



Does language shape silent gesture? ☆

Şeyda Özçalışkan^{a,*}, Ché Lucero^b, Susan Goldin-Meadow^b

^a Georgia State University, United States

^b University of Chicago, United States



ARTICLE INFO

Article history:

Received 25 June 2014

Revised 26 November 2015

Accepted 2 December 2015

Keywords:

Gesture
Cross-linguistic differences
Language and cognition
Motion events

ABSTRACT

Languages differ in how they organize events, particularly in the types of semantic elements they express and the arrangement of those elements within a sentence. Here we ask whether these cross-linguistic differences have an impact on how events are represented nonverbally; more specifically, on how events are represented in gestures produced without speech (silent gesture), compared to gestures produced with speech (co-speech gesture). We observed speech and gesture in 40 adult native speakers of English and Turkish ($N = 20$ /per language) asked to describe physical motion events (e.g., running down a path)—a domain known to elicit distinct patterns of speech and co-speech gesture in English- and Turkish-speakers. Replicating previous work (Kita & Özyürek, 2003), we found an effect of language on gesture when it was produced with speech—co-speech gestures produced by English-speakers differed from co-speech gestures produced by Turkish-speakers. However, we found no effect of language on gesture when it was produced on its own—silent gestures produced by English-speakers were identical in how motion elements were packaged and ordered to silent gestures produced by Turkish-speakers. The findings provide evidence for a natural semantic organization that humans impose on motion events when they convey those events without language.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Languages differ in how they organize the semantic components of an event, and these organizational preferences influence both the types and the arrangement of semantic elements conveyed in speech and co-speech gesture. Here we ask whether language-specific differences observed in speech have an effect beyond online production¹—in particular, we ask whether language-specific differences influence nonverbal representation of events in gesture when those gestures are produced without speech, that is, in silent gesture. If the semantic organization of events in a particular language can influence offline nonverbal representations, the arrangement of semantic elements in silent gesture should be similar to the arrangement of semantic elements in speech and co-speech gesture. If, however, the semantic organization of events in

a particular language is *not* easily mapped onto offline nonverbal representations of events, the arrangement of semantic elements in silent gesture may differ from the arrangement in speech and co-speech gesture, and perhaps be similar across speakers of different languages. We study this question by observing the gestures speakers produce when describing motion events, a domain characterized by strong cross-linguistic differences in types of semantic elements expressed and how those semantic elements are arranged within a sentence. We ask whether gestures that *do* and *do not* accompany speech display these cross-linguistic differences.

Spatial motion, a domain that displays wide variation as well as patterned regularities across the world's languages in how it is expressed, offers a unique arena in which to examine cross-linguistic variability in gesture. Previous work (Talmy, 1985, 2000) identified the 'motion event' as a basic building block of language and cognition, and proposed a set of motion elements that appear to be universal. Take, for example, a simple motion scene, such as a *baby crawling into a room*. Many languages make it possible to refer to the *figure* (baby) separately from the *ground* she traverses (room), to trace its *path* (into), or to comment on the *manner* with which she moves (crawling). However, languages also vary systematically in how they express each element type, displaying for the most part a binary split across the world's languages (Talmy, 2000). Speakers of English—a satellite-framed language—use a *conflated strategy* in speech; they typically express

* This research was supported by a grant from the March of Dimes Foundation (#12-FY08-160) to both authors and NIDCD (R01 DC00491) to SGM. We thank Andrea Pollard, Vasthi Reyes, Christianne Ramdeen and Burcu Sancar for help with data collection, transcription and coding.

* Corresponding author at: Georgia State University, Department of Psychology, PO Box 5010, Atlanta, GA 30302-5010, United States.

E-mail address: seyda@gsu.edu (Ş. Özçalışkan).

¹ We borrowed the term 'online' from Slobin's (1996) thinking-for-speaking hypothesis, and coined the term 'offline' to highlight the contrast between gestures produced when speaking (*online*) and gestures produced when not speaking (*offline*).

manner and path components in a compact description with manner in the verb (*crawl*) and path outside the verb (*into*), both expressed within a single clause, as in ‘baby crawls into the house.’ In contrast, speakers of Turkish—a verb-framed language—use a *separated strategy* in speech, with path in the verb in one clause (‘girer’ = enter), and manner outside of the verb and, importantly, in a subordinate separate clause (‘sürünerek’ = crawl), as in ‘bebek eve girer sürünerek’ = baby house-to enters by crawling; Turkish-speakers often express only the path, omitting manner entirely (Allen et al., 2007; Özçalışkan, 2009; Özçalışkan & Slobin, 1999). In addition to these differences in type and packaging of motion elements, the two languages also differ in where the primary motion element (i.e., the main verb, be it manner or path) is placed within a sentence; the motion verb is typically situated at the end of a sentence in Turkish (‘Bebek ev-e GİRER’ = baby house-to ENTERS; Figure-Ground-MOTION), but in the middle of the sentence in English (Figure-MOTION-Ground, ‘Baby CRAWLS into house’). Turkish and English thus provide a strong contrast in how motion events are described, allowing us to examine the effects of language on thinking.

The thinking-for-speaking hypothesis proposed by Slobin (1996) postulates that speakers’ conceptualization of an event is influenced by the categorical distinctions available in their language, but only during online production of the language. Recent work examining the effects of language on perceiving and remembering motion events across structurally different languages suggests an effect of language on thinking when the cognitive tasks are accompanied by verbalization of the event, but no effect of language on thinking when verbalization was not allowed. For example, when participants were asked to compare an original event to a new event that differed either in manner or path of motion, they showed a bias for manner or path (depending upon their language) when the task involved verbal description of the event (either describing the event in their native language, or inferring the meaning of novel motion verbs), but did not show the bias when the task was nonverbal and thus did not involve language (Gennari, Sloman, Malt, & Fitch, 2002; Hohenstein, 2005; Naigles & Terrazas, 1998; see Özçalışkan, Lucero, & Goldin-Meadow, under review; Özçalışkan, Stites, & Emerson, in press, for a review).

Our focus here is on gesture, which is, by definition, nonverbal. However, it is now well known that the gestures speakers produce along with their speech (i.e., co-speech gestures) often mirror patterns found in speech (Gullberg, Hendricks, & Hickmann, 2008; Kita & Özyürek, 2003). In terms of the motion events that are our focus here, English- and Turkish-speakers produce co-speech gestures that mirror the patterns in their speech and thus differ from one another. More specifically, English-speakers display the *conflated* pattern characteristic of spoken English in their co-speech gestures, synthesizing manner and path components into a single gesture (e.g., wiggle fingers while moving the hand from left to right to convey running along a left-to-right path; Fig. 1B₁). In contrast, Turkish-speakers display the *separated* pattern characteristic of spoken Turkish in their co-speech gestures, producing one gesture for manner (e.g., wiggle fingers in one spot to convey running) and another for path (move hand left to right to convey moving along a left-to-right path; Kita & Özyürek, 2003), and often express only path of motion in gesture (Fig. 1A₁; Özçalışkan, in press).

