

## Event endings in memory and language

Miguel Santin, Angeliek van Hout & Monique Flecken

To cite this article: Miguel Santin, Angeliek van Hout & Monique Flecken (2021) Event endings in memory and language, *Language, Cognition and Neuroscience*, 36:5, 625-648, DOI: [10.1080/23273798.2020.1868542](https://doi.org/10.1080/23273798.2020.1868542)

To link to this article: <https://doi.org/10.1080/23273798.2020.1868542>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 17 Jan 2021.



[Submit your article to this journal](#)



Article views: 1044



[View related articles](#)



[View Crossmark data](#)

## Event endings in memory and language

Miguel Santin <sup>a</sup>, Angeliek van Hout <sup>a</sup> and Monique Flecken <sup>b\*</sup>

<sup>a</sup>University of Groningen, Groningen, The Netherlands; <sup>b</sup>Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

### ABSTRACT

Memory is fundamental for comprehending and segmenting the flow of activity around us into units called “events”. Here, we investigate the effect of the movement dynamics of actions (ceased, ongoing) and the inner structure of events (with or without object-state change) on people’s event memory. Furthermore, we investigate how describing events, and the meaning and form of verb predicates used (denoting a culmination moment, or not, in single verbs or verb-satellite constructions), affects event memory. Before taking a surprise recognition task, Spanish and Mandarin speakers (who lexicalise culmination in different verb predicate forms) watched short videos of events, either in a non-verbal (probe-recognition) or a verbal experiment (event description). Results show that culminated events (i.e. ceased change-of-state events) were remembered best across experiments. Language use showed to enhance memory overall. Further, the form of the verb predicates used for denoting culmination had a moderate effect on memory.

### ARTICLE HISTORY

Received 17 July 2020  
Accepted 9 December 2020

### KEYWORDS

event cognition; memory; object-state change; event culmination; cross-linguistic analysis

### Introduction

The ability to segment the continuous flow of activity around us into discrete units, called “events”, is fundamental to human perception and memory. Segmentation relies on detecting boundaries that delimit when one of such units ends and a new one begins (Radvansky & Zacks, 2017; Richmond & Zacks, 2017). When a boundary is perceived, the working memory representation of the current event (the “working model”) is printed into episodic memory, creating a memory representation with information about the timing, space and participants in the event. The specific way in which an event ends is also part of this representation. Little is known about how differences in the way events end can lead to differences in how well they are remembered.

In addition to experiencing events through the senses, i.e. by silently watching them unfold in front of our eyes, we often experience them also through language, i.e. by describing them as they happen, or have happened, or by listening to others talking about events. When describing an event, a speaker has to gather information on the event’s participants (people and objects), their roles in the event, and the event’s spatial and temporal structure, thereby constructing an event representation (called “the message” in language production) similar to the working

models constructed in perception (Gerwien & von Stutterheim, 2018; Levelt, 1989). As in perception, event endings are an important dimension in the linguistic encoding of events (Beavers, 2008; Kennedy & Levin, 2008; Krifka, 1998; Vendler, 1957). To date, however, there is no consensus on how describing events, and the specific variation in the language used in these descriptions (both in terms of meaning and form), may or may not influence how events and their endings are represented in memory. The present study investigates (i) the effect of the movement dynamics of actions (ceased or ongoing) and the inner structure of events (with or without object-state change) on people’s event memory, and (ii) the role of language in event memory, in particular, the effect of verb semantics and form. The findings are relevant for event cognition theories, as well as for our understanding of the language-cognition interface.

### Event perception and memory

In perception, an event boundary is detected whenever an ongoing change in one or multiple participants (people or objects), or in the spatial and temporal structure of the event, has progressed to the extent that it becomes hard for the perceiver to predict what will

**CONTACT** Monique Flecken  monique.flecken@mpi.nl  Max Planck Institute for Psycholinguistics, Wundtlaan 1, Nijmegen 6525 XD, The Netherlands  
\*Secondary affiliation: University of Amsterdam

 Supplemental data for this article can be accessed <https://doi.org/10.1080/23273798.2020.1868542>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

happen next (Richmond & Zacks, 2017; Zacks et al., 2007; Zacks, 2020). This is especially true when the speed and acceleration of the hand movements of a person change (Malaia, 2014; Zacks, Kumar, et al., 2009). In a event segmentation study, Zacks, Kumar, et al. (2009) found that people strongly converged in detecting boundaries of fine-grained activities when the movement dynamics of the hands of an actor (i.e. their speed and acceleration) changed when folding the laundry or building a lego house at a table. This means that an event boundary was perceived whenever an action of the actor ceased, and transitioned to a new one, regardless of the type of actions the actor engaged in. Other studies have investigated the effect of changes in other dimensions (such as time, location, character, intention and causation) on event segmentation patterns (e.g. Huff et al., 2014; Zacks, Speer, et al., 2009). For example, change along more than one dimension of an event correlated positively with segmentation patterns and people's memory performance; the more dimensions changed, the higher the consistency in event segmentation across individuals and the better their event memory (Huff et al., 2014).

Upon the detection of an event boundary, people move their working model of the current event to episodic memory (Kurby & Zacks, 2008; Richmond & Zacks, 2017; Swallow et al., 2011). This then frees working memory space to allow the formation of a new working model on the basis of new incoming sensory information, the previous working model and other associated memory representations. Importantly, both the information available to the viewer prior to a perceived event boundary (i.e. information about space, participants and the temporal dynamics of the activity) and the specific event ending itself, constitute an important part of this representation in episodic memory (Kurby & Zacks, 2008; Swallow et al., 2011). This is supported by the finding that people's memory about what happened in a given room is hindered when they cross a doorway and enter a new room in a virtual reality setting (crossing a spatial boundary can be conceived as an event boundary; cf. Horner et al., 2016; Lawrence & Peterson, 2016; Pettijohn & Radvansky, 2016). In addition, causal inferences derived from earlier experiences with similar actions can distort people's overall memory about an event (Strickland & Keil, 2011; Papenmeier, Brockhoff, et al., 2019). For example, when people watched an event in which the typical cause-effect moment is cut out from the video clip (e.g. in a football-playing event, when a person is about to kick the ball, the scene changes to show the flight of the ball, omitting the moment of contact between the person's foot and the ball), they were

likely to falsely remember having seen the part of the event that would have led to the specific outcome (e.g. the moment of contact). These studies show that the ending of an event can influence how its content is represented in memory.

In addition to the importance of event endings in perception and memory, it has been argued that the goal and intention of agents' actions are critical for understanding and learning the underlying hierarchical structure of novel events. For example, prior exposure to the goal of a sequence of activities is critical for learning novel actions in infants, children and adults (e.g. Monroy et al., 2017; van Elk et al., 2011). Furthermore, many studies have reported a bias towards attending to, and remembering information related to spatial goals reached in motion events, in contrast to other dimensions of the same event (such as the source from which the motion originated, or the manner of motion, e.g. Bunger et al., 2013; Lakusta & Landau, 2005; Regier & Zheng, 2007; Wagner et al., 2008; for reviews Levine et al., 2019). In sum, the way in which an activity ends seems to play an important role in how an event is learned, perceived and represented in memory.

Although the special status of event boundaries and goals of action in perception and memory has received much attention in event cognition research, less attention has been paid to how people perceive and remember the inner temporal structure and progress of an event. Recent work has proposed that keeping track of how event participants (people and objects) interact with each other, and inflict change upon each other, is essential for apprehending the progress of unfolding events (Altmann & Ekves, 2019; Hindy et al., 2012; Radvansky, 2017). In particular, maintaining a representation of how an object changes (e.g. the change of state in a potato when it is being peeled) and the degree to which it changes (partially or fully), is important for conceptualising the type of action and the temporal progression of the event towards its ending. Studies on the internal structure of events have found that when people interpret sentences describing events that involve object state-change (e.g. "the chef peeled a potato"), people track change in an object not only by forming a mental representation of its new state (a peeled potato), but also by keeping the representation of its original (intact) state activated (a potato with skin before being peeled) (Hindy et al., 2012; Solomon et al., 2015). Similarly, Sakarias and Flecken (2019) showed that when people watched videos of change-of-state events (e.g. a person peeling a tangerine), people allocated more visual attention towards the affected objects as compared to objects in events that did not involve any conspicuous change of state (e.g. a person measuring a box).

Importantly, change-of-state events have a natural endpoint which depends on the physical characteristics of an object, e.g. the culmination moment of the action of peeling a potato or a tangerine is determined by the amount of peel left on it, and after which peeling can no longer continue. When the progress of an event coincides with the intention of an agent to reach this natural endpoint, it can be said that the event has culminated. Research on the mental representation of event culmination has only just started. The different patterns of attention established in Sakarias and Flecken (2019) suggest that people perceive an essential distinction between change-of-state events which develop towards a culmination point on the one hand, and events without a similar state change or culmination on the other. Furthermore, Ji and Papafragou (2020) showed that people can use the notion of event culmination (referred to as “boundedness” in their paper) as an abstract feature for categorising events.<sup>1</sup>

Previous research on the inner structure of events has focused on attention and categorisation. At present, it is largely unknown to what extent specific event features or dimensions may influence how well people remember events presented in isolation (i.e. in short video clips of single events), in particular, change in the movement dynamics of actions and the degree of change of state in objects. Furthermore, it is not known how culminating events that reach their natural endpoint in the shape of a fully changed object, are represented in memory, in comparison to events that do not involve an inherent endpoint, or in comparison to events that progress toward culmination but do not reach it. To our knowledge, Sakarias and Flecken (2019) is the only study that investigated memory for events with and without a change of state in an object. This study found that, although the distinction between the two types of events was indeed reflected in different visual attention patterns, it was not reflected in people’s recognition memory. Following up on their study, here we focus on the effect that movement dynamics may exert on the memory representation of change-of-state events and events with no state change. We examine in more detail to what extent these event features (ceased versus ongoing actions, and state change versus no change of state) and their interaction affect event memory.

### **Events in language**

Being engaged in the act of speaking while watching an event triggers specific patterns in visual attention allocation, and subsequent memory representation of the

event (e.g. Slobin, 1996; Engemann et al., 2015; Bungler et al., 2016). This attention process, called “thinking for speaking”, is guided by the meaning of the specific linguistic forms and constructions that speakers are planning to use in their description of the event. An event description creates a “situation model” in the comprehender (Zwaan et al., 1995), i.e. an event representation highly specific to the linguistic description (in terms of meaning and form). For example, the sentences “the eagle is in the sky” or “the eagle is in the nest” render a representation of an eagle with spread or contained wings, respectively (Stanfield & Zwaan, 2001; Zwaan et al. (2002)). Furthermore, linguistic labels can be useful aids for memory; it has been shown that the semantic specificity of a word is responsible for different levels of depth of processing (more specific labels trigger more in-depth attention), which can affect later memory retrieval ( Craik & Tulving, 1975). Studies that tested memory for events, contrasting linguistic description versus non-linguistic encoding demands, have presented mixed results. On the one hand, for example, Papafragou et al., 2008 and Sakarias & Flecken, 2019 found that the encoding conditions impacted memory performance; on the other hand though, Filipović, 2010 and Gennari et al., 2002 found that different encoding conditions did not have an impact on memory. Thus, it is still unclear under which conditions, and to what extent, language may aid memory retrieval, and under which circumstances there is no effect of “speaking for memory”.

To look more specifically at the latter question – what property of language does or does not support memory? – it is important to discuss the prominent role of verb predicates in establishing the core structure of events in language and its relation to cognition. The information lexicalised in a verb constitutes the central part of an event because it specifies the particular type of interaction and relationship in which event participants (people and objects) engage (Beavers et al., 2010; Pustejovsky, 1991; Rappaport Hovav & Levin, 2010; Talmy, 2000; 2016; Vendler, 1957, among many others). Furthermore, it has been argued that the particular event structure as expressed by verb predicates can affect how events are mentally represented (e.g. George et al., 2014; Göksun et al., 2010, 2017; Gerwien & von Stutterheim, 2018; Slobin, 2006). For example, Skordos et al. (2020) recently found that the semantics of motion verbs (lexicalising manner or path of motion) interacted with how well people remembered changes in event features related to either the manner or the goal of motion; showing that the meaning of a verb predicate presented to or produced by people can affect event memory.

