Speaking but not Gesturing Predicts Motion Event Memory Within and Across Languages

Marlijn ter Bekke (Marlijn.terBekke@mpi.nl)
Radboud University, Nijmegen, The Netherlands
Wundtlaan 1, 6525XD Nijmegen, The Netherlands

Aslı Özyürek (Asli.Ozyurek@mpi.nl)
Center for Language Studies & Donders Center for Cognition, Radboud University, Nijmegen, The Netherlands
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Wundtlaan 1, 6525XD Nijmegen, The Netherlands

Ercenur Ünal (Ercenur.Unal@ozyegin.edu.tr)
Center for Language Studies, Radboud University, Nijmegen, The Netherlands
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Özyeğin University, Istanbul, Turkey
Nişantepe Mahallesi Orman Sokak 34794 Çekmeköy, Istanbul, Turkey

Abstract

In everyday life, people see, describe and remember motion events. We tested whether the type of motion event information (path or manner) encoded in speech and gesture predicts which information is remembered and if this varies across speakers of typologically different languages. We focus on intransitive motion events (e.g., a woman running to a tree) that are described differently in speech and co-speech gesture across languages, based on how these languages typologically encode manner and path information (Kita & Özyürek, 2003; Talmy, 1985). Speakers of Dutch (n = 19) and Turkish (n = 22) watched and described motion events. With a surprise (i.e. unexpected) recognition memory task, memory for manner and path components of these events was measured. Neither Dutch nor Turkish speakers' memory for manner went above chance levels. However, we found a positive relation between path speech and path change detection: participants who described the path during encoding were more accurate at detecting changes to the path of an event during the memory task. In addition, the relation between path speech and path memory changed with native language: for Dutch speakers encoding path in speech was related to improved path memory, but for Turkish speakers no such relation existed. For both languages, co-speech gesture did not predict memory speakers. We discuss the implications of these findings for our understanding of the relations between speech, gesture, type of encoding in language and memory.

Keywords: Motion events; Memory; Cross-linguistic differences; Co-speech gesture

Introduction

People frequently perceive, remember and communicate about events. The relations between these different cognitive processes are not well-understood. In this study, we ask whether the way a visually perceived event is described relates to how it is remembered. How exactly an event is described, varies across typologically different languages. In addition, within languages there is also variation: two speakers of the same language may perceive the same event, but describe it differently. Importantly, in describing events people not only use speech but also co-speech gestures that describe main components of events. These gestures also vary both across and within languages. How does the way one speaks and gestures about events predict one's memory for various aspects of events?

Many of the events people see in their daily lives involve motion, because the world around us is constantly moving. Two crucial components of motion events are the manner of motion (e.g., running) and the path that the motion follows (e.g., to the tree). Whether people mention the manner or path during a motion event description is strongly affected by the language they speak. Verb-framed languages (e.g., Turkish, Greek, Spanish) typically encode path in the main verb and can optionally add manner of motion, for example in subordinate verbs or in adverbial phrases (see example sentence (1) from Turkish below; Talmy, 2000). By contrast, satellite-framed languages (e.g., Dutch, English, Russian) typically encode manner in the main verb and path in a variety of other structures, such as prepositional phrases (see example sentence (2) from Dutch below). A crucial difference between verb-framed and satellite-framed languages is that speakers of satellite-framed languages typically mention both path and manner information, while speakers of verb-framed languages regularly omit manner information (Slobin, 2003).

(1)

Kadın	*(koş-arak)*	*ağac-a*	*yaklaş-ıyor*
Woman	(run-Connective)	tree-Dative	approach-Present
Noun phrase	(Verb)	Noun phrase	Verb
Figure	(Manner)	Ground	Path

(2)

De vrouw	*rent*	*naar*	*de boom*
The woman	runs	to	the tree
Noun phrase	Verb	Preposition	Noun phrase
Figure	Manner	Path	Ground