Our question is whether the effect that language has on co-speech gesture—an online effect of language on thinking—can also be found offline, that is, when speakers are asked to abandon speech and use only gesture to describe a motion event. In other words, does language have an effect on silent gesture? Previous work on the impact of cross-linguistic differences in word order

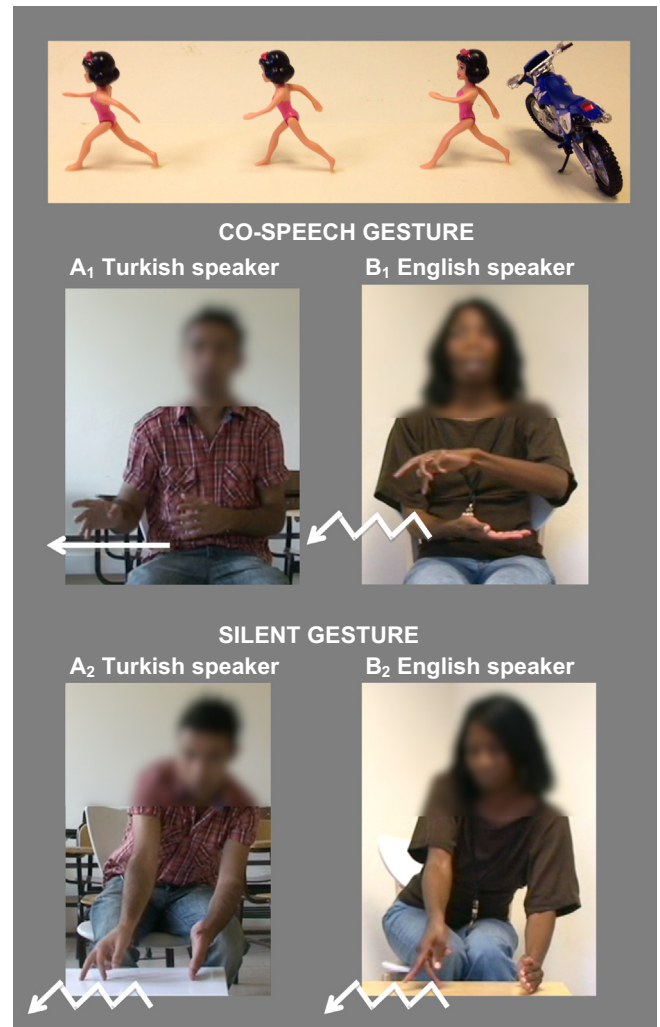


Fig. 1. Example stimulus scene of a girl running away from a motorcycle (top) and its description in co-speech gesture (A₁, B₁) and silent gesture (A₂, B₂) by speakers of Turkish (A pictures on left) and English (B pictures on right). In **co-speech gesture**, English speakers preferred to express manner (walking fingers) and path (trajectory away from speaker) simultaneously within a single gesture (B₁), and Turkish speakers preferred to express path (trajectory towards speaker’s right with both hands) by itself, omitting manner entirely (A₁). In **silent gesture**, English and Turkish speakers both preferred to express manner and path simultaneously within a single gesture (A₂, B₂); the upward facing right palm in B₁ and the sideways facing left palm in A₂ and in B₂ represent the ground (i.e., the motorcycle); the participant did not produce a gesture for the ground in A₁.

on silent gesture has found no evidence for an offline effect of language. For example, speakers of English, Turkish, Spanish, and Chinese displayed the word order patterns characteristic of their respective languages (either subject-verb-object, SVO, or subject-object-verb, SOV) when speaking, but when asked to produce gestures without speech, their gestures did not display the same cross-linguistic differences and, in fact, all followed the same order, SOV (Goldin-Meadow, So, Özyürek, & Mylander, 2008; see also Gibson et al., 2013; Hall, Mayberry, & Ferreira, 2013; Langus & Nespors, 2010; Meir, Lifshitz, Ilkbasaran, & Padden, 2010; Schouwstra & de Swart, 2014). We explore the generality of this finding by extending the work to a second set of cross-linguistic differences—how manner and path are organized within a sentence. In addition, unlike previous studies, we include analyses of co-speech gesture vs. silent gesture, allowing a within-modality contrast of online vs. offline effects of language on thinking.

2. Methods

2.1. Sample

Participants were 20 adult native English-speakers ($M_{age} = 43$ [SD = 13], range = 23–71, 15 females) and 20 adult native Turkish-speakers ($M_{age} = 26$ [SD = 7], range = 20–46, 10 females). The English and Turkish data were collected in metropolitan areas in the United States and in Turkey, respectively.² Participants received monetary compensation.

2.2. Procedure

2.2.1. Data collection

Participants were videotaped while describing 12 three-dimensional scenes that depicted motion along three different types of paths—4 to a landmark (e.g., *run into house*), 4 over a landmark (e.g., *flip over a beam*), and 4 from a landmark (e.g., *run away from motorcycle*)—with various manners (e.g., *run, flip, crawl*). For each type of path, 3 scenes depicted movement across a spatial boundary (*into, out, over*) and 1 depicted movement that did not cross a boundary (*towards, along, away from*); see Table 1 for a list of all manners and paths used in the 12 scenes. Each scene was pre-constructed on a 5 × 15 inch white foam board and contained a landmark and three stationary identical dolls (named Eve in English, Oya in Turkish) with varying postures to capture three snapshots of a continuous motion with manner and path—all glued to the foam board allowing for uniform presentation of materials across groups (see Fig. 1, top). Participants were told that Eve appears three times in each scene, but that they should think of Eve's movement as a single continuous motion and describe it as such. Participants were shown each scene one at a time in counter-balanced order³ and asked to describe the scenes twice: (1) in speech while using their hands as naturally as possible, thus producing *co-speech gesture*; (2) in gesture using only their hands without any speech, thus producing *silent gesture*. Participants described all scenes first in speech and then in silent gesture; order was not counterbalanced because we were concerned that producing silent gestures first would call attention to gesture and thus affect the naturalness of participants' co-speech gestures.⁴ Each participant completed two practice trials before describing the scenes in speech, and two before describing the scenes in silent gesture.