Turning to the description of change-of-state events with a natural culmination moment, information about how a particular event ends constitutes an important dimension for how this event is described in language. Languages have designated means to convey this type of information, with verb choice playing a major role. Lexical-semantic analyses of verb phrases make an important distinction between “telic” predicates, which describe events with a natural endpoint (e.g. *peel a potato*), and “atelic” predicates, which are descriptions of homogeneous events without such natural endpoint (e.g. *carry a potato*) (Beavers, 2008; Dowty, 1979; Hay et al., 1999; Jackendoff, 1996; Kennedy & Levin, 2008; van Hout, 1996; Vendler, 1957; Verkuyl, 1993). Importantly, languages differ as to how the result of an event is typically lexicalised in verbal predicates. Talmy (2000) classified languages in two typological categories: verb-framed and satellite-framed languages, depending on whether the outcome of an event (a specific resultant relationship between the participants in an event) is commonly encoded in the verb itself (e.g. *to break something*) or in a satellite marker outside the main verb (e.g. *to tear something off*). Several studies to date have shown that languages differ as to how typically speakers rely on single verbs or verbs plus satellites to express the core structure of an event (e.g. Bunger et al., 2016; Filipović, 2010; Gennari et al., 2002; Papafragou et al., 2008; Slobin, 1996; Soroli & Hickmann, 2010). Verb-framed languages like Spanish typically use single verbs which lexicalise either the manner or the result of an event. As a consequence, Spanish event descriptions frame events either in terms of a particular change brought about by an action (the result of an event), or in terms of a particular manner of action, leaving the other dimension unspecified (see Table 1). In contrast, satellite-framed languages like English and Mandarin typically encode both manner and result in the form of verb-satellite constructions in which the main verb lexicalises the manner of action and the satellite encodes the result of the event (see Table 1):

**Table 1.** Lexicalisation of manner and result across language types.

*Differences in the lexicalisation of manner of action and the result of events in verb predicates of verb-framed and satellite-framed languages*

	Lexicalisation of manner of action	Lexicalisation of event result
<b>English</b> (satellite-framed)	Mary <u>hit</u> the wall.	Mary knocked <u>down</u> the wall.
<b>Spanish</b> (verb-framed)	María <u>golpeó</u> la pared. “María <u>hit</u> the wall.”	María <u>derrumbó</u> la pared. “María <u>brought down</u> the wall.”
<b>Mandarin</b> (satellite-framed)	Mǎlì <u>dǎ</u> le qiáng. “Mǎlì <u>hit</u> the wall.”	Mǎlì <u>dǎ-dǎo</u> le qiáng. “Mǎlì <u>hit down</u> the wall.”

It is important to keep in mind that categorising a language as verb-framed or satellite-framed does not imply that the other linguistic pattern is not available in the language, or that it is never used by speakers. Rather, it means that speakers show a relative preference for using one or the other pattern (Talmy, 2000, 2016). For example, Spanish is classified as a verb-framed language, and is considered to lack verb-satellite constructions (although a handful of exceptions can be attested<sup>2</sup>). English is considered to be a verb-framed language (Talmy, 2000), but has quite a number of single verbs that encode the result of an event: *break, solve, fill, demolish*, etc. In this regard, Spanish follows the verb-framed pattern more strictly than English follows the satellite-framed pattern. A clearer example of a satellite-framed language is Mandarin Chinese, because it lexicalises more consistently the result of events in satellite markers and manner of action in single verbs (Talmy, 2016). In contrast to English, Mandarin follows the satellite-framed pattern more consistently (Berthele, 2013; Chen, 2018; Stefanowitsch, 2013).<sup>3</sup>

These differences in verb lexicalisation patterns have been shown to guide speakers’ allocation of attention while watching scenes of events for producing event descriptions (“thinking for speaking”, e.g. Papafragou et al., 2008; Soroli & Hickmann, 2010; Flecken et al., 2015). Furthermore, these typological differences are reflected in event segmentation patterns (e.g. Gerwien & von Stutterheim, 2018; Wolff et al., 2009) and memory (e.g. Fausey et al., 2010; Filipović, 2010). However, not only verb predicates, but also other grammatical features of a language can affect speakers’ event representations. In a recent study on the representation of change-of-state events in memory, Sakarias and Flecken (2019) found that the use of Estonian case marking on the direct object (partitive or accusative case) to specify the degree of change in an affected object (partitive signals no or partial change while accusative signals full change) enhanced participants’ memory for change-of-state events, in comparison to speakers of Dutch where this distinction is not grammatically marked. Here, we investigate to what extent language use may affect event memory, in particular, the role that the semantics and form of verb predicates may have on the memory representation of event culmination.

## Present study

The central aim of the present study was to investigate the effect of two event dimensions on event memory: the movement dynamics of actions and object-state change. On the one hand, with respect to event

cognition theories, we asked three questions. (a) To what extent do changes in the movement dynamics of actions (events ending with ceased actions vs. those ending with ongoing actions) lead to stronger representations of event endings in memory? (b) Are events in which objects undergo a change of state remembered better than those without such a change? (c) To what extent is the actual culmination of change-of-state events especially salient in memory? On the other hand, with respect to event memory and language, we also asked three questions. (a) To what extent does the use of language to describe events during encoding strengthen their representation in memory? (b) How does the lexicalisation of culmination in the semantics of verb predicates enhance the representation of this aspect of events in memory? (c) How does the use of different verb predicate forms (single verb or verb-satellite construction) to lexicalise the culmination moment of events influences memory?

To address these two sets of questions, we compared the memory performance of speakers of two typologically different languages, Spanish and Mandarin, in two different experiments. In Experiment 1, we tested how accurately speakers from both languages remembered the endings of short video clips portraying simple everyday activities after they encoded them doing a non-verbal probe-recognition task. In Experiment 2, we assessed the extent to which the explicit use of language to describe such events during encoding influenced the memory representations of the endings of the same stimuli in two additional groups of Spanish and Mandarin speakers. In particular, we analyzed how the lexical-semantics and form of verb predicates influenced the memory representation of culmination in change-of-state events.

Event segmentation theory predicts that the detection of an event boundary triggers an update in episodic memory (Zacks et al., 2007). Boundary detection happens, amongst other things, upon the perception of change in the dynamics of an agent's movement (Zacks, Kumar, et al., 2009). Here, we presented participants with videos of single, fine-grained activities (e.g. cutting an apple in half, spreading butter on a cracker), and studied the extent to which changes in movement dynamics (at the cessation of an actor's action at the ending of a video) may lead to a stronger representation of the progress of an event in memory, as compared to when movement dynamics remain uniform (during the ongoing stage of an action at video offset). Cessation of an action involves a notable change in the speed and acceleration of the movement features of agents and objects; in contrast, actions that are perceived as ongoing, progress with continuous speed and

acceleration. Thus, we expected that stimuli showing a ceased action at video offset would be remembered better than those still showing an ongoing action at its offset.

In addition, given the salience of object-state change in event representation (Altmann & Ekves, 2019), we hypothesised that events that involve a clear and substantial object-state change would be represented more strongly in memory, than events in which objects are not clearly and substantially affected. Therefore, we expected that stimuli in which objects change of state would be remembered better than those in which objects did not change visibly.

Further, in analogy to the previously reported cognitive salience of event goals (in particular, motion event endpoints), we hypothesised that the ending of culminated change-of-state events (in which actions progress towards a natural endpoint delimited by the physical characteristics of the affected objects) would be especially salient in memory, that means, even more salient than the ending of ceased events that lack a clear object-state change (in which actions progressed without a conspicuous goal or endpoint). In other words, we expected that people would remember particularly well the ending of stimuli showing events involving, both, object-state change and ceased actions.

In experiment 2, we investigated how the lexical-semantics and form of verb predicates used to describe events might influence the memory representation of culmination in change-of-state events. Given Talmy's (2000, 2016) classification of Spanish as a verb-framed language (i.e. a language in which event results are typically lexicalised in single verbs) and Mandarin as a satellite-framed language (i.e. a language in which results are typically lexicalised in a satellite marker), we recruited native speakers of these languages to compare the effect on memory of using different verb predicate forms to denote culmination (single verbs or verb-satellite constructions).

In line with depth-of-processing theory ( Craik & Tulving, 1975) and thinking-for-speaking theory (Slobin, 1996), we developed three hypotheses about the role of language in event memory. First, we hypothesised that the use of language to describe events would strengthen memory representation because describing scenes requires that people pay more attention to what they watch. We thus expected overall better memory performance of both language groups in the verbal encoding condition (Experiment 2) as compared to the non-verbal encoding condition (Experiment 1). Second, we hypothesised that the lexical-semantic contents of verb predicates that denote specific features of events would notably strengthen such features in

memory. In particular, we were interested to see if the use of verb predicates that express a culmination moment would give extra support to event memory as it emphasises this feature in event representation. Thus, we expected that the culmination of change-of-state events would be remembered better after people described such events using verb predicates that lexicalise a culmination moment than when the verb predicates lexicalised other aspects of events. Third, single verbs lexicalise less complex event structures (either manner of action *or* the result of events, not both) than verb-satellite constructions (which lexicalise both, manner in the main verb and result in a satellite marker). We thus hypothesised that the use of single verbs expressing a culmination moment might lead to an especially strong memory representation of this dimension in comparison to the use of verb-satellite constructions, which convey the same dimension plus the additional manner of action. Thus, for Experiment 2 we expected that describing videos of ceased change-of-state events with single verbs expressing the culmination moment of events (by Spanish speakers) would result in better memory of event endings, than describing them with verb-satellite constructions (by Mandarin speakers).

## Experiment 1: recognition memory task after non-verbal encoding of events

### Method

#### Participants

A group of Spanish ( $N = 20$ ; mean age = 25,9; 16 females, 4 males) and a group of Mandarin ( $N = 21$ ; mean age = 26,0; 9 females, 12 males) native speakers were recruited in Groningen and Nijmegen in the Netherlands. All participants were university students and reported to speak English above B1 level.<sup>4</sup>

#### Materials & design

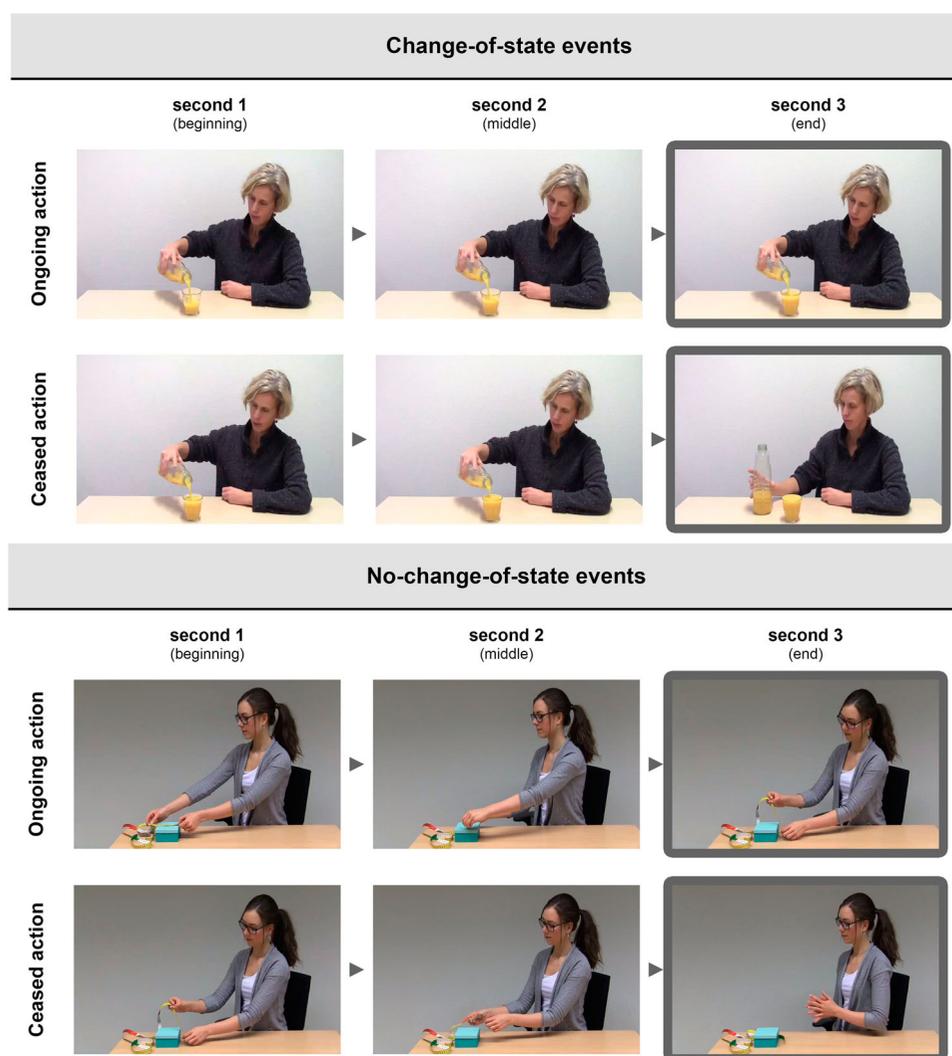
The experimental stimulus consisted of 48 three-second long videos<sup>5</sup> comprising two different versions of 24 different events in which an actor sitting at a table performed an action on an object. Half of these 24 events were change-of-state events ( $N = 12$ ): events in which the action of the agent led to a substantial change in the attributes of an affected object with the intention to reach a full result (e.g. to pour orange juice into a cup to fill it up; Figure 1A). The other half of the events showed no-change-of-state events ( $N = 12$ ): events in which the actions did not produce a visibly salient or substantial change in the physical attributes of the objects (e.g. to measure a small box with a

measuring tape; Figure 1B). All 24 events were filmed with two different types of movement dynamics at video offset: either an *ongoing action* (i.e. the action of the actor was still in progress at video offset), or a *ceased action* (i.e. the actor stopped performing the action at video offset). The total set of 48 videos thus depicted two different types of events (change-of-state and no-change-of-state, Appendix A), each with two different movement dynamics at video offset (ongoing and ceased actions).

