

## Going places in Dutch and mandarin Chinese: conceptualising the path of motion cross-linguistically

Yiyun Liao, Monique Flecken, Katinka Dijkstra & Rolf A. Zwaan

To cite this article: Yiyun Liao, Monique Flecken, Katinka Dijkstra & Rolf A. Zwaan (2019): Going places in Dutch and mandarin Chinese: conceptualising the path of motion cross-linguistically, *Language, Cognition and Neuroscience*, DOI: [10.1080/23273798.2019.1676455](https://doi.org/10.1080/23273798.2019.1676455)

To link to this article: <https://doi.org/10.1080/23273798.2019.1676455>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 17 Oct 2019.



[Submit your article to this journal](#)



Article views: 530



[View related articles](#)



[View Crossmark data](#)

## Going places in Dutch and mandarin Chinese: conceptualising the path of motion cross-linguistically

Yiyun Liao<sup>a</sup>, Monique Flecken<sup>b,c</sup>, Katinka Dijkstra<sup>a</sup> and Rolf A. Zwaan<sup>a</sup>

<sup>a</sup>Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, the Netherlands; <sup>b</sup>Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands; <sup>c</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, the Netherlands

### ABSTRACT

We study to what extent linguistic differences in grammatical aspect systems and verb lexicalisation patterns of Dutch and mandarin Chinese affect how speakers conceptualise the path of motion in motion events, using description and memory tasks. We hypothesised that speakers of the two languages would show different preferences towards the selection of endpoint-, trajectory- or location-information in Endpoint-oriented (not reached) events, whilst showing a similar bias towards encoding endpoints in Endpoint-reached events. Our findings show that (1) groups did not differ in endpoint encoding and memory for both event types; (2) Dutch speakers conceptualised Endpoint-oriented motion focusing on the trajectory, whereas Chinese speakers focused on the location of the moving entity. In addition, we report detailed linguistic patterns of how grammatical aspect, verb semantics and adjuncts containing path-information are combined in the two languages. Results are discussed in relation to typologies of motion expression and event cognition theory.

### ARTICLE HISTORY

Received 5 October 2018  
Accepted 24 September 2019

### KEYWORDS


Motion events; cross-linguistic analysis; the path of motion; grammatical aspect; verb lexicalisation patterns


### 1. Introduction

Motion through space is one of our most fundamental bodily and perceptual experiences. A motion event is a complex construct which consists of people, objects, spatial information, and temporal change. Take the following scenario as an example: *Being late for work, you run out of the house, manoeuvre through the people on the street, dash towards your working place, and eventually rush into your office.* This motion event contains the following components: a figure (i.e. a moving person or object; e.g. *you*), a path of motion (i.e. Source, Trajectory, Location, Endpoint; e.g. *out of, through, on, into*), a ground (an object that functions as a reference point for the path of motion, e.g. *the house, the people, the street, your office*), a manner of motion (specific features of the figure's motion, such as gait, speed, etc.; e.g. *to jump, to run, to manoeuvre, to dash*; Talmy, 1985; 2000), and the temporal contour of motion (whether the motion is ongoing or has ceased, marked verbally; e.g. the *-ing* or *-ed* form of *jumping, jumped*; Comrie, 1976; Smith, 1991). In addition, any clause in the story above can be classified as, either a one-state situation/activity that does not involve fundamental qualitative change

(e.g. *manoeuvre through the people on the street*), or a two-state situation (accomplishment or achievement) that involves a change of location (also called a boundary-crossing event: *rush into the office*) (Croft, 2012; Klein, 1994; Vendler, 1967). How we conceptualise the different motion components in motion events of different types and articulate them in various languages is an intriguing question that has been explored extensively by researchers from the fields of psychology and linguistics (Beavers, Levin, & Tham, 2010; De Knop & Gallez, 2011; Filippo-Enrico, 2008; Ibarretxe-Antuñano, 2004; Ji, Hendriks, & Hickmann, 2011; Ochsensbauer & Hickmann, 2010; Slobin, 1996, 2004, 2008; Talmy, 1985, 1991, 2000).

In this study, we are particularly interested in one of those components, i.e. the path of motion. We ask to what extent, and how, cross-linguistic differences in grammar (grammatical aspect) and semantics (verb lexicalisation patterns) influence how speakers conceptualise and describe the path of motion in a motion event. The path of motion is a complex concept that contains more than one element, i.e. it consists of the Source, Trajectory and Endpoint of motion (FROM xx, VIA xx, TO xx; Jackendoff, 1983). In a broader sense, the Location of

**CONTACT** Yiyun Liao  liao@essb.eur.nl

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/23273798.2019.1676455>.

© 2019 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-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

motion (AT xx) also belongs to the path of motion (Talmy, 2000). The path of motion represents the “core schema” of a motion event (Talmy, 2000), and it has been reported that endpoints in boundary-crossing events, events in which a spatial endpoint is reached by an entity in motion, e.g. a car entering a garage, are universally salient and prioritised over other types of path information in motion encoding (in particular, the source; Lakusta & Landau, 2005; Papafragou, 2010; Regier & Zheng, 2007; Stefanowitsch & Rohlde, 2004). This “goal-bias” is in line with Event Segmentation theory (Zacks, Speer, Swallow, Braver, & Reynolds, 2007), which proposes that people perceive event boundaries when change in an event accumulates, as with a change of location when an endpoint is reached in a boundary-crossing event. Event boundaries are an important anchoring point for people’s representations of events, and they are likely to be part of what people consider a “reportable” event when asked to describe events (e.g. Gerwien & Stutterheim, 2018). Nevertheless, cross-linguistic differences have been reported in relation to the encoding of the path of motion, and endpoints specifically. For example, speakers of languages from different typological families, showing variation in whether the path of motion is typically described in the verb root (verb-framed languages, e.g. the path verb *salió* “exited”- in the Spanish sentence *la botella salió de la cueva flotando* “the ball exited the cave floating”), or outside of the verb in satellites (satellite-framed languages, e.g. verb particles or prefixes; the particle *out* in the English sentence *a boy is walking out of a house*; Talmy, 2000), distribute their attention differently when viewing and describing boundary-crossing events (Papafragou, Hulbert, & Trueswell, 2008), and they also memorise motion events in a distinct way (Gennari, Sloman, Malt, & Fitch, 2002).

Complementary to Talmy’s typology, Carroll, Weimar, Flecken, Lambert, and von Stutterheim (2012) and Flecken, Carroll, Weimar, and Von Stutterheim (2015) proposed that speakers of verb-framed vs satellite-framed languages differ in their conceptualisation of the path of motion, *beyond* how endpoints are encoded: They observed that, in verb-framed languages, when stimuli depicted motion only oriented *towards* an endpoint (critically, endpoints that were not actually reached), speakers (of French, in this case) construed the motion events by focusing on the location of the entity in motion (“a woman walks *on the road*”). They argued that the spatial concepts to conceptualise motion paths in such languages are in fact derived from the entity in motion; this is evidenced through the abundance of path verbs that exist that say something about the orientation of the entity in space, and

its proximity and distance towards a spatial endpoint (e.g. the French path verbs *se diriger vers* “to head towards”, *s’approcher* “to approach”, *s’avancer vers* “to approach towards”). When there is no evident endpoint in an event, speakers quite frequently use manner verbs, while maintaining the focus on the entity. Hence, the motion event is conceptualised as a property of the moving entity, rather than as motion directed along a path, towards a potential endpoint. Speakers of satellite-framed languages, on the other hand, typically construe motion events with a focus on features of the ground, combining manner verbs with adjuncts or particles describing endpoints (e.g. *to a building, into a building*) or trajectory-information (e.g. *along the river, down the street*), rather than locations (e.g. *on the street*).

In addition, research has studied cross-linguistic differences in the conceptualisation of motion paths in relation to grammatical aspect. It was found that for motion events that show orientation towards an endpoint that is not reached, speakers of languages with grammaticalized markers of imperfective or progressive aspect (Modern Standard Arabic, English, Russian; e.g. *-ing* in English) tended to focus on the ongoing phase, hence, the trajectory (e.g. *a person walking along a road*) of a motion event, thereby defocusing the potential endpoint of the event. On the other hand, speakers of languages (e.g. Dutch and German) that lack progressive aspect preferred to take a holistic view on an event, with a higher likelihood of including the event’s endpoint in a description (e.g. *a woman walking to a house*). For boundary-crossing events (from now on referred to as Endpoint-reached events), speakers encoded endpoints regardless of language background (Athanasopoulos & Bylund, 2013; von Stutterheim, Andermann, Carroll, Flecken, & Schmiedtová, 2012).

There is one recent study that takes into account these two types of typological features and studies both Endpoint-oriented and Endpoint-reached motion events (Georgakopoulos, Härtl, & Sioupi, 2019). They compared three languages, English, German, and Greek, in which English and German are both satellite-framed languages, but which are not in the same group of aspectual/non-aspectual languages (English is an aspectual language, German does not have grammatical markers to express aspect). English and Greek are both aspectual languages but not in the same group of satellite-/verb-framed languages (English is satellite-framed, Greek is verb-framed). Their analysis of motion conceptualisation for Endpoint-oriented events reports a similarly frequent encoding of endpoints in German compared to English, whereas both German and English speakers mentioned endpoints significantly more often than Greek speakers (in total 94 endpoint

expressions in Endpoint-oriented events: German  $N = 42$ ; English  $N = 39$ ; Greek  $N = 13$ ). Hence, Georgakopoulos et al. (2019) speculate that verb lexicalisation patterns have a stronger impact on speaker's motion conceptualisation than grammatical aspect, and that speakers of satellite-framed languages are more likely to conceptualise motion events in terms of their endpoint than speakers of verb-framed languages. This study underlines the importance and necessity of combining the two typological features, i.e. verb lexicalisation patterns and aspectual systems, in a cross-linguistic comparison of motion event conceptualisation. This is the approach that we follow in the present study.

Here, we study how speakers of Dutch and Mandarin Chinese, languages that differ in terms of verb lexicalisation patterns as well as grammatical aspect, conceptualise the path of motion in both Endpoint-reached and Endpoint-oriented (not-reached) events. Concerning verb semantics, Dutch is a typical satellite-framed language, whereas Chinese is often characterised as sharing features of a satellite-framed and a verb-framed language (Beavers et al., 2010; Ji & Hohenstein, 2017; see detailed description below). Regarding aspect, Chinese is an aspectual language, in which markers of both the progressive as well as the perfective aspect are frequently used (Klein, Li, & Hendriks, 2000; Li & Thompson, 1981; Xiao & McEnery, 2004). Dutch encodes progressive aspect to some extent as well, though it is not considered to be an aspectual language (Flecken, 2011). We are thus comparing languages that differ typologically, in complex ways, making it an interesting test case for the study of motion conceptualisation. We study (1) how the path of motion is *conceptualised* in language production, i.e. what element of the path of motion in a visually depicted event do people select for verbalisation? That is, do they refer to the (potential) Endpoint, Trajectory or Location, as shown in videos of Endpoint-reached and Endpoint-oriented events? During conceptualisation speakers construct a so-called "message" of the event, which contains the core content of what they are going to say (before the retrieval of the actual words); at this stage the process of information *selection* happens, involving the selection of the component(s) of the path of motion that the utterance is centred on (Levelt, 1989). We also analyse (2) how the path of motion is *described*, in terms of the linguistic means used: what verbs and adjunct types are used and how are they combined in descriptions of events of the two types? In addition, to specifically investigate the relation between verb semantics, aspect and endpoint conceptualisation, we analyse (3) to what extent the available aspectual markers in Chinese are combined with

different verb types (manner verbs, path verbs, serial verb constructions, see below) when speakers choose to mention endpoints. Besides their language production patterns, we are interested in participants' memory of the endpoints of motion events. As another window on potential endpoint encoding differences across the two languages, we administered a surprise post-verbalisation event memory task, in which participants were tested on their memory representation of the endpoints in Endpoint-oriented events.