2.2.2. Data coding

We transcribed all *speech* produced in the co-speech gesture condition and segmented it into sentence-units. Each sentence-unit contained at least one verb and associated arguments and subordinate clauses (e.g., 'Eve runs away from the motorcycle'; 'Oya motorsikletten uzaklaşır' = Oya motorcycle-from moves-away; 'Oya motorsikletten uzaklaşır koşarak' = Oya motorcycle-from moves-away running). We also transcribed all *gestures* that accompanied each sentence unit in the co-speech gesture condition, and that were produced on their own in the silent gesture condition. Gesture was defined as communicative hand movements that had an identifiable beginning and end; we included only gestures

Table 1

List of motion event types used in the study.

Item	Type of path	Type of manner	Event description
MOTION TO LANDMARK			
1	INTO a landmark	Run	Run into house
2	INTO a landmark	Crawl	Crawl into house
3	INTO a landmark	Climb	Climb into treehouse
4	TOWARD a landmark	Walk	Walk towards crib
MOTION OVER LANDMARK			
5	OVER a landmark	Crawl	Crawl over carpet
6	OVER a landmark	Jump	Jump over hurdle
7	OVER a landmark	Flip	Flip over beam
8	ALONG a landmark	Crawl	Crawl along tracks
MOTION FROM LANDMARK			
9	OUT of a landmark	Run	Run out of house
10	OUT of a landmark	Roll	Roll out of tunnel
11	OUT of a landmark	Crawl	Crawl out of house
12	AWAY from landmark	Run	Run away from motorcycle

that conveyed characteristic actions or features associated with the stimulus scenes.

Sentence-units were further coded along two dimensions. (1) Packaging of different types of motion elements: *conflated*, where manner and path are both conveyed within a single spoken clause or within a single gesture; or *separated*, where manner and path are conveyed in separate spoken clauses or in separate gestures. A sentence-unit was classified as *separated* if it contained manner-only (e.g., 'she is running,' 'koşuyor' = running), path-only (e.g., 'she is moving away from the motorcycle,' 'motorsikletten uzaklaşır' = motorcycle-from moves-away), or manner and path, each conveyed in a separate clause (which was expressed only once in English, but frequently in Turkish, 'motorsikletten uzaklaşır koşarak' = motorcycle-from moves-away running). (2) Ordering of semantic elements: *Figure-Ground-MOTION* or *Figure-MOTION-Ground*. We classified spoken utterances according to the placement of the primary motion element—the main verb of the sentence unit, which typically conveyed manner in English and path in Turkish. When produced, path in English (typically conveyed in the satellite) and manner in Turkish (typically conveyed in secondary clauses) were always contiguous with the main verb. Similarly, we classified gesture strings according to the placement of the motion gesture—either a manner gesture alone, a path gesture alone, a manner + path conflated gesture, or a sequential manner gesture followed by a path gesture (or vice versa) within a single sentence-unit; sequential manner-path gestures were almost always contiguous (88%; in the 3 exceptions, the two motion gestures were separated by a gesture for ground and thus were excluded from the ordering analysis).⁵ Participants who conveyed more than one semantic element in gesture within the same sentence-unit typically combined their motion gesture with a gesture for the ground, frequently omitting a gesture for the figure (see Fig. 2A₁–A₂); the ground element was typically expressed by a stationary open palm either facing sideways or upward (see left

⁵ In some cases, participants produced both a separated and a conflated gesture within the same sentence-unit (i.e., a mixed pattern, see Özyürek, Furman, & Goldin-Meadow, 2015), which accounted for roughly 6% of sentence-units across conditions and languages, $M_{range} = 0.55$ –1.20; all of these instances were excluded from our analysis for packaging as they cannot be classified as separated or conflated. The majority of the sentence-units that showed a mixed packaging pattern were also excluded from our ordering analysis (100% for English, 11 instances, 79% for Turkish, 11 instances), because they either conveyed only motion information (e.g., a manner gesture followed by a conflated manner-path gesture) or did not show a consistent order (e.g., manner gesture followed by a ground gesture and then a path gesture). The few instances with mixed packaging pattern that showed a consistent order in Turkish (3 instances, 21% of the mixed category) were all included in the ordering analysis.

² The data are part of a larger project examining effects of blindness on gesture production (Özçalışkan et al., under review); sighted participants were matched to blind participants in each culture in the parent study, which led to the differences in age and gender across the two language groups in our study.

³ Counterbalancing was done in 3 blocks, each containing 4 items (one of each of the 3 different path types and one non-boundary crossing event).

⁴ The co-speech gesture condition was followed by two other language tasks unrelated to the goals of the current analysis—one eliciting narratives, and the other eliciting metaphors; as a result, the silent gesture condition never immediately followed the co-speech gesture condition, making it unlikely that responses in the co-speech condition had a direct impact on responses in the silent gesture condition.