The video stimuli were counterbalanced across two lists of 24 videos so that each event appeared only once on a list showing either an ongoing or a ceased action at video offset. In each list, six change-of-state events and six no-change-of-state events showed an ongoing action at video offset. The other six videos of each type showed a ceased action at video offset. Furthermore, 18 additional videos were used as fillers ( $N = 12$ ) and practice items ( $N = 6$ ); these were the same on each list. These videos showed events involving only one participant (e.g. a person sleeping with their head on the table), or events in which two actors interacted with each other (e.g. to put a hat on someone's head). Thus, each list contained 36 items in total: 12 change-of-state events (6 ceased and 6 ongoing) and 12 no-change-of-state events (6 ceased and 6 ongoing), plus 12 filler events.

As the non-verbal encoding task involved a probe-recognition task, a set of 11 still images was created using screenshots of some of the filler and practice items to avoid extra exposure to the critical stimuli. These images were used as probes in half of the encoding trials ( $N = 18$ ; out of which 8 matched with filler events, 5 mismatched with change-of-state events and 5 mismatched with no-change-of-state events). The purpose of this task was to keep the attention of participants engaged on the events they were watching. To avoid inducing special attention to the offset of the videos (i.e. the critical feature to be tested in the subsequent memory task), these screenshots were taken from the mid-part of the videos.

For the crucial memory task after the encoding stage, screenshots of the final frames of the critical stimuli were used as recognition probes in both experiments. Images showed the events either as ceased or ongoing. Items for the memory task were counterbalanced on different lists of 24 images each so that, for each event on a particular encoding list, the recognition probes either matched or mismatched with the actual video ending. In each list, half of the probes served as matching cues (i.e. pictures showing ceased events matched with videos that had an actual ceased ending, and pictures showing ongoing events matched with videos



**Figure 1.** A. Second-by-second example of a change-of-state event (filling a glass with juice) in a version with an ongoing and a ceased action at the offset of the video. B. Second-by-second example of a no-change-of-state event (measuring a box) in a version with an ongoing and a ceased action at the offset of the video.

that had an actual ongoing ending), and the other half of the probes served as mismatching cues (i.e. pictures of ceased events mismatched with videos that had an actual ongoing event ending, and pictures of ongoing events mismatched with videos that had an actual ceased ending). Thus, the combination of all counterbalanced lists of videos and images covered all possible combinations of encoding conditions and recognition cues for the memory task (a detailed table of the stimuli lists can be consulted on the online supplementary materials).

### Procedure

Before the start of the experiment, all participants filled out a brief linguistic background survey (with questions about their nationality, sex, age, mother tongue, and self-rated English proficiency), and gave explicit

consent to collect and use their data for the purposes of the present study. The encoding and recognition tasks were programmed in E-prime 2.0 and run on a Windows XP laptop. The screen of the notebook was placed at approximately 50 cm viewing distance. During the development of the experiment, the clarity of the instructions of the test was revised by two native speakers of both languages before we proceeded to the data collection phase. The instructions were presented on screen in participants' native language. Participants could ask questions before the experiment started.

The procedure involved three stages. First, an encoding stage in which video clips of events were shown whilst participants engaged in a probe-recognition task. Then a distractor stage during which their verbal working memory was tested with an oral digit span

task. And finally, a recognition stage in which we tested participants' memory about the ending of the videos they had watched in the encoding stage. Participants were not told beforehand that their memory would be tested.

### **First stage: non-verbal encoding task**

In the encoding stage, participants were presented with videos from one of the four stimuli lists. Before the onset of every video a fixation cross appeared for 1000 ms. All videos were presented in fully randomised order. After some of the videos a probe-recognition question would follow ( $N = 18$ ). In these trials, at the offset of the video, there was a 1000 ms delay after which a screenshot taken from the middle part of the filler videos was presented on screen. Whenever a probe-recognition question would appear, the task was to judge as fast as possible whether or not the screenshot on screen was taken from the last video they had just watched by pressing either a "yes" or "no" designated key on the keyboard. After participants gave a response, and also right after a video finished playing (in trials in which no probe-recognition question followed), an empty screen was displayed for 3000 ms to separate the trials. After this delay a new fixation cross appeared on screen and a new trial started automatically. Participants were not given feedback on their performance. Before starting the experiment, participants practiced the probe-recognition task in a practice block of six trials during which they received feedback on their performance in the encoding task and were allowed to ask questions. After the experiment started, it was not interrupted until all trials were completed.

### **Second stage: distractor task**

After the encoding stage, an oral digit span task was conducted to create a time lag between the encoding and recognition tasks, to test participants' verbal working memory, and to clean their working memory from language traces that might be useful for remembering events. To this end, a version of the digit span task as revised by Woods et al. (2011) was used. Participants listened to lists of numbers and had to repeat the numbers in the exact same, forward order. Different from conventional digit span tasks, the digit span scoring system developed by Woods and colleagues uses a fixed number of 14 trials and goes on until all trials are performed. The length of the items was increased by one digit each time a correct answer was produced and decreased by one digit each time two consecutive wrong answers occurred in items with the same digit length. This version of the digit span task was especially useful to control memory workload

across participants: participants performed the same number of trials and with a subject-tailored level of difficulty.

### **Third stage: recognition memory task**

In the final stage of the experiment, participants were surprised with a recognition task in which their memory about the endings of the videos was tested. Participants judged whether or not a screenshot showing either a ceased or an ongoing action correctly portrayed the actual ending frames of the corresponding videos they had watched during the encoding stage (similar to the recognition task in Gennari et al., 2002).<sup>6</sup>

At the start of the third stage, the memory task was explained in a brief training for which filler items illustrated what it meant that a screenshot matched or mismatched the final scene of a video and participants received feedback on their performance. In the recognition task, a fixation cross appeared for 1000 ms before an image appeared. When it appeared on screen, participants had to judge as quickly as possible whether or not it depicted the actual ending of the corresponding video they had watched during the encoding stage. The "m" key of the laptop keyboard was marked with a green sticker and stood for "yes, the image matches", and the "z" key was marked with a red sticker and stood for "no, the image does not match". When one of these buttons was pressed, the image disappeared from the screen and an empty screen was displayed for 500ms in between trials.

## **Results: Experiment 1**

### **Non-verbal encoding task**

As the goal of the non-verbal encoding task was to maintain participants' attention to the events displayed in the videos, their responses were only analyzed descriptively. Correct responses were coded as 1 and incorrect responses as 0. Mean accuracy ratios of participants showed that overall accuracy was similarly high across groups (in Spanish speakers,  $M = 0.97$ ,  $SD = 0.03$ , in Mandarin speakers,  $M = 0.96$ ,  $SD = 0.04$ ). Only two participants (a Spanish and a Mandarin speaker) made more than one mistake in the probe-recognition encoding task. These results indicate that the encoding task succeeded in engaging Spanish and Mandarin speakers' attention to the events portrayed in the videos.

### **Digit span task**

An independent samples  $t$ -test comparing the mean span scores of Mandarin and Spanish speakers indicated there was no difference between the verbal working

memory of Mandarin ( $M = 6.12$ ,  $SD = 1.1$ ) and Spanish speakers ( $M = 6.4$ ,  $SD = 1.0$ );  $t(39) = 0.917$ ,  $p = 0.364$ . Both groups showed a similar performance on the digit span task.

### Recognition task: non-verbal encoding

The results of the recognition task were analyzed with a mixed effects logistic regression model in R (version 3.4.3) using the GLMER function in the LME4 package (Bates et al., 2015; R Core Team, 2019). Model selection was performed based on a forward stepwise selection of the random intercepts (by subjects and items) and slopes (for all fixed factors included in the model) that improved the akaike information criterion (AIC) score of the models. When the AIC score was similarly good for two or more models, the less complex model was chosen for parsimony and to avoid overfitting (Bates et al., 2018; Matuschek et al., 2017). To increase the convergence likelihood of more complex models, when they failed to converge, the optimiser “bobyqa” was used instead of the default optimiser in the LME4 package (Barr et al., 2013). All models had recognition accuracy as the binary dependent variable (1 for correct responses and 0 for incorrect responses) and included the same fixed effects and interactions following the hypotheses of the experiment. Contrasts regarding all fixed factors were sum-coded. The fixed factors included: (a) Movement Dynamics (2 levels: ceased and ongoing actions), (b) Event type (2 levels: change of state and no change of state), and (c) Language (2 levels: Spanish and Mandarin). In addition, all models included a three-way interaction between these three factors (Movement Dynamics, Event type and Language). The random-effects structure of the final model consisted of a random intercept for Participants and a random slope for Movement Dynamics type by Participant.<sup>7</sup>

The output of the final model (Table 2) and its effects plot (Figure 2) revealed an interaction between Movement Dynamics and Event Type ( $\beta = 0.26$ ,  $SE = 0.06$ ,  $p < .001$ ). On average, participants from both languages recognised ceased change-of-state events more

accurately than the other event conditions. Furthermore, the model showed a three-way interaction between Movement Dynamics, Event Type and Language ( $\beta = 0.16$ ,  $SE = 0.06$ ,  $p < .02$ ). Participants who speak Spanish had an even higher accuracy at recognising ceased change-of-state events than Mandarin participants (Spanish: 72.5% and Mandarin: 60.3%). Accuracy performance split by cue type (matching / mismatching) is provided in Appendix B.

### Discussion: Experiment 1

In Experiment 1 we investigated the role of the movement dynamics of actions and object-state change in event memory. We found that ceased actions did not lead to enhanced memory accuracy. This is different from an event segmentation study (Zacks, Kumar, et al., 2009), in which it was found that changes in the movement dynamics of actions performed by an agent (as when the speed and acceleration of a hand movement decreases prior to starting movement in a different direction) correlates with the detection of event boundaries. Our results suggest that mere cessation of an action does not necessarily lead to an update of events in episodic memory. Further, events involving object-state change were not remembered better than events without a change of state in an object. Altmann and Ekves (2019) have proposed that tracking how objects change during the progress of events is crucial in event cognition. Our results suggest that the cognitive salience of object-state change during event perception does not in isolation lead to a stronger representation of events in episodic memory. Nevertheless, the combination of cessation and object-state change, resulting in change-of-state events which progress to a culmination moment, showed to be especially salient in event memory. We interpret this in line with previously reported biases in attention and memory towards goals and endpoints (Lakusta & Landau, 2012; Levine et al., 2019; Ünal et al., 2019). We will discuss the implications of these findings for event cognition theories and their relationship with language in the General Discussion.

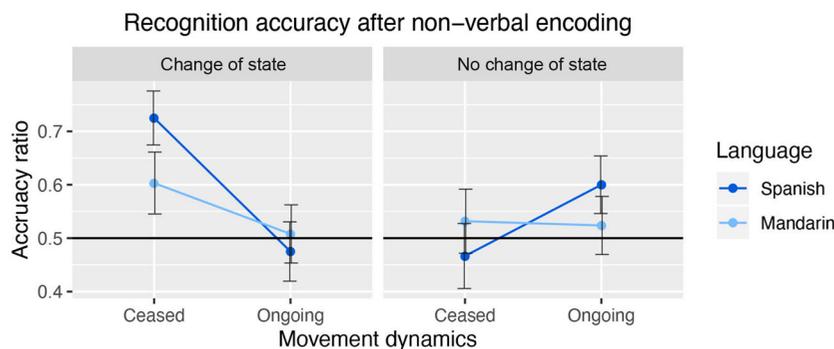
Surprisingly, the results indicate that event culmination was more strongly represented in the memory of Spanish speakers, compared to Mandarin speakers. Furthermore, the results of the digit span task indicated that this difference does not derive from differences in working memory ability between the groups. We discuss different possible interpretations for this effect in the General Discussion.

In Experiment 2 we examined the role of language on event memory by investigating to what extent the

**Table 2.** Recognition accuracy in Experiment 1.

Fixed effects model on recognition accuracy in Experiment 1. Formula in R:  $glmer(Accuracy \sim MovDyn * EvType * Language + (1 + MovDyn | Participant.ID), data = data, family = binomial)$

	Estimate	z value	Pr(> z )
(Intercept)	0.24	3.254	0.001
MovDyn	0.12	1.152	0.249
EvType	0.11	1.694	0.090
Language	0.05	0.792	0.428
MovDyn: EvType	0.26	3.911	<0.001**
MovDyn: Language	0.01	0.121	0.903
EvType: Language	0.05	0.739	0.459
MovDyn: EvType: Language	0.16	2.460	0.013**



**Figure 2.** Recognition accuracy for the four conditions across language groups when participants encoded events during a non-verbal probe-recognition task.

explicit use of language, and the semantics and form of the verb predicates used to describe events, influences event memory.

## Experiment 2: recognition memory after verbal encoding of events

### Method

#### Participants

Another group of Spanish ( $N = 22$ , mean age = 26;6, 12 females, 10 males) and a group of Mandarin ( $N = 21$ , mean age = 25;8, 6 females, 15 males) native speakers were recruited in Groningen and Nijmegen in the Netherlands. All participants were university students and reported to speak English above B1 level.

#### Materials & design

The same stimuli and general experimental set-up as in Experiment 1 were used with one change: this time participants performed a verbal description task during the encoding stage. Here, the set of images used in Experiment 1 for the non-verbal encoding of events was not used in the encoding stage of Experiment 2.