We hypothesise, first of all for Endpoint-reached events, that speakers of both languages will most frequently select the endpoint for verbalisation, given the saliency of goals and boundaries that are reached or crossed (Athanasopoulos & Bylund, 2013; Papafragou, 2010; von Stutterheim et al., 2012). Second of all, for Endpoint-oriented events, Chinese and Dutch speakers may show differences in the frequency of endpoint mentioning, given the differences in the aspectual systems and in verb lexicalisation patterns: Dutch is hypothesised to conceptualise events in which endpoints have to be inferred as more goal-oriented than Chinese speakers. Moreover, Chinese and Dutch speakers will show differences in terms of the frequency of selection of trajectory vs location information for motion construal, as Dutch is a typical satellite-framed language, whereas Chinese exhibits features of a verb-framed language. We hypothesise that Dutch speakers, using manner verbs predominantly, focus more on features of the trajectory traced, compared to Chinese speakers, who in turn will more often concentrate on the location of the moving entity in space. Regarding the second question, we expect that, in Endpoint-reached events, speakers of Dutch and speakers of Chinese will adopt different ways of describing endpoints: Chinese predominantly uses serial verb constructions (see details in section 2), whereas path verbs and manner verbs plus satellites are possible options. Dutch, on the other hand, predominantly uses satellites to describe endpoints (in combination with manner verbs). As for Endpoint-oriented events, Dutch will again follow a typical satellite-framed pattern, in which manner verbs are combined with satellites encoding path information. It is unclear what patterns Chinese will exhibit exactly because no previous studies have systematically investigated the conceptualisation of Endpoint-oriented events in Chinese. We can expect satellite-framed patterns with manner verbs and path satellites, as well as serial verb constructions, or single path verbs. Concerning the use of aspect in Chinese endpoint descriptions, we expect the perfective aspectual marker *le* to frequently be combined with path verbs and serial verb constructions (see Li & Thompson, 1981). Our line of reasoning is that path

verbs and serial verb constructions often contain information on endpoints of motion, and a combination with the perfective aspect highlights and asserts the completion of the action – in this case the act of motion –, and with that the realization of the boundary crossing (endpoint reached) in these events (see Klein et al., 2000). Furthermore, we expect the progressive aspect (*zai*; Xiao & McEnery, 2004) to be combined with manner verbs and satellite constructions, predominantly, highlighting the ongoingness of an activity (manner verbs highlight the manner of action, and do not make explicit reference to endpoints). However, we do not exclude the possibility that perfective aspect is used with manner verbs (see e.g. the use of the perfective *le* to describe one-state situations in Klein et al., 2000), and that progressive aspect is used with path verbs (e.g. *zai guo malu* “PROG cross street”) and serial verb constructions (e.g. *zai chao [...] zou-qu* “PROG towards [...] walk-go”). Meanwhile, we should be aware of the fact that although Chinese is classified as an aspectual language, both aspect markers under investigation (perfective *le* and progressive *zai*) are not obligatory and aspectual meaning can also be obtained through context (see details in section 2).

Regarding the memory task: Previous studies have reported that the use of language in a verbal event encoding task can influence subsequent memory of the events (Athanasopoulos & Bylund, 2013; Papafragou & Selimis, 2010). Specifically, overt verbal encoding requires attention to the event elements to be mentioned, which in turn enhances the likelihood that this information is committed to a memory representation of the event. Therefore, we expect participants’ memory of endpoints to reflect their verbalisation patterns, that is, if Dutch participants displayed more mentions of the endpoints in endpoint-oriented events than Chinese participants, they should show enhanced memory of endpoints, and/or speeded judgements on this task (reflected in RTs).

## 2. Comparing Dutch and mandarin Chinese

### 2.1. Verb lexicalisation patterns

Following Talmy’s typology, Dutch is a typical satellite-framed language with a rich vocabulary of manner verbs. The manner of motion is usually conveyed through the verb root while the path of motion is expressed outside of the verb root, through particles or prepositions (e.g. *uitlopen* “walk out of”, *rijden naar* “drive to”). Path verbs (e.g. *arriveren* “arrive”, *oversteken* “cross”) are also used to express the path of motion, but the language lacks a wide variety of those (Slobin, 2004; Talmy, 2000).

Chinese presents a more complex case. It makes use of serial verb constructions in which two or more verbs appear together in a simple sentence, for example, the verbal construction *zou-guo* in *ta zou-guo gongyuan* “he is walking across a park”. The literal translation of *zou-guo* in this sentence is “walk-cross” in English. The first element *zou* is a manner verb that means “walking” and the second element *guo* can be used as a path verb referring to the trajectory of the motion in isolation. A central question in the abundant discussions concerning Chinese in this typology is the status of the second element in a serial verb construction, i.e. it is debated whether it is the main verb or just a verb complement (Chen & Guo, 2009; Kan, 2010; Lamarre, 2005; Liu, 2014; Shen, 2003; Slobin, 2004; Tai, 2003; Tai & Su, 2013; Talmy, 2000; Xu, 2013). Talmy considered the second element to be a verb complement and classified Chinese as a typical satellite-framed language (see similar opinion in Lamarre, 2005; Liu, 2014; Shen, 2003). Tai (2003) and Tai and Su (2013), on the other hand, considered the second element, that often represents the “result” of an action, as the main verb and claimed that Chinese is a verb-framed language. Slobin (2004), however, proposed that in serial verb languages, the manner verb is on a par with the (second) directional verb in semantic and syntactic prominence, and therefore Chinese should belong to a third language type: an equipollently-framed language (hence an E-language) (see similar opinion in Chen & Guo, 2009; Kan, 2010; Xu, 2013). According to Slobin (2004), E-languages express both manner and path in “equipollent” elements that are equal in formal linguistic terms and significance. It is in addition worth mentioning that Chinese also makes use of the typical satellite-framed pattern with manner verbs followed by directional prepositions, for example, *zou-xiang* “walk-towards” in motion event descriptions. In addition, it also uses single path verbs to express direction of motion (e.g. *shang* “ascend”, *xia* “descend”, *qian-jin* “approach”) and the number of path verbs in Chinese is larger than in a typical satellite-framed language, such as English (Xu, 2013). Ji et al. (2011) compared English and Chinese speakers in a caused motion event description task (e.g. a boy pushing a suitcase down the hill), and they discovered that in Chinese, serial verb constructions (e.g. *tui-shang* “pull-ascend”, *gun-xia* “roll-descend”) were used most often (70%), while single path verbs (e.g. *shang* “ascend”, *guo* “cross”, *jin* “enter”) were used around 30% of the time. Although satellite-framed patterns are also an option to express the direction of motion in caused motion events (e.g. *ba xiangzi tui-xiang dongxue* “BA suitcase **push-towards** cave”), they were not found in this description task. This is likely caused



by the fact that the videos used in that study all showed boundary-crossing events, in which goals are prominent and other elements of the path of motion less so. Interestingly, the paper reports that Chinese exhibits both satellite- (e.g. English) and verb-framing (e.g. French) properties, regardless of whether the second element in the verb compound was identified as a verb or a satellite. Specifically, its satellite properties were mainly shown in the frequently used BA construction (42%) combined with main verbs encoding manner of action as in “push” (e.g. *ba xiaochē tui-xia shānpō* “BA car push-descend/down hill”), just like in English (e.g. *push the car down the hill*). Meanwhile, its verb-framing properties were visible from the fact that Path information was frequently encoded in a single path verb (30%) while manner components were encoded in a subordinated ZHE clause (e.g. *ta lā zhē yī-liàng yīngérchē guō jiē* “he pull ZHE (pulling) one pram cross street”). This is a pattern that can also be observed in, for example, French, a typical verb-framed language (e.g. *il traverse la rue en tirant la poussette* “he is crossing the street pulling the pram”) (Hickmann & Hendriks, 2010).

Based on the observations discussed above, we adopt the view in Ji et al. (2011) that this language is of a mixed type, containing features of satellite-framed languages and verb-framed languages (see similar views in Beavers et al., 2010; Ji & Hohenstein, 2017; see also Shi & Wu, 2014, claiming that historically Chinese was a typical verb-framed language which is now in the process of transforming into a satellite-framed language).

## 2.2. The aspectual systems

The expression of grammatical aspect in Dutch and in Chinese also exhibits cross-linguistic differences. Chinese has a progressive aspectual marker *zai* (e.g. *ta zai zōu xiāng tūshūguān* “he is walking towards a library”) and a perfective aspectual marker *le* (e.g. *ta zōu xiāng le tūshūguān* “he walked towards a library”) (Klein et al., 2000; Li & Thompson, 1981; Smith, 1991; Xiao & McEnery, 2004). As a progressive aspectual marker, *zai* is similar to the English progressive marker *-ing*. They both offer us a viewpoint on the internal temporal structure or contour of an event. Like a magnifying glass or a spotlight, they allow us to focus on the intermediate ongoing phases of a situation, leaving the initial part and the final point of the situation unspecified. Conversely, as a perfective marker, *le* builds up an external viewpoint that enables us to view the situation as a whole or as a completed event from an outside perspective; thus, the internal structure of the situation is defocused. The progressive marker *zai* can be used to describe goal-oriented/directed motion (as

in *ta zai zōu xiāng tūshūguān* “he is walking towards a library”). The perfective *le* can be used to describe goal-reached motion (as in *ta zōu-jìn le tūshūguān* “he walked into a library”). However, it should be noted that unlike English, in which tense and aspect are obligatorily marked on the verb, Chinese is more flexible in using the available aspectual markers. Temporal information can often be contextually inferred (e.g. the interpretation of the following example sentence without any aspectual markers *ta zōu xiāng tūshūguān* “he walks towards a library” is progressive by default; the sentence *ta zōu-jìn tūshūguān* “he walks into a library” is perfective by default; see Bohnemeyer & Swift, 2004 for an explanation of the relation between default aspect and the telicity of a predicate). Hence, the aspectual markers are not obligatorily used in Chinese. In Dutch, the *aan het*-construction is used to express progressive aspect (Flecken, 2011). However, it is rarely used to express directed motion (see \**Oscar is naar de bibliotheek aan het lopen* “he is walking towards the library”). There is no designated grammatical marker of perfective aspect in Dutch.

To summarise our characterisation, Chinese exhibits features of both satellite-framed and verb-framed languages. Dutch, on the other hand, is a typical satellite-framed language. In terms of aspect, Chinese is an aspectual language, with markers encoding an aspectual opposition (the progressive *zai* and the perfective *le*), whereas Dutch mainly encodes progressive aspect for activities that are atelic (e.g. *Max is aan het wandelen* “Max is taking a stroll”). Considering these differences, Dutch and Chinese provide an intriguing test case for shedding light on the cross-linguistic comparison of the path information encoding.

## 3. Experiment

### 3.1. Method

The experiment consisted of three parts and was conducted in the Erasmus Behavioural Lab, Erasmus University Rotterdam. Participants first performed an event description task without being informed in advance of the subsequent memory task. Next, they completed a surprise memory task, which was presented on the screen. Finally, the participants completed a linguistic-background questionnaire on paper.

### 3.2. Participants

Sixty-one participants (30 native speakers of Dutch and 31 native speakers of Chinese) participated in the experiment. Each participant described 20 video clips in one

sentence and thus each participant created 20 sentences. The participants of the two language groups were from educational backgrounds of college level or above (with one exception in the Chinese group who was a high school student). The Dutch participants were first or second-year bachelor students from the Department of Psychology, Education, and Child Studies at the Erasmus University Rotterdam with a mean age of 19.63 (Range 18–24 years old;  $SD=1.45$  years, 28 females and 2 males). They were all born in the Netherlands and were Dutch native speakers. Chinese participants were students (high school, bachelors, masters or Ph.D. level) in the Netherlands with a mean age of 24.52 (Range 17–44 years old;  $SD=4.51$  years old; 24 females and 7 males). They were all born in China and were native mandarin Chinese speakers. The average time they had been residing in the Netherlands was 23.1 months (Range 1–96 months;  $SD=28.60$  months). The majority of the Chinese participants did not speak any Dutch (29 out of 31). Two Chinese participants had learned Dutch for an average duration of 3.25 years, but they did not speak Dutch with their family or friends. Participants from both language groups had learned English more than 10 years before the time of testing (Chinese: Range 8–18 years,  $M=14.4$  years,  $SD=2.30$  years; Dutch: Range 5–19 years,  $M=10.6$  years,  $SD=3.28$  years). This, however, is inevitable since English is used in the school curriculum in both countries. Most of the participants spoke English at school. Instead, they used their native languages with their family and friends, and to think, express emotion, talk to themselves and dream (self-report). Overall, participants from both language groups were representative of typical speakers of their native languages. All experimental instructions were provided in writing, in the participant's native language, in order to provide a monolingual experimental environment. All participants received research credits or a monetary reward for their participation. Event description data from one Dutch participant and one Chinese participant were excluded due to over 30% incomplete recordings (technical failures), leaving a final sample of 29 Dutch and 30 Chinese participants in the analyses of the linguistic data.

