Events are commonly conveyed and understood through language. Language plays an important role in event studies. In this dissertation, I have looked at language-based event conceptualization from a cross-linguistic perspective (Chapter 3). Moreover, I have examined language-based event comprehension by focusing on two grammatical morphemes in Chapter 2, 4, 5. Furthermore, I have investigated the effects of two non-linguistic factors on the use of two pairs of directional prepositions in two different languages, targeting motion event endpoint conceptualization (Chapter 4 and Chapter 5). All the chapters from this dissertation contribute to our understanding of how language plays a role in event studies and how we convey and understand events through language.

Conveying and Understanding Events Through Language

Yiyun Liao

Conveying and Understanding Events Through Language

Yiyun Liao

Conveying and Understanding Events Through Language

Yiyun Liao

The research presented in this dissertation was financially supported by the China Scholarship Council (CSC).

ISBN: 978-94-6361-672-0

Cover design by: Xingping Huang and Yiyun Liao

Layout by: Yiyun Liao

Printed by: Optima | https://www.ogc.nl

Conveying and Understanding Events Through Language

Het overdragen en begrijpen van gebeurtenissen door middel van taal

Thesis

to obtain the degree of Doctor from the

Erasmus University Rotterdam

by command of the

rector magnificus

Prof.dr. A.L. Bredenoord

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

Thursday 28 April 2022 at 13.00 hrs

by

Yiyun Liao

born in Hunan, China.

Erasmus University Rotterdam

Doctoral Committee:

Promotors:	Prof.dr. R.A. Zwaan
	Prof.dr. K. Dijkstra
Other members:	Prof.dr. H.A.E.M. Pecher
	Prof.dr. T. Richter
	Dr. B.B. de Koning

Contents

送给我的奶奶和爸妈

Chapter 1

General Introduction

As I look out of the window of the school library, a van is coming from the entrance of the campus, a group of students are chatting and walking towards the sports center, two girls are playing tennis in the tennis field, and a young man just parked his bike in the parking area. As I keep observing, some events are still ongoing (e.g., the two girls are still playing tennis), whereas the others have already ended (e.g., the van has left and that specific group of students as well). At the same time, new events, being performed by other people, constantly come into my sight (e.g., a young man is riding a scooter into the campus). In fact, these short 5 minutes are just an epitome of our daily life.

Events

When we open a history book, we see introductions to all kinds of events, such as the 1911 Revolution in China (Xinhai Revolution), World War 1 (1914-1918), World War II (1939-1945), the 1954-1968 civil rights movement in the US, the September 11 attacks in the US, and so on. Turning on the TV, we observe that there is a European Cricket Championship match between Portugal and the Netherlands, a chef is cooking in London, and that there is a serious storm happening in the Liaoning Province in China. As a matter of fact, our human history is a collection of events as well as each person's entire life. It seems so natural to us to understand the world and our life in terms of a series of connected and separated events. However, on second thought, what is an event? How do we actually define an event?

The American philosopher Quine (1985) has proposed that events are just like objects. The main difference is that objects are distinguished through spatial boundaries, whereas events have boundaries both in space and in time. For instance, when we talk about a book, we identify it by its location (e.g., the book is on the table). However, when we talk about reading a book, we take the event of "reading a book" as having a beginning and an end in time (i.e., to start, continue, or finish reading a book) (see similar illustrations in Zacks & Tversky, 2001). The location where a particular event of "reading a book" happens is also used to distinguish the event, for instance, reading a book in a library or on a train.

There are also numerous ways of categorizing events. For example, events can be identified in terms of their emotional values, for instance, positive, neutral, or negative events. Alternatively, events can be classified according to their contents, such as games, competitions, exams, weddings, and so on. Moreover, events can either be concrete or abstract. For instance, "Bubbles keep coming up in a beer glass" denotes an event that we can clearly see with our eyes and probably we can even hear the sound of popping bubbles, which are surely concrete experiences. "Ideas keep bubbling up in her mind", however, describes an abstract event that we cannot really touch, see, or hear. In this case, our experiences with the concrete event of bubbles coming up in a beer glass facilitate our understanding of the event of ideas forming like bubbles in someone's mind.

We also often use spatial concepts to understand other abstract concepts, such as time.

For instance, we talk about time in terms of motion: *time flies*, *time passes by*, *my birthday is in the coming week*, etc. We use dimensional expressions to specify the time, such as at *7 o'clock* and *in an hour*. We use the word "journey" to describe the progression of our life, both having a beginning, middle phases, and an end. Motion through space is one of our most fundamental bodily experiences. Accordingly, motion events are also considered one of the most basic types of events that we experience and encounter in our daily life (e.g., we have been observing the motion of everything around us since the day we were born, and what really signifies the start of our life journey is the first time we started to stand up and walk on our own). It is not surprising that motion events are also often used metaphorically for understanding the other more abstract events (e.g., *we are moving towards a better society*). Our bodily experiences are one of the most direct ways we use to make sense of the world.

Given the pervasiveness and the importance of motion events in our daily life, in this dissertation, my main focus is on investigating the understanding of simple motion events, such as situations like a person/a vehicle moving along a specific trajectory with a possible destination in the near distance. We know perfectly well what situations are considered motion events. However, what is the definition of a motion event, then? Many cognitive scientists, linguists, and psycholinguists have provided some definitions. Here, I adopt the one that was proposed by the linguist Talmy (2000), given that his definition is the most relevant to the current dissertation and is used the most often in cross-linguistic comparison studies (one of our chapters is a cross-linguistic study: **Chapter 3**).

Talmy (2000) has proposed that a motion event is mainly composed of the following semantic elements: figure, motion, path of motion, manner of motion, and ground. Take the following motion event as an example: *the ball rolled off the table*. In this motion event, *the ball* is the moving figure, *rolled* contains the information about the motion of the figure and the manner in which the figure moves, *off* encodes the information about the moving trajectory (the path of motion), and *the table* is the reference object with respect to which the moving trajectory of the figure is characterized. One extra component we would like to add is the temporal aspect of a motion event. We can infer from the verb suffix *-ed* (i.e., perfective aspectual maker in English) in *rolled* that this motion event is completed and has reached its temporal endpoint. In the following section, I will focus on elaborating another important part of this dissertation: language and its importance in event studies.

Language and events

In the opening of this general introduction, I used language to describe what I was seeing when looking outside the window of the school library. Just now, I also used a sentence (i.e., *the ball rolled off the table*) to instantiate a motion event. I believe that it is not difficult for the readers to imagine the depicted events themselves. We have to admit that, either through writing or speaking, language is one of the major ways we use to communicate and to share experiences with others. Using just a set of sentences (the combination of words and

grammar), we are able to reconstruct past and ongoing experiences, including the specific people, action, location, and temporal information of an event (i.e., happening or completed). We can manipulate all this information easily and effortlessly (i.e., by changing a few words or the syntactic structures). We can even create novel situations and stories that people have never heard of or seen before through language (e.g., sci-fi novels).

Given its primary role in conveying information about events, language has been taken as an important window into how we understand events. For instance, language comprehension studies provide important theoretical insights on how the semantic components of an event are processed and how an event model is built during language comprehension (e.g., Event-Indexing model in Zwaan, Langston, & Graesser, 1995). Moreover, language production studies offer us a platform on which we can study how an event is perceived and conceptualized through examining the usage patterns of certain linguistic expressions (Ericsson & Simon, 1993). Specifically, under what circumstances a certain linguistic expression is used to describe an event largely reflects the extent to which an event is conceptualized. Two types of linguistics cues are of special importance to the current dissertation: ***grammatical aspect*** and ***directional prepositions***.