#### Procedure

As in Experiment 1, all participants filled out a linguistic background survey and gave their consent to collect and use their data. The experiment was programmed using E-prime 2.0 and run on the same Windows XP laptop. During the development of the experiment, the clarity of the instructions was once more revised by two native speakers of each language before we proceeded to the data collection phase. Instructions were presented on the screen in the participants' native language. Participants were only allowed to ask questions before the actual experiment commenced.

For the verbal encoding stage, participants engaged in an event description task in which they watched the

videos from one of four stimulus lists. Once again, videos were presented in fully randomised order. Before the onset of each video a fixation cross appeared for 1000 ms. At the end of every video, the video disappeared and an icon of a microphone appeared for 3000 ms to indicate that a description needed to be produced in their mother tongue. Participants were instructed to provide descriptions by answering "What happened?" once the icon of a microphone appeared on screen (in Spanish: "¿Qué sucedió?"; in Mandarin: "Fāshēngle shénme?"). To motivate participants to be as informative as possible, they were told that their descriptions would later be used to test whether other native speakers could correctly identify the events they had described. Descriptions were recorded using the microphone of the laptop. An empty screen was displayed for 500ms in between trials after which a new trial would start automatically.

Before starting the encoding stage, participants practiced the event description task in a practice block of six trials during which they were reminded to only provide descriptions after the offset of the video, to produce full sentences and to keep their descriptions short and precise. During the practice trials participants were allowed to ask questions to the experimenter. After the experiment started it was not interrupted until all trials were completed.

## Results: Experiment 2

### Digit span task

Both groups had a similar performance in the digit span task. An independent samples  $t$ -test conducted to compare the mean digit span scores of Mandarin and Spanish speakers confirmed that there was no difference between the verbal working memory of Mandarin ( $M = 5.82$ ,  $SD = 0.59$ ) and Spanish speakers ( $M = 6.04$ ,  $SD = 0.93$ );  $t(41) = 0.889$ ,  $p = 0.378$ .

### Event description task

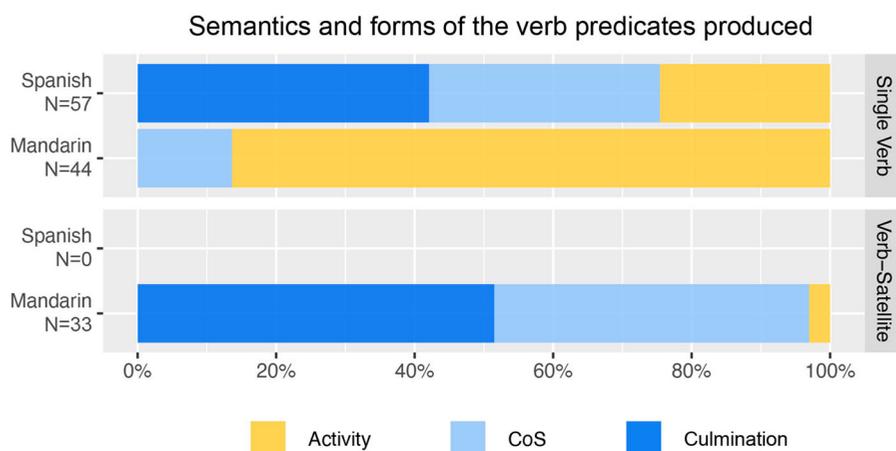
Only the descriptions of the experimental stimuli (change-of-state and no-change-of-state events) were transcribed and then coded (filler items were not analyzed). Descriptions without verbs were excluded given that the goal of the experiment was to examine the influence of the semantics of verb predicates on memory. In Spanish 4 out of 528 utterances lacked a verb predicate, and in Mandarin 11 out of 504 lacked verb predicates.

The coding of the verb predicates used in the event descriptions was done on two dimensions: form and meaning. As for the coding of *forms*, a native speaker linguist of each language identified and classified the verb predicates as either single verbs or verb-satellite constructions, following Talmy (2000). This classification was informed by theories proposed in Li (2013) and Tai (1984) on Mandarin, and García del Real (2015) and Slobin (1996) on Spanish. When verb predicates in the Mandarin descriptions were resultative verb compounds, they were coded as a verb-satellite construction (verb + satellite marker); for example, “dào-mǎn” (“pour-full”) in Mandarin. When verb predicates did not contain a satellite marker, they were coded as single verb; for example “llenar” (“fill”) in Spanish and “dào” (“pour”) in Mandarin.

As for analysing the semantics expressed by the different verb predicates (either single verbs or verb-satellite constructions), following a widely used verb classification model originally proposed by Vendler (1957) and further developed by Rappaport Hovav (2016), three naive native speaker informants per language independently classified all verb types in one of three semantic categories. *Activity verbs* are verbs which denote actions that occur without necessarily producing a change of state in an object (e.g. in Mandarin, the single verb “jiǎn” (roughly translated as “cut”) denotes the action of opening and closing a pair of scissors even if there is no object to cut or if the blades are dull). *Change-of-state (CoS) verbs* are verbs which denote a change of state regardless of the degree to which an object is affected, that is, regardless of whether the action reached the culmination point or not (e.g. in Spanish, the single verb “cortar” (“cut”) expresses that a cut is made in an object without asserting if the object was divided in two separate pieces). Finally, *Culmination verbs* denote a change-of-state event which must reach culmination (e.g. in Mandarin, the verb-satellite construction “jiǎn-kāi” (roughly translated as “cut-open”) conveys that an object was divided in half by using scissors).<sup>8</sup>

The semantic classification was based on the responses of the coders to the following three

questions for each verb predicate: (A) do objects that can possibly be affected by the action denoted by the verb predicate necessarily undergo a change of state?, (B) can the action denoted by the verb predicate occur without producing a change of state in an object?, and (C) does the action denoted by the verb predicate require that the object undergoes a change of state to a culmination point (to the point where the action denoted by the verb can no longer continue)? Question A served as a first filter to classify part of the verbs as Activity verbs; a “no” to this question indicated that the semantics of the verb predicate does not involve a CoS in the affected object. Question B served to confirm the answer to question A; a “yes” to this question indicated that the semantics of the verb predicate does not involve a CoS and hence the verb is an Activity verb. Question C served to classify the remaining non-Activity verbs either as CoS or Culmination verbs; a “yes” to this question indicated that the semantics of the verb includes the culmination moment of a change-of-state event (i.e. a bounded change of state), a “no” indicated that the semantics of the verb expresses a change of state without asserting a culmination moment (i.e. an unbounded change of state). Differences in the initial verb classification made by informants were resolved in a subsequent meeting in which they shared with each other their motivation for classifying a verb either as Activity, CoS or Culmination verb. Cases in which coders did not reach an unanimous agreement were resolved on the basis of the majority (2 out of 3 coders needed to agree upon a classification; 7 verbs in Spanish and 3 verbs in Mandarin). For a list of the final semantic classification of the verb predicates in Spanish and Mandarin consult the online supplementary materials. The coding of the verb predicates, according to their form and meaning, revealed that overall Spanish speakers exclusively produced single verbs, whereas Mandarin speakers produced single verbs and verb-satellite constructions (Figure 3). As expected, these patterns are in line with the typology literature (Talmy, 2000, 2016). While Spanish speakers only used single verbs to express the culmination moment of events (24 different types), Mandarin speakers used only verb-satellite-constructions to lexicalise this aspect of events (17 different types). Again, these results are in line with the Talmy’s typology and they show that culmination is expressed in different verb predicate forms in the two languages. This thus supports our selection of Spanish and Mandarin as proxies for the two forms in which the culmination moment of events can be lexicalised in verb predicates,



**Figure 3.** Proportion of the semantics of the different single verb and verb-satellite constructions types produced by Spanish and Mandarin speakers.

either by means of single verbs or verb-satellite constructions, and the effect of use of these forms on memory.

Importantly, in both languages, verb predicates that express culmination were produced frequently for describing videos which showed ceased change-of-state events (in Spanish 40.1% of trials,  $N = 55$ ; and in Mandarin 50% of trials,  $N = 20$ ) in comparison to the other event conditions (Figure 4A and B). Moreover, it is worth mentioning that overall Spanish speakers produced more culmination verb predicates (counting tokens,  $N = 137$ ) than Mandarin speakers ( $N = 40$ ). This suggests that, even though Mandarin speakers had the linguistic means to denote event culmination in verb predicates (in the form of verb-satellite constructions), Spanish speakers used more verb predicates that lexicalised culmination (in the form of single verbs).

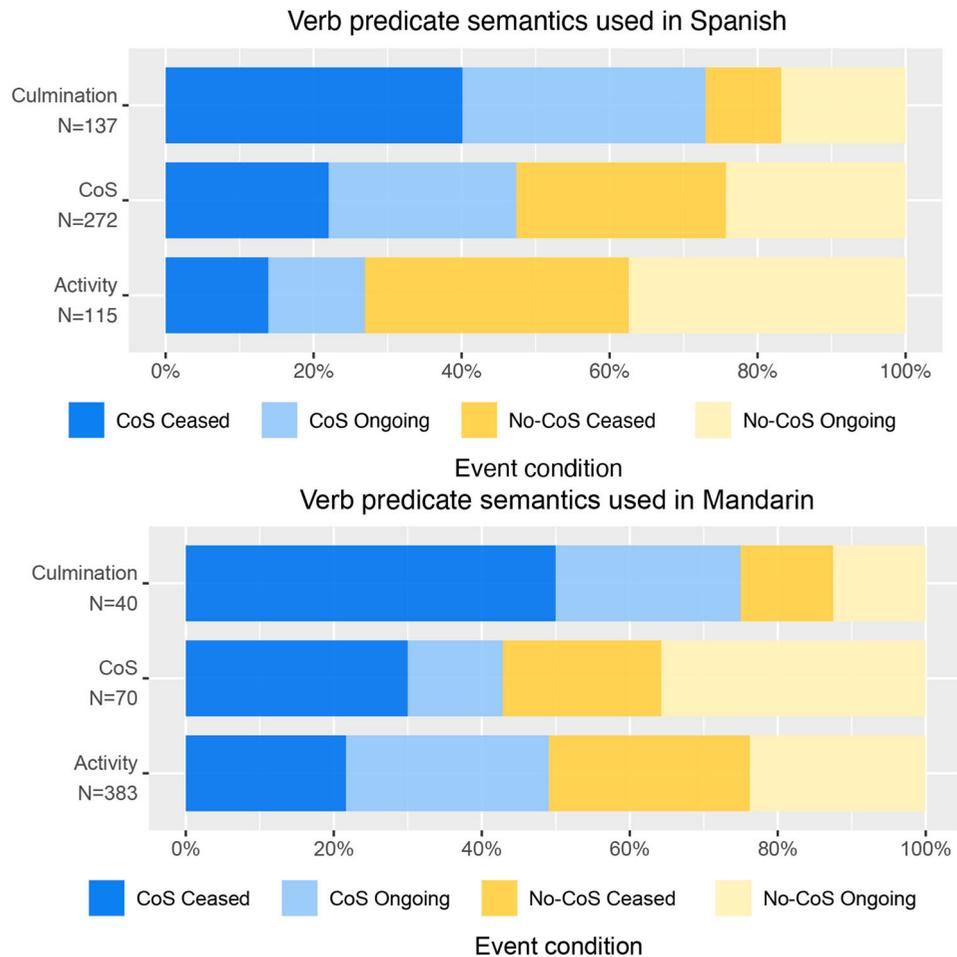
### **Recognition task: non-verbal vs. verbal encoding**

First, to assess the general effect of language use on the memory representation of the ending of events, a mixed effects logistic regression analysis was conducted on the combined data collected in Experiment 1 and 2. This allowed us to compare the overall recognition performance of participants after the non-verbal and verbal encoding of events. Then, with the objective of assessing the influence of the semantics and forms of the verb predicates produced on the representation of event culmination in memory, data collected in Experiment 2 was analysed separately with another mixed effects logistic regression model. Both analyses were done in R using the GLMER function in the LME4 package. As in the previous regression analysis, model selection was performed in a forward stepwise selection of the random intercepts (by subject and items) and random slopes, for all fixed factors included in the model, that improved

the AIC score of the models. For parsimony and to avoid overfitting, when the AIC score of two models was similarly good, the less complex model was chosen over the more complex one, however, to increase the convergence likelihood of more complex models, the optimiser “bobyqa” was used when models failed to converge. In both regression analyses, recognition accuracy was again the binary dependent variable (1 for correct responses and 0 for incorrect responses).

For the regression analysis combining the data of Experiment 1 and 2, all models included the same fixed factors and interactions. All fixed factors were sum-coded and included: (a) Encoding Type (2 levels: verbal and non-verbal encoding), (b) Movement Dynamics (2 levels: ceased and ongoing actions), (c) Event type (2 levels: change of state and no change of state in objects) and (d) Language (2 levels: Spanish and Mandarin). In addition, a four-way interaction between Encoding Type, Movement Dynamics, Event Type, and Language was included. The random effects structure of the final model included random intercepts for Participants and a random slope for Movement Dynamics by Participant.<sup>9</sup>

The analysis of the combined data (Table 3) and the effects plot of the final model (Figure 5) revealed a main effect of the Encoding Type on accuracy ( $\beta = 0.21$ ,  $SE = 0.05$ ,  $p < .001$ ). Participants in the verbal encoding experiment performed better in the recognition task than participants in the non-verbal encoding experiment. Different from Experiment 1, the model revealed a main effect of Movement Dynamics ( $\beta = 0.22$ ,  $SE = 0.08$ ,  $p < .01$ ) and a main effect of Event Type ( $\beta = 0.09$ ,  $SE = 0.04$ ,  $p < .05$ ), however, an interaction between Movement Dynamics and Event Type ( $\beta = 0.29$ ,  $SE = 0.04$ ,  $p < .001$ ) was once more revealed. No interaction with language was found in this joint



**Figure 4.** A. Verb semantics of the Spanish descriptions of events in the four different conditions. B. Verb semantics of the Mandarin descriptions of events in the four different conditions.

model. Altogether, the results indicate that the explicit use of language enhanced the overall representation of events in memory of participants of both languages.