### 3.3. Materials

The experiment was programmed using the E-Prime 2.0 software (Psychology Software Tools, Inc., Pittsburgh, PA, USA). The items of the event description task consisted of 40 video clips that were filmed and edited by von Stutterheim and colleagues at Heidelberg University. The items were used in similar studies, such as Athanasopoulos & Bylund, 2013; von Stutterheim et al., 2012; and

Flecken, Carroll, & von Stutterheim, 2014. The video clips showed real-life events, each of 6 s in length. There was a blue screen with a centred fixation cross in between each item. Participants were instructed that they could start to describe each video clip when they had recognised “what was happening” in the video. Descriptions were recorded with an external voice recorder. The blue screen between each video was shown for 8 s, leaving ample time for participants to verbalise their description. The stimulus set contained 20 motion events that can be classified into two types (10 each):

*Endpoint-oriented (not reached) events:* motion events that displayed an entity in motion (a vehicle or person) along a specific trajectory in the direction of a visible Endpoint location (e.g. a village, church, playground), which was crucially not reached by the end of the 6 s video clip (see an example of this event in Figure 1).

*Endpoint-reached events:* motion events in which the entity in motion was depicted as reaching a goal or destination (e.g. walking into a church; driving into a garage; see an example of this event in Figure 2).

In addition, there were 20 fillers that were not motion events; the videos showed either static scenes (e.g. a bicycle parked at a lamppost; a dog sitting and panting on the grass) or causative event scenes (e.g. a woman knitting a scarf; a man folding a paper airplane).

Materials for the memory test consisted of pictures of the 10 Endpoint-oriented motion events and 6 fillers. All pictures were screenshots from the previously seen videos. Critical pictures were screenshots taken from the 10 Endpoint-oriented video clips. The pictures were manipulated, such that the visible Endpoints of 6 of the items were removed using Photoshop, leaving a natural scene (see Figure 3). The other 4 items were left unchanged. The 6 fillers were screenshots from filler videos, amongst which 3 had certain objects removed that had appeared in the videos previously (e.g. a bicycle in the static scene of it being parked at a lamp post, cosmetics in a video of a woman putting on make-up); the other 3 were left unchanged. The order of the videos in the elicitation task and the pictures in the memory task were randomised across participants. This was done to cancel out potential distance effects.

### 3.4. Procedure

For the event description task participants were seated in front of a computer in a quiet room and were asked to read the following instructions on the screen (von Stutterheim et al., 2012):

You will see a set of 40 video clips showing everyday events that are not in any way connected to each other. Each clip lasts 6 seconds. Before each clip starts,



**Figure 1.** Screenshot of an Endpoint-oriented motion event: a car driving on a road towards a village/houses.

a blue screen with a white fixation cross will appear. Please focus on this fixation cross. Your task is to tell “what is happening” in each video clip, using a complete sentence. You may begin to speak as soon as you recognize what is happening in the clip. It is not necessary to describe the video clips in detail (e.g., “the sky is blue”). Please focus on the event that is happening only.

After reading the instructions, participants were told to stay close to the recorder that was placed in front of them on the table. They could press the spacebar to start the task if they did not have any further questions for the experimenter. Their descriptions were recorded with a recorder connected with E-prime. Recordings



**Figure 2.** Screenshot of an Endpoint-reached motion event: a man walking into a church.





**Figure 3.** Example of an item in the memory task: Endpoint-oriented motion event with potential endpoint removed.

were automatically saved as .wav files on the computer. Each session took about 20 min. After the last video clip had ended, participants saw a screen directly instructing them on the memory task:

Now, you will see screenshots of some of the videos you saw earlier. Please decide as quickly as you can whether the picture shown on the screen is exactly the same as what you saw in the video earlier. Press YES or NO on the button box in front of you. Note: some of these screenshots were directly taken from the previous videos, but some are not. Please observe carefully and make your judgment quickly.

Participants pressed the spacebar to start the experiment if they did not have further questions. After making a decision, the experiment proceeded to the next trial immediately. Accuracy scores and reaction times were logged for analyses. Each session took around 5–7 min. After this part, participants filled in a language background questionnaire. The experimental procedure took around 40 min in total. All the instructions were presented in the participants' native language (Dutch or Chinese).

#### 4. Data coding

Both Dutch and Chinese recordings were transcribed by native speakers. Incomplete or missing recordings of sentences in both languages were coded as missing values and excluded from our statistical analyses

(0.021% out of 1170 sentences in total, 0.006% in the Dutch data and 0.015% in the Chinese data). This resulted in a total of 573 data points in Dutch (287 in the Endpoint-oriented event type, and 286 in the Endpoint-reached event type) and 583 data points in Chinese (289 in the Endpoint-oriented event type and 294 in the Endpoint-reached event type).

Data were coded following the coding scheme elaborated below. For each coding category, its presence was coded as “1”, otherwise a “0” was entered in the relevant column (binary data coding). Both Chinese and Dutch data were coded by a native speaker and a second researcher, independent of one another. Points of disagreement were discussed and in most cases resolved.<sup>1</sup>

##### 4.1. Path of motion (endpoint, trajectory, and location-only)

We first coded Path information in the motion event descriptions in both languages, distinguishing, first of all, utterances that included reference to an **Endpoint** object (irrespective of whether the endpoint was described as reached or not, e.g. the *house* mentioned in *into a house*, the *playground* referred to in *to(wards) a playground*), regardless of additional, other types of path information mentioned in the same utterance (e.g. *walk [across the road] towards a car*).<sup>2</sup> Utterances encoding **Trajectory** were sentences containing trajectory information (and no endpoint information),

irrespective of whether in addition location information was mentioned (e.g. *over een weg [op het platteland]/ yan-zhe yi-tiao xiaolu [zai jiaowai]* “along a road [in the country side]”). Lastly the category **Location-only** included references to a location as the only path element (e.g. *op een weg/zai lu shang* “on a road”). In our videos, there were no obvious source locations. Hence, source information was not considered in the current study. The coding scheme reflects the differences in the viewpoints that people can take during conceptualisation: a maximal, holistic viewpoint, including an (inferred) Endpoint and potentially other path elements, to a minimal viewpoint that only locates the entity in motion in space. Examples of a maximal (holistic) viewpoint in Dutch and in Chinese are *hij loopt over een weg op het platteland naar een kerk* and *zai jiao wai ta yan-zhe yi-tiao xiaolu zou xiang yi-ge jiaotang*, respectively, “he walks along a road in the countryside to a church”. Examples of a minimal viewpoint, only locating the entity in motion in space, in Dutch and in Chinese are *hij loopt op straat* and *ta zou zai lu shang*, respectively, “he walks on a road”.

#### 4.2. Verb-adjunct combinations

We then coded the types of verbs used and the combinations of verb types and adjunct types, only considering the proportion of sentences containing descriptions of path of motion in both languages (0.87 vs 0.86: Chinese (509/583) vs Dutch (491/573)). In Chinese, we coded for three types of verbs, including path verbs, manner verbs and serial verb constructions, whereas in Dutch we coded the former two types. The path verbs<sup>3</sup> coded were those which were used independently as predicates and which encoded path information in a main clause. Examples of Chinese path verbs are *jìn* “enter”, *guo* “cross”, *jing-guo* “pass-by”, *qian-jin* “approach”, *shang* “ascend”, *dao* “arrive”, etc. Examples of Dutch path verbs are *arriveren*, “arrive”, *oversteken* “cross”, etc. Manner verbs were the only verb in a main clause, encoding the manner of a moving Figure, such as the gait or the speed of motion. Examples of manner verbs in both languages are *kai/rijden* “drive”, *zou/lopen* “walk”, *pao/rennen* “run”, and *pa/beklimmen* “climb”. A serial verb construction<sup>4</sup> is a special verbal construction that exists in Chinese but not in Dutch. Examples of such verbal constructions are *kai-jin* “drive-enter”, *kai-guo* “drive-pass”, *zou-jin* “walk-enter”, *zou-qu* “walk-go” and *pa-shang* “climb-ascend”. Taking all possible combinations into account, we coded the data regarding the following categories: the combination of manner verb with either Endpoint adjuncts, Location-only adjuncts, or Trajectory adjuncts (MaEnd, MaLoc,

MaTra in figures), the combination of path verbs with either Endpoint adjuncts, Location-only adjuncts, or Trajectory adjuncts (PaEnd, PaLoc, PaTra in figures), and the combination of serial verb constructions with either Endpoint adjuncts, Location-only adjuncts, or Trajectory adjuncts (SvcEnd, SvcLoc, SvcTra in figures).

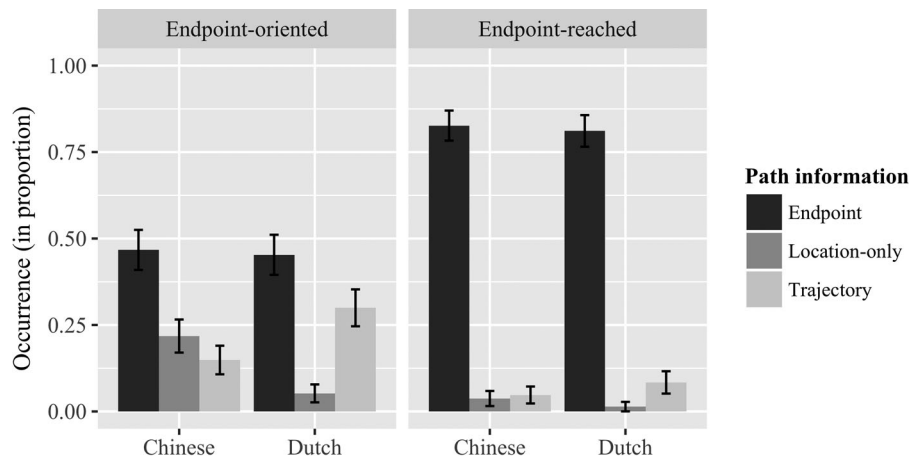
#### 4.3 Verb type-aspect combinations when describing endpoints

In addition, we coded the aspect markers (progressive and perfective) that occurred in the Chinese Endpoint description data (total number of utterances containing Endpoint: 378/583, 243 in the Endpoint-reached events and 135 in the Endpoint-oriented events). We coded *zai* in Chinese as the progressive, and *le* in Chinese as the perfective. We were interested in the extent to which the aspect markers were combined with the different verb types in the Chinese Endpoint descriptions. We thus counted the occurrence of manner verbs with either the perfective marker or the progressive marker (MaPerf, MaProg in figures), path verbs with the perfective marker or the progressive marker (PaPerf, PaProg in figures), and serial verb constructions with either the perfective or the progressive (SvcPerf, SvcProg in figures). We also counted the use of each verb type when no aspect markers were used (MaOnly, PaOnly, SvcOnly in figures).

### 5. Analysis and results<sup>5</sup>

#### 5.1. Conceptualising the path of motion

To test the effects of Language and Event type on the encoding of each path element, we set up separate mixed-effect binomial logistic regression models for each Path type in R<sup>6</sup> (R Core Team, 2016) using the *glmer* function implemented in the package *lme4* (Bates, Mächler, Bolker, & Walker, 2015). We included subjects and video clips (stimulus items) in our model as random intercepts. Both Language and Event type (fixed factors) were sum coded. The dependent variable in each of the models was the respective path element mentioned yes (1) or no (0). Figure 4 shows the proportion of occurrence of each type of path information (Endpoint, Location-only, and Trajectory) in both languages and for both event types (**Endpoint: Endpoint-oriented events:** Chinese  $N = 135/289$ , Range 0.20–0.80 vs Dutch  $N = 130/287$ , Range 0.10–0.90, **Endpoint-reached events:** Chinese  $N = 243/289$ , Range 0.50–1.00 vs Dutch  $N = 232/287$ , Range 0.50–1.00; **Location-only: Endpoint-oriented events:** Chinese  $N = 63/289$ , Range 0.00–0.70 vs Dutch  $N = 15/287$ , Range 0.00–0.38,



**Figure 4.** Selection of Path components (Endpoint, Location-only and Trajectory) in Chinese and Dutch utterances for Endpoint-oriented and Endpoint-reached events.

*Endpoint-reached events:* Chinese  $N = 11/289$ , Range 0.00–0.29 vs Dutch  $N = 4/287$ , Range 0.00–0.10; **Trajectory:** *Endpoint-oriented events:* Chinese  $N = 43/289$ , Range 0.00–0.63 vs Dutch  $N = 86/287$ , Range 0.00–0.60; *Endpoint-reached events:* Chinese  $N = 14/289$ , Range 0.00–0.20 vs Dutch  $N = 24/287$ , Range 0.00–0.30). See Table 1 below for examples of each path component in each language.