Grammatical aspect

Grammatical aspect is the morphosyntactic marking of the verb for conveying information on the temporal status of an event. During language-based event comprehension, the use of progressive aspect highlights the progression of an event (e.g., the action of drawing conveyed in the sentence *he is drawing a picture*), whereas the use of perfective aspect leads the attention to the endpoint of an event (e.g., the completed picture implied in the sentence *he drew a picture*) (e.g., Madden & Zwaan, 2003).

One important fact is that not all languages have grammaticalized aspectual markers. For instance, unlike in English where progressive aspect is often used to describe ongoing events (e.g., *he is eating an apple*), in Dutch, most ongoing events are described with the unmarked simple present tense (e.g., *hij eet een appel* 'he eats an apple'). This typological difference among languages is found by some researchers to have an impact on motion event endpoint conceptualization (e.g., Athanasopoulos & Bylund, 2013; von Stutterheim, Andermann, Carroll, Flecken, & Schmiedtová, 2012).

Directional prepositions

Directional prepositions are often used together with the main verb to constitute a verbal phrase (e.g., *walk to school*). Whereas grammatical aspect specifies the temporal status of an event, directional prepositions usually contribute to the telicity of an event. A telic event is an event that has an inherent endpoint (e.g., *to eat an apple*, *to draw a picture*, *to run three miles*). An atelic event is an event that does not have an inherent endpoint (e.g., *to sleep*, *to run*, *to sing*).

Directional prepositions can accordingly be classified into two types: telic directional prepositions (e.g., *to*) and atelic prepositions (e.g., *towards*) (Krifka, 1998). An event that is described with a telic directional preposition is considered telic (e.g., *he is walking to the bus station*). Similarly, an event that is described with an atelic directional preposition is atelic (e.g., *he is walking towards the bus station*). The first example implies that the endpoint of the walking event is the bus station. The second example does not imply that the bus station is the endpoint. He may never reach the bus station, or simply walk past the bus station. Most languages have these two types of directional prepositions. For instance, Dutch has the telic directional preposition *naar* and the atelic directional preposition *richting*. The English equivalents of the two Dutch directional prepositions are *to* and *towards*.

Given that the use of directional prepositions implies whether an event boundary is identified, studying the choice between a telic and an atelic directional preposition in a language production task provides a window into motion event endpoint conceptualization.

An overview of the dissertation

In this dissertation, I mainly focus on event comprehension and event conceptualization in relation to the use of the above-mentioned two linguistic cues (i.e., *grammatical aspect* and *directional prepositions*), with a special focus on motion events. **Chapter 2** is a review article and provides an overview of a series of studies that have investigated the role of *grammatical aspect* in language-based event comprehension and in social cognition (e.g., problem-solving, decision making). **Chapter 3** is a cross-linguistic study in which we have compared Dutch speakers and mandarin Chinese speakers in the way they describe and conceptualize motion events, targeting the path of motion. This chapter also provides data on the usage patterns of *grammatical aspect* and *directional prepositions* in motion event descriptions in Dutch and in mandarin Chinese. **Chapter 4** investigates the semantic meaning of *two Dutch directional prepositions* (i.e., *naar* 'to' and *richting* 'towards'). Moreover, in this chapter, we have studied motion event endpoint conceptualization by examining the effects of two non-linguistic factors (i.e., the intention of the actor and the speech context) on the choice between the two Dutch directional prepositions. **Chapter 5** is an extension of **Chapter 4** by studying the effect of the two same non-linguistic factors on the choice between *two English directional prepositions* (i.e., *to* and *towards*) in a motion event description task, targeting motion event endpoint conceptualization. The effect of *grammatical aspect* on event comprehension (the sensitivity to the two non-linguistic factors) and its consequent effect on motion event endpoint conceptualization, which was not studied in **Chapter 4**, was examined in this chapter. **Chapter 6** discusses and summarizes the main findings of all the studies conducted in this dissertation (**Chapter 2** – **Chapter 5**). We have also discussed the limitations of this dissertation and pointed out directions for future research in this chapter.

What follows is a more detailed introduction to the main studies presented in this dissertation (**Chapter 2** – **Chapter 5),** including the relevant theoretical backgrounds, research questions,

study designs, and formed hypotheses.

Chapter 2: The role of grammatical aspect in and beyond event comprehension

Chapter 2 provides a review of a series of language comprehension studies that have investigated how grammatical aspect plays a role in and beyond event comprehension. By reviewing these studies, we were interested in examining the extent to which grammatical aspect modulates event representations during language comprehension and the extent to which this modulation effect goes beyond text comprehension and extends to higher cognitive processes which rely on text comprehension (social cognition such as problem-solving and decision-making).

We categorized those studies into three types based on the processing level that was studied in those studies: sentence-level processing, discourse-level processing, and social cognition that is based on text comprehension. We expected different roles of grammatical aspect regarding these three types of processing. For sentence-level processing, events are represented at a micro level. We expect that the use of grammatical aspect modulates the representations of the event structure per se, for instance, the action performed in the event, the location where the event occurs, the possible instrument used to perform the action, etc.

For discourse-level processing, events are represented at a macro level. The relations between events, that is, how events are related and developed within a certain discourse, are the main focus of this level of processing. We expect that the use of grammatical aspect should affect the availability of the depicted event, including the information about the protagonist, in a later context. Moreover, it should also affect how event models are updated in a later context and how events are segmented within this discourse.

Regarding social cognition that is based on text-based event comprehension, we expect that the effect of grammatical aspect on event representations does not necessarily extend to social cognition, such as problem-solving and decision-making. This is because there are many other factors that can outweigh the effect of grammatical aspect on affecting the process of solving a problem or making a decision.

Chapter 2 mainly focuses on the role of grammatical aspect in event comprehension during language comprehension. However, not all languages have grammaticalized grammatical aspect and languages differ in the extent grammatical aspect is used to describe events. In **Chapter 3**, we have conducted a cross-linguistic study and examined the extent to which Dutch speakers and mandarin Chinese speakers differ in the way they conceptualize and describe motion events. The typological difference in the use of grammatical aspect between the two languages is considered in this study.

Chapter 3: A cross-linguistic study of motion event description and conceptualization between Dutch and mandarin Chinese speakers

Despite the fact that language generally is the most prevalent way humans use to describe events, cross-linguistic differences in how events are encoded in different languages remain. As we have mentioned previously, not all languages have grammaticalized aspectual markers when specifying the temporal status of an event. This typological difference seems to affect people's conceptualization of a motion event. Specifically, in motion event description tasks, researchers discovered that speakers of non-aspectual languages (e.g., German, Dutch) were more likely to mention the endpoint of a directional motion event (e.g., *the car is driving* ***to the village***), compared to speakers of aspectual languages (e.g., English speakers; *the car is driving along the road*) (e.g., von Stutterheim et al., 2012).