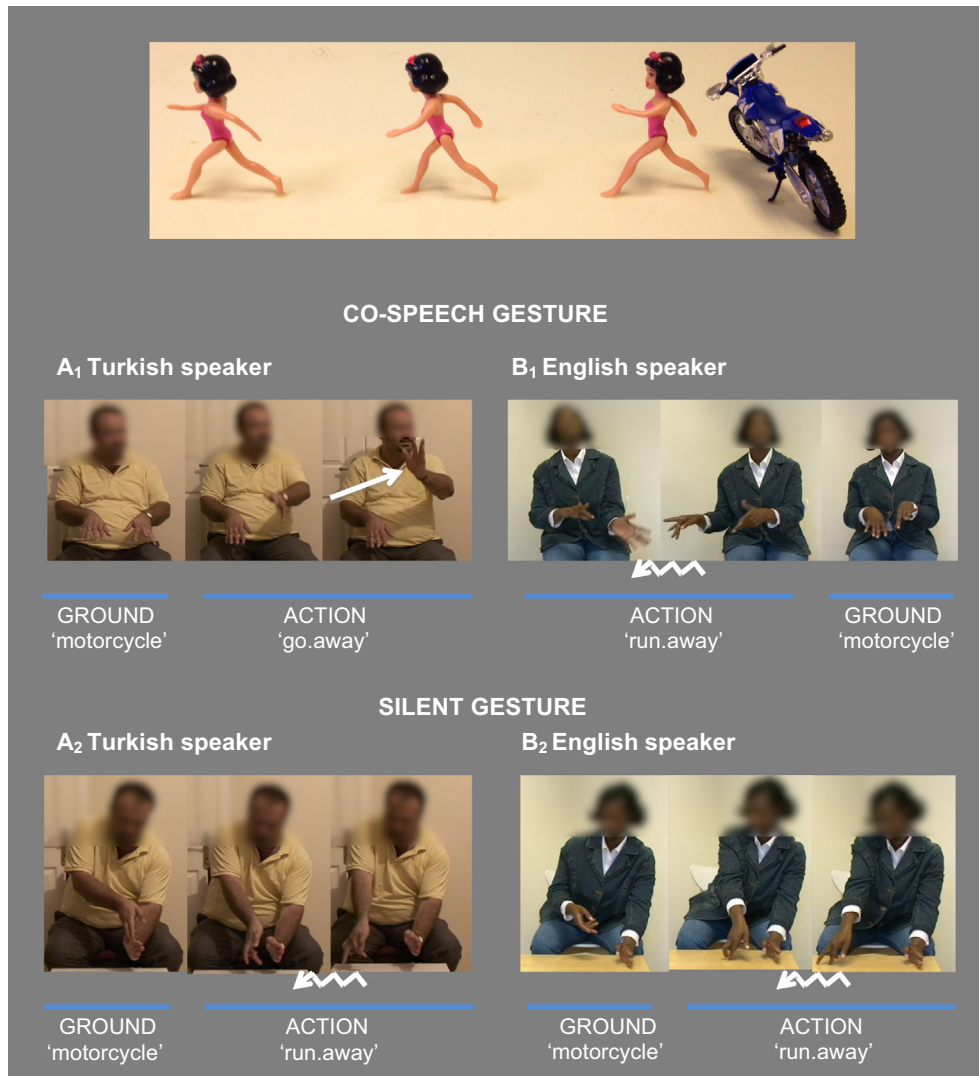


Fig. 2. Example stimulus scene of running away from motorcycle (top) and its description in co-speech gesture (A₁, B₁) and silent gesture (A₂, B₂) by speakers of Turkish (A pictures on left) and English (B pictures on right). In **co-speech gesture**, English speakers preferred to express motion (run.away) first, followed by ground (motorcycle); Turkish speakers preferred to express ground (motorcycle) first, followed by motion (go.away). In **silent gesture**, both English and Turkish speakers preferred to express ground (motorcycle) first, followed by motion (run.away); the downward facing palm in the first scene of A₁ and the last scene of B₁, and the sideways facing left palm in all three scenes of A₂–B₂ indicate ground (i.e., the motorcycle).

palm in Fig. 1A₂–B₂ and Fig. 2A₂–B₂). Reliability was assessed with an independent coder; agreement between coders was 94% for identifying gestures, 97% for describing gesture form, and 100% and 95% for coding motion elements in speech and gesture, respectively.

2.2.3. Data analysis

We analyzed the data by fitting generalized linear mixed-effect models with a Poisson distribution function.⁶ Language (English, Turkish) was a between-subjects and within-items factor. Ordering (figure-motion-ground, figure-ground-motion), Packaging (separated, conflated), and Output channel (speech, co-speech gesture, silent gesture) were within-subject and within-item factors. Path-Type (to, over, from) was a within-subject and between-item factor. We treated Subject ($N = 40$) and Scene ($N = 12$) as random effects, including random intercepts for both in all analyses.⁷ To control

⁶ We conducted the analyses using R (R Core Team, 2013) and the glmer() function in the lme4 library (Bates, Maechler, Bolker, & Walker, 2014).

⁷ We used the “Maximal” approach (Barr, Levy, Scheepers, & Tily, 2013) and included random slopes for Subject and for Scene where the data were able to support the complexity of these slope estimations (Barr, 2013).

for effects of type of path, we included PathType as a fixed effect in all models. Our procedure was the same for all statistical tests. We first fit a model that included our three primary factors (Language, Output channel, and either Packaging or Ordering) to the data. We then fit a reduced model that excluded one of the factors to the same data. Finally, we compared the relative goodness of fit of the models using a likelihood ratio test via the anova() command. This procedure compares the relative fits (expressed as log-likelihoods) of the two models to test whether the factor removed in the reduced model is statistically significant. Comparing the fits of the models in this way provides a chi-square statistic, degrees of freedom, and a p -value, all of which we report for each test.

3. Results

3.1. Packaging motion elements

We first examined interactions among factors for packaging, and found a significant Language (Turkish, English) \times Packaging (Separated, Conflated) \times Output channel (Speech, Co-speech Gesture, Silent Gesture) interaction ($\chi^2 = 77.10$, $df = 2$, $p < .001$).

We also found a significant Language \times Packaging interaction for each Output channel: Speech ($\chi^2 = 11.55$, $df = 1$, $p < .001$); Co-speech Gesture ($\chi^2 = 16.99$, $df = 1$, $p < .001$); Silent Gesture ($\chi^2 = 15.29$, $df = 1$, $p < .001$).

3.1.1. Speech and Co-speech Gesture

In Speech (Fig. 3A), Turkish-speakers produced significantly more Separated packaging responses than English-speakers ($\chi^2 = 14.00$, $df = 1$, $p < .001$). A Separated response could be a path on its own (Turkish: *ayrılıyor*; English: she is going away), manner on its own (Turkish: *koşuyor*; English: she is running), or manner in one clause and path in another (Turkish: *motorsikletten ayrılıyor koşarak*; English: she is going away from the motorcycle by running). Conversely, English-speakers produced significantly more Conflated responses than Turkish-speakers (English: she runs away from motorcycle; Turkish: *motorsikletin yanından koşuyor*, with manner and path in the same clause) ($\chi^2 = 17.73$, $df = 1$, $p < .001$).

As expected, we found the same pattern in Co-speech Gesture (Fig. 3B). Turkish-speakers produced significantly more Separated responses than English-speakers ($\chi^2 = 5.47$, $df = 1$, $p = .02$). A separated response could be a gesture for manner on its own (e.g., wiggling fingers), a gesture for path on its own (e.g., moving the hand forward), or separate gestures for manner and path both within the same sentence-unit (e.g., wiggling fingers followed by the hand moving forward).⁸ Conversely, English-speakers produced significantly more Conflated responses than Turkish-speakers (e.g., wiggling the fingers, manner, while moving the hand forward, path) ($\chi^2 = 4.44$, $df = 1$, $p = .04$).⁹ We thus found the recognized cross-linguistic differences for our participants in both speech and co-speech gesture.

To determine whether these cross-language differences varied by type of path (TO, FROM, OVER), we asked whether there was a significant Language \times Packaging \times PathType interaction, first for Speech and then for Co-speech Gesture. The 3-way interaction was significant for Speech ($\chi^2 = 22.42$, $df = 2$, $p < .001$), but Turkish-speakers produced more Separated responses than English-speakers, and English-speakers produced more Conflated responses than Turkish-speakers, for each PathType (Fig. 3A), and the crucial Language \times Packaging interaction was significant for TO ($\chi^2 = 10.22$, $df = 1$, $p = .001$); FROM ($\chi^2 = 13.23$, $df = 1$, $p < .001$), and OVER ($\chi^2 = 3.94$, $df = 1$, $p = .05$). The 3-way interaction was not significant for Co-speech Gesture ($\chi^2 = 0.34$, $df = 2$, $p = .85$), indicating no variation across PathType (Fig. 3B).

