden, Roseberry, Göksun, Hirsh-Pasek, & Golinkoff, 2013). Finally, in a study where 2.5-year-old English-, Japanese-, and Spanish-speaking children were tested to construe novel verbs, they highly relied on the path information regardless of their language-specific encoding of manner and path information (Maguire et al., 2010). Overall, these studies suggest that children prefer salient information in the environment such as *goal – directed* actions or the *path* information that leads to the goal. Taken together, the prominence of path, namely “path bias”, in both speech and gesture could be in part due to the characteristics of Farsi, but most likely also has universal origins in the nature of the representational system. As such, it introduces a new constraint on models of language-gesture interaction that otherwise expect a close correspondence between the utterance and the accompanying gestures.

The present study also investigated the relationship between word order and gesture sequence, a question mostly ignored in the past research. The little work that has explored the issue has demonstrated that the SOV word order emerges even in the gestures of speakers of languages without the SOV constructs, suggesting that there is a cognitive preference for this order (Futrell et al., 2015; Gibson et al., 2013; Goldin-Meadow et al., 2008). Recently, Özçaliskan et al. (2016a) tested whether the order of motion elements (figure – ground – action) found in a particular language affects the way speakers of that language represent the motion events in their gestures with both spontaneous speech production and in gesture-only condition. They found that English- and Turkish-speakers displayed cross-linguistic differences in the way they ordered motion elements (figure – MOTION – ground vs. figure –ground – MOTION) in their speech and gestures, only during online production of language. The cross-linguistic differences did not appear

during offline language condition when they produced only gestures without speech. Our results also indicated that gesture sequences followed the same order as their linguistic counterparts during speaking. In describing motion events, manner gestures that were expressed as adverbs in speech occurred before path gestures that were mainly expressed as a combination of preposition and light verbs at the final part of the sentence. This finding is compatible with the role of language-specific encoding on gesture use as claimed by the interface model.

To summarise, we found evidence in favour of the influence of language on gesture production as postulated by the interface model. However, our results also revealed an important constraint on this relationship, namely the predominance of path gestures regardless of the accompanying linguistic construct. These findings call for closer inspection of factors involved in language-gesture interaction.

Notes

1. “Charkhidan” in Farsi is intransitive.
2. All the results reported in this paper were the same with the Friedman chi-square test.

Acknowledgements

We would like to thank everyone in the Language and Cognition Lab at Koç University for their helpful comments in this research, and Aylin Küntay for her constructive feedback on the findings.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was partially supported by BAGEP Award (Turkish Science Academy Young Investigator Award) given to Tilbe Göksun.

References

- Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, 5(4), 307–331. doi:10.1207/s15427633scc0504_2
- Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes*, 15(6), 593–613. doi:10.1080/016909600750040571
- Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M.S. (2007). Language-specific and universal influences in children’s syntactic packaging of manner and path: A

- comparison of English, Japanese, and Turkish. *Cognition*, 102 (1), 16–48. doi:10.1016/j.cognition.2005.12.006
- Berman, R. A., & Slobin, D. I. (1994). Narrative structure. In R. A. Berman & D. I. Slobin (Eds.), *Relating events in narrative: A crosslinguistic developmental study* (pp. 39–84). Hillsdale, NJ: Lawrence Erlbaum.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43(1), 1–22. doi:10.1006/cogp.2001.0748
- Butterworth, B., & Hadar, U. (1989). Gesture, speech and computational stages: A reply to McNeill. *Psychological Review*, 96, 168–174. doi:10.1037/0033-295X.96.1.168
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41(1), 83–121. doi:10.1016/0010-0277(91)90033-Z
- Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York, NY: Academic Press.
- De Ruiter, J. P. (2007). Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture*, 7 (1), 21–38. doi:10.1075/gest.7.1.03rui
- Feiz, P. (2011). Traveling through space in Persian and English: A comparative analysis of motion events in elicited narratives. *Language Sciences*, 33(3), 401–416. doi:10.1016/j.langsci.2010.10.010
- Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, 31(1), 41–81. doi:10.1006/cogp.1996.0012
- Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. In L. Gleitman & B. Landau (Eds.), *Acquisition of the lexicon* (pp. 333–375). Cambridge: The MIT Press.
- Folli, R., Harley, H., & Karimi, S. (2005). Determinants of event type in Persian complex predicates. *Lingua*, 115(10), 1365–1401. doi:10.1016/j.lingua.2004.06.002
- Futrell, R., Hickey, T., Lee, A., Lim, E., Luchkina, E., & Gibson, E. (2015). Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition*, 136, 215–221. doi:10.1016/j.cognition.2014.11.022
- Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56, 165–193. doi:10.1016/0010-0277(95)00661-H
- Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological Science*, 24(7), 1079–1088. doi:10.1177/0956797612463705
- Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge: Cambridge University Press.
- Gleitman, L. R., & Papafragou, A. (2013). Relations between language and thought. In D. Reisberg (Ed.), *Handbook of cognitive psychology* (pp. 504–523). New York, NY: Oxford University Press.
- Göksun, T., Lehet, M., Malykhina, K., & Chatterjee, A. (2015). Spontaneous gesture and spatial language: Evidence from focal brain injury. *Brain and Language*, 150, 1–13. doi:10.1016/j.bandl.2015.07.012
- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, 64, 257–283. doi:10.1146/annurev-psych-113011-143802
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, 105(27), 9163–9168. doi:10.1073/pnas.0710060105
- Grimshaw, J. (1981). Form, function, and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 165–182). Cambridge, MA: The MIT Press.
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495–514. doi:10.3758/PBR.15.3.495
- Ibarretxe-Antuñano, I. (2004a). Motion events in Basque narratives. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 89–111). Hillsdale, NJ: Lawrence Erlbaum.
- Ibarretxe-Antuñano, I. (2004b). Language typologies in our language use: The case of Basque motion events in adult oral narratives. *Cognitive Linguistics*, 15(3), 317–349.
- Ibarretxe-Antuñano, I. (2009). Path salience in motion events. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & S. Özcaliskan (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 403–414). New York, NY: Psychology Press.
- Karaduman, A. N., Çatak, E. N., Bahtiyar, S., & Göksun, T. (March 2015). *The role of gestures in describing motion in English and Turkish*. Poster presented at the International Conference on Psychological Science, Amsterdam, Netherlands.
- 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(1), 16–32. doi:10.1016/S0749-596X(02)00505-3
- Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, 22(8), 1212–1236. doi:10.1080/01690960701461426
- 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.
- Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, 96, 1–33. doi:10.1016/j.cognition.2004.03.009
- Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development*, 3 (3), 179–197. doi:10.1080/15475440701360168
- Landau, B., & Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition? *Behavioral and Brain Sciences*, 16(2), 255–265. doi:10.1017/S0140525X00029927

- Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S., ... Sanchez-Davis, B. (2010). A developmental shift from similar to language-specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition*, 114(3), 299–319.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2000). *Language and gesture* (Vol. 2). Boston, MA: Cambridge University Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.
- McNeill, D., & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge: Cambridge University Press.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge: Harvard University Press.
- Özçalışkan, S., Lucero, C. & Goldin-Meadow, S. (2016a). Does language shape silent gesture? *Cognition*, 148, 10–18. doi:10.1016/j.cognition.2015.12.001
- Özçalışkan, S., Lucero, C., & Goldin-Meadow, S. (2016b). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science*, 27, 737–747. doi:10.1177/0956797616629931
- Özçalışkan, Ş., & Slobin, D. I. (2003). Codability effects on the expression of manner of motion in Turkish and English. In A. S. Özsoy, D. Akar, M. Nakipoğlu-Demiralp, E. Erguvanlı-Taylan, & A. Aksu-Koç (Eds.), *Studies in Turkish linguistics* (pp. 259–270). Istanbul: Boğaziçi University 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). London: Lawrence Erlbaum.
- Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, 44(4), 1040–1054. doi:10.1037/0012-1649.44.4.1040
- Özyürek, A., Kita, S., Allen, S. E., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from crosslinguistic variations and similarities. *Gesture*, 5(1–2), 219–240. doi:10.1075/gest.5.1.15ozy
- Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, 34(6), 1064–1092. doi:10.1111/j.1551-6709.2010.01107.x
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184. doi:10.1016/j.cognition.2008.02.007
- Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n'roll: The representation of motion in language and cognition. *Cognition*, 84(2), 189–219. doi:10.1016/S0010-0277(02)00046-X
- Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, 98(3), B75–B87. doi:10.1016/j.cognition.2005.05.005
- Papafragou, A., & Selimis, S. (2010a). Event categorization and language: A cross-linguistic study of motion. *Language and Cognitive Processes*, 25(2), 224–260. doi:10.1080/01690960903017000
- Pinker, S. (1989). *Learnability and cognition*. Cambridge, MA: The MIT Press.
- Pouw, W. T., De Nooijer, J. A., Van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.00359
- Pruden, S. M., Göksun, T., Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Find your manners: How do infants detect the invariant manner of motion in dynamic events? *Child Development*, 83(3), 977–991. doi:10.1111/j.1467-8624.2012.01737.x
- Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Infant categorization of path relations during dynamic events. *Child Development*, 84(1), 331–345. doi:10.1111/j.1467-8624.2012.01843.x
- Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: The MIT Press.
- Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195–217). Oxford: Oxford University Press.
- Slobin, D. I. (2004). The many ways to search for a frog. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Vol. 2. Typological and contextual perspectives* (pp. 219–257). Mahwah, NJ: Lawrence Erlbaum.
- Talmy, L. (1983). How language structures space. In H. Pick & L. Acredolo (Eds.), *Spatial orientations: Theory, research, and application* (pp. 225–282). New York, NY: Plenum Press.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and semantic description* (pp. 36–149). Cambridge: Cambridge University Press.
- Talmy, L. (1991). Path to realization: A typology of event conflation. In *Proceedings of the seventeenth annual meeting of the Berkeley linguistics society* (pp. 480–519). Berkeley, CA: Berkeley Linguistics Society.
- Talmy, L. (2000). *Toward a cognitive semantics: Vol. II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.
- Verkerk, A. (2013). Scramble, scurry and dash: The correlation between motion event encoding and manner verb lexicon size in Indo-European. *Language Dynamics and Change*, 3(2), 169–217. doi:10.1163/22105832-13030202
- Verkerk, A. (2014). Diachronic change in Indo-European motion event encoding. *Journal of Historical Linguistics*, 4(1), 40–83.
- Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.