We found that the frequency of **Endpoints** was significantly different across event types, but not across languages (*Event type*  $\beta = -1.212$ ,  $SE = 0.376$ ,  $z = -3.222$ ,  $p = .001$ ; *Language*  $\beta = 0.054$ ,  $SE = 0.132$ ,  $z = 0.405$ ,  $p = .686$ ). There was no interaction between the two factors (*Event type\*Language*  $\beta = 0.006$ ,  $SE = 0.081$ ,  $z = 0.079$ ,  $p = .937$ ). Thus, endpoints were mentioned more frequently in Endpoint-reached events, than Endpoint-oriented events, regardless of language, as hypothesised. In terms of references to **Location-only**, there was a main effect of Language (*Language*  $\beta = 0.803$ ,  $SE = 0.262$ ,  $z =$

$3.066$ ,  $p < .01$ ) and a main effect of Event type ( $\beta = 1.141$ ,  $SE = 0.452$ ,  $z = 2.522$ ,  $p < .05$ ). No interaction between Language and Event type was found ( $\beta = 0.242$ ,  $SE = 0.183$ ,  $z = 1.326$ ,  $p = .185$ ). Dutch participants showed lower frequency of mentioning Location-only information than Chinese participants, as hypothesised. Moreover, Location-only references were obtained less frequently in Endpoint-reached events, than Endpoint-oriented events. Regarding the mentioning of **Trajectory** information, results show main effects of Language and Event type (*Language*  $\beta = -0.468$ ,  $SE = 0.152$ ,  $z = -3.068$ ,  $p < .01$ ; *Event type*  $\beta = 0.952$ ,  $SE = 0.365$ ,  $z = 2.610$ ,  $p < .01$ ). The interaction was not significant ( $\beta = -0.108$ ,  $SE = 0.111$ ,  $z = -0.967$ ,  $p = .334$ ). Dutch speakers mentioned Trajectory information more frequently compared to Chinese speakers; it was encoded less frequently when participants described Endpoint-reached events than when they described Endpoint-oriented events.

To shed light on the frequency of mentioning the three path elements within each of the two languages, we ran separate multinomial logistic regression models via the mlogit package (Croissant, 2018) in R for each language.<sup>7</sup> Event type (fixed factor) was dummy coded and Path type was the dependent variable.<sup>8</sup> First, in the **Chinese** data, there was a significant effect of Event type ( $\chi^2 = 85.699$ ,  $p < .001$ ): In Endpoint-oriented events<sup>9</sup>, Endpoints were mentioned significantly more often than both Location-only and Trajectory information (*Location-only vs Endpoint: Intercept*  $\beta = -0.762$ ,  $SE = 0.153$ ,  $z = -4.995$ ,  $p < .001$ ; *Trajectory vs Endpoint: Intercept*  $\beta = -1.144$ ,  $SE = 0.175$ ,  $z = -6.534$ ,  $p < .001$ , respectively). In the Endpoint-reached events [see note 9], Chinese speakers also mentioned Endpoint significantly more often than both Location-only and Trajectory, but the difference was larger in this event type than in Endpoint-oriented events (*Location-only vs*

**Table 1.** Examples of path component descriptions in Chinese and Dutch.

	Chinese	Dutch
Endpoint	<i>Liang-wei nvshi zou xiang yi-ge fangzi.</i> "two women walk towards a house" (Endpoint-oriented) <i>Yi-liang che kai-jin cheku.</i> "a car drive-enter garage" (Endpoint-reached)	<i>Twee vrouwen lopen naar een huis.</i> "two women walk to a house" (Endpoint-oriented) <i>De auto rijdt de garage in.</i> "the car drives the garage in" (Endpoint-reached)
Location-only	<i>Yi-liang baise kache kai zai xiangcun daolu shang.</i> "a white truck drive on village road"	<i>Er rijdt een auto op de weg.</i> "there drives a car on the road"
Trajectory	<i>Yi-liang che kai-guo le lumian.</i> "a car drive-cross PERF road"	<i>Er fietst een meisje over het fietspad.</i> "there cycles a girl along the cycle path"

*Endpoint: Intercept*  $\beta = -3.095$ ,  $SE = 0.308$ ,  $z = -10.041$ ,  $p < .001$  and *Trajectory vs Endpoint: Intercept*  $\beta = -2.854$ ,  $SE = 0.275$ ,  $z = -10.384$ ,  $p < .001$ , respectively). In the **Dutch** data, we also found a significant effect of Event type ( $\chi^2 = 71.283$ ,  $p < .001$ ). Similar to the Chinese group, speakers of Dutch also mentioned Endpoints significantly more often than Location-only and Trajectory information in both Endpoint-oriented events (*Location-only vs Endpoint: Intercept*  $\beta = -2.160$ ,  $SE = 0.273$ ,  $z = -7.919$ ,  $p < .001$ ; *Trajectory vs Endpoint: Intercept*  $\beta = -0.413$ ,  $SE = 0.139$ ,  $z = -2.973$ ,  $p < .01$ , respectively) and Endpoint-reached events (*Location-only vs Endpoint: Intercept*  $\beta = -4.060$ ,  $SE = 0.504$ ,  $z = -8.052$ ,  $p < .001$  and *Trajectory vs Endpoint: Intercept*  $\beta = -2.269$ ,  $SE = 0.214$ ,  $z = -10.580$ ,  $p < .001$ , respectively). The differences between the frequency of mentioning Endpoints vs Trajectory and Location-only information were larger in the latter compared to the former event type.

## 5.2. Verb and adjunct types used

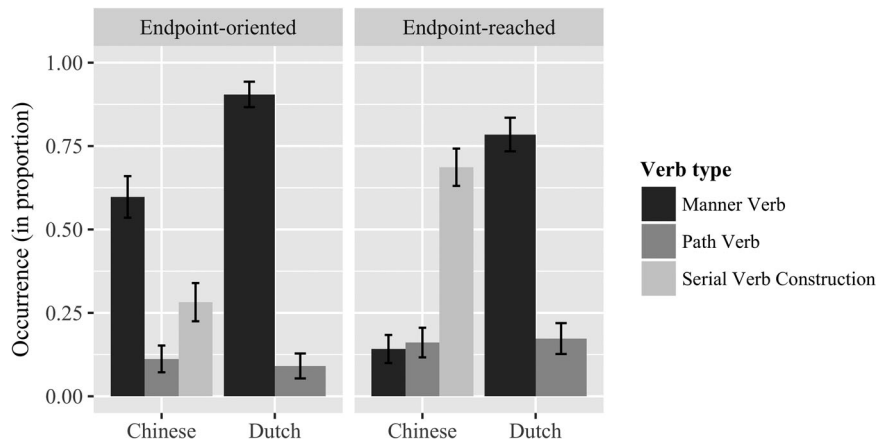
We again analysed the effects of Language and Event type on the use of each verb type with binomial mixed effect logistic regression models.<sup>10</sup> Subjects and video clips (stimulus items) were included in each model as random intercepts. Language and Event type were sum coded. The dependent variable in each of the models was the respective Verb type used (Manner verbs, Path verbs), yes (1) or no (0). Serial verb constructions were only used in Chinese and thus not compared cross-linguistically. Figure 5 shows the verb types used in relation to the two event types, in both languages, only considering sentences containing descriptions of path of motion (**Manner verbs: Endpoint-oriented events:** Chinese  $N = 144/241$ , Range in proportion 0.00–1.00 vs Dutch  $N = 209/231$ , Range 0.56–1.00, **Endpoint-reached events:** Chinese  $N = 38/268$ , Range 0.00–0.57 vs Dutch  $N = 204/260$ , Range 0.20–1.00; **Path verbs: Endpoint-oriented events:** Chinese  $N = 27/241$ , Range 0.00–0.83 vs Dutch  $N = 21/231$ , Range 0.00–0.44, **Endpoint-reached events:** Chinese  $N = 42/268$ , Range 0.00–0.56 vs Dutch  $N = 45/260$ , Range 0.00–0.70; **Serial verb constructions: Endpoint-oriented events:** Chinese  $N = 68/241$ , Range 0.00–0.86; **Endpoint-reached events:** Chinese  $N = 184/268$ , Range 0.20–1.00).

For **manner verbs**, there was a significant main effect of Language (*Language*  $\beta = -1.388$ ,  $SE = 0.157$ ,  $z = -8.862$ ,  $p < .001$ ) and Event type ( $\beta = 0.725$ ,  $SE = 0.181$ ,  $z = 4.007$ ,  $p < .001$ ). The interaction was also significant ( $\beta = 0.415$ ,  $SE = 0.085$ ,  $z = 4.899$ ,  $p < .001$ ). Dutch speakers used manner verbs significantly more often than Chinese speakers, especially so in Endpoint-reached events, compared to Endpoint-oriented events. Moreover, speakers

of Chinese used manner verbs significantly more often to describe Endpoint-oriented events, than Endpoint-reached events. Speakers of Dutch did not differ in their use of manner verbs between event types. Regarding the use of **path verbs**, there was no main effect of Language (*Language*  $\beta = 0.084$ ,  $SE = 0.186$ ,  $z = 0.449$ ,  $p = .653$ ), nor of Event type ( $\beta = -0.300$ ,  $SE = 0.257$ ,  $z = -1.169$ ,  $p = .243$ ). The interaction was also not significant ( $\beta = 0.109$ ,  $SE = 0.097$ ,  $z = 1.124$ ,  $p = .261$ ). Speakers of the two language groups did not differ in their use of path verbs.

Comparing patterns closely within each language, two logistic regression models were set up.<sup>11</sup> First, in the **Chinese** group, we analysed the dependent variable “verb type” with 3 levels: manner verb, path verb and serial verb construction. We built a multinomial logistic regression model to statistically test the effect of Event type on the choice of verb type, and to discover the typical patterns used in each event type for the Chinese group. Event type was dummy coded. A significant effect of Event type was found ( $\chi^2 = 123.620$ ,  $p < .001$ ). Within Endpoint-reached events, Chinese speakers used serial verb constructions more often than both manner verbs and path verbs (*serial verb construction vs manner verb: Intercept*  $\beta = 1.577$ ,  $SE = 0.178$ ,  $z = 8.852$ ,  $p < .001$ ; *serial verb construction vs path verb: Intercept*  $\beta = 1.454$ ,  $SE = 0.169$ ,  $z = 8.583$ ,  $p < .001$ ). There was no difference in the proportion of use of path verbs and manner verbs (*path verb vs manner verb: Intercept*  $\beta = 0.124$ ,  $SE = 0.223$ ,  $z = 0.555$ ,  $p = .578$ ). When describing Endpoint-oriented events, Chinese speakers were more likely to use manner verbs than the other two types of verbs (*path verb vs manner verb: Intercept*  $\beta = -1.674$ ,  $SE = 0.210$ ,  $z = -7.982$ ,  $p < .001$ ; *serial verb construction vs manner verb: Intercept*  $\beta = -0.750$ ,  $SE = 0.147$ ,  $z = -5.099$ ,  $p < .001$ ). Moreover, serial verb constructions were used significantly more often than path verbs (*serial verb constructions vs path verbs: Intercept*  $\beta = 0.924$ ,  $SE = 0.228$ ,  $z = 4.061$ ,  $p < .001$ ). The binomial mixed-effect logistic regression model on **Dutch** data contained the dependent variable (verb type) with two levels (manner verb and path verb). Event type (fixed factor) was sum coded. Random effects included subjects and video clips (stimulus items). There was no effect of Event type on the choice of verb types among Dutch speakers (*Event type*  $\beta = -0.490$ ,  $SE = 0.359$ ,  $z = -1.366$ ,  $p = .172$ ). For both Endpoint-oriented events and Endpoint-reached events, Dutch speakers were significantly more likely to use manner verbs than path verbs (*Endpoint-oriented: manner verb vs path verb Intercept*  $\beta = 3.639$ ,  $SE = 0.697$ ,  $z = 5.220$ ,  $p < .001$ ; *Endpoint-reached: manner verb vs path verb Intercept*  $\beta = 2.658$ ,  $SE = 0.646$ ,  $z = 4.115$ ,  $p < .001$ ).