3.1.2. Silent Gesture

Our primary question was whether, when asked to communicate without speech, Turkish- and English-speakers would display in Silent Gesture the differences found in their Speech and Co-speech Gestures. They did not (Fig. 3C). Instead, both groups produced more Conflated responses than Separated responses in Silent Gesture (as seen in the significant main effect for Packaging, DiffM: -9.38 , $\chi^2 = 46.15$, $df = 1$, $p < .001$). There was no detectable difference between English- and Turkish-speakers in their predominant Conflated responses ($\chi^2 = 0.55$, $df = 1$, $p = .46$), but English-speakers did produce more Separated responses than Turkish-speakers ($\chi^2 = 5.57$, $df = 1$, $p = .02$); note, however, that this effect was small and reversed the pattern found in Speech and Co-speech Gesture (where Turkish-speakers produced more

Separated responses than English-speakers, Fig. 3A and B). To determine whether the Silent Gesture pattern held across type of path, we asked whether both Turkish- and English-speakers produced more Conflated responses than Separated responses for each PathType, and found a significant main effect of Packaging for TO ($\chi^2 = 29.62$, $df = 1$, $p < .001$); FROM ($\chi^2 = 14.58$, $df = 1$, $p < .001$); and OVER ($\chi^2 = 19.53$, $df = 1$, $p < .001$).

3.2. Ordering of semantic elements

We again began by examining interactions among factors, this time for Ordering of semantic elements. We found a Language (Turkish, English) \times Ordering (Figure-Ground-MOTION, Figure-MOTION-Ground) \times Output channel (Speech, Co-speech Gesture, Silent Gesture) interaction ($\chi^2 = 72.14$, $df = 2$, $p < .001$). We also found a significant Language \times Ordering interaction for each Output channel: Speech ($\chi^2 = 34.71$, $df = 1$, $p < .001$), Co-speech Gesture ($\chi^2 = 6.89$, $df = 1$, $p = .01$), and Silent Gesture ($\chi^2 = 8.01$, $df = 1$, $p = .004$).

3.2.1. Speech and Co-speech Gesture

In speech (Fig. 4A), Turkish-speakers produced more Figure-Ground-MOTION responses than did English-speakers ($\chi^2 = 85.51$, $df = 1$, $p < .001$), and English-speakers produced more Figure-MOTION-Ground responses than Turkish-speakers ($\chi^2 = 29.07$, $df = 1$, $p < .001$). In Co-speech Gesture (Fig. 4B), Turkish-speakers also produced more Figure-Ground-MOTION responses than English-speakers ($\chi^2 = 6.43$, $df = 1$, $p = .01$), and English-speakers produced more Figure-MOTION-Ground responses than Turkish-speakers, although this difference was not significant ($\chi^2 = 2.61$, $df = 1$, $p = .11$). To determine whether these cross-language differences varied by type of path, we asked whether there was a significant Language \times Ordering \times PathType interaction for Speech and for Co-speech Gesture. The 3-way interaction was not significant for either Speech ($\chi^2 = 4.94$, $df = 2$, $p = .08$) or Co-speech Gesture ($\chi^2 = 3.04$, $df = 2$, $p = .22$), indicating no variation across PathType for either Output channel.

3.2.2. Silent Gesture

Turning next to Silent Gesture (Fig. 4C), again we found that Turkish- and English-speakers did not display in Silent Gesture the differences found in their Speech and Co-speech Gesture. Both groups produced more Figure-Ground-MOTION responses than Figure-MOTION-Ground responses (as seen in the significant main effect of Ordering, DiffM: -7.80 , $\chi^2 = 31.63$, $df = 1$, $p < .001$). There was no detectable difference between English- and Turkish-speakers in their predominant Figure-Ground-MOTION responses ($\chi^2 = 0.10$, $df = 1$, $p = .75$), but English-speakers did produce the relatively rare Figure-MOTION-Ground response more often than Turkish-speakers ($\chi^2 = 6.37$, $df = 1$, $p = .01$); this effect (although small) might reflect an influence of English on participants' silent gestures. To determine whether the Silent Gesture pattern held across type of path, we asked whether both Turkish- and English-speakers produced more Figure-Ground-MOTION responses than Figure-MOTION-Ground responses for each PathType, and found a significant main effect of Ordering for TO ($\chi^2 = 14.92$, $df = 1$, $p < .001$); FROM ($\chi^2 = 7.52$, $df = 1$, $p = .01$); and OVER ($\chi^2 = 16.06$, $df = 1$, $p < .001$).

4. Discussion

Our study asked whether the organization of motion events found in a particular language affects the way speakers of that language represent the events nonverbally, not only online (i.e., in gestures produced along with speech, co-speech gesture) but also

⁸ In Co-speech Gesture, Turkish-speakers produced a gesture for manner alone in 1.55 (SD = 2.35) responses, path alone in 9.45 (SD = 5.93) responses, and sequential path-manner in 1.00 (SD = 1.95) responses; comparable means for English-speakers were: 1.80 (SD = 2.63), 6.35 (SD = 5.82), and 0.25 (SD = 0.55).

⁹ Separated responses outnumbered Conflated responses when collapsing across language groups in both Speech and Co-speech Gesture; this main effect of Packaging was significant for Co-speech Gesture (DiffM: 4.53, $\chi^2 = 12.67$, $df = 1$, $p < .001$), but not for Speech (DiffM: 2.65, $\chi^2 = 0.82$, $df = 1$, $p = .36$).