If speakers of different languages describe the same motion event differently, do they also remember the event differently? Prior work found no cross-linguistic differences in how speakers of verb-framed and satellite-framed languages remember manner and path (Engemann et al., 2015; Gennari et al., 2002; Papafragou et al., 2002; Papafragou, Hulbert, & Trueswell, 2008, but see Filipović, 2011 for differences using complex motion events). However, these studies simply compared speakers of verb-framed and satellite-framed languages at the group level, without considering the variation within languages in terms of which motion event information is described. It remains unknown whether which information a speaker mentions in a particular motion event description may predict their later memory for that information, regardless of their native language. For example, if a speaker described the path of a motion event, do they remember that path better? In addition, how these specific descriptions might interact with native language to predict memory also remains unclear. For example, does describing path have a different effect on path memory for speakers of verb-framed languages compared to speakers of satellite-framed languages?

It is plausible that the information encoded in linguistic descriptions predicts memory performance for two reasons. First, it could be that the description is a window into the mental representation of the event: if a speaker describes the path, this might indicate that the speaker has mentally represented the path of the event. Therefore, the speaker may be more likely to remember the path (Papafragou et al., 2002). Second, it could be that the verbal description functions as an additional format in which the event is encoded in memory. This way, the description itself might be remembered and thus aid memory for the components encoded in the description (Papafragou et al., 2002). Indeed, it appears that what exactly is said in a motion event description is important for memory: speakers who described a path of motion later remembered this path better (Billman, Swilley, & Krych, 2000).

When investigating the link between descriptions and memory, it is important to keep in mind that language is multimodal (Vigliocco, Perniss, & Vinson, 2014). In fact, descriptions of events are often accompanied by iconic co-speech gestures. For example, while saying “The woman ran to the tree”, a speaker might wiggle one’s index and middle fingers in an inverted V-shape across space from left to right. Co-speech gestures can represent path, manner, or both in one gesture (Figure 1). Importantly, co-speech gestures accompanying motion event descriptions differ both across and within languages. In terms of cross-linguistic differences, the *form* of motion event co-speech gestures differs between speakers of verb-framed and satellite-framed languages (Kita & Özyürek, 2003). However, it is yet unknown whether there are cross-linguistic differences between speakers of verb-framed and satellite-framed languages in terms of *how often* they gesture about path and manner, and whether this relates to their memory for path and manner. In addition, co-speech gesture production also differs within languages. Within speakers of a language, one element of motion might be gestured more often than another element for different events. Therefore, both speech and co-speech gesture need to be taken into account to see how differences within and across languages in motion event descriptions relate to motion event memory.

Figure 1: Gestures can represent only path (A), only manner (B) or both manner and path (C)

Indeed, prior work shows that gestures are related to event memory. For example, producing co-speech gestures when describing motion and action events leads to better memory for these events (Cook, Yip, & Goldin-Meadow, 2010). In addition, the specific action event information conveyed in gesture predicts the information later remembered (Koranda & MacDonald, 2015). These results are in line with research on the enactment effect, which shows that reading descriptions of action events and performing these actions leads to better memory for the descriptions that does only reading (for review, see e.g., Cohen, 1989). The involvement of the motor system could lead to richer memory representations, or to stronger memory representations (Madan & Singhal, 2012). These studies point to the importance of taking co-speech gestures into account when investigating the relation between motion event descriptions and memory.

The Present Study

The main aim of the present study was to investigate whether the speech and co-speech gestures that speakers use to describe motion events predict their memory, and whether cross-linguistic differences in speech and gesture lead to cross-linguistic differences in memory. To test these questions, Dutch and Turkish speakers watched and described motion events, after which their surprise recognition memory for manner and path was tested. We had the following predictions:

(a) In general, we expected that encoding a motion event component in speech would predict better memory for that component.
(b) Similarly, we expected that encoding a motion event component in gesture would predict better memory for that component.
(c) Cross-linguistically, we expected Dutch speakers to encode manner more often in speech and gesture than Turkish speakers, due to the optional encoding of manner in Turkish. As a result, we expected Dutch speakers to have better memory for manner.

Method

Participants. Data were collected from 19 adult native speakers of Dutch (15 females, M_{age} = 23) and 22 adult native speakers of Turkish (16 females, M_{age} = 21). Dutch speakers received monetary compensation for their participation. Turkish speakers received course credit for their participation.