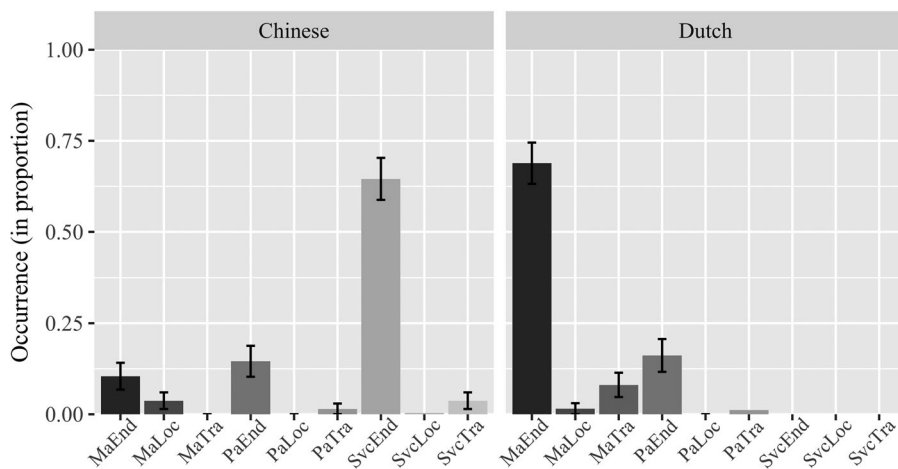


**Figure 5.** Verb types used by Chinese and Dutch participants in all Path descriptions in Endpoint-oriented and Endpoint-reached events: Manner verb, Path verb, Serial verb construction.

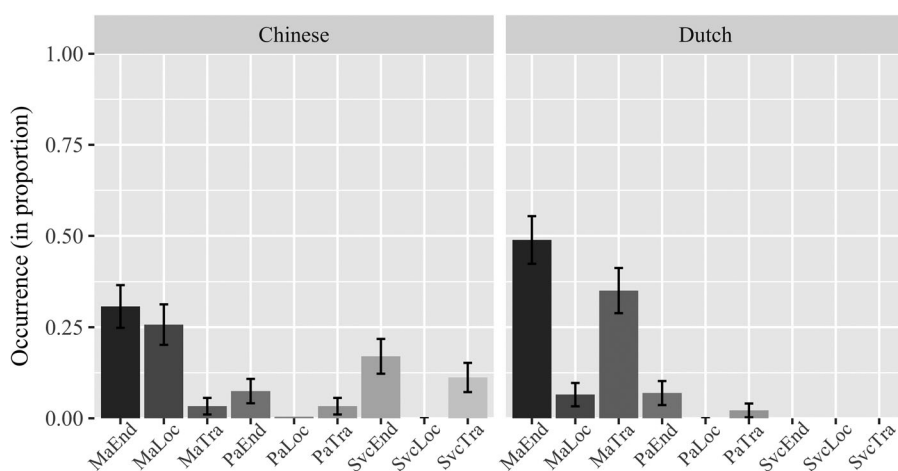
Binomial mixed-effect logistic regression models were built for each verb and adjunct combination that existed in the languages.<sup>12</sup> Language and Event type (fixed factors) were sum coded; the dependent variable in each of the models was the respective combination used, yes (1) or no (0). Subjects and video clips (stimulus items) were included in each model as random intercepts. Figures 6 and 7 show the combinations of verb and adjunct types found in descriptions in both languages, in Endpoint-reached events and Endpoint-oriented events, only considering sentences containing descriptions of path of motion. Typical examples of each verb and adjunct combination in each language are given in Table 2.

In terms of the combination of **manner verbs with Endpoint adjuncts**, there was a significant main effect of Language ( $Language \beta = -1.171, SE = 0.166, z = -7.053, p < .001$ ), but no effect of Event type ( $\beta = 0.052, SE = 0.257, z = 0.202, p = .840$ ). The interaction between

Language and Event type was significant ( $\beta = 0.695, SE = 0.091, z = 7.645, p < .001$ ). Speakers of Dutch used manner verbs with Endpoint adjuncts significantly more often than speakers of Chinese, especially so in the Endpoint-reached events. For **path verbs in combination with Endpoint adjuncts**, there was no significant effect of Language ( $Language \beta = 0.064, SE = 0.204, z = 0.315, p = .753$ ), nor Event type ( $\beta = -0.613, SE = 0.320, z = -1.916, p = .055$ ) and no interaction either ( $\beta = 0.085, SE = 0.119, z = 0.711, p = .477$ ). Thus, speakers of the two languages did not differ in their use of path verbs combined with Endpoint adjuncts. Turning to **manner verbs in combination with Location-only adjuncts**, we found a significant main effect of Language ( $Language \beta = 0.762, SE = 0.274, z = 2.781, p < .01$ ) and Event type ( $\beta = 1.142, SE = 0.449, z = 2.542, p < .05$ ). No interaction effect was detected ( $\beta = 0.281, SE = 0.185, z = 1.521, p = .128$ ). Chinese speakers used manner verbs plus Location-only adjuncts significantly more often



**Figure 6.** Verb and Path combinations used by Chinese and Dutch participants in encoding Path information in Endpoint-reached events: Manner verb (Ma), Path verb (Pa), Serial verb construction (Svc); Endpoint (End), Location (Loc), Trajectory (Tra).



**Figure 7.** Verb and Path combinations used by Chinese and Dutch participants in encoding Path information in Endpoint-oriented events: Manner verb (Ma), Path verb (Pa), Serial verb construction (Svc); Endpoint (End), Location (Loc), Trajectory (Tra).

than Dutch speakers, and this combination was used most frequently in Endpoint-oriented events. Location-only was not combined with path verbs in either language. Regarding the combination of **manner verbs with Trajectory adjuncts**, there was a main effect of Language ( $\beta = -1.698$ ,  $SE = 0.236$ ,  $z = -7.201$ ,  $p < .001$ ) and also a main effect of Event type ( $\beta = 1.098$ ,  $SE = 0.378$ ,  $z = 2.904$ ,  $p < .01$ ). Specifically, speakers of Dutch were more likely to use manner verbs plus Trajectory adjuncts than speakers of Chinese, and this combination was more frequent in descriptions of Endpoint-oriented events than Endpoint-reached events.

To take a closer look at the patterns within each language group separately, we built two multinomial logistic regression models, with Event type as the fixed factor (dummy coded), one for each language.<sup>13</sup> The dependent variable in each model was verb and adjunct combination. There was a significant effect of Event type in the **Chinese** group ( $\chi^2 = 178.02$ ,  $p < .001$ ). The most frequent patterns in descriptions of Endpoint-oriented events in Chinese were manner verb

plus Endpoint and manner verb plus Location-only (with a proportion of 0.31 [ $N = 74/241$ , Range 0.00–0.89] and 0.26 [ $N = 62/241$ , Range 0.00–0.70], respectively), and no significant difference in their proportion of use was found ( $\beta = -0.177$ ,  $SE = 0.172$ ,  $z = -1.028$ ,  $p = .304$ ). The most typical pattern in descriptions of Endpoint-reached events in Chinese was the combination of serial verb construction plus Endpoint (with a proportion of 0.65,  $N = 173/268$ , Range 0.20–1.00). There was also a significant effect of Event type in the **Dutch** group ( $\chi^2 = 71.341$ ,  $p < .001$ ). In Dutch, the most typical pattern in descriptions of Endpoint-reached events was manner verb plus Endpoint (with a proportion of 0.69,  $N = 179/260$ , Range 0.10–1.00). However, the most typical patterns in the descriptions of Endpoint-oriented events were manner verb plus Endpoint and manner verb plus Trajectory (with a proportion of 0.49 [ $N = 113/231$ , Range 0.12–1.00] and 0.357 [ $N = 81/231$ , Range 0.00–0.75], respectively). The former pattern occurred significantly more often than the latter one ( $\beta = -0.333$ ,  $SE = 0.146$ ,  $z = -2.287$ ,  $p < .05$ ).

**Table 2.** Examples of verb-path combinations in Chinese and Dutch.

	Chinese	Dutch
Manner verb + Endpoint (MaEnd)	<i>Yi-ge nanshi zou xiang yi-liang che.</i> "a man walk towards a car"	<i>Een hond rent naar een deur.</i> "a dog runs to a door"
Manner verb + Location (MaLoc)	<i>Liang-ge ren zou zai lu shang.</i> "two people walk on road"	<i>Een paardrijder op een paard loopt op een pad.</i> "A horse rider on a horse walks on a path"
Manner verb + Trajectory (MaTra)	<i>Yi-ge nanshi yan-zhe tizi zai pa.</i> "a man along-ZHE ladder PROG climb"	<i>Een man loopt over de straat.</i> "a man walks along the street"
Path verb + Endpoint (PaEnd)	<i>Yi-ge ren jin le yi-dong lou.</i> "a person enter PERF a building"	<i>Een kind gaat een speeltuin binnen.</i> "a child goes a playground inside"
Path verb + Trajectory (PaTra)	<i>Yi-ge zhongnian nanzi zai shang-lou.</i> "a middle-age man PROG ascend-stairs"	<i>Een man steekt de straat over.</i> "a man crosses the street"
Serial verb construction + Endpoint (SvcEnd)	<i>Yi-ge ren pao-jin yi-dong dalou.</i> "a person run-enter a building"	–
Serial verb construction + Trajectory (SvcTra)	<i>Yi-ge ren zou-guo yi-dong lou.</i> "a person walk-cross a building"	–

### 5.3. Verb and aspect markers used in Chinese descriptions of endpoints

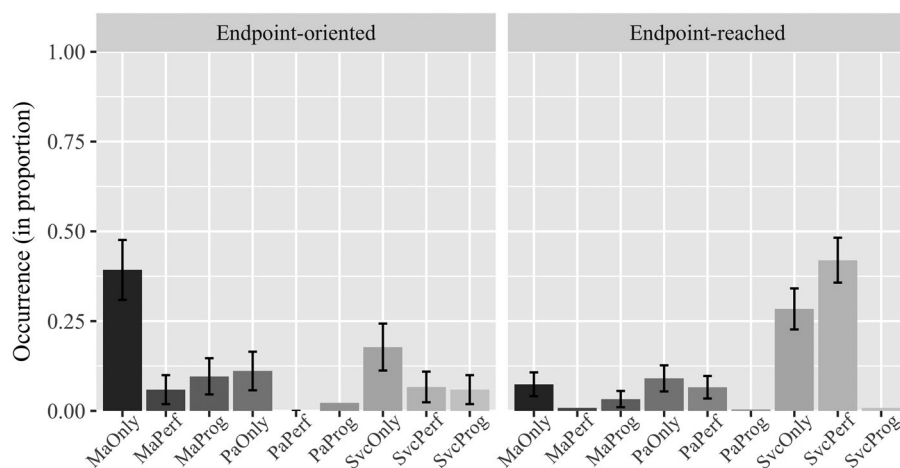
The overall proportion of use of progressive and perfective aspect in Chinese endpoint descriptions was 0.05 ( $N = 12/243$ , Range 0.00–0.40) and 0.49 ( $N = 120/243$ , Range 0.00–1.00), respectively, for Endpoint-reached events, and 0.18 ( $N = 24/135$ , Range 0.00–1.00) and 0.13 ( $N = 17/135$ , Range 0.00–0.50), respectively, for Endpoint-oriented events. We conducted multinomial logistic regression models<sup>14</sup> to test differences between event types regarding verb and aspect combinations (manner verb +/- aspect, path verb +/- aspect or serial verb construction +/- aspect). Event type was dummy coded. Figure 8 presents the frequency of all combinations of aspect markers and verb types in Chinese Endpoint descriptions, for the two event types. Typical examples of each verb and aspect combination are illustrated in Table 3.

We found a significant difference between event types ( $\chi^2 = 134.49$ ,  $p < .001$ ). For Endpoint-oriented events, the most frequent pattern in describing endpoints was manner verbs without aspect (with a proportion of 0.39,  $N = 53/135$ , Range 0.00–1.00). Serial verb constructions and path verbs (both without aspect) were the second and third most frequent patterns found (with a proportion of 0.18 [ $N = 24/135$ , Range 0.00–1.00] and 0.11 [ $N = 15/135$ , Range 0.00–0.75], respectively). Manner verbs (without aspect) were used significantly more often than serial verb constructions and path verbs (both without aspect; *serial verb construction vs manner verb* Intercept  $\beta = -0.792$ ,  $SE = 0.246$ ,  $z = -3.220$ ,  $p = .001$ ; *path verb vs manner verb* Intercept  $\beta = -1.262$ ,  $SE = 0.293$ ,  $z = -4.316$ ,  $p < .001$ ), while the latter two did not differ significantly (*serial verb*

*construction vs path verb* Intercept  $\beta = 0.470$ ,  $SE = 0.329$ ,  $z = 1.428$ ,  $p = .153$ ). The progressive marker was combined with manner verbs and satellites, as well as with serial verb constructions and path verbs but each only with a proportion of 0.10 ( $N = 13/135$ , Range 0.00–0.50), 0.06 ( $N = 8/135$ , Range 0.00–0.50), and 0.02 ( $N = 3/135$ , Range 0.00–0.50). In Endpoint-reached events, the most frequent patterns in encoding endpoints were serial verb constructions plus the perfective aspect and serial verb constructions without aspect (with a proportion of 0.42 [ $N = 102/243$ , Range 0.00–1.00] and 0.28 [ $N = 69/243$ , Range 0.00–0.86], respectively). The former one was used significantly more often than the latter one (*serial verb construction plus perfective vs serial verb construction* Intercept  $\beta = 0.391$ ,  $SE = 0.156$ ,  $z = 2.508$ ,  $p < .05$ ). The perfective marker was also found in combination with path verbs and manner verbs but only with a proportion of 0.07 ( $N = 16/243$ , Range 0.00–0.33) and 0.01 ( $N = 2/243$ , Range 0.00–0.13), respectively.