There is another typological difference among languages that seems to also affect motion event conceptualization, specifically the way in which the path of motion information is encoded. For instance, languages, such as English and Dutch, typically encode the path information outside of the verb root (e.g., verb particles or prefixes; *the boy is walking* ***out of*** *the classroom*). Such languages are often classified as satellite-framed languages. On the other hand, languages, such as Spanish and French, often encode the path information in the verb root (e.g., *la botella* ***salió*** *de la cueva flotando* "the ball **exited** the cave floating"). Such languages are categorized as verb-framed languages (Talmy, 2000).

In **Chapter 3,** we compared mandarin Chinese speakers and Dutch speakers in the way they conceptualize the path information of a motion event. Two types of motion events were investigated, specifically ***endpoint-oriented motions events*** and ***endpoint-reached motion events***. For endpoint-oriented motion events, the figure in motion (person/vehicle) was designed to move along a trajectory in the direction of a clear location (e.g., a village, a playground), but this location was not reached at the endpoint of the video clip (e.g., a van driving towards a village; two girls walking towards a house). For endpoint-reached motion events, the figure in motion was designed to reach a goal/destination (e.g., a car driving into a garage; a man walking into a church). We designed a language production task followed by a surprising memory task. Participants had to describe these videos, and after that, they were tested about their memory of what they had seen before (specifically on the memory of a potential endpoint of an endpoint-oriented motion event).

Mandarin Chinese and Dutch differ in both their lexicalization patterns in encoding motion events and in their aspectual systems. First, Chinese has features of both satellite-framed languages and verb-framed languages, whereas Dutch is a typical satellite-framed language. Second, Chinese speakers use grammaticalized aspectual markers (e.g., the progressive marker *-zai* and the perfective marker *-le*) to describe directional motion events, whereas Dutch speakers do not. As we have mentioned before, Dutch speakers often use the unmarked simple present tense to describe ongoing directional motion events. Besides, there

is no perfective marker in Dutch.

We were particularly interested in (1) How the path of motion was ***conceptualized*** by mandarin Chinese speakers and by Dutch speakers and (2) How motion events were ***described*** (with a special focus on the descriptions that contain the path information of a motion event) by mandarin Chinese speakers and by Dutch speakers (i.e., the combination of verbs and adjuncts and the combination of verbs and aspectual markers).

Given the typological differences between mandarin Chinese and Dutch, we expected that (1) for ***endpoint-oriented motion events***, there should be a difference in the frequency of the mention of the endpoints between Chinese and Dutch speakers: Dutch speakers should mention endpoints more often than Chinese speakers, and consequently, Dutch speakers should remember the endpoint information better than Chinese speakers in the surprising memory task. Moreover, there should be a difference in the selection of trajectory versus location information for motion events conceptualization. (2) ***For endpoint-reached motion events***, both Chinese and Dutch speakers should mention the endpoint more often than the often path elements (i.e., trajectory and location), given the salience of the endpoint information in a boundary-crossing event.

Furthermore, we expected that (3) for both types of motion events, differences in the use of verb types in the combination of adjunct types between mandarin Chinese speakers and Dutch speakers should occur, and (4) that in Chinese, the perfective marker (*-le*) should be used more frequently with path verbs and serial verb constructions, whereas the progressive marker (*-zai*) should be used more frequently with manner verbs.

A serendipitous finding from **Chapter 3** concerns the use of two Dutch directional prepositions (i.e., *naar* 'to' and *richting* 'towards'). We found that the two prepositions were both used frequently by Dutch participants when they were describing a directional motion event. However, the two prepositions differ in their semantic implications in relation to the certainty of the referential object being the endpoint of the motion event. The choice between the two prepositions implies how a motion event endpoint is conceptualized during language production. Hence, in **Chapter 4**, we focus on examining the factors that affect the choice between the two prepositions in a motion event description task, targeting motion event endpoint conceptualization.

Chapter 4: Examining motion event endpoint conceptualization: the choice between two Dutch directional prepositions

Naar is a goal preposition, which means that the use of *naar* implies the reference object is the endpoint of a motion event (e.g., *hij loopt naar de school* 'he walks to the school'). On the other hand, the use of *richting* does not designate the reference object as the endpoint of a motion event but only the direction of motion (e.g., *hij loopt richting de school* 'he walks towards the school'). In **Chapter 4**, we were interested in confirming this semantic difference

between the two directional prepositions in a language comprehension task (Experiment1). Moreover, we were interested in examining the factors that would affect the choice between the two directional prepositions in a language production task, targeting motion event endpoint conceptualization (Experiment 2).

In Experiment 1, we designed a slider-dragging task. On the left side of the slider, a figure was presented (either a human or a vehicle), and on the right side of the slider, a referential object was presented. Participants were asked to perform a language comprehension task, that is, to read sentences that either used the preposition *richting* or the preposition *naar* (e.g., *de man loopt richting/naar de kerk* 'the man walks towards/to the church'), and then to locate the figure on the slider based on the sentences they had just read. The control condition was that the referential object was reached as the endpoint of the motion event (e.g., *de man loopt de kerk in* 'the man walks into the church). Given the semantic differences between the two prepositions, we hypothesized that when *naar* was used in the description of a motion event, participants should locate the figure closer to the referential object, compared to when *richting* was used to describe a motion event. For the control condition, participants should locate the figure the closest to the referential object.

In Experiment 2, we investigated the factors that would have an impact on the choice between the two Dutch prepositions. Particularly, we examined how ***the actor's goal*** **(*Intention*)** and ***the interlocutor's social status*** **(*Interlocutor*)** affect the choice between the two directional prepositions in a language production task. We expected that (1) if the goal of the actor was clearly shown in the referential scenario, participants should choose to use the preposition *naar* more often when describing the referential scenario, compared to when the goal of the actor was not clearly shown. The use of *richting* should show an opposite pattern. Moreover, we hypothesized that (2) if the social distance between the interlocutor and the speaker was larger (i.e., when the interlocutor was informed as a police officer), participants should use *richting* more often, compared to when the social distance between the interlocutor and the speaker was small (i.e., when the interlocutor was informed as a friend of the speaker). The use of *naar* should exhibit an opposite pattern. In Experiment 3, we increased the strength of the Intention and expected to find a stronger effect of Intention.

In **Chapter 4**, we investigated motion event endpoint conceptualization by examining the factors that affect the choice between two Dutch prepositions (i.e., *naar* and *richting*). **In Chapter 5**, we switched our attention to their English equivalents (i.e., *to* and *towards*) for the purpose of extending the findings in **Chapter 4** in another language.

Chapter 5: Linguistic and non-linguistic cues in motion event endpoint conceptualization: the choice between two English directional prepositions

As we have mentioned previously, Dutch speakers only use the unmarked simple present to describe an ongoing directional event (e.g., *hij* ***loopt*** *naar het station* 'he **walks** to the station').

However, in English, there is a grammaticalized progressive marker that is the most used in English to express ongoing events, including ongoing directional events (e.g., he *is walking* to the station). The simple present tense in English can also be used to express ongoing events but brings a different interpretation to the depicted event, compared to the use of the English progressive marker.

Therefore, in **Chapter 5**, our first goal was to extend the findings of **Chapter 4** when English directional prepositions were studied. Our second goal was to go beyond **Chapter 4** by investigating the effect of grammatical aspect on the sensitivity to the two factors (i.e., Intention and Interlocutor) during event comprehension, and its consequent effect on motion event endpoint conceptualization.