Materials. Target events presented in the study phase consisted of 16 silent video clips that depicted a female actor moving with respect to a landmark object along a particular path with a particular manner (e.g., a woman hopped to a cactus). Each clip was 2500ms long. Each clip was created by combining four spontaneous manners of motion (run, hop, twirl, tiptoe) with four motion paths (to, into, from, out of). Sixteen additional video clips of transitive events served as fillers (e.g., a woman biting an apple).

In the memory phase, half of the events had a change to either the manner (e.g., a woman tiptoed instead of hopped to a cactus) or the path (e.g., a woman hopped from instead of to a cactus) of motion (Figure 2). The other half of the events remained the same. Of the 15 filler events, half remained the same and half involved an object change (e.g., a woman biting a banana).

Procedure. Each participant was tested in a quiet room at their university campus in their native language by a native speaker together with a confederate who served as an addressee.

In the study phase, participants saw 16 target and 16 filler events. Each trial started with a fixation screen of 1000ms, followed by the event shown for 2500ms. Then a gray screen appeared, during which participants described "what happened in the video" to the addressee. Participants' speech and gestures were videotaped for later coding. The memory task was presented immediately after the study phase. The memory task was a surprise for the participants, because this way the prospect of the memory task could not affect the production results. During the memory task, participants saw another set of events and for each event indicated whether they had seen this exact video before by pressing a button. In both study and memory phases, each participant saw the events in different randomized order.

Figure 2: Example of a manner change (hop became tiptoe; left panel) and a path change (to became from; right panel)

Coding. Descriptions of target events were coded for the presence of path and manner information in speech and gesture using ELAN software (Lausberg & Sloetjes, 2009) by a native speaker of the relevant language. In speech, manner information was coded as present if how the motion was performed was encoded with a manner verb (e.g., rennen; running – mostly in Dutch) or a manner verb subordinated to a path verb via a connective (e.g., koşarak; run-Connective – mostly in Turkish). Path information was coded as present if the change of location with respect to something was encoded with prepositions or spatial/directional nouns (e.g., naar (to), içine (inside)) or path verbs (e.g., gir (enter), yaklaş (approach)).

In gesture, manner information was coded as present if speakers produced a gesture representing the motion in a non-linear way. Gestures could represent the manner from a third person perspective (e.g., for twirling, a manner gesture could involve the index finger turning in circles) or could be an enactment of the figure's posture during the movement (e.g., for running, a manner gesture could involve moving the arms up and down). Path information was coded as present if speakers deliberately traced the change of location with a body part chosen to represent the figure. Path gestures could trace the change of location in the lateral axis (either with a correct or incorrect direction) or in the sagittal axis (moving towards or away from the body). Points to the location of the landmark were not coded as path gestures. Gestures could either include one motion element (manner-only or path-only) or a combination of both elements.

Results

Data were analyzed with generalized binomial linear mixed effects modelling (glmer) with crossed random intercepts for Subjects and Items using *lme4* package (Bates et al., 2015) in *R* (R Core Team, 2018). This mixed effects approach allowed us to take into account the random variability that is due to having different participants and different items.

Speech and gesture production

First, we tested whether there were cross-linguistic differences in how often path and manner components of motion events were mentioned in speech (Figure 3). We excluded three trials (two Dutch) in which the addressee talked and affected the speaker's speech production. A glmer model that tested the effects of Language (Turkish, Dutch) and Component (Path, Manner) on binary values for mention in speech (0 = no, 1 = yes) at the item level revealed only a main effect of Component ($\beta = 3.41$, $SE = 1.50$, $z = 2.28$, $p = .02$). Speakers mentioned Manner ($M = 0.97$) more often than Path ($M = 0.72$). No other effects or interactions were significant. Furthermore, the proportion of mention of path and manner components in speech by Turkish and Dutch speakers were similar per specific types of path or manner (Table 1).