### 5.4. Memory task

For the surprise memory task, we first calculated each participant's d-prime score (detection sensitivity, Macmillan & Creelman, 1991). An independent  $t$  test was conducted on the d-prime scores. We found no significant difference in the d-prime scores between groups (Chinese: Range 0.28–4.65,  $M = 2.08$ ,  $SD = 1.23$ ; Dutch: Range 0.00–3.28,  $M = 2.10$ ,  $SD = 0.93$ ;  $t(59) = 0.051$ ,  $p = 0.960$ ). We then did an independent  $t$  test on the RTs of all accurate YES/NO responses. Zero to 3500 ms was determined as a reasonable cut-off based on the density of the RT data. RTs that were longer than 3500 ms were excluded from analyses. We found that



**Figure 8.** The combination of aspect markers and different verb types in Chinese Endpoint descriptions for Endpoint-oriented and endpoint-reached events: zero aspect marker combined (MaOnly, PaOnly, SvcOnly), with the perfective marker (MaPerf, PaPerf, SvcPerf), with the progressive marker (MaProg, PaProg, SvcProg).

**Table 3.** Examples of verb-aspect combinations in Chinese.

	Chinese
Manner verb + Perfective (MaPerf)	<i>Yi-wei nanshi zou xiang le yi-liang lanse de che.</i> "a man walk towards PERF a blue car"
Manner verb + Progressive (MaProg)	<i>Yi-ge xiaohai zai zou xiang youlechang.</i> "a kid PROG walk towards playground"
Path verb + Perfective (PaPerf)	<i>Yi-ge nande jin le yi-ge jiaotang.</i> "a man enter PERF a church"
Path verb + Progressive (PaProg)	<i>Yi-ge nanshi zai guo malu.</i> "a man PROG cross street"
Serial verb construction + Perfective (SvcPerf)	<i>Yi-liang qiche kai-jin le cheku.</i> "a car drive-enter PERF garage"
Serial verb construction + Progressive (SvcProg)	<i>Zhe-ge nvshi zai xiang dalou zou-qu.</i> "this woman PROG towards building walk-go"

Dutch participants were significantly faster than Chinese participants in giving accurate responses (Chinese:  $M = 1757$  ms,  $SD = 670$  ms, Range 827.75 ms–3216.00 ms; Dutch:  $M = 1483$  ms,  $SD = 494$  ms, Range 850.67 ms–1984.00 ms;  $t(293.67) = 4.223$ ,  $p < .001$ ). However, the same pattern was also found for filler items (Chinese:  $M = 1663$  ms,  $SD = 634$  ms, Range 810.67 ms–2484 ms; Dutch:  $M = 1356$  ms,  $SD = 489$  ms, Range 676.33 ms–1891 ms;  $t(266.27) = 4.550$ ,  $p < .001$ ).

## 6. Discussion

Our motion event description task set out to discover (1) how the path of motion, the core schema of a motion event, is *conceptualised* in Chinese and Dutch (what path information is selected for encoding?), and (2) how the path of motion is *described* in Chinese and Dutch, in terms of the linguistic means used (what verb and adjunct types are used and combined in path descriptions?). Third, we were specifically interested in the extent to which different aspectual markers were used and combined with verb types in Chinese endpoint descriptions. In addition, we explored memory of endpoints in a surprise memory task, administered post verbalisation. Speakers of the two languages viewed and described two types of video stimuli. One type showed Endpoint-oriented events in which a moving entity moved along a trajectory (e.g. a road, a street) with a potential but not reached Endpoint at its end (woman walking along a street towards a bus stop). The other video type depicted Endpoint-reached events with similar scenarios, but the Endpoints were all reached by the end of the video clips (man entering a building).

### 6.1. Conceptualising the path of motion

We hypothesised that in Endpoint-reached events, speakers of both languages would prefer to mention

endpoints over other types of path information (trajectory and location). We indeed found a strong preference for mentioning goals of Endpoint-reached motion events in both languages. This pattern did not occur when endpoints were displayed as not reached in the video clips. For events in which entities were only moving towards an endpoint, the trajectory or location of the motion was conceptualised and described as the path of motion most frequently.

Speakers in our experiment seemed to distinguish between these two event types, in line with classifications made in situation type theories (Croft, 2012; Klein, 1994; Vendler, 1967) and in line with Event Segmentation theory (Zacks et al., 2007). According to Klein (1994), situations are categorised into zero-state situations (e.g. *a tree is a plant*), one-state situations (e.g. *she is sleeping*) and two-state situations (e.g. *she left*). In our study, the Endpoint-oriented events are one-state situations that involve no (substantial) qualitative changes (e.g. *two women are walking along/on a path towards a house in the distance*), whereas Endpoint-reached events are in nature boundary-crossing events that indicate a change of location, and thus represent two-state situations. Event Segmentation theory proposes that event boundaries are perceived when specific features of an event change substantially: Studies have shown that the event boundaries that people detect, when asked to segment ongoing activity into individual events, include changes in spatial location (for example, Zacks, Speer, & Reynolds, 2009; Zwaan & Radvansky, 1998). When a change in spatial location is prominent as in our Endpoint-reached events, event boundaries (the endpoints) are highly salient and are thus likely to be mentioned when people are asked to verbally report on the events. Therefore, speakers of both languages exhibited a strong preference for encoding Endpoint information when the video clips showed reached endpoints. This is similar to the goal (over source) bias reported in previous studies in relation to motion event conceptualisation (e.g. Lakusta & Landau, 2005; Papafragou, 2010). However, when the endpoints had to be inferred and were not depicted as reached, people in the present study did not predominantly construe the events as two-state situations. To form a reportable "unit" of the motion event they were watching, speakers selected other elements in the scene to complement the motion verbs. Specifically, in these cases, other elements of the path of motion became "anchoring points" for the motion event, as the path of motion represents the "core" of an event of motion (Talmy, 2000). Utterances describing motion events without *any* path information are under-informative from a communicative perspective; and indeed, utterances without



adjuncts containing path information at all are rare in the present data set (e.g. *a man is walking*).

As for Endpoint-oriented events, we hypothesised that speakers of Chinese and speakers of Dutch would show different preferences in terms of mentioning endpoints, given cross-linguistic differences in the use of grammatical aspect and verb lexicalisation patterns. However, our results showed similar frequencies regarding endpoint mentioning in both Endpoint-reached and Endpoint-oriented events in Chinese and Dutch. In addition, we hypothesised that for endpoint-oriented events, speakers of Dutch would react more accurately and/or faster than speakers of Chinese in our surprise endpoint recognition memory task. Our results did not support this hypothesis either. There was no significant difference in the detection sensitivity between the two language groups. We did find that Dutch participants were significantly faster than the Chinese group in giving accurate responses. However, the same pattern was also found for filler items. It seems that Dutch participants were generally faster than the Chinese group in making accurate choices, possibly due to the fact that the Dutch participants had more experience participating in psycholinguistic experiments than participants in the Chinese group. Our hypotheses were based on von Stutterheim's work (2006, 2012), in which speakers of aspectual languages were shown to be less likely to mention endpoints, compared to speakers of non-aspectual languages (for the same Endpoint-oriented events used in the present study). We do not observe this same pattern, likely due to the fact that the two languages under investigation do not clearly fall into the one or the other language cluster in terms of aspect; whereas Chinese marks both progressive and perfective aspect, the markers are not used across the board in descriptions of motion events (see results in section 5.3). In addition, although Dutch speakers do not typically use progressive aspect to describe motion, the available markers are used frequently in other event types, e.g. causative actions (knitting a scarf, peeling potatoes) or activities (playing football). In this sense, Dutch speakers are used to marking an aspectual viewpoint on an event. The cross-linguistic comparison is thus by no means straightforward and it is not clear whether Dutch and Chinese would behave similarly to previously investigated aspectual and/or non-aspectual languages. Georgakopoulos et al. (2019) reported that verb lexicalisation patterns should have an impact on the frequency of reference to goals or endpoints in motion events. Specifically, speakers of satellite-framed languages should be more likely to encode endpoints than speakers of verb-framed languages. We are currently not able to support this claim on the basis of the

present data from Dutch and Chinese. Again, this could be due to the difficulty of classifying Chinese into one or the other language cluster: We find that although Chinese has an abundance of path verbs, the data show a typical satellite-framed pattern (a manner verb followed by a satellite), when endpoints were depicted as not reached and had to be inferred in the events. Given motion events with reached endpoints, Chinese mainly uses serial verb constructions with a perfective marker (see results in section 5.3). The typical syntactic framing patterns in the descriptions thus differ across event types in Chinese. This corresponds to what is mentioned as "split conflation" in Talmy (2000) and in Levin and Rappaport Hovav (2019).

As for Endpoint-oriented events, we also hypothesised that Chinese and Dutch would show divergent patterns in terms of the frequency of selection of trajectory vs location information for motion construal. We indeed find a significant difference here: Dutch speakers encoded features of the ground traversed (the trajectory of motion) much more often than Chinese speakers. On the other hand, speakers of Chinese were found to encode the location of moving entity without any additional path elements much more frequently. These Location-only references in Chinese often followed a manner verb. This same pattern was found in verb-framed languages, such as French, Italian and Arabic (Carroll et al., 2012; Flecken et al., 2015; von Stutterheim, Bouhaous, & Carroll, 2017). Flecken et al. (2015) and von Stutterheim et al. (2017) argue that this is driven by the abundant presence of path verbs in verb-framed languages: whereas path verbs are typically used to describe changes in location and directed motion, manner verbs are often used when a change in spatial location is *not* evident (as is the case in our Endpoint-oriented events). When manner verbs are used, however, the event is conceptualised in terms of the characteristics of the Figure in motion (is she skating, running, hopping?) rather than about the changes in space she is engaged in; the manner verb asserts a property of the figure, which is then combined with an adjunct, locating the figure in motion in space. If a speaker of a verb-framed language wants to describe goal-directed motion and provide information on the Figure's manner of motion at the same time, the typically reported pattern is of a "division" of information across two utterances, e.g. *a woman is walking on the road, and is heading for a bus stop*. It has recently been demonstrated that these differences in motion event conceptualisation between satellite-framed and verb-framed languages also lead to differences in event segmentation patterns: in the verb-framed language French, a video of, for example, a woman walking along a road, then turning

right up a flight of stairs, was segmented into two single events, with the first one providing information on manner of motion (combined with a reference to a spatial location, e.g. *walk on the street*), and the second one containing directed motion information, e.g. *ascend stairs* (Gerwien & Stutterheim, 2018). As we have mentioned above, Chinese exhibits characteristics of verb-framed languages (Beavers et al., 2010; Ji et al., 2011; Ji & Hohenstein, 2017). Chinese also shows the pattern that when manner verbs are used alone to describe a motion event, a locative adjunct is preferred over a trajectory adjunct by speakers. As a typical satellite-framed language, Dutch descriptions typically contain a manner verb followed by trajectory-information, instead of a locative adjunct alone. The current description pattern thus hints at a different role for manner verbs in motion conceptualisation in Chinese, compared to Dutch.

### 6.2. Verb and adjunct types used

We hypothesised that in Endpoint-reached events, speakers of Chinese would adopt serial verb constructions predominantly to describe endpoints, whereas speakers of Dutch would predominantly make use of manner verbs plus satellites to describe Endpoint. Among the descriptions of Endpoint-reached events, we found that speakers of Chinese indeed mainly used serial verb constructions, whilst both single manner verbs and single path verbs were used occasionally. The latter two types did not differ in their proportion of occurrence. On the other hand, speakers of Dutch used manner verbs plus satellites most often. Speakers of the two languages indeed presented different verb lexicalisation patterns when encoding Endpoint-reached events. Our findings are similar to what has been found in Ji et al. (2011) in which they compared English and Chinese for caused motion events.