In general, all of the studies were conducted in the context of language-based event comprehension (language comprehension tasks) or language-based event conceptualization (language production tasks). I aim to understand how language plays a role in event comprehension and in event conceptualization. I also aim to understand how we comprehend and conceptualize events in general.

Chapter 2

The role of grammatical aspect in and beyond event comprehension

This chapter has been submitted for publication:

Liao, Y., Dijkstra, K., & Zwaan, R.A. (submitted). The role of grammatical aspect in and beyond event comprehension

Abstract

Grammatical categories contribute fundamentally to language-driven event comprehension. There is an abundance of literature regarding the role of grammatical aspect in relation to event comprehension. Moreover, a growing body of literature examines whether the effect of grammatical aspect on event comprehension extends to social cognition, such as decision-making. We review those studies from three main perspectives: the role of grammatical aspect in sentence-level event comprehension, the role of grammatical aspect in discourse-level event comprehension, and the role of grammatical aspect in higher cognitive processes that are based on language-driven event comprehension. Our major conclusions are (1) grammatical aspect has a robust effect on the representations of the internal structure of an event (i.e., the temporal development of an event) during sentence comprehension; (2) grammatical aspect plays a significant role in the construction and updating of the situation models during discourse comprehension; (3) the effect of grammatical aspect does not necessarily extend to social cognition; (4) we should be aware that grammatical aspect interacts with lexical aspect, world knowledge, other linguistic factors, and contextual factors; (5) most of the studies focused only on the English language, but a more diverse set of languages should be investigated in this future.

Keywords: Grammatical aspect; sentence comprehension; discourse comprehension; social cognition

Introduction

We participate in and are surrounded by events every day. We constantly learn about what other people did in the past and what they are doing at present. This information comes to us via stories told by friends and writers past and present. Language is an effective tool to communicate about our past and current experiences. It allows us, as speakers or writers, to construct a series of events without much effort. We can manipulate how we portray these events in subtle ways. For instance, we can use a prepositional phrase (e.g., *into a pond*) to indicate a change in space or a temporal phrase (e.g., *an hour later*) to mark a time shift. As producers of language, we include such cues in our verbal output to steer the comprehension process. The role of the comprehender is to interpret these cues effectively to construct a mental representation of the events that are described. How we comprehend an event depicted in language reflects to a large extent how we understand an event in real life. Therefore, an important focus of language comprehension research has targeted the subject of event comprehension (e.g., Radvansky & Zacks, 2014; Zacks, 2020; Zwaan & Radvansky, 1998).

An event has both spatial and temporal characteristics. It is often viewed as a segmented piece of our experiences with the world, which has a beginning and an end in time (Zacks & Tversky, 2001). Our current review focuses on how the temporal contour of an event is represented during language comprehension. In particular, we target the specific linguistic cue that grammatically marks the temporal development of an event: grammatical aspect. It is the morphosyntactic marking of the verb to express the internal temporal structure of an event (e.g., beginning, ongoing, or completed; Comrie, 1976; Dowty, 1979; Langacker, 1987; Smith, 1997) and has caught many cognitive psychologists' attention in language comprehension research. Linguists tend to make a general distinction between imperfective aspect and perfective aspect (Langacker, 1987). Imperfective aspect is used to target the initial and intermediate stages of an event, that is, the progressive states of an event (e.g., the sentence *the woman is writing a letter* focuses on the process of writing the letter, possibly including the protagonist, the materials used to write, the action of writing). In contrast, perfective aspect is used to only focus on the end state of an event (e.g., the sentence *the woman wrote a letter* signifies the completion of the writing process, and the end state is the finished letter).

The majority of language comprehension research focuses either on sentence-level comprehension or discourse-level comprehension. This also applies to studies that have investigated the role of grammatical aspect in event comprehension. In particular, with regard to event comprehension at the sentence level, the primary focus is on the representations of a single event and the event structure per se. Event comprehension at the discourse level, on the other hand, is tantamount to understanding the relations and connections between events (Zwaan, 2008). Therefore, in the present review, we mainly classify and discuss studies in terms of those two research scopes (i.e., sentence vs. discourse comprehension).

Moreover, we review the results of the studies mainly from the perspectives of two major theoretical frameworks in language comprehension research, specifically situation models and mental simulations. Whereas the former is often applied to discourse comprehension, as it focuses primarily on the relations among events, the latter is mainly applied to sentence comprehension, as it focuses primarily on the internal structure of events. In the following section, we briefly discuss the two types of theories and how their research focuses differ.

Situation-models and mental simulation theories

Over the past four decades, situation-model theories have developed rapidly, which has led to a significant contribution to language comprehension research (Johnson-Laird, 1983; Kintsch, 1998; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983; Zwaan, 2016; Zwaan & Madden, 2005; Zwaan & Radvansky, 1998). The main idea behind situation-model theories is that understanding a string of connected sentences does not occur merely as a result of extracting abstract linguistic propositions from the text itself but involves constructing a mental representation of a state of affairs that is described in the text, a situation model, by combining text-based information with background information from the comprehender's long-term memory. The basic units of a situation model are not linguistic constructions but event representations (Zwaan & Radvansky, 1998).

In the event-indexing model (Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998), events are connected by at least five dimensions: space, time, protagonist, causation, and intentionality. In the process of narrative comprehension, new information about an event is encountered with the incoming of new clauses. The new information is either incorporated into the already-existing situation model based on their overlap on any of the five dimensions or leads to the updating of the current situation model if any change is detected in one of the five dimensions in the new information (Zwaan, 2016). Take the following narrative as an example: *A man was wandering in the park. He saw some sheep and wild deer in the park. An hour later, he left the park and entered a pub on the street*. The first two sentences are very likely to be included in the same situation model, given that they overlap in terms of both the temporal and spatial dimensions (i.e., still at the time of wandering and in the park). When the last sentence is encountered, there are obvious changes in both time and space (i.e., *an hour later* and from the park to a pub on the street). Consequently, in order to accommodate these changes, the information that depicts the man wandering in the park and what he saw there is outdated and needs to be discarded. The current situation model is, therefore, updated.

The primary focus of situation-model theories is on the relations between events, which is a macrolevel of event representation (Zwaan, 2008). Unlike situation-model theories, mental simulation theories target the representation of the event structure per se, which is a microlevel of event representation. In particular, according to mental simulation theories, an event representation is not an abstract and independent existence but is embedded in our sensorimotor experiences (Barsalou, 1999). During language comprehension, we simulate

elements of the text, for example, an object's orientation (Stanfield & Zwaan, 2001), shape (Zwaan, Stanfield, & Yaxley, 2002), size (de Koning, Wassenburg, Bos, & van der Schoot, 2017), color (Connell & Lynott, 2009; Hoeben Mannaert, Dijkstra, & Zwaan, 2017; Zwaan & Pecher, 2012), and even body movement (Borghi & Scorolli, 2009; Gentilucci, Benuzzi, Bertolani, Daprati, & Gangitano, 2000; Zwaan & Taylor, 2006) by activating our previous sensorimotor experiences that are stored in modality-specific systems of the brain. Whereas situation-model theories target mainly discourse comprehension, mental simulation theories may target the comprehension of a single word, phrase, or sentence.