**Table 3.** Recognition accuracy in Experiment 1 and 2.

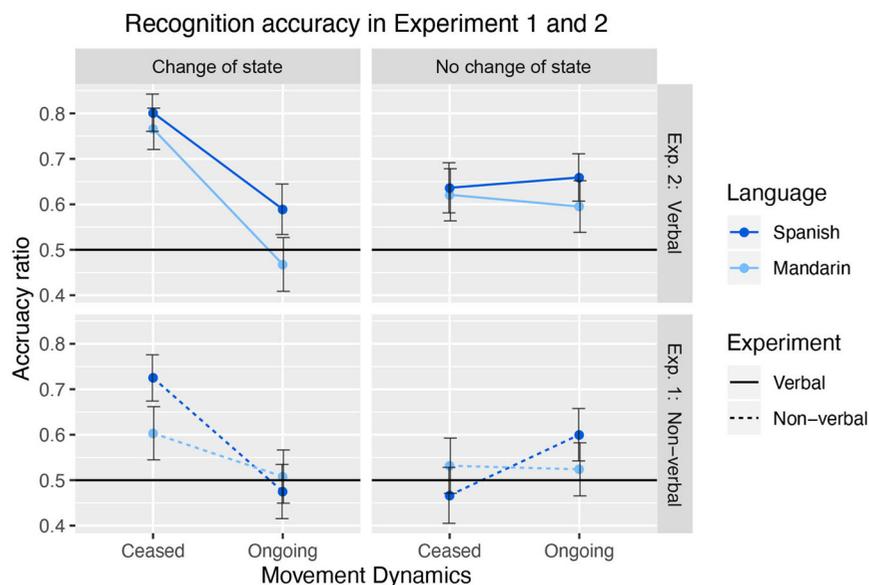
Model output of the analysis on combined recognition accuracy data from Experiment 1 and 2. Formula in R:  $glmer(Accuracy \sim Experiment * MovDyn * EvType * Language + (1 + MovDyn | Participant.ID), data = data, family = binomial, control = glmerControl(optimizer = "bobyqa"))$ .

	Estimate	z value	Pr(> z )
(Intercept)	0.46	8.235	<0.001**
Experiment	0.21	3.843	<0.001**
MovDyn	0.22	2.621	0.008**
EvType	0.09	1.995	0.045**
Language	0.08	1.523	0.127*
Experiment: MovDyn	0.09	1.122	0.262
Experiment: EvType	-0.01	-0.325	0.745
EvType: MovDyn	0.29	5.928	<0.001**
Experiment: Language	0.02	0.447	0.654
MovDyn: Language	-0.02	-0.277	0.781
EvType: Language	0.04	0.982	0.326
Experiment: MovDyn: EvType	0.02	0.470	0.638
Experiment: MovDyn: Language	-0.03	-0.435	0.663
Experiment: EvType: Language	-0.00	-0.021	0.983
MovDyn: EvType: Language	0.07	1.598	0.109
Experiment: MovDyn: EvType: Language	-0.09	-1.843	0.065

Furthermore, the combination of ceased actions and object-state change showed to especially enhance the recognition accuracy of the ending of events (in the verbal experiment: Spanish: 80.2% and Mandarin: 76.6%, and in the non-verbal experiment: Spanish: 72.5% and Mandarin: 60.3%). Ceased actions and changes of state in objects were shown to have a slight to moderate effect on memory. Accuracy performance split by cue type (matching / mismatching) is provided in Appendix B.

### Recognition task: influence of verb predicate semantics and form

To analyze the influence of verb predicates that denote culmination, and their form, on the representation of this event dimension in memory in comparison to other verb predicates across language groups, the recognition accuracy data of Experiment 2 was analyzed in isolation. Verb predicates that denote activities or unbounded changes of state were collapsed in a category we named "other". Furthermore, as Spanish



**Figure 5.** Recognition accuracy in Experiment 1 (non-verbal encoding) and Experiment 2 (verbal encoding) for the four event conditions across language groups.

speakers only used single verbs to denote culmination, and Mandarin speakers only did so in verb-satellite constructions, language type reflected the verb predicate form used to denote culmination. Once more, in this analysis the binary dependent variable was recognition accuracy (1 for correct responses and 0 for incorrect responses), and the fixed effects structure was kept the same in all models. The fixed factors included: (a) Movement Dynamics type (2 levels: ceased and ongoing actions), (b) Event type (2 levels: change-of-state and no-change-of-state events), (c) Verb Predicate Semantics type (2 levels: culmination and “other”), and (d) Language type (2 levels: Spanish and Mandarin). To centre the analysis on testing the particular prediction guiding the experiment (namely, that the ending of ceased change-of-state events would be remembered better when described with single verbs that denote culmination, as compared to other forms and semantics), the contrasts of the factors for Movement Dynamics and Event factors were dummy-coded and their reference level set to ceased actions and change-of-state events, respectively. On the other hand, contrasts of fixed factors for Verb Predicate Semantics type and Language type were sum-coded. The use of these contrasts served to assess the main effect of culmination verb predicates in Spanish (in the form of single verbs) against the reference level at the intercept of the model (i.e. the mean performance of participants on ceased change-of-state events). If single verbs lexicalising culmination in Spanish especially enhanced memory for that event dimension, we would expect to find a two-way interaction between Verb Predicate Semantics and Language, because the regression

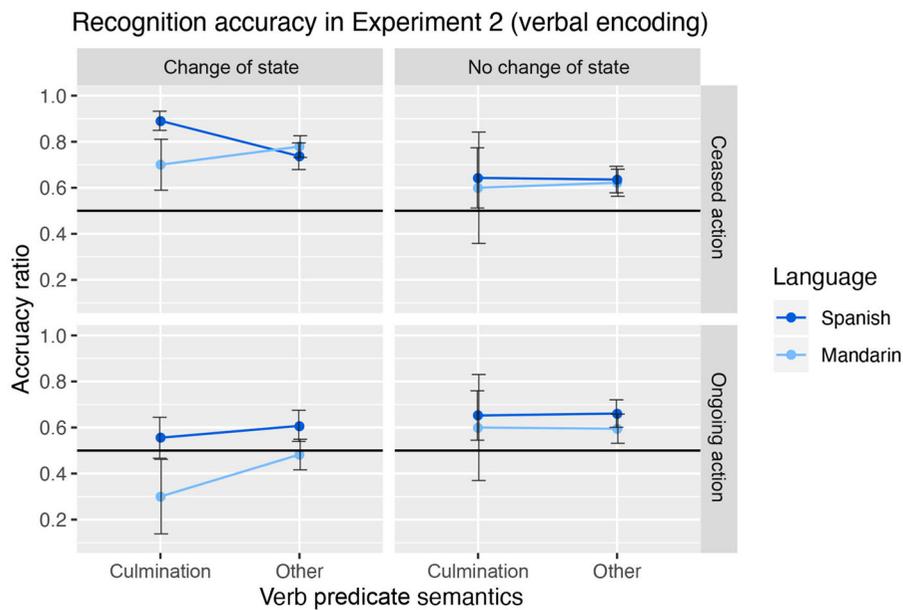
compares the effect of the different fixed factors and interactions to the reference level: ceased change-of-state events. An interaction between all fixed factors was included to specify a dependence among all of them in the regression analysis. The random structure of the final model had random intercepts for Participants and a random slope for Movement Dynamics by Participant.<sup>10</sup>

The results of the analysis (Table 4) and the effects plot (Figure 6) revealed a *simple effect* of Movement Dynamics ( $\beta = -1.49$ ,  $SE = 0.36$ ,  $p < .001$ ) and Event Type ( $\beta = -0.84$ ,  $SE = 0.35$ ,  $p < .02$ ), indicating that ongoing actions and events without change of state were less accurately recognised than ceased change-

**Table 4.** Recognition accuracy in Experiment 2.

*Fixed effects model on recognition accuracy in Experiment 2. Formula in R: glmer (Accuracy ~ MovDyn \* EvType \* Semantics \* Language + (1 + MovDyn | Participant.ID), data = data, family = binomial, control = glmerControl (optimizer = “bobyqa”))*

	Estimate	z value	Pr(> z )
(Intercept)	1.42	6.294	<0.001**
MovDyn: Ongoing	-1.49	-4.079	<0.001**
EvType: No-CoS	-0.84	-2.387	0.016**
Semantics	0.14	0.763	0.445
Language	0.25	1.166	0.243
Ongoing: No-CoS	1.56	3.097	0.001**
Ongoing: Semantics	-0.43	-1.463	0.143
No-CoS: Semantics	-0.14	-0.419	0.675
Ongoing: Language	0.17	0.471	0.637
No-CoS: Language	-0.13	-0.387	0.698
Semantics: Language	0.39	2.018	0.043**
Ongoing: No-CoS: Semantics	0.49	0.977	0.328
Ongoing: No-CoS: Language	-0.20	-0.413	0.679
Ongoing: Semantics: Language	-0.25	-0.851	0.394
No-CoS: Semantics: Language	-0.28	-0.800	0.423
Ongoing: No-CoS: Semantics: Language	0.11	0.230	0.817



**Figure 6.** Recognition accuracy for the four event conditions when participants produced verb predicates lexicalising culmination and other semantics across language groups.

of-state events. Additionally, the model showed an interaction between Movement Dynamics and Event Type ( $\beta = 1.56$ ,  $SE = 0.50$ ,  $p < .01$ ). Taken together, these results suggest that, overall, ceased actions and change of state had a positive effect on recognition accuracy when language is used to describe events (which aligns with moderate effect found in the previous model combining the data of both experiments, but not with the results of the non-verbal experiment alone). With respect to our question about the influence of the semantics and form of verb predicates on memory, the model showed a moderate interaction between Verb Predicate Semantics and Language ( $\beta = 0.39$ ,  $SE = 0.19$ ,  $p < .05$ ). This suggests that the semantics of the verb predicates used to describe events may have had a different impact on recognition accuracy depending on their form (either single verbs by Spanish speakers of verb-satellite constructions by Mandarin speakers). In the case of ceased change-of-state events, single verbs that denote the culmination moment of events seem to have enhanced participants' representation of this event feature in memory as compared to the culmination semantics conveyed by verb-satellite constructions (Spanish: 90.2% and Mandarin: 71.4%) and verb predicates of both languages that lexicalise a different semantic feature of events (Spanish: 75.7% and Mandarin: 80.3%).

### Discussion: Experiment 2

Results of Experiment 2 showed that explicitly describing events strengthened event memory overall, and in

particular, people's memory for events with ceased actions (regardless of event type), and for change-of-state events (regardless of movement dynamics at the ending of the videos). Furthermore, the combination of these event features, in culminated change-of-state events, led to an enhanced memory accuracy in both language groups. As for the effect of using particular verb predicates on the memory representation of culminated change-of-state events, the results indicated that the lexical semantics of verbs did not have an effect on event memory. This finding contrasts with what we predicted based on the depth-of-processing ( Craik & Tulving, 1975) and thinking-for-speaking (Slobin, 1996): here, there were no differences in memory for verb predicates that expressed the inherent culmination moment of change-of-state events as compared to those with different semantics. Instead, the results suggest that the form of verb predicates may interact with their semantics. It seems that using single verbs to lexicalise the culmination moment of change-of-state events (by Spanish speakers) enhanced the representation of event endings in memory in comparison to the use of verb-satellite constructions (by Mandarin speakers).

### General discussion

In the present study, using a surprise recognition task, we investigated in which way different movement dynamics, and different types of events influenced people's memory for event endings. We varied the movement dynamics of actions at video offset (ceased

versus ongoing actions), and event type (events in which the change in state of an object clearly progresses towards a particular culmination moment, versus events without a conspicuous change of state in objects). We furthermore asked to what extent the use of language, or not, and the semantics of verb predicates (encoding culmination or not) and their form (single verb or verb-satellite construction) affected how accurately people remembered event endings. We hypothesised that videos showing ceased actions, events involving a change-of-state in an object, and/or events showing culmination (the combination of both dimensions) would lead to enhanced memory. Additionally, we hypothesised that the explicit use of language to describe events would enhance memory overall. We were specifically interested in how verb predicates that lexicalise the culmination moment of events, and the different linguistic forms of these verb predicates (single verbs versus satellite constructions) may influence people's memory of culminated events. The most important findings are, first of all, that event culmination is a particularly well represented dimension in event memory, second, that the verbal encoding of events enhances memory, and third, that the form of the verbal predicates used to denote event culmination modulated recognition accuracy.