We hypothesised that in Endpoint-oriented events, Dutch would exhibit features of a typical satellite-framed language and would mostly use satellites to describe path information (including Endpoint, Trajectory and Location-only). However, we were unclear as to what patterns Chinese would use to encode the path of motion, as no previous studies have systematically studied this. We considered the use of satellites, serial verb constructions and path verbs as possible options. We found that in Dutch, satellites were used most often (encoding either endpoints or trajectory information, combined with manner verbs), and path verbs were infrequently used to describe the path of motion (0.91 [ $N = 209/231$ , Range = 0.56–1.00] vs 0.09

[ $N = 21/231$ , Range 0.00–0.45] proportion of use respectively). Therefore, Dutch showed characteristics of a typical satellite-framed language. In Chinese, all three types of verbal constructions were used (0.60 [ $N = 144/241$ , Range 0.00–1.00] vs 0.11 [ $N = 27/241$ , Range 0.00–0.83] vs 0.28 [ $N = 68/241$ , Range 0.20–1.00] of use of manner verbs, path verbs, and serial verb constructions), and manner verbs occurred mainly with endpoints or location-references (proportion of 0.31 [ $N = 74/241$ , Range 0.00–0.89] and 0.26 [ $N = 62/241$ , Range 0.00–0.70], respectively) in this type of event. This latter pattern is hypothesised to be a typical verb-framed pattern, as discussed in the previous section. The former pattern is considered to be typical of satellite-framed languages. In addition, serial verb constructions (predominantly without additional path adjuncts) and single path verbs were used as well. Overall, Chinese showed use of all three available options and exhibited both satellite- and verb-framing properties.

In sum, Dutch showed features of a typical satellite-framed language in both event types, whereas Chinese is a mixed type with features of satellite-framed and verb-framed languages. Our findings are similar to what was reported in Ji et al. (2011), in which English and Chinese descriptions of caused motion events were analysed. The authors concluded that Chinese was different from English (a typical satellite-framed language) in that it exhibited both satellite- and verb-framing properties. Our study thus supports their conclusion on the typology of Chinese, for the domain of voluntary motion with varying degrees of goal-orientation.

### 6.3. Verb types and aspect markers used in Chinese descriptions of endpoints

This section specifically targets the use of the aspect markers (*zai* and *le*) and their combinations with different verb types in Chinese references to endpoints. An interrelation between aspect and endpoints has been reported before for motion events in which an endpoint is referable but not reached (Athanasopoulos & Bylund, 2013; von Stutterheim et al., 2012). Moreover, researchers have claimed that verb lexicalisation patterns can affect the way people construe endpoints as well (Georgakopoulos et al., 2019). Chinese is considered an aspectual language but use of aspect is not obligatory. Besides, Chinese exhibits features of both satellite-framed languages and verb-framed languages in terms of verb lexicalisation patterns. Therefore, by following the approach in Georgakopoulos et al. (2019), we tried to unravel the typical patterns in Chinese endpoint

descriptions, taking into account both typological features. In doing so, we aim to gain a deeper insight into the interrelation of aspect, verb semantics, and endpoint conceptualisation in Chinese.

We found that the patterns differed across Endpoint-oriented and Endpoint-reached events. For Endpoint-oriented events, the typical pattern in Chinese was the use of a manner verb without aspect followed by a directional preposition (e.g. *zou xiang* “walk to/towards”). The progressive marker *zai* occurred more often than the perfective *le* in this pattern. Serial verb constructions were also used (e.g. *chao xx zou-qu* “towards xx walk-go”), mainly without aspect, but occasionally with the progressive aspect. Moreover, a single path-verb pattern was also observed in descriptions of inferable endpoints (e.g. *qu* “go”), but no aspect was used. In sum, when the endpoints were depicted as not reached, Chinese exhibits a typical satellite-framed language pattern (a manner verb followed by a directional preposition), without the use of aspect markers, in line with the patterns obtained for typical non-aspectual languages (von Stutterheim et al., 2012).

When endpoints were shown as reached in the video clips, the typical pattern found in Chinese was the use of a serial verb construction followed by a perfective marker (e.g. *zou-jin le* “walk-enter PERF”). Single path verbs (e.g. *jin* “enter”) were also used in combination with the perfective marker *le*. Chinese thus shows a preference to use the perfective aspect in case of reached endpoints, so-called boundary crossing events, different from the typical description patterns obtained for cases in which goals are not reached. This again underlines the previous statement that the available aspect markers in Chinese are not used across the board; they are used frequently (especially in boundary-crossing events), but use is optional.

#### **6.4. Implications for typologies of motion expression and theories of event cognition**

The present study provides a fine-grained picture of how people conceptualise and describe events, going beyond previous work in several aspects. First of all, we investigated a language pair (Chinese and Dutch) that has not been compared in the domain of motion conceptualisation, looking at both grammatical aspect and verb lexicalisation patterns, before. The two languages do not fall into a clear category of aspectual/non-aspectual languages, nor of satellite-framed/verb-framed languages, providing new insights relevant to typologies of motion expression. Second of all, we included two event types, endpoint-reached and endpoint-oriented events, offering a detailed analysis of motion

conceptualisation and description cross-linguistically. Importantly, we included a focus on events in which there are *no* clear goals or boundaries, different from most previous motion description work which is heavily centred on understanding how people deal with goals and boundary-crossing events almost exclusively. We thus investigated in more detail how people conceive of the *content* of a motion event, i.e. how are other elements that are part of the event leading up to its boundary conceptualised during production? What information do people select to form an event “unit” (Gerwien & Stutterheim, 2018), when the event boundary is *not* salient? We found that other elements of the path of motion became anchoring points for event conceptualisation, i.e. people select either the Figure’s location in space (*on the street*), or the trajectory traced by the Figure (*along the road*) for encoding. Importantly, we found cross-linguistic differences therein, showing diversity with respect to how people conceptualise the inner structure of a motion event. In sum, we extend the cross-linguistic research on motion expression which has predominantly reported a general goal-bias (compared to sources), in relation to boundary-crossing events (e.g. Papafragou, 2010).

Importantly, our findings can also inform event cognition theories because we go beyond a strong and exclusive focus on event boundaries (motion event endpoints): Existing theories describe how important event boundaries are during perceptual processing and in memory representation of events (Zacks et al., 2007). However, they do not describe how people process and weight the various other elements that comprise an event, prior to its boundary. Here, we show that people are sensitive to those other elements of the path of motion depicted in a motion event. People hardly produce utterances without any path information (e.g. using only a manner verb to describe manner of motion: *a woman is walking*), underlining the fact that the path of motion is the core schema of a motion event (Talmy, 2000), and suggesting that manner of motion (in isolation) may not be represented strongly in our event cognition. Interestingly, the cross-linguistic differences we obtained in the conceptualisation and expression of trajectory and location of motion show that speakers of different languages may take different viewpoints when encoding certain events. When event boundaries (endpoints) are not salient, and the Figure in the event is moving without a clear goal in sight, speakers of both Chinese and Dutch seemingly take the same perspective, by highlighting the manner of motion using manner verbs. However, they combine them with different path-elements, implying different foci of attention: Chinese speakers, resembling speakers

of verb-framed languages, followed a pattern in which motion through space without a clear boundary-crossing entails localising the moving entity in space (e.g. *a man is cycling in the countryside*). In this pattern, the manner of motion becomes more like a statement regarding a property of the figure in motion (i.e. the man is *cycling*, not *running*, in the countryside), rather than an instance of directed motion. In Dutch, such events are conceptualised as instances of movement along a trajectory (manner of motion verbs combined with trajectory information, e.g. cycling across the street, along the road, over the hill). Indeed, an eye-tracking study comparing French and German speakers shows that, when endpoints are not salient, the use of manner verbs in the two groups goes hand in hand with different patterns of visual attention allocation during speech planning: Speakers of the verb-framed language French showed a more pronounced degree of attention allocated to the figure in motion, as compared to German speakers (Flecken et al. 2015). In their descriptions, French speakers also showed a preference to combine manner verbs with the mentioning of the location of the figure. In all, the present study thus underlines that the production of seemingly similar verbal material (manner verbs) in different languages can be guided by different conceptualisation patterns, as reflected in online attention allocation, and the path components these verbs are combined with in the descriptions (see also Carroll et al., 2012; von Stutterheim et al., 2017). This renders the interesting possibility that certain principles and processing routines in event cognition in general may differ across speakers of different languages. For example, Gerwien and Stutterheim (2018) were able to show differences in the granularity of event segmentation between French and German, because of differences in the languages' verb lexicalisation and motion conceptualisation patterns.

It is important to note that the current findings of cross-linguistic differences were obtained in settings with very specific task demands (scene conceptualisation during an event description task). One important open question is, how do people perceive and conceptualise endpoint-oriented events when they are not required to speak about them? In our comparison of production patterns, we find a strong influence of the respective linguistic means (both semantic and grammatical) available in the languages tested on path component selection; the question is, what saliency do people attribute to trajectory and location in a motion event (when endpoints are not salient), outside the realm of speaking? Future experiments should consider using non-linguistic tasks that tag the saliency (as reflected in perception, attention, and memory) of various elements of the path of motion for endpoint-oriented events. A combination of

linguistic tasks with other more sensitive measures (e.g. eye tracking, reaction-time-based or neurophysiological measures) can be very helpful for enhancing our understanding of the relation between language and event cognition. One way to move forward would be to tap into post-verbalisation memory of location and trajectory of motion, linking more closely path-component selection in production (and cross-linguistic differences therein) and memory. In sum, as mentioned above, we find differences in event conceptualisation (path component selection) between Chinese and Dutch speakers when describing endpoint-oriented events. Important remaining questions are, can we capture these as differences in online event processing during speaking, in memory after speaking, or beyond speaking, during event processing without language in a nonlinguistic task?

Interestingly, even though Chinese and Dutch speakers showed the same conceptualisation patterns for endpoint-reached events, they packaged the information into different linguistic means and structures. The question is, what consequences do different information packaging styles have for online processing of events, both during information uptake for event description, but also more globally, outside a speaking context, when taking into account habitual patterns in information packaging and distribution? An intriguing test case concerns differences in the use of aspect and the consequences this has for online event processing. There is increasing evidence that aspectual marking entails viewpoint selection, reflected in eye movement patterns, brain potentials and also mental simulation during event comprehension (e.g. Huette, Winter, Matlock, Ardell, & Spivey, 2014; Madden & Zwaan, 2003). For example, the frequent use of the perfective aspect in Chinese in describing reached endpoints (different from Dutch) might lead to different online processing of this spatial information in Chinese speakers, compared to Dutch speakers: does the highlighting of the action as completed with verb aspect influence attentional focus to event endpoints, during event processing (and when?)? Also, when taking into account the verbal structures used to describe motion events with reached endpoints, what cognitive processing pattern underlies the production of serial verb constructions in Chinese, as compared to verb + satellite constructions, frequent in Dutch? Because of the tight packaging of manner and endpoint information in one construction, it is possible that Chinese speakers, when planning to use serial verb constructions, pay attention to the two elements at the same time/to the same extent, while Dutch speakers may first (or predominantly) pay attention to manner, given that that is the



information encoded in the verb. The information encoded in the verb is crucial for defining event structure and roles (see e.g. Slobin, 2006; Vendler, 1967), and should therefore hypothetically lead to enhanced attention allocation to the respective event dimension during early phases of sentence planning. These subtle differences in information packaging could be reflected in differences in the time course of processing and attention allocation towards visual endpoints. Previous work has studied event segmentation under the dependency of serial verb constructions and found that when people were primed with a serial verb construction, their segmentation of subsequent visual input was more coarse-grained, compared to the use of coordinated sentences (two single verbs) (Defina, 2016a). Also, a study of co-speech gestures suggests that serial verb constructions reflect single conceptual events (Defina, 2016b) – to what extent is this similar to, or different from, satellite-framed packaging of manner information in the verb, and path information in a satellite? These issues deserve detailed investigation in the future, not only studying language production, but also language comprehension accompanied with visual input, and visual processing in the absence of overt language use, comparing participants under different conditions (different event types, primed with language, or without), and from different language backgrounds.