Grammatical aspect and language comprehension research

Languages have various means to present the internal temporal structure of an event, such as lexical morphemes (e.g., *start*, *continue*, and *finish*), constructions (e.g., the *aan-het* construction in Dutch), and grammatical aspect (e.g., the progressive marker *-ing* in English). Therefore, grammatical aspect is not the only way of expressing the temporal development of an event. There are also cross-linguistic differences in the way the temporal structure is typically encoded. For instance, English is considered an aspectual language, given that the ongoingness of an event is grammatically marked by a progressive marker (i.e., the verb inflection *-ing*; see *he is walking to the station*) in English. However, Dutch is often considered a non-aspectual language (Flecken, 2011). The expression of ongoingness in Dutch is often not marked. It has progressive constructions, such as the *aan-het* construction (e.g., *hij is aan het wandelen in het park* 'he is taking a walk in the park'), but the use of this construction is situation-type constrained (Flecken, 2011). For example, when expressing an ongoing directional motion event in Dutch, the *aan-het* construction is often not used. Using only the simple present tense suffices to express an ongoing directional motion event (e.g., *hij loopt naar het station* 'he walks to the station'; see Flecken, 2011; von Stutterheim, Carroll, & Klein, 2009).

Why is grammatical aspect important and has it received so much attention in language comprehension research, then? First, despite the fact that not all languages grammatically mark an event's temporal development, there are still many languages, including English, that employ grammatical aspect as the major way of expressing the temporal development of an event. Not surprisingly then, language comprehension studies involving English sentences or narratives have focused on studying aspectual markers. Second, language comprehension research demonstrates that both lexical morphemes (e.g., nouns and verbs) and grammatical morphemes (e.g., aspectual markers) play an equivalently important role in language processing. Whereas lexical morphemes provide elements of a situation (e.g., objects, actions), grammatical morphemes structure those elements into a coherent situation. One important function of grammatical aspect is to specify what aspect of a situation should be focused on and what perspective should be used to view the situation (Morrow, 1986). Third, the subtle difference in aspectual framing (e.g., whether an event is framed as ongoing or completed) appears to have a great impact on how the depicted event is remembered and

recalled at a later time (e.g., Magliano & Schleich, 2000), which also has consequences for how people predict future events (e.g., Ferretti, Rohde, Kehler, & Crutchley, 2009). This means that aspectual markers are especially important for discourse comprehension. Given the above-mentioned reasons, it is not surprising that grammatical aspect has obtained much attention in language comprehension research.

It should be noted that a distinction between grammatical aspect and lexical aspect is often made in the linguistic literature (e.g., Comrie, 1976; Vendler, 1967). Lexical aspect targets the telicity of an event, that is, whether an event has an inherent boundary or not. Consider for example, the contrast between *the woman is writing a letter* and *the woman is singing*. The former sentence describes a telic event with a specified event boundary, that is, the finished product of a letter, whereas the latter sentence describes an atelic activity that does not have a clear endpoint. Unlike lexical aspect, grammatical aspect modulates the way the internal temporal structure of an event is being viewed, namely through an internal viewpoint provided by imperfective aspect versus an external viewpoint provided by perfective aspect (Smith, 1991; Vendler, 1967). The sentence *the woman is writing a letter* expresses a telic event that has an event boundary (i.e., to write a letter) but is temporally unbounded and is viewed internally (i.e., the progress of the event of writing a letter; Declerck, 2007). The focus of the current review is on grammatical aspect, but sometimes lexical aspect will also be discussed when it interacts with grammatical aspect in modulating an event representation.

The objective of the review

Researchers have conducted a considerable number of empirical studies to examine the role of grammatical aspect in event comprehension (see, for example, Carreiras, Carriedo, Alonso, & Fernández, 1997; Ferretti, Kutas, & McRae, 2007; Madden & Zwaan, 2003; Morrow, 1985, 1990). Studies differ in their unit of analysis (i.e., sentence-level processes vs. discourse-level processes), research target (e.g., protagonist, location, action), grammatical aspectual pairs investigated (e.g., past progressive vs. past tense, or past progressive vs. past perfect), languages investigated (e.g., English, French, and Mandarin Chinese), and the factors that interact with grammatical aspect during language comprehension (e.g., lexical aspect, world knowledge, and context). Some studies have even extended their research on grammatical aspect to investigate its effect on social cognition (see, for example, Fausey & Matlock, 2011; Sherrill, Eerland, Zwaan, & Magliano, 2015).

To the best of our knowledge, no review has been conducted to classify and summarize those studies by covering and discussing all the above-mentioned topics. A comprehensive review could, therefore, offer a clear picture of how grammatical aspect has been investigated from many different angles in language comprehension research. Another objective of the review is to clearly answer the question regarding the effect of grammatical aspect on event comprehension and whether this effect goes beyond event comprehension and extends to social cognition, like decision-making. In the following section, we will explain the

methodology we used to include studies and the standards of the eligibility of the studies included in this review article.

Methodology

Data collection

Our first step was to search for relevant studies through online databases. Formal searches were conducted in the Web of Science and Scopus databases using the search term "grammatical aspect" or "verb aspect". For both databases, our searches were restricted to the title, abstracts, and keywords sections of a paper, and we set no limit to publication years. Moreover, we excluded studies that had the terms "second language" and "acquisition" in the title, abstracts, or keywords sections[1]. We found 192 publications in the Web of Science and 294 publications in Scopus. Our searches were performed on 12 February 2021.

At this stage, we did not restrict studies to language comprehension studies only. By doing so, we aimed to ensure that all studies relevant to grammatical aspect were included first. Since we were certain that second language studies and acquisition studies were not within the scope of the current review article, we excluded them at this stage to save time and energy for the next abstract screening stage.

Abstract screening

Our second step was to review all the abstracts of the studies found in the two databases. The goal was to find all the studies that empirically investigated grammatical aspect within the scope of language comprehension research. Therefore, any studies that focused on a theoretical analysis of grammatical aspect or compared cross-language differences in their aspectual systems were excluded. Furthermore, language production studies on grammatical aspect were excluded.

To increase our efficiency and reduce human bias in selecting relevant studies, we adopted the ASReview software (van de Schoot et al., 2021) to facilitate the abstract screening process. ASReview has been recently developed by a research team from Utrecht University to aid abstract screening via incorporating active learning using artificial intelligence. It provides an interactive interface in which the reviewer selects relevant studies, and a machine learning model is trained based on the reviewer's choices to predict what abstract will be presented next. This next presented abstract is always chosen by the model based on its degree of relevance to the selected studies. This helps to increase the speed of abstract reviewing and avoid random selection processes on the part of the writers of the reviewer, which is likely to

1 The search syntax we used in Web of Science is: TOPIC: ("grammatical aspect" OR "verb aspect") NOT TOPIC: ("second language") NOT TOPIC: (acquisition), and the search syntax we used in Scopus is: (TITLE-ABS-KEY ("grammatical aspect" OR "verb aspect") AND NOT TITLE-ABS-KEY ("second language") AND NOT TITLE-ABS-KEY (acquisition)).

be biased. Most importantly, the abstract reviewing process is stored as a project in ASReview and is reproducible.

Before reviewing abstracts in ASReview, we removed duplicates from the publications we found in the databases through the literature managing software Zotero. After removing duplicates, we obtained 379 publications (out of the 486 publications). We then imported all these publications into ASReview and started the abstract reviewing process. After reviewing around 120 publications (out of the 379 publications), we obtained 51 relevant publications. The rest of the publications were all irrelevant[2] according to our criteria (see the full ASReview results in supplementary material 1, and supplementary material 2 is the ASReview project that has recorded the whole abstract screening process).