Studies on event segmentation have shown that changes in the movement dynamics (speed and acceleration) strongly correlate with the segmentation of long stretches of ongoing activities into event units at a fine-grained level (Hard et al., 2006; Zacks, Kumar, et al., 2009). As event segmentation theory posits that the detection of an event boundary leads to the commitment of perceived events to episodic memory (Zacks et al., 2007), we hypothesised that change in movement dynamics of actions would contribute to the updating of event representations in episodic memory. Therefore, the cessation of a fine-grained activity at the offset of short video clips (the ending of an event) should lead to a better memory of the last frames of the stimuli, compared to videos that show an ongoing activity at the offset. However, the results of the present study suggest that this memory update may not be directly influenced by just change in the movement dynamics of actions. Events with ceased actions (in which the speed and acceleration of event participants came to a stop) were not remembered better than ongoing ones, at least when people encoded events without describing them. Thus, the update of events in episodic memory, which takes place upon the detection of event boundaries, may be sensitive to a more complex, multidimensional set of cognitive foundations that regulate what people conceive as an event.

Further, it has been proposed that attending to and keeping track of how objects change is essential for apprehending the progression of events (Altmann & Ekves, 2019; Sakarias & Flecken, 2019; Solomon et al., 2015). People not only represent the current, altered state of an affected object but they also keep the original state of the object activated in memory (presumably, to be able to track how an specific object changed state). In this regard, we hypothesised that activities involving a change of state in an object should lead to stronger memory representations than those in which there is no object change. However, the results of the present study indicate that, when people did not describe events, events with a clear change of state in an object were not remembered better than those in which objects did not undergo a change of state. This is reminiscent of the findings reported by Sakarias and Flecken (2019) who found an asymmetry between participants' gaze patterns recorded during online event processing and their subsequent memory of events: Although during encoding people paid more attention to the development and ending of events that involved a state change in objects, compared to events that lacked such state change, their memory performance for such event types was not enhanced in the subsequent recognition task. The finding that object-state change did not enhance event memory (in our study and Sakarias & Flecken, 2019), in spite of the salience of such change during the perception of ongoing events, may be due to the involvement of two different memory mechanisms for event cognition (as outlined by Altmann & Ekves, 2019; Radvansky & Zacks, 2017; Zacks, 2020). When encoding an activity in progress, people keep track of changes in the situation by continuously updating a *working model* of the unfolding event in working memory. Upon detecting an event boundary, that working model is committed to episodic memory (freeing space in working memory to track the new incoming sensory input in a new working model). This means that, when apprehending and keeping track of events as they unfold, people engage attention and working memory resources to form a mental representation of the event in progress. Only when a boundary is detected, is the event model in episodic memory updated (e.g. Richmond & Zacks, 2017; Zacks, 2020). It is these representations in episodic memory, and not the representation of events in working memory, that our recognition memory task (and the one reported in Sakarias & Flecken, 2019), tapped into. Even though, as hypothesised, participants may have engaged more attentively during perception of object-state change than during perception of events without it, their working memory model was committed to

episodic memory only at the ending (i.e. when the agent's action ceased), which may have led to the result that we attested: memory was consistently enhanced only for ceased change-of-state events (events that reached culmination). In line with this, it has been found that people, despite having paid attention to causal events in videoclips, frequently reported seeing parts of events that were actually missing (Papenmeier, Brockhoff, et al., 2019; Strickland & Keil, 2011), suggesting that visual attention during encoding does not directly map to the representation of events in memory.

The most important finding of our study was that event culmination has a special status in the representation of events. While cessation and object-state change in isolation did not particularly boost event memory, the combination of both features (namely, change-of-state events that progressed until a culmination moment was reached) showed to be especially salient in memory. This was the case regardless of whether or not language was used during encoding. This is in line with previous studies that reported a bias towards the goal and endpoint of moving agents, over manner and source of motion (e.g. Lakusta & Landau, 2012; Lakusta & Carey, 2015; Papafragou, 2010; Regier & Zheng, 2007, for reviews cf. Levine et al., 2019; Ünal & Papafragou, 2019). Further, it has been shown that identifying the goal of an action is not only important for conceptualising an action or event in general (e.g. Ambrosini et al., 2013; van Elk et al., 2011), or for learning novel actions (e.g. Monroy et al., 2017, 2018), but also for the detection of event boundaries (Bläsing, 2015; Levine et al., 2017; Zacks et al., 2001). Recently, Ji and Papafragou (2020) also concluded that the culmination moment of an event is a relevant dimension of an event. Using a categorisation paradigm, they found that when people were specifically trained to generalise the notion of event culmination, when presented with a new set of stimuli they were able to distinguish events that had a culmination moment from those that did not. In sum, conceiving the intended goal of an agent's action has been shown to be essential for predicting, segmenting, comprehending and organising the flow of sensory information. Our findings contribute to this literature in showing that the cognitive bias towards goals and endpoints of events is also reflected in a strong memory representation for change-of-state events that reach an inherent goal or culmination moment.

The enhanced representation of event culmination in memory can be explained by two different event cognition theories: event segmentation (Zacks et al., 2007) and event indexing (Zwaan et al., 1995). On the one

hand, differences in the inner structure of the event types we tested may have impacted the predictability of agents' intentions, which may have led to the attested differences in event memory. On the other hand, the co-occurrence of changes across multiple event dimensions may have led to a strong representation of culminated events in memory. Event segmentation theory posits that the update of event models in episodic memory is a process that relies on the spiking of prediction error, i.e. a sudden increase in prediction error related to the updating of working models in episodic memory (cf. Richmond & Zacks, 2017). As mentioned above, it is possible that events which progressed towards an inherent culmination moment (i.e. change-of-state events in our study) allowed people to create clear expectations about the event's end goal and the actor's intentions, helping them to predict the course of actions (and keeping prediction error relatively low while these events unfolded). In contrast, when events lacked a clear culmination moment (i.e. events without a change of state in our study), people may have had problems forming similarly clear expectations about the goal of the events, hindering their ability to predict their course of action (and keeping prediction error at a higher level during the processing of these events, as compared to those in which the state change in objects followed a predictable path towards an inherent goal-state). This difference in people's expectations about the goal (or culmination moment) towards which events progressed could have led to a spike of different intensity in prediction error at the ending of both event types, which could have had a different influence on episodic memory updating. Future research could investigate this possibility by comparing people's ability to predict upcoming stages of events (e.g. Huff et al., 2014) during the mid-phases of fine-grained event units of both types (events which involve a change of state in objects and those without a conspicuous change of state in objects) and at the ending of such events. This would allow to assess the extent to which differences in the predictability of events relate to people's memory performance.

Alternatively, event indexing theory (Zwaan et al., 1995; Zwaan & Radvansky, 1998) posits that events can be decomposed along five indexes: time, location, character, intention and causation. The more dimensions change in a given situation, the harder it becomes to integrate it as part of the current situation model. Furthermore, this increase in the difficulty to maintain the current event as part of the same situation model has been found to correlate with: (1) the consistency with which people detect an event boundary, (2) the memory of people about the event, and even (3) the

strength of people's prediction error (cf. Huff et al., 2014). In our study, event culmination (operationalised as the cessation of change-of-state events) yielded change in *three* different dimensions: location (change in the movement dynamics), causation (object-state change), and intention (achievement of an agent's goal). It is possible that this type of event led to better memory because it was the only condition in which more than one dimension changed, resulting in a more substantial change at video offset than in any of the other conditions, as event indexing theory would predict. Future studies could shed light on the underlying mechanisms of this effect, for example, by comparing how well people remember the culmination of change-of-state events (with change in two dimensions: movement dynamics and object state) against the culmination of change-of-location events (with change in only movement dynamics but not in object state).

Turning to the discussion of the influence of language on event memory, the results of our study suggested an enhancement of people's memory of events due to their explicit use of language during encoding. When people actively described events, memory performance was better overall as compared to when participants did not use language during encoding. Moreover, event type and movement dynamics became important individual factors that improved memory performance; ceased actions and events involving object-state change led to better event memory when language was explicitly used. As hypothesised, it is possible that the attention required to describe events sharpened the overall focus of participants to event details. This may have implications for event cognition methodology: for example, in classic event segmentation tasks (Newtson, 1973) it is possible that results may in part have been driven by the influence of language use, and not merely by non-linguistic cognitive processes.

The existence of an interface between language and cognition is not a new idea, even if it is not yet fully understood. It has been attested in perceptual categorisation (cf. Lupyan (2016) for a review), attention (Slobin's (1996) thinking-for-speaking theory), memory ( Craik and Tulving's (1975) depth-of-processing theory), and also event segmentation (e.g. Gerwien & von Stutterheim, 2018; Papenmeier, Maurer, et al., 2019). Papenmeier, Maurer, et al. (2019) found that information lexicalised in language can help people detect fine-grained event boundaries, and Gerwien and von Stutterheim (2018) showed that specific characteristics of the language that people speak can guide their performance in an event segmentation task. Furthermore, Sakarias and Flecken (2019) showed that the use of language to describe events in general, and the use of language-

specific means to describe particular features of events (in particular, event endings and the end-state of objects), can enhance event memory. All these studies, including the present study, suggest that language plays an important role in event cognition (as put forth recently in Zacks [2020]). However, there are other studies that conclude that there is no or only little effect of language on event cognition (e.g. Athanassopoulos & Bylund, 2013; Bungler et al., 2016; Montero-Melis et al., 2017; Ji & Papafragou, 2020). These conclusions are based on the lack of effects in tasks that rely on verbal interference. Nevertheless, even in those studies, in similar conditions in which language is not interfered with, language effects on event cognition do emerge, suggesting that language indeed enhances cognitive representations and processes.

In order to examine whether the use of verb predicates with a particular semantics and form had an influence on memory for event culmination, we analyzed the verb predicates produced by Spanish and Mandarin speakers with respect to form and semantics. As expected, in line with Talmy's (2000, 2016) event-framing theory, Spanish speakers lexicalised the presence of a culmination moment in single verbs, while Mandarin speakers did so mostly using verb-satellite constructions. Hence, Spanish and Mandarin showed to be adequate proxies for different ways of lexicalising the culmination moment of events in verb predicates (single verbs versus verb-satellite constructions). Contrary to what we hypothesised given the depth-of-processing theory (Craik & Tulving, 1975) and thinking-for-speaking theory (Slobin, 1996), the particular lexical-semantics of the verb predicates did *not* seem to have an effect on event memory. The use of verb predicates that denoted event culmination did not particularly enhance the memory representation of this dimension as compared to the use of verb predicates with other semantics. Instead, the results suggest that lexical-semantics may interact moderately with the form of verb predicates. The use of single verbs (by Spanish speakers) to lexicalise the culmination moment of events seems to have especially enhanced the representation of this event dimension in participants' memory in comparison to the use of verb-satellite constructions (by Mandarin Speakers). This is in line with our hypothesis about the different effect of single verbs and verb-satellite constructions on the memory of event culmination. Lexicalising the culmination moment of events in single verbs (in Spanish) gives rise to a less complex event structure (namely, just a change-of-state event) than the use of verb-satellite constructions (in Mandarin) which encode a more complex event structure as they (as specifying both a manner subevent and a change-

of-state subevent). The former seems to have led to a better memory for event culmination. This finding connects to Fausey et al.'s (2010) language-memory interface study, in which cross-linguistic differences in memory performance were interpreted as the result of the use of different linguistic structures to frame events. In that study, differences in how well people remembered agents involved in accidental events could be found not only across, but also within language groups by simply manipulating how participants described events (either by explicitly mentioning the agent involved or not).

We can think of three possible explanations as to why we did not find an effect of the lexical-semantics of verb predicates on event memory. One possible explanation is that the memory enhancement driven by the semantics of words in depth-of-processing theory ( Craik & Tulving, 1975) may be limited to the presentation and retrieval of isolated words. In our study, the stimuli used were visual and more complex (i.e. videos of unfolding events in which a person interacted with an object), and the dimension that was selected as an indicator of people's event memory performance was only a small part of those events (i.e. the final frames of the events at video offset). Furthermore, people's event descriptions contained a broader range of semantic information besides the meaning of the verb predicates. The combination of all these factors in the present study could have diminished the effectiveness of verb predicates as individual semantic cues for retrieving the ending of the presented events from episodic memory. Follow-up research may investigate whether the memory enhancement effect provided by the semantics of linguistic cues holds as the language these cues are embedded in becomes increasingly complex. A second explanation may relate to the semantic categorisation of verbs performed by informants in our study. To our knowledge, our verb classification represents a first attempt to establish a method for assessing whether the lexical semantics of verbs convey an unbounded change of state (CoS verbs), a bounded change of state (Culmination verbs), or no change of state at all (Activity verbs). The use of a different method for classifying verbs may lead to different results. One last explanation for the lack of influence of semantics is that "thinking for speaking" during encoding (i.e. focused attention on the particular event dimensions of events to be described, cf. Slobin, 1996) does not necessarily lead to a stronger representation of these event dimensions in episodic memory. As pointed out above, attention allocation during event processing may not be directly linked to episodic memory (Altmann & Ekses, 2019; Radvansky & Zacks, 2017; Zacks, 2020). While we

found a "speaking for memory" effect (a memory advantage in the verbal condition over the non-verbal one), we did not find a "thinking for speaking" effect. So, "thinking for speaking" may be different from "speaking for memory". This could also explain why certain prior studies did not find a clear symmetry between online attention allocation whilst preparing to speak, and subsequent memory performance (e.g. Papafragou et al., 2002; Sakarias & Flecken, 2019). This asymmetry is reminiscent of other studies which have found that, despite paying full attention to events, people's memory of them is not faithful, but moulded by inferences that originate from their world knowledge (e.g. Loftus & Palmer, 1974; Papenmeier, Brockhoff, et al., 2019; Strickland & Keil, 2011).