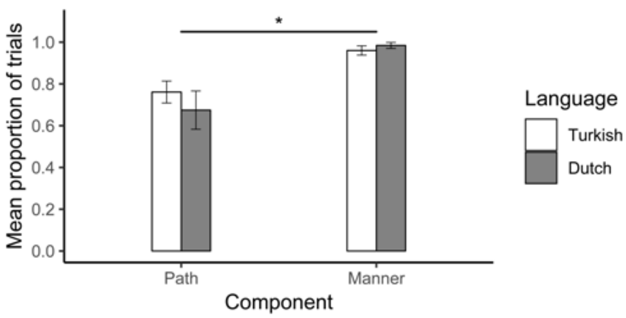

Figure 3: Event components encoded in speech. Error bars represent the standard error around the mean proportion of trials in which a component is mentioned per participant.

Table 1: Proportions of event components encoded in speech for each manner and path type, separated by language.

		Language	
		Turkish	Dutch
Manner	hop	0.99	0.98
	run	0.99	1.00
	tiptoe	0.91	1.00
	twirl	0.95	0.96
Path	to	0.67	0.56
	from	0.61	0.51
	into	0.95	0.90
	out of	0.80	0.73

Next, we tested whether there were cross-linguistic differences in how often speakers gestured about path and manner components while describing motion events (Figure 4). We excluded the same three trials that were excluded from the speech data analyses. A glmer model that tested the effects of Language (Turkish, Dutch), Component (Path, Manner) and their interaction on binary values for whether a component was encoded in gesture in an event description (0 = no, 1 = yes) revealed only a main effect of Language ($\beta = -1.59$, $SE = 0.49$, $z = -3.24$, $p < .01$). Turkish speakers ($M = 0.48$) gestured more often about both elements than Dutch speakers ($M = 0.28$). No other effects or interactions were significant. These patterns were replicated in a follow-up analysis that selected only the trials in which speakers gestured, and thus eliminated the possibility that differences in gesture rates hide cross-linguistic differences in what speakers of Dutch and Turkish prefer to gesture about.

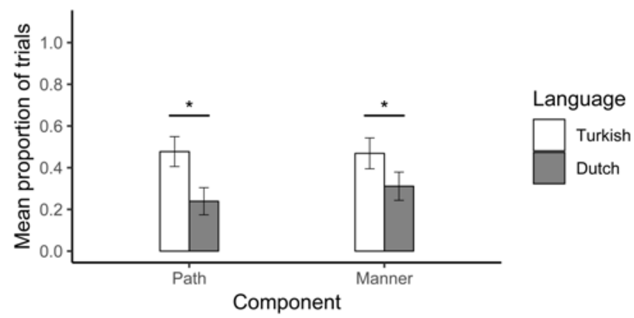

Figure 4: Event components encoded in gesture. Error bars represent the standard error around the mean proportion of trials in which a component is gestured per participant.

Memory performance

Beginning with filler events, Dutch ($M = 0.99$) and Turkish ($M = 0.95$) speakers had similar memory accuracy, indicating that the language groups were comparable in general memory performance. Furthermore, collapsed across language groups, memory for No change items ($M = 0.78$, $SD = 0.15$, $t(40) = 11.98$, $p < 0.001$) and Path changes ($M = 0.68$, $SD = 0.26$, $t(40) = 4.29$, $p < 0.001$) were significantly higher than chance level. However, memory for Manner changes ($M = 0.40$, $SD = 0.26$, $t(40) = -2.39$, $p = 0.99$) did not differ from chance level. This suggests that the participants may have simply been guessing when there was a Manner change. In addition, looking at the distribution of Manner change detection accuracy, it was clear that almost all participants had poor manner memory. It was thus not the case that some participants' memory was very poor, while other participants' memory was good. Therefore, we did not further attempt to predict manner memory using speech, gesture and language, because we did not want to predict guessing behavior.