Full-text screening

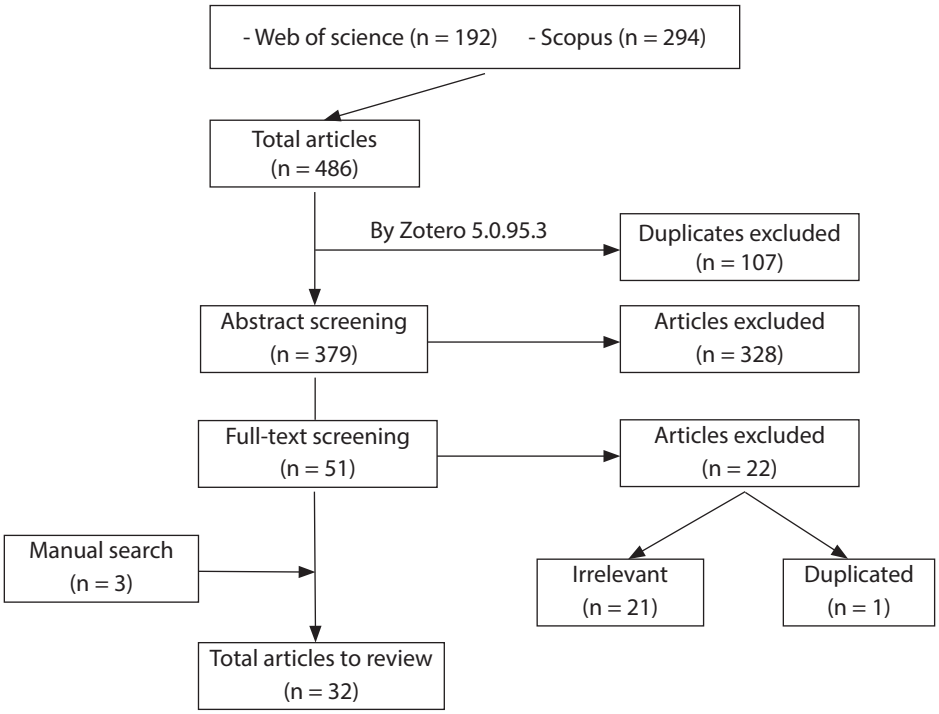
Our last step of the studies' selection was to do a full-text screening for the 51 publications obtained after the abstract screening process. At this step, we reviewed all the 51 publications by carefully reading their full texts. We removed articles that were still duplicates (n = 1) or were actually irrelevant (e.g., reviews, theoretical articles, acquisition studies with only children's data, production studies, and studies that only involve verb tense) but skipped away under our nose during the abstract reviewing process (n = 21).

After we read the full texts of all the 51 publications, we obtained 29 publications for the current review article. Moreover, we added three publications that we know are relevant but were not found when we did the searches in the databases. In total, we included 32 publications in this review article (see a full list of chosen articles in supplementary material 3). Figure 1 presents a flow diagram of the studies' step-by-step selection and screening processes. The structure of this flow diagram was adapted from the one provided in Tawfik et al. (2019).

2 This means that after reviewing around 32% of the whole dataset, we already obtained all relevant studies, which corresponds to the simulation results provided by ASReview: https://asreview.readthedocs.io/en/latest/guides/simulation_study_results.html. All 379 publications were reviewed to ensure all relevant studies were found.

Figure 1.

A flow diagram of the studies' selection and screening processes.



The current review

All the 32 publications were classified into three sections for discussion based on whether they 1) target sentence-level processes (*n* = 16), or 2) focus on discourse-level processes (*n* = 9), or 3) involve language-driven event comprehension but also go beyond it (i.e., cognitive processes, such as problem-solving and decision-making; *n* = 7). Sub-topics of each section were defined based on the following themes: which part of an event representation was investigated during sentence processing (e.g., endpoint, instrument, location, action) or which event information was targeted during narrative comprehension (e.g., the protagonist or other event-related information), the languages investigated, and the aspectual pairs studied (for an overview of the classification of all publications, see supplementary material 3).

Sentence-level processes

Studies that have investigated event comprehension at the sentence level have been mainly devoted to examining how the internal structure of an event is represented during sentence comprehension. According to Moens and Steedman (1988), the basic event temporal structure covers three parts, the initial stage, the intermediate stage, and the result state (Zwaan, 2008). An abundance of empirical evidence has demonstrated that grammatical

aspect constrains the mental simulation of an event (including both sensory and motor properties) via its function in specifying the current temporal stage of the event (e.g., ongoing or completed) (e.g., Madden & Therriault, 2009; Madden & Zwaan, 2003; Matlock, 2011). In the following subsections, we provide a review of how these studies have reached this conclusion regarding the modulation effect of grammatical aspect on event representations during sentence comprehension. We classify and discuss those studies in terms of the similarity of their research content, for instance, whether they target the representation of a whole event's temporal stage or a certain event feature, such as the instrument used to perform an action, the location where an event occurs, or the action that is performed in an event.

Ongoing versus completed stages of an event

An event can be depicted as either ongoing or completed through the use of different aspectual markers (e.g., *The man was making a fire* vs. The *man made a fire*). Several studies have demonstrated that during sentence comprehension, the visual representations of an event vary depending on which stage of the event is being specified by the aspectual markers (i.e., ongoing or completed). A representative study was conducted by Madden and Zwaan (2003). They employed the sentence-picture verification task that was originally developed by Stanfield and Zwaan (2001) and Zwaan et al. (2002).

Specifically, Madden and Zwaan (2003) first presented participants with a sentence that was either in its perfective form (e.g., *The man made a fire*) or in its imperfective form (e.g., *The man was making a fire*). Then, a picture that either showed the end state of the depicted event (e.g., the presence of the fire) or the intermediate state of the event (e.g., the man putting logs in the fireplace but no presence of the fire) was shown. Subsequently, participants answered "yes" or "no" in response to whether the picture matched the sentence. The correct answer would always be "yes" to the critical items, given that on a global level, the referred event in the picture always matched the event depicted in the sentence. Longer response latencies were expected when, for instance, a perfective sentence was followed by a picture that showed the intermediate state of the depicted event. Indeed, Madden and Zwaan (2003) found that when perfective sentences were presented, participants responded slower when answering "yes" to the pictures that showed the intermediate state of the event compared to the pictures that showed the end state of an event. In contrast, when the pictures were preceded with imperfective sentences, the matching effect was not detected. That is, there was no significant difference in the reaction times between in-process pictures and completed pictures.

Madden and Zwaan (2003) explained that the lack of the matching effect of imperfective aspect did not necessarily mean that participants did not simulate the event's ongoing stage after reading an imperfective sentence. They postulated that, instead, it might be due to the difficulty of capturing the appropriate intermediate state of the event. Whereas a clear end state of an event would be easy to demonstrate (e.g., the appearance of the fire), it would be nearly impossible to catch an ongoing state of the event through a picture stimulus

that could ideally match everybody's mental simulation of an event's progressive stage. Moreover, the authors also mentioned that the picture stimuli used in this study to express the completion status of an event actually also provided cues for the ongoing state of the event, for example, the presence of the agent, the instrument, and the location (e.g., the man holding a match next to the fireplace). All these cues provided in the completed picture could increase the chance that participants interpreted it as presenting an ongoing event still.