## 7. Conclusion

The current study sheds light on the extent to which cross-linguistic differences between Dutch and mandarin Chinese affect their speakers' conceptualisation of the path of motion in motion events. We first investigated motion event conceptualisation patterns in speakers of the two language groups, by targeting their selection of Endpoint, Location-only or Trajectory information when describing motion events. We then took both typological features (aspect and verb lexicalisation patterns) into account and analysed the use of different linguistic patterns, in terms of combinations of verb semantics, path-elements encoded and aspect, in motion event descriptions. Specific hypotheses were based on the typology of verb lexicalisation patterns and grammatical aspect systems. We found that in terms of the frequency of encoding of goals or endpoints, speakers of the two languages showed similar patterns. In terms of how endpoints were described, speakers exhibited cross-linguistically similar patterns for Endpoint-oriented events, but different ones for Endpoint-reached events. However, regarding the selection of path elements other than Endpoints (in Endpoint-oriented events), speakers of Dutch preferred to

encode information concerning the trajectory of motion (e.g. a car drives along the road), whilst speakers of Chinese often chose to locate the entity in motion in space (a woman walks on the street). The former presents a typical satellite-framed pattern, while the latter shows the same pattern found in verb-framed languages, such as French, Italian and Arabic (von Stutterheim et al., 2017). We conclude that cross-linguistic differences in motion conceptualisation between Dutch and mandarin Chinese can be detected given the present design, using events with different degrees of goal orientation, and taking into account both verb lexicalisation patterns and grammatical aspect usage. By reporting detailed comparisons of the linguistic patterns (considering both aspectual systems and verb lexicalisation patterns) in both languages, for two event types, we provide an in-depth understanding of the relation between linguistic typology and event conceptualisation in language production.

## Notes

1. The verb "to park" was excluded from analysis for both languages because it conveys features of manner (not all moving entities can park) and features of path (a change of location) at the same time (1.038% of the data were excluded: 12 cases out of 1156 sentences in total).
2. Utterances that did not mention a specific endpoint location but implied one through an action (e.g. supermarket was implied in "going shopping"; gas station was implied in "go filling in gas") were considered as endpoint mentioning. Utterances that implied a specific endpoint location (e.g. *hij loopt naar binnen* "he goes inside") or an unspecific endpoint (e.g. *ta xiang qian zou* "he towards front walk") were considered as endpoint mentioning as well.
3. The deictic verb "go" (*gaan* in Dutch and *qu* in Chinese) was coded as a path verb in both languages. It should be noted that in both languages, the deictic verb can be followed by a spatial location (e.g. go to a supermarket) or an action (e.g. go shopping). The deictic verb "go" was also coded as a path verb for the latter case in which the directed motion meaning is bleached/less evident.
4. Subordinate ZHE clauses (for example, *qi zhe zixingche jin mendong* "ride ZHE bike enter gate") were not coded as part of serial verb constructions in this study. We coded the verb/verbs following these subordinate clauses. In this example, the verb *jin* "enter" was coded as path verb. The utterance *qi zhe zixingche zou xiang mendong* "ride ZHE bike walk towards gate" the verb was coded as containing a manner verb; if the verbal construction following the ZHE clause was a serial verb, the utterance was coded as containing a serial verb construction (e.g. *qi zhe zixingche zou-jin mendong* "ride ZHE bike walk enter gate").
5. We are happy to share our data (transcriptions, coded data, memory data, experiment and analysis scripts) upon request.

6. `glmer (Endpoint ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial)`  
`glmer (Location_Only ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial)`  
`glmer (Trajectory ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial).`
7. `mlogit(Path~1|EventType, data = PathC2, reflevel = 1) #reflevel 1 = endpoint`  
`mlogit(Path~1|EventType, data = PathD2, reflevel = 1) #reflevel 1 = endpoint.`
8. All the dependent variables in the multinomial logistic regression models were dummy coded. When one category of a dependent variable was coded as 0, the other categories of the dependent variable were all coded as 0. So if there were  $N$  categories, there were  $N-1$  dummy variables.
9. We got these within-event-type statistics via the intercepts information after running the binomial/multinomial logistic regression models.
10. `glmer (MV ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial)`  
`glmer (PV ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial, control = glmerControl(optimizer="bobyqa",optCtrl = list(maxfun = 100000))).`
11. `mlogit(VerbType~1|EventType, data = VerbC2, reflevel = 1) #reflevel 1 = MV`  
`glmer(VerbType~EventType + (1|PP) + (1|Stimulus), data = VerbD3, family = binomial).`
12. `glmer (MaEnd ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial)`  
`glmer (PaEnd ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial)`  
`glmer (MaLoc~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial)`  
`glmer (MaTra ~ Language + EventType + (1|PP) + (1|Stimulus), data = data, family = binomial) # The model that included the interaction did not converge.`  
`glmer (PaTra ~ Language*EventType + (1|PP) + (1|Stimulus), data = data, family = binomial).`
13. `mlogit(VerbAdjunct~1|EventType, data = ChAll3, reflevel = 1) #reflevel 1 = MV_endpoint`  
`mlogit(VerbAdjunct~1|EventType, data = DuAll3, reflevel = 1) #reflevel 1 = MV_endpoint.`
14. `mlogit(VerbAspect~1|EventType, data = Csyntax4, reflevel = 1) #reflevel 1 = MV.`

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by China Scholarship Council: [grant number 201606020119]; National Social Science Fund of China: [grant number 18CYY004].

## References

Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic

- comparison of English and Swedish speakers. *Cognitive Science*, 37, 286–309.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Beavers, J., Levin, B., & Tham, S. W. (2010). The typology of motion expressions revisited. *Journal of Linguistics*, 46, 331–377.
- Bohnemeyer, J., & Swift, M. (2004). Event realization and default aspect. *Linguistics and Philosophy*, 27, 263–296.
- Carroll, M., Weimar, K., Flecken, M., Lambert, M., & von Stutterheim, C. (2012). Tracing trajectories: Motion event construal by advanced L2 French-English and L2 French-German speakers. *Language, Interaction and Acquisition*, 3, 202–230.
- Chen, L., & Guo, J. (2009). Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of Pragmatics*, 41, 1749–1766.
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge, England: Cambridge University Press.
- Croft, W. (2012). *Verbs: Aspect and causal structure*. Oxford, England: Oxford University Press.
- Croissant, Y. (2018). mlogit: Multinomial logit models. R package version 0.3-0. Retrieved from <https://CRAN.R-project.org/package=mlogit>
- Defina, R. (2016a). *Events in language and thought: the case of serial verb constructions in Avatime* (Doctoral dissertation). Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands.
- Defina, R. (2016b). Do serial verb constructions describe single events? A study of co-speech gesture in Avatime. *Language*, 92, 890–910.
- De Knop, S., & Gallez, F. (2011). Manner of motion: A privileged dimension of German expressions. *International Journal of Cognitive Linguistics*, 2, 25–40.
- Filippo-Enrico, C. (2008). Manner of motion saliency: An inquiry into Italian. *Cognitive Linguistics*, 19, 533–570.
- Flecken, M. (2011). What native speaker judgments tell us about the grammaticalization of a progressive aspectual marker in Dutch. *Linguistics*, 49, 479–524.
- Flecken, M., Carroll, M., & von Stutterheim, C. (2014). Grammatical aspect influences motion event perception: Findings from a cross-linguistic nonverbal recognition task. *Language and Cognition*, 6, 45–78.
- 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. *Modern Language Journal*, 99, 100–122.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79.
- Georgakopoulos, T., Härtl, H., & Sioupi, A. (2019). Goal realization: An empirically based comparison between English. *German, and Greek. Languages in Contrast*, 19, 280–309.
- Gerwien, J., & Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, 180, 225–237.
- Hickmann, M., & Hendriks, H. (2010). Typological constraints on the acquisition of spatial language in French and English. *Cognitive Linguistics*, 21, 189–215.
- Huettenlocher, J., Winter, B., Matlock, T., Ardell, D. H., & Spivey, M. (2014). Eye movements during listening reveal spontaneous grammatical processing. *Frontiers in Psychology*, 5, 410.

- Ibarretxe-Antuñano, I. (2004). Language typology in our language use: The case of Basque motion events in adult oral narratives. *Cognitive Linguistics*, 15, 317–349.
- Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.
- Ji, Y., Hendriks, H., & Hickmann, M. (2011). The expression of caused motion events in Chinese and in English: Some typological issues. *Linguistics*, 49, 1041–1077.
- Ji, Y., & Hohenstein, J. (2017). Conceptualising voluntary motion events beyond language use: A comparison of English and Chinese speakers' similarity judgments. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, 195, 57–71.
- Kan, Z. (2010). Motion-event typology in mandarin Chinese revisited. *Contemporary Linguistics*, 2, 126–135.
- Klein, W. (1994). *Time in language*. London, England: Routledge.
- Klein, W., Li, P., & Hendriks, H. (2000). Aspect and assertion in mandarin Chinese. *Natural Language & Linguistic Theory*, 18, 723–770.
- Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, 96, 1–33.
- Lamarre, C. (2005). The linguistic encoding of motion events in Chinese: With reference to cross-dialectal variation. *Typological studies of the linguistic expression of motion events*, (pp. 3–33). COE-ECS Working Paper.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levin, B., & Rappaport Hovav, M. (2019). Lexicalization patterns. In R. Truswell (Ed.), *Oxford handbook of event structure* (pp. 395–425). Oxford, England: Oxford University Press.
- Li, N., & Thompson, S. (1981). *Mandarin Chinese*. Berkeley, CA: University of California Press.
- Liu, L. (2014). How Chinese codes path of a motion: A typological study of motion events in modern Chinese. *Chinese Teaching in the World*, 3, 322–332.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.
- Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, 31, 663–672.
- Ochsenbauer, A. K., & Hickmann, M. (2010). Children's verbalizations of motion events in German. *Cognitive Linguistics*, 21, 217–238.
- Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, 34, 1064–1092.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108, 155–184.
- Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes*, 25, 224–260.
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Regier, T., & Zheng, M. (2007). Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science*, 31, 705–719.
- Shen, J. (2003). The resultative construction in Chinese: A typological perspective. *Chinese Teaching in the World*, 3, 17–23.
- Shi, W., & Wu, Y. (2014). Which way to move: The evolution of motion expressions in Chinese. *Linguistics*, 52, 1237–1292.
- Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani, & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195–220). Oxford, England: Clarendon Press.
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist, & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspective* (pp. 219–257). Mahwah, NJ: Lawrence Erlbaum Associates.
- Slobin, D. I. (2006). What makes manner of motion salient? Exploration in linguistic typology, discourse, and cognition. In M. Hickmann, & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 59–81). Amsterdam, the Netherlands: John Benjamins.
- Slobin, D. I. (2008). Relations between paths of motion and paths of vision: A cross-linguistic and developmental exploration. In V. M. Gathercole (Ed.), *Routes to language: Studies in honour of Melissa Bowerman* (pp. 197–221). Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith, C. (1991). *The parameter of aspect*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Stefanowitsch, A., & Rohde, A. (2004). The goal bias in the encoding of motion events. In G. Radden, & K. U. Panther (Eds.), *Studies in linguistic motivation* (pp. 249–267). Berlin, Germany: Mouton de Gruyter.
- Tai, J. H. (2003). Cognitive relativism: Resultative construction in Chinese. *Language and Linguistics*, 4, 301–316.
- Tai, J. H., & Su, S. (2013). Encoding motion events in Taiwan sign language and mandarin Chinese: Some typological implications. In G. Cao, H. Chappell, R. Djamouri, & T. Wiebusch (Eds.), *Language and linguistics monograph series 50-breaking down the barriers: Interdisciplinary studies in Chinese linguistics and beyond* (pp. 79–98). Taipei, Taiwan: Institute of Linguistics, Academia Sinica.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and lexical description: Vol. 3. Grammatical categories and the lexicon* (pp. 36–149). Cambridge, England: Cambridge University Press.
- Talmy, L. (1991). Path to realization: A typology of event conflation. *Proceedings of the Berkeley Linguistics Society*, 17, 480–519.
- Talmy, L. (2000). *Toward a cognitive semantics: Vol. II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.
- Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., & Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics*, 50, 833–867.
- von Stutterheim, C., Bouhaous, A., & Carroll, M. (2017). From time to space. The impact of aspectual categories on the construal of motion events: The case of Tunisian Arabic and modern Standard Arabic. *Linguistics*, 55, 207–249.
- von Stutterheim, C., & Carroll, M. (2006). The impact of grammaticalised temporal categories on ultimate attainment in advanced L2-acquisition. In H. Byrnes (Ed.), *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment* (pp. 40–53). Georgetown, WA: Georgetown University Press.
- Xiao, R., & McEnery, T. (2004). *Aspect in mandarin Chinese: A corpus-based study*. Amsterdam, the Netherlands: John Benjamins.

- Xu, Z. (2013). An empirical study of motion expressions in mandarin Chinese. *English Language and Literature Studies*, 3, 53–67.
- Zacks, J., Speer, N., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, 138, 307–327.
- Zacks, J., Speer, N., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin*, 133, 273–293.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185.