For predicting path memory, path mentions in speech that only used unspecific verbs (e.g., to advance) that do not indicate or imply the spatial relation between the figure and the landmark were analyzed together with no mention trials and were contrasted to path mentions with prepositions, spatial/directional nouns or path verbs. Because these unspecific path verbs could be used regardless of the trajectory of motion we reasoned that they would not aid memory. Following a similar reasoning for gestures, we analyzed path gestures in the sagittal axis together with no gesture trials and contrasted them to path gestures in the lateral axis with the correct direction. Path gestures in the lateral axis with the incorrect direction were excluded from the analyses because they might even hinder memory.

A glmer model tested the effects of Path in speech (0 = no mention, 1 = mention), Path in gesture (0 = no gesture, 1 = path gesture), Language (Turkish, Dutch) and Condition (No change, Path change) on binary values for whether an item was remembered (0 = no, 1 = yes). The best-fitting model revealed a main effect of Condition as well as an interaction between Condition and Path in speech ($\beta = 1.26$, $SE = 0.51$, $z = 2.46$, $p = .01$): for No change items, speakers had similar accuracy regardless of whether Path was mentioned in speech; for Path changes accuracy was higher if Path was mentioned in speech than if it was not. There was also an interaction between Path in speech and Language ($\beta = 1.57$, $SE = 0.58$, $z = 2.71$, $p < .01$): Dutch and Turkish speakers had similar accuracy when they did not mention Path in speech, but Dutch speakers had higher accuracy than Turkish speakers when they mentioned Path in speech (Figure 5). No other main effects or interactions were significant. Notably, there were no effects or interactions involving the factor Path in gesture. Thus, contrary to our expectations, gesturing about path did not predict better memory for path of motion. We turn to the significance of these findings below.

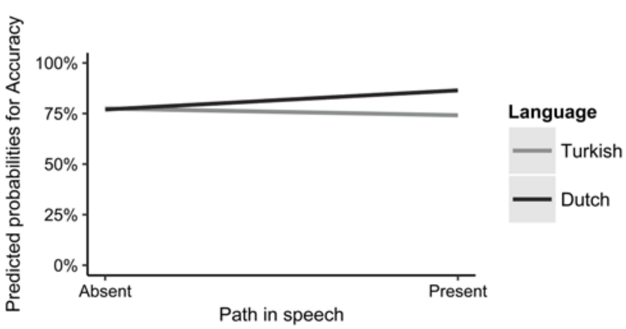

Figure 5: Interaction between Path in speech and Language for path memory accuracy, as predicted by the glmer model

Discussion

We tested whether the speech and gesture used to describe a particular motion event predicts memory for that motion event, looking at variation across and within languages. Our study has five key findings. First, speakers of Turkish did not omit the manner more often than speakers of Dutch. Second, speakers of both Dutch and Turkish had chance level memory for manner of motion. Third, speakers who mentioned path in their speech were later more accurate at detecting changes to this path. Fourth, path mention in speech was positively related to path memory for Dutch speakers, but not for Turkish speakers. Finally, we found that speaking but not gesturing predicts memory for path information.

Regarding the production results, we did not replicate the classic typological finding that speakers of verb-framed languages omit the manner more often than speakers of satellite-framed languages (Slobin, 2003). Instead, we found that speakers of both Dutch and Turkish almost always mentioned the manner of motion. A possible explanation can be found in the stimuli used in the present study. In an attempt to increase manner memory, we used manners that were rather salient (tiptoe, twirl, hop, run). It is plausible that because these manners were so salient, speakers of Turkish deemed it important to mention them. This interpretation is in accordance with the finding that speakers of Greek, a verb-framed language, mention the manner of motion much more often when it is not inferable for the listener compared to when it is inferable (Papafragou, Massey, & Gleitman, 2006). Although the cross-linguistic difference in manner omission has been reported many times, our findings show that within-language encoding flexibility makes it possible that under certain conditions (e.g., for some events), such cross-linguistic differences can be diminished.