It is important to note that the number of verb predicates that participants produced to explicitly highlight the culmination moment of events was low in both language groups. This is a shortcoming of free elicited-production tasks, like the one we used. To better understand the effects of "speaking for memory", it would be pertinent to create a design that tests *both* online attention allocation during encoding and subsequent memory in a verbal and non-verbal encoding context, while systematically "feeding" participants with particular verb predicates or linguistic structures controlled for complexity and lexical-semantics, in the verbal variant of the experiment. In addition, given the complexity of the design of the experiments we conducted, the sample sizes used are fairly low to provide robust results. Future studies could benefit from larger samples of speakers.

Surprisingly, the results of the non-verbal experiment showed that the culmination of change-of-state events was more salient in the memory of Spanish speakers than Mandarin speakers. This difference may be rooted in group differences with respect to the understanding of the task, the amount of experience in performing experimental tests, etc. Such explanations seem unlikely, however, as the instructions for participants were thoroughly checked and revised by native speakers of both languages, the educational profile and age of the groups was similar, and none of the participants had experience in participating in psycholinguistic studies. Instead, we believe that people's native language, as a means of enculturation (the process whereby people adopt the relevant beliefs and behaviour patterns of a culture; cf. Silvey et al., 2015), may have habituated participants from the two language groups to represent event culmination with a different strength in memory. Depending on the event dimension that people's native language habitually requires them to lexicalise (e.g. a result feature in Spanish single verbs, cf. Göksun

et al., 2010, 2017), such habituation may have reinforced the relevance of this event dimension in event memory. As discussed by Talmy (2000, 2016), and as supported by the coding of the verb predicates that people produced in our study, it seems that Spanish main verbs often express the result of events (either change of state or culmination moment). In contrast, Mandarin main verbs often express the manner of action and not the result of events. Moreover, Mandarin speakers use verb-satellites to denote event results only when they want to emphasise the achievement of a particular culmination moment; frequently event culmination remains implicit and follows as an inference from a morphologically simple, manner verb (Chen, 2018; Tai, 1984). This cross-linguistic asymmetry in the semantic content that single verbs habitually encode might have led one group, but not the other, to have an especially strong memory representation of event culmination, even in the absence of explicit use of language.

Interestingly, the performance of Mandarin speakers attested in the verbal and non-verbal experiments of the present study was similar to the pattern attested in a parallel study using the same stimuli and design in Dutch (Santín et al., 2020), a satellite-framed language as well. Similar to the Mandarin speakers in the present study, Dutch participants' memory of event culmination was poor when events were encoded non-verbally; however, when events were encoded verbally, their memory for event culmination was specifically enhanced. The similarity of the memory patterns found in these two satellite-framed languages, and the differences with the memory pattern attested in Spanish (a verb-framed language), supports the idea that language-specific effects may underlie these differences on event memory.

Considering the evolution of language as a reflection of what people over time have found relevant to discriminate and lexicalise (Silvey et al., 2015; Tamariz & Kirby, 2016), this study has made clear that event culmination is an important dimension of events. The presence and achievement of culmination is expressed through a wide variety of designated means in different languages (Beavers et al., 2010). How verb predicates express an event culmination moment has been thoroughly discussed in linguistic theory for many years as the telicity of event descriptions (Dowty, 1979; Jackendoff, 1996; Hay et al., 1999; Kennedy & Levin, 2008; Krifka, 1998; McNally, 2017; van Hout, 2016; Vendler, 1957; Verkuyl, 1993). In linguistics, reaching event culmination is analyzed as the result of the interaction of two different semantic dimensions: the inherent structure of events (i.e. the temporal contour along which an event develops which may or may not have a natural

endpoint) and the viewpoint structure of events (i.e. the temporal perspective from which an event is linguistically portrayed). An inference of event culmination arises when the verb predicate lexicalises the result of an action (a telic verb predicate such as *to cut a piece of paper in two*), and, moreover, the tense-aspect marking on the verb presents an event as a finished action (perfective aspect, e.g. *she cut the paper*). Thus, the interpretation that is brought about by these two linguistic dimensions is that the resultative event described by the telic verb predicate progressed until its natural endpoint was reached (e.g. a cutting event progressed till a piece of paper separates in two pieces). In our study, the video stimuli operationalised these two linguistic event dimensions by means of varying event type and movement dynamics. First, the stimuli showed events with a different inherent temporal development: events that progress towards a culmination moment (i.e. change-of-state events) and events that do not (i.e. events without object-state change). Second, stimuli showed events with a different temporal structure: videos of events in which a process comes to an end at video offset (i.e. ceased actions) and videos in which events do not come to an end at their offset (i.e. ongoing actions). Our study thus operationalised parallel distinctions in linguistic theory (perfective aspect and telicity) on the one hand, and event cognition theory (movement dynamics and object-state change) on the other. We would like to advocate the use of approaches closely integrating linguistic and cognitive neuroscientific theories, as the one taken in the present study, to help advance our knowledge and understanding of how events are construed in language and cognition. Furthermore, they may be useful to better understand how our experience with events through language can affect event cognition.

## Notes

1. However, there is a different way of interpreting the results of their experiments. Their results show that people performed well at abstracting the notion of culmination only when the target stimuli (i.e. videos showing an event progressing towards a culmination moment) were highlighted with a red frame. Nevertheless, when the target stimuli were not highlighted, people had a hard time abstracting the notion of culmination. This asymmetry suggests that the performance of people was influenced by the experimental setup and may not truthfully reflect a natural sensitivity to event culmination.
2. A recent corpus study by Métairy et al. (2020) suggests that resultative constructions (i.e. verb-satellite constructions) can indeed be attested in Spanish in the very particular case of nomination verbs (i.e. verbs that denote a change in someone's status, e.g. "elegir" ("elect")),

“proclamar” (“proclaim”), and “coronar” (“crown”). For example: “Henry fue proclamado rey” / “Henry was proclaimed King”.

3. The existence of a few exceptions in which satellites encoding event results can be used as single verbs has led some linguists to classify Mandarin as a third type of language: so-called equipollently framed languages (cf. Beavers et al., 2010; Slobin 2006). In the present study we follow Talmy’s (2010, 2016) classification of Mandarin as a satellite-framed language.
4. As most international students in the Netherlands have no to very little knowledge of Dutch language. We only asked participants about their proficiency in English because it is the common language they use for their daily interactions and because it was the language used by the experimenter to interact with participants.
5. 42 of these videos were taken from Sakarias and Flecken (2019). Six extra videos were recorded using a new actor but keeping the same characteristics of the recording set-up. Materials (event videos) developed by Sakarias and Flecken (2019) can be downloaded from <https://osf.io/uuxtq/>
6. Presenting a single cue rather than two alternatives in the recognition task (as in other memory studies, e.g. Sakarias & Flecken, 2019) allowed us to probe with more precision the representation of events in memory.
7. Note that the final model does not include a random intercept for Items (event-videos). Including this random intercept in the model reduced the AIC score, hence we favored the current model.
8. Note that the verb predicates were coded without reference to tense-aspect marking on the verb or in the sentence, this is because the focus of the present study was on the lexical semantics of single verbs and verb-satellite constructions, specifically on whether or not they express the culmination moment of an event. Tense-aspect marking encodes a different temporal dimension: whether or not a given action is ongoing (imperfective aspect) or ceased (perfective aspect). In the present study, Mandarin speakers scarcely used perfective aspect markers (either “le” or “ba”) to describe ceased events ( $N=43$  for ceased change-of-state events, and  $N=15$  for no change-of-state events). Spanish speakers used perfective aspect to describe ceased events to an even lesser extent ( $N=10$  for ceased change-of-state events, and  $N=7$  for no change-of-state events). Hence, we did not analyse the use of aspect further.
9. Note that the final model does not include a random intercept for Items (event-videos). Including this random intercept in the model reduced the AIC score, hence we favored the current model.
10. Note that the final model does not include a random intercept for Items (event-videos). Including this random intercept in the model reduced the AIC score, hence we favored the current model.

## Acknowledgements

We thank Muqing Li, Hanjing Yu, Hongying Peng, Amanda Fredes, Ivan Moreno and Miguel Souto for their help in coding the linguistic data.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Data deposition

The materials and data that support the findings of this study are openly available for reviewing purposes OSF at: <https://osf.io/ypxkm/>

## ORCID

Miguel Santin  <http://orcid.org/0000-0003-1674-7224>

Angeliek van Hout  <http://orcid.org/0000-0002-4971-7243>

## References

- Altmann, G. T. M., & Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review*, 126(6), 817–840. <https://doi.org/10.1037/rev0000154>
- Ambrosini, E., Reddy, V., de Loopier, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking ahead: Anticipatory gaze and motor ability in infancy. *PLoS ONE*, 8(7), e67916. <https://doi.org/10.1371/journal.pone.0067916>
- Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, 37(2), 286–309. <https://doi.org/10.1111/cogs.12006>
- 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. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models. *ArXiv*. <https://arxiv.org/abs/1506.04967v2>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Beavers, J. (2008). Scalar complexity and the structure of events. In J. Dölling, T. Heyde-Zybatow, & M. Schäfer (Eds.), *Event structures in linguistic form and interpretation*. De Gruyter. <https://doi.org/10.1515/9783110925449.245>
- Beavers, J., Levin, B., & Wei Tham, S. (2010). The typology of motion expressions revisited. *Journal of Linguistics*, 46(2), 331–377. <https://doi.org/10.1017/S0022226709990272>
- Berthele, R. (2013). *Disentangling manner and path: Evidence from varieties of German and Romance* (pp. 55–75). John Benjamins.
- Bläsing, B. E. (2015). Segmentation of dance movement: Effects of expertise, visual familiarity, motor experience and music. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01500>
- Bunger, A., Papafragou, A., & Trueswell, J. C. (2013). Event structure influences language production: Evidence from structural priming in motion event description. *Journal of Memory and Language*, 69(3), 299–323. <https://doi.org/10.1016/j.jml.2013.04.002>

- Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2016). How children and adults encode causative events cross-linguistically: Implications for language production and attention. *Language, Cognition and Neuroscience*, 31(8), 1015–1037. <https://doi.org/10.1080/23273798.2016.1175649>
- Chen, J. (2018). "He killed a chicken, but it didn't die": An empirical study of the lexicalization of state change in Mandarin monomorphemic verbs. *Chinese Language and Discourse. An International and Interdisciplinary Journal*, 9(2), 136–161. <https://doi.org/10.1075/cld.17007.che>
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in Episodic Memory. *Journal of Experimental Psychology: General*, 104(3), 268–294. <https://doi.org/10.1037/0096-3445.104.3.268>
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. D. Reidel Publishing Company.
- 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(S1), 209–213. <https://doi.org/10.1007/s10339-015-0696-7>
- Fausey, C. M., Long, B. L., Inamori, A., & Boroditsky, L. (2010). Constructing agency: The role of language. *Frontiers in Psychology*, 1. <https://doi.org/10.3389/fpsyg.2010.00162>
- Filipović, L. (2010). Thinking and speaking about motion: Universal vs. Language-specific effects. In G. Marotta, A. Lenci, L. Meini, & F. Rovai (Eds.), *Space in Language: Proceedings of the Pisa International Conference* (pp. 235–348). Editrice Testi Scientifici. <http://www.edizioniets.com/scheda.asp?n=9788846729026>
- 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. <https://doi.org/10.1111/j.1540-4781.2015.12181.x>
- García del Real, I. (2015). *Acquisition of tense and aspect in Spanish* [Unpublished doctoral dissertation]. Universidad del País Vasco.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79. [https://doi.org/10.1016/S0010-0277\(01\)00166-4](https://doi.org/10.1016/S0010-0277(01)00166-4)
- George, N. R., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Carving the world for language: How neuroscientific research can enrich the study of first and second language learning. *Developmental Neuropsychology*, 39(4), 262–284. <https://doi.org/10.1080/87565641.2014.906602>
- Gerwien, J., & von Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, 180, 225–237. <https://doi.org/10.1016/j.cognition.2018.07.008>
- Göksun, T., Aktan-Erciyes, A., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Event perception and language learning: Early interactions between language and thought. In F. N. Ketrez, A. C. Küntay, Ş. Özçalışkan, & A. Özyürek (Eds.), *Trends in language acquisition research* (Vol. 21, pp. 179–198). John Benjamins Publishing Company. <https://doi.org/10.1075/tilar.21.12gok>
- Göksun, T., Hirsh-Pasek, K., & Michnick Golinkoff, R. (2010). Trading spaces: Carving up events for learning language. *Perspectives on Psychological Science*, 5(1), 33–42. <https://doi.org/10.1177/1745691609356783>
- Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition*, 34(6), 1221–1235. <https://doi.org/10.3758/BF03193267>
- Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in "degree achievements." *Proceedings of the 9th semantics and linguistic theory conference*, 127–144. <http://dx.doi.org/10.3765/salt.v9i0.2833>
- Hindy, N. C., Altmann, G. T. M., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object state-changes on event processing: Do objects compete with themselves? *Journal of Neuroscience*, 32(17), 5795–5803. <https://doi.org/10.1523/JNEUROSCI.6294-11.2012>
- Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., & Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition*, 154, 151–164. <https://doi.org/10.1016/j.cognition.2016.05.013>
- Huff, M., Meitz, T. G. K., & Papenmeier, F. (2014). Changes in situation models modulate processes of event perception in audiovisual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1377–1388. <https://doi.org/10.1037/a0036780>
- Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory*, 14(2), 305–354. <https://doi.org/10.1007/BF00133686>
- Ji, Y., & Papafragou, A. (2020). Is there an end in sight? Viewers' sensitivity to abstract event structure. *Cognition*, 197, 104197. <https://doi.org/10.1016/j.cognition.2020.104197>
- Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements. In L. McNally & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics, and discourse* (pp. 156–182). Oxford University Press.
- Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar* (Vol. 70, pp. 197–235). Springer Netherlands. [https://doi.org/10.1007/978-94-011-3969-4\\_9](https://doi.org/10.1007/978-94-011-3969-4_9)
- Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 12(2), 72–79. <https://doi.org/10.1016/j.tics.2007.11.004>
- Lakusta, L., & Carey, S. (2015). Twelve-month-old Infants' encoding of goal and source paths in agentive and non-agentive motion events. *Language Learning and Development*, 11(2), 152–175. <https://doi.org/10.1080/15475441.2014.896168>
- Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, 96(1), 1–33. <https://doi.org/10.1016/j.cognition.2004.03.009>
- 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. <https://doi.org/10.1111/j.1551-6709.2011.01220.x>
- Lawrence, Z., & Peterson, D. (2016). Mentally walking through doorways causes forgetting: The location updating effect and imagination. *Memory*, 24(1), 12–20. <https://doi.org/10.1080/09658211.2014.980429>
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
- Levine, D., Buchsbaum, D., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Finding events in a continuous world: A developmental account. *Developmental Psychobiology*, 61(3), 376–389. <https://doi.org/10.1002/dev.21804>