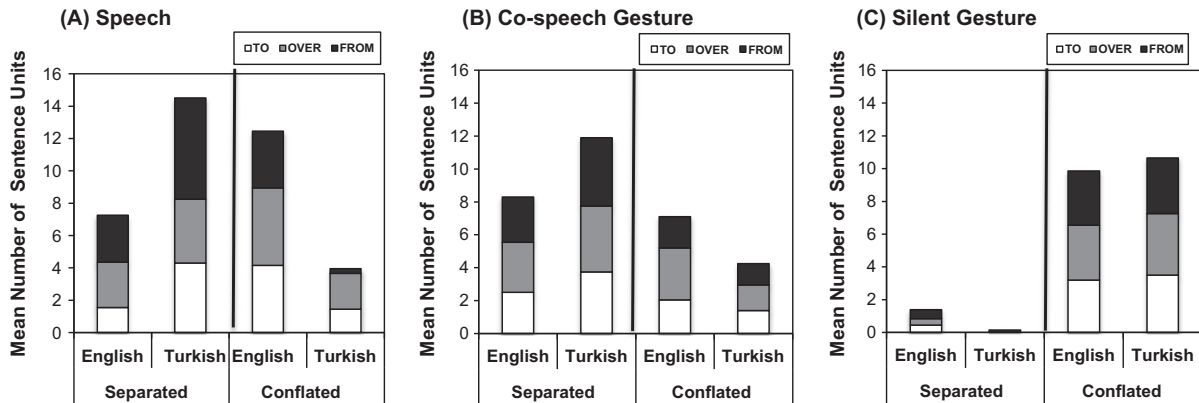


Fig. 3. Mean number of sentence units with separated (manner-only, path-only, manner-path) or conflated (manner + path) motion elements in speech (A), in gesture with speech (co-speech gesture, B) and in gesture without speech (silent gesture, C). Turkish and English participants differ in both speech and co-speech gesture, but not silent gesture; these patterns hold for each of the three path types (TO, OVER, FROM).

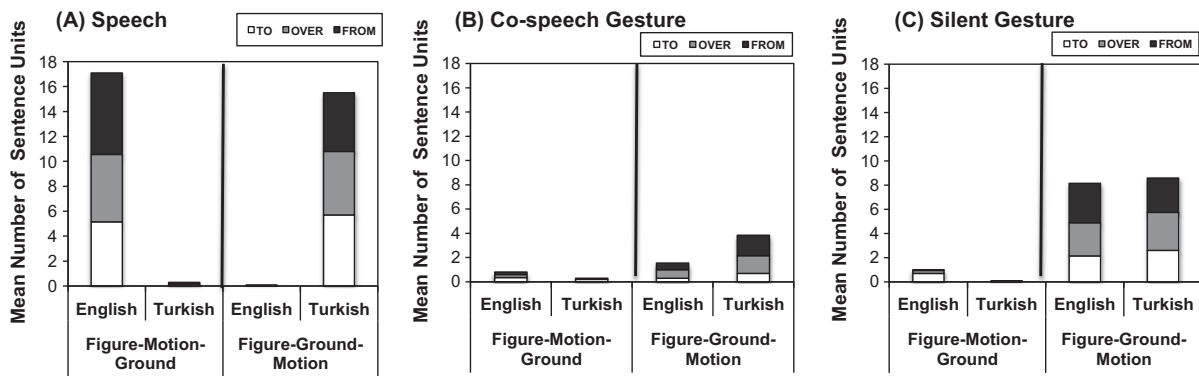


Fig. 4. Mean number of sentence units that follow Figure-Ground-MOTION or Figure-MOTION-Ground orders in speech (A), in gesture with speech (co-speech gesture, B) and in gesture without speech (silent gesture, C). Turkish and English participants differ in both speech and co-speech gesture, but not silent gesture; these patterns are the same for each of the three types of paths (TO, OVER, FROM). Note that participants produced more speech responses than gesture responses; they often produced several spoken sentences per trial (all of which were analyzed), but typically produced only one gesture string in their silent gestures; most of the participants' co-speech gestures were not combined with other gestures and thus could not be analyzed for order.

offline (i.e., in gestures produced instead of speech, silent gesture). We found that English- and Turkish-speakers displayed cross-linguistic differences in the way they packaged motion elements (conflated vs. separated) and ordered semantic elements (figure-MOTION-ground vs. figure-ground-MOTION) in their speech and co-speech gestures. However, these cross-linguistic differences did not appear in their silent gestures. In fact, both English- and Turkish-speakers displayed the same packaging of motion elements (conflated, the English pattern) and used the same ordering of semantic elements (figure-ground-MOTION, the Turkish pattern) in their silent gestures. Our results thus provide no evidence for an offline effect of language on nonverbal thinking and, in fact, suggest a possible natural semantic organization that humans impose on motion events when conveying them nonverbally in gesture.

4.1. The ordering of semantic elements

Our data on the ordering of semantic elements replicate previous work (Goldin-Meadow et al., 2008), showing that speakers of English, Turkish, Chinese, and Spanish all produce silent gestures that follow the SOV order, comparable to the Figure-Ground-MOTION order in our silent gesturers. It is the English-speakers in our data who are interesting as they abandoned the Figure-MOTION-Ground (SVO) order that they used exclusively in speech

and replaced it almost as exclusively with Figure-Ground-MOTION (SOV) in silent gesture.¹⁰ We also included order analyses of co-speech gesture and found that neither Turkish- nor English-speakers produced many strings containing gestures for figure, motion, and ground (4.2 Turkish-speakers, 2.4 English-speakers, as might be expected given that speakers typically produce only one gesture per spoken clause, McNeill, 1992). The few co-speech gesture strings the Turkish-speakers did produce mirrored their speech--more SOV than SVO. But the English-speakers' co-speech gestures did not--although they produced more SVO strings than Turkish-speakers (.85 vs. .30), this number was smaller than the number of SOV strings they produced (.85 vs. 1.55), thus mirroring their silent gesture pattern. However, the extremely small numbers of strings that both groups produced in co-speech gesture makes it impossible to draw conclusions from these patterns.

¹⁰ Gershkoff-Stowe and Goldin-Meadow (2002) found that the silent gesturers in their study (who were all English-speakers) frequently produced the Figure-Ground-MOTION ordering that we found in our study. However, their participants also produced the Ground-Figure-MOTION ordering, which we rarely observed. Unlike this earlier study, which varied both the figure and the ground in every item, our study varied only the ground and kept the figure constant across all items. This aspect of our design probably downplayed the salience of the figure, which may then have had an impact on where gestures for the figure were positioned relative to gestures for the ground; note, however, that both were placed before MOTION.

Our results thus showed convergence on a Figure-Ground-MOTION gesture pattern in both languages when speakers were not speaking. Why do speakers resort to this order, placing ground before motion, even when the word order in their native language shows the opposite pattern? A possible explanation could be that *figure* and *ground* represent entities, which are cognitively easier to express and understand than relational information (Gentner, 1982), which is represented by *motion*. Placing the anchors to which a *motion* relates (i.e., its *figure* and/or *ground*) early in a string might then ease the processing burden for the string as a whole. This strategy might be particularly effective in silent gesture, where there is no grammatical marking to guide processing. Another possibility could be that the Figure-Ground-MOTION order reflects the default way we conceptualize order of events in the world—a hypothesis that needs to be tested in future studies with young infants who have yet to acquire linguistic representations of events in their native language.