Our study was the first to directly compare memory for manners and paths, where the path and manner changes did not involve object changes, but manner and path changes for intransitive events (unlike e.g., Bunger, Trueswell, & Papafragou, 2012 investigating instrumental motion). The finding that path is remembered better than manner is in accordance with a previously reported developmental path bias in terms of categorization (Konishi et al, 2016; Pruden et al., 2012, 2013). It is possible that path is remembered better than manner because it is related more to intentionality or goal-directedness of the motion (Pourcel, 2004). Such a relation between intentionality and memory of motion is also found when comparing memory for goal paths (e.g., to) versus source paths (e.g., from). Goals are remembered better than sources, possibly because they are more informative about the figure's intentions (Lakusta & Landau, 2012; Papafragou, 2010). Notably, this goal-source asymmetry exists only for animate figures, who can have intentions (Lakusta & Landau, 2012).

Interestingly, while speakers were not successful at remembering the manner, they did almost always describe the manner. This dissociation indicates that in terms of manner, there is no strong correspondence between speech and memory. However, this overall comparison is based on data that is averaged across different participants, items, and languages. It is still possible that when these factors are taken into account, one might find a subtle relation between mentioning manner in speech and remembering manner. In future research, this can be tested if manner memory accuracy is increased to above chance level. Nevertheless, this overall dissociation between manner mention in speech and manner memory is still quite striking. This suggests that there are at least partly different criteria for which motion event information is important to describe to another person and for which motion event information is important to remember.

For describing motion events to another person, manner of motion may be important when it is salient and not inferable. By contrast, for remembering motion events, path of motion may be important because it relates to the intentions of the figure.

In terms of the relation between descriptions and memory, we found that speakers who described a path in speech were more accurate at detecting changes to that path. This is consistent with a previous finding that speaking about path predicts better memory for path (Billman et al., 2000). It is also consistent with prior findings from other domains, demonstrating relations between how speakers describe and remember visual stimuli (e.g., eye-witness memory, Marsh, Tversky, & Hutson, 2005; picture recognition, Zormpa et al., 2018). Whether this relation between path speech and path change detection is causal is a question for further research.

In addition, the relation between path speech and path memory differed cross-linguistically: for Dutch speakers only, speaking about path predicted better memory for path. For Turkish speakers, path memory was similar regardless of whether path had been mentioned. This result might be due to cross-linguistic differences in how path was mentioned. For example, while Dutch speakers mentioned path in prepositions, Turkish speakers mentioned path mainly in verbs. Perhaps these are differentially related to memory. Another cross-linguistic encoding difference is that while Dutch speakers almost always used path prepositions that indicate the spatial relation between the figure and the landmark, Turkish speakers sometimes used unspecific verbs (e.g., to advance) to describe the path. Thus, if a Turkish speaker wants to mention path, specifically mentioning the relation to the landmark is optional. This greater optionality may have resulted in a weaker link between linguistically encoding the relation to the landmark in speech and remembering it. Further research is necessary to investigate these speculations. Either way, this interaction indicates that when linking typological differences to cognition, it is important to move from studying main effects of native language to investigating more subtle interactions of native language and descriptions.

Finally, we found no relation between co-speech gesture and memory. Importantly, path gestures typically co-occur with path speech. Therefore, this lack of a relation between gesture and memory can be interpreted to mean that path memory is equally accurate for speakers who speak and gesture about path, compared to speakers who only speak about path. The lack of a relation between path gesture and memory was surprising, given that previous research has shown a link between gesture production and event memory (Cook et al., 2010; Koranda & MacDonald, 2015). However, these studies differ from ours in one important respect: while we used motion events only, they either collapsed motion events with actions (Cook et al., 2010) or used actions only (Koranda & MacDonald, 2015). Perhaps the different memory results can be attributed to the differences between iconic co-speech gestures that describe actions versus gestures that describe paths of motion events. For example, action gestures might involve motor simulation more strongly than tracing path gestures (Hostetter & Alibali, 2008).