- Levine, D., Hirsh-Pasek, K., Pace, A., & Michnick Golinkoff, R. (2017). A goal bias in action: The boundaries adults perceive in events align with sites of actor intent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(6), 916–927. <https://doi.org/10.1037/xlm0000364>
- Li, C. (2013). Mandarin resultative verb compounds: Simple syntax and complex thematic relations. *Language Sciences*, 37, 99–121. <https://doi.org/10.1016/j.langsci.2012.11.001>
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585–589. [https://doi.org/10.1016/S0022-5371\(74\)80011-3](https://doi.org/10.1016/S0022-5371(74)80011-3)
- Lupyan, G. (2016). The centrality of language in human cognition. *Language Learning*, 66(3), 516–553. <https://doi.org/10.1111/lang.12155>
- Malaiya, E. (2014). It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass*, 8(3), 89–98. <https://doi.org/10.1111/lnc3.12071>
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>
- McNally, L. (2017). On the scalar properties and telicity of degree achievements. In O. Fernández-Soriano, E. Castroviejo, & I. Pérez-Jiménez (Eds.), *Linguistik Aktuell/linguistics today* (Vol. 239, pp. 174–192). John Benjamins Publishing Company. <https://doi.org/10.1075/la.239.09mnc>
- Métairy, J., Lauwers, P., Enghels, R., Taverniers, M., & Van Peteghem, M. (2020). A micro-typological perspective on resultative secondary predicates: The case of nomination verb constructions. *Language Sciences*, 78, 101253. <https://doi.org/10.1016/j.langsci.2019.101253>
- Monroy, C. D., Gerson, S. A., & Hunnius, S. (2017). Toddlers' action prediction: Statistical learning of continuous action sequences. *Journal of Experimental Child Psychology*, 157, 14–28. <https://doi.org/10.1016/j.jecp.2016.12.004>
- Monroy, C. D., Gerson, S. A., & Hunnius, S. (2018). Translating visual information into action predictions: Statistical learning in action and nonaction contexts. *Memory & Cognition*, 46(4), 600–613. <https://doi.org/10.3758/s13421-018-0788-6>
- Montero-Melis, G., Eisenbeiss, S., Narasimhan, B., Ibarretxe-Antuñano, I., Kita, S., Kopecka, A., Lüpke, F., Nikitina, T., Tragel, I., Florian Jaeger, T., & Bohnemeyer, J. (2017). Satellite- vs. Verb-framing underpredicts nonverbal motion categorization: Insights from a large language sample and simulations. *Cognitive Semantics*, 3(1), 36–61. <https://doi.org/10.1163/23526416-00301002>
- Newton, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, 28(1), 28–38. <https://doi.org/10.1037/h0035584>
- Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, 34(6), 1064–1092. <https://doi.org/10.1111/j.1551-6709.2010.01107.x>
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184. <https://doi.org/10.1016/j.cognition.2008.02.007>
- Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition*, 84(2), 189–219. [https://doi.org/10.1016/S0010-0277\(02\)00046-X](https://doi.org/10.1016/S0010-0277(02)00046-X)
- Papenmeier, F., Brockhoff, A., & Huff, M. (2019). Filling the gap despite full attention: The role of fast backward inferences for event completion. *Cognitive Research: Principles and Implications*, 4(1), 3. <https://doi.org/10.1186/s41235-018-0151-2>
- Papenmeier, F., Maurer, A. E., & Huff, M. (2019). Linguistic information in auditory dynamic events contributes to the detection of fine, not coarse event boundaries. *Advances in Cognitive Psychology*, 15(1), 30–40. <https://doi.org/10.5709/acp-0254-9>
- Pettijohn, K. A., & Radvansky, G. A. (2016). Walking through doorways causes forgetting: Event structure or updating disruption? *Quarterly Journal of Experimental Psychology*, 69(11), 2119–2129. <https://doi.org/10.1080/17470218.2015.1101478>
- Pustejovsky, J. (1991). The syntax of event structure. *Cognition*, 41(1–3), 47–81. [https://doi.org/10.1016/0010-0277\(91\)90032-Y](https://doi.org/10.1016/0010-0277(91)90032-Y)
- Radvansky, G. A. (2017). Event segmentation as a working memory process. *Journal of Applied Research in Memory and Cognition*, 6(2), 121–123. <https://doi.org/10.1016/j.jarmac.2017.01.002>
- Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, 17, 133–140. <https://doi.org/10.1016/j.cobeha.2017.08.006>
- Rappaport Hovav, M. (2016). Grammatically relevant ontological categories underlie manner/result complementarity. *Proceedings of Israel Association for Theoretical Linguistics: 32nd Annual Meeting*, 77–98. <https://www.iatl.org.il/wp-content/uploads/2017/12/4RappaportHovavPaper.pdf>
- Rappaport Hovav, M., & Levin, B. (2010). Reflections on manner/result complementarity. In M. Rappaport Hovav, E. Doron, & I. Sichel (Eds.), *Lexical semantics, syntax, and event structure* (pp. 21–38). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199544325.003.0002>
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Regier, T., & Zheng, M. (2007). Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science*, 31(4), 705–719. <https://doi.org/10.1080/15326900701399954>
- Richmond, L. L., & Zacks, J. M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences*, 21(12), 962–980. <https://doi.org/10.1016/j.tics.2017.08.005>
- Sakarias, M., & Flecken, M. (2019). Keeping the result in sight and mind: General cognitive principles and language-specific influences in the perception and memory of resultative events. *Cognitive Science*, 43(1), e12708. <https://doi.org/10.1111/cogs.12708>
- Santín, M., Hobbelenk, C., Flecken, M., & van Hout, A. (2020). Resultative event representations in Dutch children and adults: Does describing events help memory? *Proceedings of the 44th Boston University Conference on Language Development*, 508–521. <http://www.lingref.com/bucl44/BUCL44-41.pdf>
- Silvey, C., Kirby, S., & Smith, K. (2015). Word meanings evolve to selectively preserve distinctions on salient dimensions. *Cognitive Science*, 39(1), 212–226. <https://doi.org/10.1111/cogs.12150>
- Skordos, D., Bungler, A., Richards, C., Selimis, S., Trueswell, J., & Papafragou, A. (2020). Motion verbs and memory for motion

- events. *Cognitive Neuropsychology*, 37(5-6), 254–270. <https://doi.org/10.1080/02643294.2019.1685480>
- Slobin D. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickmann & S. Robert (Eds.), *Typological studies in language. Space in languages: Linguistic systems and cognitive categories* (p. 59–81). Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.66.05slo>.
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (Vol. 17, pp. 70–96). Cambridge University Press.
- Solomon, S. H., Hindy, N. C., Altmann, G. T. M., & Thompson-Schill, S. L. (2015). Competition between mutually exclusive object states in event comprehension. *Journal of Cognitive Neuroscience*, 27(12), 2324–2338. [https://doi.org/10.1162/jocn\\_a\\_00866](https://doi.org/10.1162/jocn_a_00866)
- Soroli, E., & Hickmann, M. (2010). Language and spatial representations in French and in English: Evidence from eye-movements. In G. Marotta, L. Alessandro, & M. Linda (Eds.), *Space in language: Proceedings of the Pisa international conference* (pp. 581–597). Editrice Testi Scientifici.
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12(2), 153–156. <https://doi.org/10.1111/1467-9280.00326>
- Stefanowitsch, A. (2013). Variation and change in English path verbs and constructions: Usage patterns and conceptual structure. In J. Goschler & A. Stefanowitsch (Eds.), *Variation and change in the encoding of motion events* (Vol. 41, pp. 223–244). John Benjamins Publishing Company. <https://doi.org/10.1075/hcp.41.10ste>
- Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distort memory in a matter of seconds. *Cognition*, 121(3), 409–415. <https://doi.org/10.1016/j.cognition.2011.04.007>
- Swallow, K. M., Barch, D. M., Head, D., Maley, C. J., Holder, D., & Zacks, J. M. (2011). Changes in events alter how people remember recent information. *Journal of Cognitive Neuroscience*, 23(5), 1052–1064. <https://doi.org/10.1162/jocn.2010.21524>
- Tai, J. (1984). Verbs and times in Chinese: Vendler’s four categories. Papers from the Parasession on Lexical Semantics of the 20th Regional Meeting of the Chicago Linguistic Society, 289–296.
- Talmy, L. (2000). *Toward a cognitive semantics, Vol II: Typology and process in concept structuring* (Vol. 2). MIT Press.
- Talmy, L. (2016). Properties of main verbs. *Cognitive Semantics*, 2(2), 133–163. <https://doi.org/10.1163/23526416-00202001>
- Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, 8, 37–43. <https://doi.org/10.1016/j.copsyc.2015.09.003>
- Ünal, E., Ji, Y., & Papafragou, A. (2019). From event representation to linguistic meaning. *Topics in Cognitive Science*, tops.12475. <https://doi.org/10.1111/tops.12475>
- Ünal, E., & Papafragou, A. (2019). How children identify events from visual experience. *Language Learning and Development*, 15(2), 138–156. <https://doi.org/10.1080/15475441.2018.1544075>
- van Elk, M., Paulus, M., Pfeiffer, C., van Schie, H. T., & Bekkering, H. (2011). Learning to use novel objects: A training study on the acquisition of novel action representations. *Consciousness and Cognition*, 20(4), 1304–1314. <https://doi.org/10.1016/j.concog.2011.03.014>
- van Hout, A. (1996). *Event semantics of verb frame alternations. A case study of Dutch and its acquisition* [Doctoral dissertation, Tilburg University]. TILDIL Dissertations Series. <https://research.tilburguniversity.edu/files/178183/72745.pdf>
- van Hout, A. (2016). Lexical and grammatical aspect. In J. L. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford handbook of developmental linguistics* (Vol. 1). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199601264.013.25>
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66(2), 143. <https://doi.org/10.2307/2182371>
- Verkuyl, H. J. (1993). *A theory of aspectuality: The interaction between temporal and atemporal structure* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511597848>
- Wagner, L., Yocom, A. M., & Greene-Havas, M. (2008). Children’s understanding of directed motion events in an imitation choice task. *Journal of Experimental Child Psychology*, 100(4), 264–275. <https://doi.org/10.1016/j.jecp.2008.03.008>
- Wolff, P., Jeon, G.-H., & Li, Y. (2009). Causes in English, Korean, and Chinese and the individuation of events. *Language and Cognition*, 1(2), 167–196. <https://doi.org/10.1515/LANGCOG.2009.009>
- Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., Hink, R. F., & Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology*, 33(1), 101–111. <https://doi.org/10.1080/13803395.2010.493149>
- Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, 29. <https://doi.org/10.1146/annurev-psych-010419-051101>
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6), 651–655. <https://doi.org/10.1038/88486>
- Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, 112(2), 201–216. <https://doi.org/10.1016/j.cognition.2009.03.007>
- Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, 138(2), 307–327. <https://doi.org/10.1037/a0015305>
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133(2), 273–293. <https://doi.org/10.1037/0033-2909.133.2.273>
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6(5), 292–297. <https://doi.org/10.1111/j.1467-9280.1995.tb00513.x>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168–171. <https://doi.org/10.1111/1467-9280.00430>