4.2. The packaging of semantic elements

The packaging of motion elements shows a similar pattern in that the cross-linguistic differences that characterize speech disappear in silent gesture, although here it is the Turkish-speakers who abandoned the *separated* pattern in their native language and took on the *conflated* pattern (see Özyürek et al., 2015, who also found *conflated* gestures in Turkish silent gesturers using animated vignettes as stimuli). But even English-speakers boosted the level of *conflated* gestures they produced in silent gesture over their levels in speech and co-speech gesture.

We suggest that the expression of motion events may be driven by a need to convey maximal information with limited effort. Turkish expresses path information in the verb, requiring an additional adjunct or subordinate clause to also express manner. This construction is syntactically complex, leading Turkish-speakers to routinely omit manner information from speech and express only path

(Özçalışkan & Slobin, 1999). Gesture allows both manner and path to be simultaneously conveyed in a relatively easily produced form, which may be why both Turkish and English silent gesturers adopt the *conflated* form almost exclusively. Indeed, Turkish-speakers have been found to convey both manner and path in speech if it is possible to convey the two within a single lexical item (Özçalışkan & Slobin, 2000). In our study, the availability of a simple gesture conveying both manner and path may have encouraged Turkish-speakers to override typological patterns and express components not typically encoded in speech or co-speech gesture.

One final point is worth making. Homesigners, deaf children whose hearing losses prevent them from acquiring speech and whose hearing parents have not exposed them to sign, create gesture systems that are, of course, produced without speech. But homesigners' gestures—which also do not constitute a conventionalized language system like the silent gestures in our study—do *not* show the conflated pattern found in the silent gestures that the hearing adults in our study produce—even though, in both cases, the gestures are produced *without* speech. The arrangements that the participants in our study display in their silent gestures cannot therefore be dictated solely by the manual modality—they must also reflect a preference on the part of the speaker (see also Goldin-Meadow, 2015; Özyürek et al., 2015).

In summary, we have shown that speakers of languages that differ in their organization of motion events do *not* rely on language-specific patterns when asked to describe these events without speaking, that is, in silent gesture. Instead, their silent gestures display the same organizational patterns. The commonalities in silent gesture that we have found across speakers of different languages suggest that silent gesture can be a window onto representations of motion events that are divorced from a particular language, a level of representation that may cut across linguistic, and perhaps even cultural, differences.

Appendix. A

Mean distribution of motion elements by semantic packaging and semantic ordering for each scene in **speech** by language

	ENGLISH		TURKISH		ENGLISH		TURKISH	
	SEP	CONF	SEP	CONF	F-M-G	F-G-M	F-M-G	F-G-M
MOTION TO LANDMARK								
<i>Run into house-</i>	0.30	1.05	1.40	0.20	1.25	0.0	0.0	1.45
<i>Crawl into house</i>	0.65	1.0	1.25	0.20	1.40	0.0	0.0	1.40
<i>Climb into house</i>	0.25	1.20	0.85	0.70	1.25	0.0	0.0	1.60
<i>Walk towards crib</i>	0.35	0.90	0.80	0.35	1.25	0.0	0.0	1.25
Mean TO	1.55 (1.79)	4.15 (1.18)	4.30 (2.13)	1.45 (1.23)	5.15 (2.94)	0.0 (0.0)	0.0 (1.18)	5.70 (0.31)
MOTION OVER LANDMARK								
<i>Crawl over carpet</i>	0.60	1.45	1.30	0.25	1.15	0.0	0.05	1.25
<i>Jump over hurdle</i>	0.85	1.35	0.90	1.0	1.35	0.05	0.0	1.50
<i>Flip over beam</i>	1.10	1.05	1.0	0.65	1.80	0.05	0.10	1.30
<i>Crawl along tracks</i>	0.25	0.95	0.75	0.30	1.10	0.0	0.0	1.05
Mean OVER	2.80 (2.61)	4.80 (1.64)	3.95 (2.16)	2.20 (1.28)	5.40 (3.44)	0.10 (0.31)	0.15 (0.37)	5.10 (1.59)
MOTION FROM LANDMARK								
<i>Roll out tunnel</i>	0.70	1.10	1.45	0.10	2.55	0.0	0.05	1.25
<i>Run out house</i>	1.10	0.55	1.90	0.05	1.30	0.0	0.05	1.20
<i>Crawl out house</i>	0.55	0.90	1.35	0.05	1.15	0.0	0.0	1.25
<i>Run away from motorcycle</i>	0.55	0.95	1.55	0.10	1.55	0.0	0.05	1.0
Mean FROM	2.90 (2.57)	3.50 (1.61)	6.25 (1.92)	0.30 (0.57)	6.55 (3.02)	0.0 (0.0)	0.15 (0.37)	4.70 (1.08)

SEP: separated, CON: conflated, F-G-M: Figure-Ground-Motion, F-M-G: Figure-Motion-Ground; standard deviations for mean path types are indicated in parentheses.

Appendix. B

Mean distribution of motion elements by semantic packaging and semantic ordering for each scene in **co-speech gesture** by language

	ENGLISH		TURKISH		ENGLISH		TURKISH	
	SEP	CONF	SEP	CONF	F-M-G	F-G-M	F-M-G	F-G-M
MOTION TO LANDMARK								
<i>Run into house</i>	0.85	0.30	1.45	0.15	0.10	0.15	0.0	0.10
<i>Crawl into house</i>	0.50	0.90	0.70	0.45	0.0	0.10	0.10	0.20
<i>Climb into house</i>	0.65	0.55	0.85	0.40	0.20	0.0	0.0	0.35
<i>Walk towards crib</i>	0.50	0.30	0.75	0.40	0.05	0.05	0.10	0.05
Mean TO	2.50 (1.70)	2.05 (1.57)	3.75 (1.62)	1.40 (1.43)	0.35 (0.44)	0.30 (0.57)	0.20 (1.18)	0.70 (0.31)
MOTION OVER LANDMARK								
<i>Crawl over carpet</i>	0.60	0.75	1.15	0.45	0.10	0.25	0.05	0.25
<i>Jump over hurdle</i>	1.05	0.95	1.10	0.30	0.05	0.25	0.00	0.30
<i>Flip over beam</i>	0.80	0.80	1.0	0.60	0.05	0.0	0.05	0.45
<i>Crawl along tracks</i>	0.60	0.65	0.75	0.20	0.05	0.20	0.00	0.45
Mean OVER	3.05 (3.22)	3.15 (2.50)	4.0 (1.89)	1.55 (1.96)	0.25 (0.95)	0.70 (0.52)	0.10 (1.34)	1.45 (0.31)
MOTION FROM LANDMARK								
<i>Roll out tunnel</i>	0.85	0.65	0.95	0.40	0.05	0.15	0.0	0.50
<i>Run out house</i>	0.70	0.30	1.10	0.35	0.0	0.0	0.0	0.40
<i>Crawl out house</i>	0.55	0.45	0.95	0.30	0.05	0.10	0.0	0.45
<i>Run away from motorcycle</i>	0.65	0.50	1.15	0.25	0.10	0.30	0.0	0.35
Mean FROM	2.75 (1.86)	1.90 (1.71)	4.15 (2.23)	1.30 (1.63)	0.20 (0.44)	0.55 (0.45)	0.0 (1.42)	1.70 (0.0)

SEP: separated, CON: conflated, F-G-M: Figure-Ground-Motion, F-M-G: Figure-Motion-Ground; standard deviations for mean path types are indicated in parentheses.

