Expression of motion events in Farsi

Nilooftar Akhavan, Nazbanou Nozari and Tilbe Göksun

Department of Psychology, Koç University, Istanbul, Turkey; Department of Neurology, School of Medicine, Johns Hopkins University, Baltimore, MD, USA; Department 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.

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; Özcalışkan, 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 yuvarland: aşağıya =
preposition encoding the path; *yuvarlandı* = 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) *çıktı* = subordinate clause encoding the manner; *çıktı* = 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, Özürek, Kita, Brown, Furman, Ishizuka, & Fujii, 2007; Berman & Slobin, 1994; Choi & Bowkerman, 1991; Özçalıskan & 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 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 & Özürek, 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 & Özürek, 2003; Krauss, Chen, & Gottfesnum, 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 & Özürek (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 & Özürek, 2003; McNeill, 2000; McNeill & Duncan, 2000). Kita and Özürek (2003) compared English and Turkish speakers’ gestural representations to investigate the language-specific encodings of motion events (see also Kita et al., 2007; Özürek, 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 multiclause 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 multiclause 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 & Özürek, 2003; Özürek & Kita, 1999; Özürek 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 (Özürek 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 disassociation 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 *davidan* “to run” is a verb that contains manner information:

```
az tappe baala davidan
lit: [From hill] up run
```

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).

```
dovan davan dor -e - [z tapp] charkhidan
lit: runningly runningly around [...] to circle
```

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):

```
lei ley konan az [...] khaarej shodan
lit: Hop Hop doing from [...] exit to become
```

“lei 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.

```
lei ley kardan az [...] khaarej shodan
lit: Hop hop do. From [...] exit to become
```

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, Malykhania, & 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.](image)

**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):
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

dor

“to run”

lit: run

(5b) Verb

charkhidan

“around”

lit: a circle around

(5c) Verb + preposition

dor charkhidan

“circle”

lit: around circle

(5d) Light verb

az bein rad shodan

“to come”

lit: from between pass to become

(5e) Light verb + preposition

aamadan

az bein rad shodan

“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) Dokhtar dor […] 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) rah miraf 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) Dokhtar bodo bodo konan be samte […] raft

lit: Girl run run doing to direction of […] went

“The girl went toward the […] while running”

(6d) 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
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), X²(2, N = 324) = 137.35, p < .001. 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, X²(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), X²(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) conflations (M = 11%, SD = 13.14), X²(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
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 before completion. On average, 43% of these descriptions were encoded by one-clause and the rest (56%) by multiclause 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:

"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).

"Dokhtar bepar be par 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.

---

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

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

**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 Multiclause) types in all of all the event descriptions that included both manner and path ($N = 260$).

<table>
<thead>
<tr>
<th></th>
<th>Path only</th>
<th>Manner only</th>
<th>P + M (separate)</th>
<th>P + M (conflated)</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-clause ($n = 158$)</td>
<td>59</td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>69</td>
</tr>
<tr>
<td>Multiclause ($n = 202$)</td>
<td>90</td>
<td>11</td>
<td>51</td>
<td>19</td>
<td>31</td>
</tr>
</tbody>
</table>

The results also revealed that the distribution of separate versus conflated gestures differed only in the condition of multiclause 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).

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

**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.

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

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 the interface model (Kita & Özyürek, 2003).
with other results from our lab, which indicate the predo-
minant use of path gestures in contrast to manner ges-
tures 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 representa-
tions, 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 speak-
ers 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 distri-
bution of separate versus conflated gestures differed in
the condition of multiclausal expressions. As expected
by the interface model, there were more separate ges-
tures 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 separ-
ated 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 partici-
pants 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 sal-
ciency 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. Ibar-
retxe-Antuñano (2009) presented six factors that contrib-
ute 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 predomi-
nance of path gestures in Farsi may be in part due to the
path-salient nature of the language. However, this expla-
nation 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 concep-
tual structures and linguistic systems (Fisher, 1996;
Grimshaw, 1981; Pinker, 1989). One example of promi-
ience 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 infor-
mation about goals of motion events accurately than

---

**References**

Movement and path in spontaneous speech. In K. R.itschky, H. Kita, & J. H. van der Linden (Eds.), 
*Understanding Motion (pp. 261–293)*. Dordrecht: Springer. 

It's not what you say but where you say it: Production and comprehension of 
spontaneous gestures. In K. Ritschky, H. Kita, & J. H. van der Linden (Eds.), 
*Understanding Motion (pp. 115–147)*. Dordrecht: Springer. 


Culture and the embodied mind. In E. de Vignemartin, 
W. D. van Berkel, H. de Weger, & J. Kolk (Eds.), 

Path and manner in spontaneous speech. In K. Ritschky, H. Kita, 
& J. H. van der Linden (Eds.), *Understanding Motion (pp. 295–325)*. Dordrecht: Springer. 


---

*Language, Cognition and Neuroscience* 801
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 in 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**