Unlike Madden and Zwaan (2003), who only presented participants with one picture stimulus at each trial (either an in-progress picture or a completed picture), Zhou, Crain, and Zhan (2014) employed a visual world eye-tracking paradigm and tested participants' eye movements toward the two kinds of pictures at the same time. With a clear comparison shown between an in-progress picture and a completed picture, Zhou et al. (2014) found that upon hearing a perfective sentence, participants significantly moved their eyes toward the completed picture area more often compared to the in-progress picture area. In contrast, upon hearing an imperfective sentence, participants were significantly more likely to move their eyes toward the in-progress picture area rather than the completed picture area. Their study demonstrated that when a clear comparison was presented between the two types of picture stimuli, participants could utilize the provided perceptual cues to unambiguously comprehend both imperfective and perfective sentences.

A series of other studies have also investigated the visual representations of motion event descriptions in which either perfective aspect or imperfective aspect was used (Anderson, Matlock, Fausey, & Spivey, 2008; Anderson, Matlock, & Spivey, 2010; Anderson, Matlock, & Spivey, 2013). Using a mouse-tracking technique, Anderson et al. discovered that for a perfective motion event description (e.g., *he jogged to the woods*), the character was placed nearer to the endpoint location (e.g., the woods) compared to an imperfective motion event sentence (e.g., *he was jogging to the woods*). In the latter case, the character was often located "in the middle" of the path of the motion event. These studies demonstrate that for a motion event that has an implied endpoint (*to* is a goal preposition; Eschenbach, Tschander, Habel, & Kulik, 2000), the use of perfective aspect highlights the completion status of the event; hence, the endpoint of the motion event is at the focus of the visual representation of this event. On the contrary, the use of imperfective aspect allocates the comprehender's attention to the ongoing phase of the event, which highlights the path/trajectory of the motion event in the visual representation of the event.

All the studies discussed above mainly focused on events with a natural endpoint (i.e., accomplishments; e.g., *make a fire*, *drink a cup of juice*, *jog to the woods*). Yap et al. (2009) argued that lexical aspect could interact with grammatical aspect in sentence processing when perceptual cues are tested, including those without natural endpoints. They employed a similar experimental paradigm as used by Madden and Zwaan (2003), but they distinguished between event types denoting accomplishments (i.e., durative and have a natural endpoint) and event types denoting activities (i.e., durative but have no specific natural endpoint; e.g.,

swim, sing). Moreover, they also presented both the completed and the ongoing picture stimuli simultaneously like Zhou et al. (2014) did in their study. They found that for events that had a natural endpoint (i.e., accomplishments), participants were significantly more accurate and quicker to respond when matching the perfective sentence to the completed picture than when matching the imperfective sentence to the in-progress picture. In contrast, for events that did not have a natural endpoint (i.e., activities), participants responded significantly more accurately and faster when matching the imperfective sentence to the in-progress picture compared to when matching the perfective sentence to the completed picture. These results suggest that there is compatibility between perfective aspect and accomplishments (i.e., the completion status of an accomplishment is implied by its inherent endpoint) and between imperfective aspect and activities (i.e., activities are processes that are inherently not bounded). Consequently, perfective aspect facilitates the process of accomplishments but brings processing cost to activities, whereas imperfective aspect helps to process activities but not accomplishments. This processing cost is significantly noticeable in the reaction times.

So far, the studies discussed above targeted the visual representations of the completion status versus the ongoing status of an event. However, only one of them has delved deeper into the relation of grammatical aspect and the degree of completion status. This was a study conducted by Kaplan, Raju, and Arunachalam (2021) that has provided implications for the extent to which completion types (i.e., a fully-completed end state vs. a partially-completed end state of an event) were perceptually represented in response to the aspectual form of the verb. Specifically, the authors used the visual world paradigm to compare participants' eye movements toward between a fully-affected object (e.g., a plate with only a few cookie crumbs on it) and a partially-affected object (e.g., a plate with part of the cookie on it) when they were processing either an imperfective sentence (e.g., *she was eating the cookie*) or a perfective sentence (e.g., *she has eaten the cookie*). The results demonstrated that when participants heard a verb in its perfect form (e.g., *has eaten* from the sentence *she has eaten the cookie*), they arrived at a fully-completed interpretation significantly more often than a partially-affected interpretation. Specifically, participants looked at a fully-affected referent more often compared to a partially-affected referent. A partially-completed interpretation of a perfect sentence was also detected, but it occurred at a later stage, one second after the play of the sentence. In contrast, when readers heard a verb in its progressive form (e.g., *was eating* from the sentence *she was eating the cookie*), there was no difference in the eye movements toward between a fully-affected referent and a partially-affected referent. These results indicate that both completion types are equally available for the interpretation of a progressive event, which corresponds to the claim made in Madden and Zwaan (2003) that any stages of an event could be relevant to the progression of the event. However, a fully-affected referent is a more default and natural option than a partially-affect referent for a visual representation of a completed event.

Instrument information

The studies discussed above targeted the role of grammatical aspect on the visual representations of an event as a whole, focusing on the completion status and the ongoing stage of the event. Some other studies, however, only focused on the visual representations of a certain event feature, for example, the instrument used by the protagonist when performing an action (e.g., an implied instrument of "hammer" in the sentence *she was fixing a birdhouse*). Instrument information is especially more relevant to the ongoing stage of an event than the completed stage of the event. Researchers have discovered that imperfective aspect activates the perceptual simulation of instrument information, whereas perfective aspect does not.

For example, Madden and Therriault (2009) investigated whether the use of grammatical aspect would affect the perceptual simulation of the instrument information and subsequently affect sentence comprehension. They replaced the instrument information in the sentence stimuli (e.g., *John was working/had worked on his laptop at home*) by either an in-use picture of it (e.g., an opened laptop) or a not-in-use picture of it (e.g., a closed laptop). They then tested the reading times of the first and the second word after the picture, respectively. They found that the reading times of each of the two words were faster when an in-use picture was employed in an imperfective sentence compared to when a not-in-use picture was used. This effect was also detected in the later sentence sensibility judgment. However, perfective sentences did not show this in-use effect. Their findings suggest that imperfective aspect activates the simulation of the instrument being in use and helps to maintain this in-use effect to facilitate further sentence processing. However, perfective aspect does not activate the simulation of an instrument in use, and hence the in-use effect is attenuated.

Madden, Dominey, and Ventre-Dominey (2017) used a similar paradigm as in Madden and Therriault (2009), but they focused on the visual representations of both the instrument and the recipient object information during sentence comprehension. For example, in the sentence *John was using a corkscrew to open the bottle in the restaurant*, a picture of the corkscrew (either in-use or not in-use), and a picture of the bottle (either in its ongoing state: half-open or in its completed state: fully open) were both used to replace the corresponding two words in the sentence. Moreover, instead of investigating how grammatical aspect affects the processing speed of the words following the pictures, they focused on the effect of grammatical aspect on the processing speed of the pictures directly. By doing so, they built a direct connection between grammatical aspect and the visual representations of the event. The study conducted by Madden et al. (2017) demonstrates that imperfective aspect tended to exhibit a congruence effect on both the in-use instrument picture and the object picture in its ongoing state, but this effect was marginal and not significant. They explained that this might be due to the less constrained temporal stages expressed by imperfective aspect (a similar point of view was mentioned in Madden & Zwaan, 2003). The significant

congruence effect came from the affected objects (the recipient object in its completed state) in the perfective sentences, which again corresponds to the view that perfective aspect highlights the end/result state of an event. The study conducted by Kaplan et al. (2021) could also provide an explanation for why the ongoing status of the affected object (e.g., a half-opened bottle) showed no significant congruence effect with imperfective aspect. That is, the representation of an ongoing event is more tolerant with the completion status of an affected object, which could allow either a half- or fully-affected representation compared to the representation of a completed event.