Appendix. C

Mean distribution of motion elements by semantic packaging and semantic ordering for each scene in **silent gesture** by language

	ENGLISH		TURKISH		ENGLISH		TURKISH	
	SEP	CONF	SEP	CONF	F-M-G	F-G-M	F-M-G	F-G-M
MOTION TO LANDMARK								
<i>Run into house</i>	0.15	0.80	0.0	0.80	0.25	0.5	0.0	0.6
<i>Crawl into house</i>	0.15	0.80	0.0	0.95	0.15	0.5	0.05	0.65
<i>Climb into house</i>	0.05	0.75	0.0	0.75	0.10	0.55	0.0	0.70
<i>Walk towards crib</i>	0.10	0.85	0.0	1.0	0.20	0.60	0.0	0.65
Mean TO	0.45 (0.89)	3.20 (1.11)	0.0 (0.0)	3.50 (0.83)	0.70 (1.13)	2.15 (1.57)	0.05 (0.22)	2.60 (1.47)
MOTION OVER LANDMARK								
<i>Crawl over carpet</i>	0.10	0.95	0.0	0.95	0.10	0.65	0.0	0.85
<i>Jump over hurdle</i>	0.10	0.75	0.0	0.95	0.05	0.70	0.0	0.75
<i>Flip over beam</i>	0.05	0.80	0.0	0.90	0.05	0.70	0.05	0.80
<i>Crawl along tracks</i>	0.15	0.85	0.0	0.95	0.05	0.70	0.0	0.75
Mean OVER	0.40 (0.94)	3.35 (1.14)	0.0 (0.0)	3.75 (0.55)	0.25 (0.55)	2.75 (1.33)	0.05 (0.22)	3.15 (1.23)
MOTION FROM LANDMARK								
<i>Roll out tunnel</i>	0.10	0.80	0.0	0.90	0.0	0.75	0.0	0.75
<i>Run out house</i>	0.15	0.85	0.05	0.80	0.0	0.9	0.0	0.70
<i>Crawl out house</i>	0.10	0.85	0.0	0.95	0.0	0.85	0.0	0.65
<i>Run away from motorcycle</i>	0.20	0.80	0.10	0.75	0.05	0.75	0.0	0.75
Mean FROM	0.55 (1.05)	3.30 (1.26)	0.15 (0.37)	3.40 (1.05)	0.05 (0.22)	3.25 (1.21)	0.0 (0.0)	2.85 (1.42)

SEP: separated, CON: conflated, F-G-M: Figure-Ground-Motion, F-M-G: Figure-Motion-Ground; standard deviations for mean path types are indicated in parentheses.

References

- Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's syntactic packaging of Manner and Path: A comparison of English, Japanese and Turkish. *Cognition*, *102*, 16–48.
- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, 328.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-6*. <<http://CRAN.R-project.org/package=lme4>>.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49–79.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol.2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.
- Gershkoff-Stowe, L., & Goldin-Meadow, S. (2002). Is there a natural order for expressing semantic relations? *Cognitive Psychology*, *45*, 375–412.
- 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.
- Goldin-Meadow, S. (2015). The impact of time on predicate forms in the manual modality: Signers, homesigners, and silent gesturers. *TopiCS*, *7*, 169–184. <http://dx.doi.org/10.1111/tops.12119>.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *PNAS*, *105*(27), 9163–9168.
- Gullberg, M., Hendricks, H., & Hickmann, M. (2008). Learning to talk and gesture about motion in French. *First Language*, *28*(2), 200–236.
- Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, *129*(1), 1–17.
- Hohenstein, J. M. (2005). Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development*, *6*(3), 403–425.
- Kita, S., & Özyürek, A. (2003). What does crosslinguistic 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.
- Langus, A., & Nespors, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, *60*, 291–318.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: The University of Chicago Press.
- Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. (2010). The interaction of animacy and word order in human languages: A study of strategies in a novel communication task. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *Proceedings of the Eighth evolution of language conference* (pp. 455–456). Singapore: World Scientific Publishing Co.
- Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science*, *9*(5), 363–369.
- Özçalışkan, Ş. (in press). When do gestures follow speech in bilinguals' description of motion? *Bilingualism: Language and Cognition*. <http://dx.doi.org/10.1017/S1366728915000796>.
- Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (under review). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science*.
- Özçalışkan, Ş., Stites, L. J., & Emerson, S. (in press). Crossing the road or crossing the mind: How differently do we move across physical and metaphorical spaces in speech or in gesture? In I. Ibarretxe-Antuñano (Ed.), *Motion and space across languages and applications*. NY: John Benjamins.
- Özçalışkan, Ş. (2009). Learning to talk about spatial motion in language-specific ways. In J. Guo et al. (Eds.), *Cross-linguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 263–276). NY: Psychology Press.
- Özçalışkan, Ş., & Slobin, D. I. (2000). 'Climb up' vs. 'ascend climbing': Lexicalization choices in expressing motion events with manner and path components. In S. Catherine-Howell, S. A. Fish, & K. Lucas (Eds.), *Proceedings of the 24th Boston University conference on language development* (pp. 558–570). Somerville, MA: Cascadilla Press.
- Özçalışkan, Ş., & Slobin, D. I. (1999). Learning 'how to search for the frog': Expression of manner of motion in English, Spanish and Turkish. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *Proceedings of the 23rd Boston University conference on language development* (pp. 541–552). Somerville, MA: Cascadilla Press.
- Özyürek, A., Furman, R., & Goldin-Meadow, S. (2015). On the way to language: Event segmentation in homesign and gesture. *Journal of Child Language*, *42*(1), 64–94. <http://dx.doi.org/10.1017/S0305000913000512>.
- R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <<http://www.R-project.org/>>.
- Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, *131*(3), 431–436.
- Slobin, D. (1996). From 'thought' and 'language' to 'thinking for speaking'. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, MA: Cambridge University Press.
- Talmy, L. (1985). Semantics and syntax of motion. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 181–238). New York: Academic Press.
- Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring*. Cambridge, MA: The MIT Press.