Either way, it appears that for the path of motion events, speech but not co-speech gesture predicts memory. There are two potential explanations of this finding. One possibility is that speech planning affects attention more than does co-speech gesture planning. Speech planning affects attention: while watching motion events to prepare for description, people look at the events in such a way that they can describe it later (Bunger et al., 2012; Flecken et al., 2015; Flecken, von Stutterheim, & Carroll, 2014; Papafragou et al., 2008; Trueswell & Papafragou, 2010). By contrast, because gestures do not follow such a strict system, their planning might have less of an impact on attention, and in turn have less of an impact on memory. Another possible explanation for why speech but not gesture predicts memory concerns the nature of speech and gesture representations. While speech is categorical and relies on discrete units, gesture is analogue and allows information to be conveyed imagistically (Cook, Yip, & Goldin-Meadow, 2012). Therefore, the verbal representation is an easier, more simplified version of the real event, compared to the gestural representation, and thus might be more useful as a memory cue.

In conclusion, the present study reveals differential contributions of speech and gesture in predicting motion event memory. Our findings underline that the relation between language and event memory is intricate and is influenced by subtle variations in how motion events are described within and across speakers of different languages.

Acknowledgments

We acknowledge support from an NWO-VICI grant awarded to A.Ö.

References

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Billman, D., Swilley, A., & Krych, M. (2000). Path and manner priming: Verb production and event recognition. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 615-620). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, *122*(2), 135-149.

Cohen, R. L. (1989). Memory for action events: The power of enactment. *Educational Psychology Review*, *1*(1), 57-80.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory and Language*, *63*(4), 465-475.

Engemann, H., Hendriks, H., Hickmann, M., Soroli, E., & Vincent, C. (2015). How language impacts memory of motion events in English and French. *Cognitive Processing*, *16*(1), 209-213.

Filipović, L. (2011). Speaking and remembering in one or two languages: Bilingual vs. monolingual lexicalization and memory for motion events. *International Journal of Bilingualism*, *15*(4), 466-485.

Flecken, M., Carroll, M., Weimar, K., & von Stutterheim, C. (2015). Driving along the road or heading for the village? Conceptual differences underlying motion event encoding in French, German, and French–German L2 users. *The Modern Language Journal*, *99*(S1), 100-122.

Flecken, M., von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: Findings from a cross-linguistic non-verbal recognition task. *Language and Cognition*, *6*(1), 45-78.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49-79.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495-514.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16-32.

Konishi, H., Pruden, S. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Categorization of dynamic realistic motion events: Infants form categories of path before manner. *Journal of Experimental Child Psychology*, *152*, 54-70.

Koranda, M., & MacDonald, M. C. (2015). Language and gesture descriptions affect memory: A nonverbal overshadowing effect. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1183-1188). Austin, TX: Cognitive Science Society.

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science*, *36*(3), 517-544.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*(3), 841-849.

Madan, C. R., & Singhal, A. (2012). Using actions to enhance memory: Effects of enactment, gestures, and exercise on human memory. *Frontiers in Psychology*, *3*, 507.

Marsh, E. J., Tversky, B., & Hutson, M. (2005). How eyewitnesses talk about events: Implications for memory. *Applied Cognitive Psychology*, *19*(5), 531-544.

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, *34*(6), 1064-1092.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*(1), 155-184.

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition*, *84*(2), 189-219.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, *98*(3), 75-87.

Pourcel, S. (2004). What makes path of motion salient?. *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society* (pp. 505-516). Ann Arbor, MI: Sheridan Books.

Pruden, S. M., Göksun, T., Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Find your manners: How do infants detect the invariant manner of motion in dynamic events?. *Child Development*, *83*(3), 977-991.

Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Infant categorization of path relations during dynamic events. *Child Development*, *84*(1), 331-345.

R Core Team. (2018). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundations for Statistical Computing. Available online at: https://www.R-project.org/

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner and S. Goldin-Meadow (Eds.), *Language in mind*. Cambridge, MA: MIT Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description*. New York, NY: Cambridge University Press.

Talmy, L. (2000). *Toward a cognitive semantics*. Cambridge, MA: MIT Press.

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, *63*(1), 64-82.

Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Phil. Trans. R. Soc. B*, *369*(1651), 20130292.

Zormpa, E., Brehm, L., Hoedemaker, R. S., & Meyer, A. S. (2018). The production effect and the generation effect improve memory in picture naming. *Memory*, *27*(3), 340-352.