Another study adopted a mouse-tracking technique and studied the effect of grammatical aspect on activating an implied instrument during sentence comprehension in Persian (Golshaie & Incera, 2020). In this study, participants first read a sentence (e.g., *Sara is slicing the zucchinis*, then a picture of an implied instrument (e.g., knife) appeared, and participants answered whether the name of the picture was present in the previously-read sentence by pressing one of the two response buttons (PRESENT and ABSENT). All critical answers should be ABSENT. They found that imperfective aspect resulted in significantly more deviations of the mouse trajectories toward the PRESENT response button than perspective aspect. This demonstrated that imperfective aspect activated the perceptual representation of the implied instrument information, which could be misleading to the participants for a short while before they realized that it was not explicitly mentioned in the sentence.

Location information

Like the instrument information, location information (i.e., the location where an event occurs) is another event feature that should be more relevant to the ongoing phase of an event than the completed phase of an event. The study conducted by Ferretti and colleagues (Ferretti et al., 2007) demonstrates that the typical location information of an event was indeed activated only when the verbal form was imperfective rather than perfective. Both behavior (Experiment 1 and 2) and electrophysiological (Experiment 3) measures were employed in this study. In Experiment 1, participants first read a verb and then named the noun (location) after the verb aloud. The verb and the noun were either semantically related (e.g., *was skating – arena*) or not (e.g., *was praying – arena*). Using this semantic-priming paradigm, they found that participants named the location noun faster when it was semantically related to the verb and when the verb was in its imperfective form (e.g., *was skating*). This relatedness-facilitation effect was not found when the verb was in its perfective form (e.g., *had skated*). This finding shows that location information is more accessible when imperfective aspect is used rather than perfective aspect. In Experiment 2, the authors used a sentence completion task and found that participants completed the imperfective sentence (e.g., *the cow was grazing _*) with locative phrases more often than when the sentence was perfective (e.g., *the cow had grazed _*). In the third experiment, they employed the Event-related Brain Potential (ERP) method to examine participants' online comprehension of sentences with location information that was highly expected with the verb (e.g., *the girl was hiding/had hidden in the closet*) and sentences

with location information that was semantically less expected (e.g., *the girl was hiding/had hidden in the field*). The N400 response in ERP indicates the semantic expectancies of a word in a given context (N400 is larger to semantically less-expected concepts). They found that only imperfective sentences triggered a larger N400 when the location information was less expected compared to a highly-expected location. Perfective sentences did not show any expectancy difference. Those results indicate that perfective aspect suppresses the activation of location information, whereas imperfective aspect makes the location information more accessible during sentence comprehension.

It should be noted that the results of this study do not necessarily imply that the location information is perceptually simulated by imperfective aspect during language comprehension, given that no visual stimuli or other perceptual stimuli were used in this study. However, the results clearly demonstrate that the typical location information is part of an event representation when the event is depicted as ongoing rather than completed. In combination with the results from the previously-discussed studies (e.g., Madden & Zwaan, 2003), a prediction can be made from Ferretti's study (2007) that the same effects of grammatical aspect on the activation of typical location information would be detected if the locative phrases/words were replaced by their pictorial forms.

Action and motor simulation[3]

The above-discussed studies mainly targeted visual representations of an event during sentence comprehension. There is evidence that not only perceptual simulations are activated during language processing, but relevant motor experiences can be simulated as well (e.g., García et al., 2019; Zwaan & Taylor, 2006). When it comes to the role of grammatical aspect in event comprehension, some studies have provided evidence that grammatical aspect simulates motor experiences during sentence comprehension.

For example, Bergen and Wheeler (2010) demonstrated an action-sentence compatibility effect (Glenberg & Kaschak, 2002) for imperfective sentences but not perfective sentences. The authors studied actions that either involved hand movements TOWARD the agent (e.g., *Nicholas is opening/has opened the drawer*) or hand movements AWAY from the agent (e.g., *Nicholas is closing/has closed the drawer*). Participants performed sentence sensibility judgments regarding the meaning of the presented sentences by pressing response buttons (YES/NO) that involved movements either AWAY from their body or TOWARD their body. The results show that only for imperfective sentences, participants responded faster when the direction of pressing the response button matched the direction of the hand movement indicated in the sentence (e.g., YES button AWAY and movement AWAY). This action-sentence

3 We should be careful with the reliability of the results of the two studies discussed in this section, given that a recent replication study failed to replicate the action-sentence compatibility effect (ACE) (Morey et al., in press). However, we decided not to further discuss this issue in the main text because ACE is not the focus of the current review.

compatibility effect was not found for perfective sentences.

The motor simulation of hand movements triggered by the use of imperfective aspect during sentence processing is also spotted to be used to predict the ending location of an action in the study conducted by Liu and Bergen (2016). Liu and Bergen (2016) used a similar paradigm as by Bergen and Wheeler (2010). However, instead of targeting the action itself, they focused on the information on the ending location of the action and whether motor simulation was combined to predict such information during sentence comprehension. Specifically, the targeted ending location of an action was either AWAY from or TOWARD the protagonist's body in the sentence stimuli (e.g., AWAY *Judith is closing the cupboard*/ TOWARD *Cheryl is pulling the door*). Sentences were either in their imperfective form (e.g., *Judith is closing the cupboard*) or perfective form (e.g., *Judith has closed the cupboard*). Like in Bergen and Wheeler (2010), participants also performed sensibility judgments regarding the meaning of the sentences by pressing the response buttons (YES or NO) that were either AWAY from or TOWARD the participant's body. The results show that participants responded faster when the ending location of the action was compatible with the location of the response button (e.g., location AWAY and YES button AWAY) only for the imperfective sentences. This compatibility effect was not found for the perfective sentences. This study demonstrates that motor simulation is not only involved in the conceptual processing of the movements of an action but is also combined in predicting the ending location of the action. Moreover, only imperfective aspect triggers the motor simulation of an action, and the ending location of an action is more relevant to an event that is depicted as ongoing rather than completed.

Action and its temporal dynamics

Simulation could also be involved in representing the temporal dynamics of an event. For instance, when motor experiences are combined in the mental simulation of an event, subjective and moment-to-moment experiences are involved when the event is being processed, and hence, this event should be treated as temporally more dynamic. Compared to a temporally less dynamic event, a temporally more dynamic event is often viewed as containing more actions/movements and having a longer duration. Several behavioral studies have provided evidence for the idea that grammatical aspect, through its effect on mental simulation, also constrains the temporal dynamics of an event.

Matlock (2011) presents evidence regarding the effect of grammatical aspect on the temporal dynamics of an event in three experiments. In Experiment 1, participants performed a sentence completion task in which they needed to complete either an imperfective adverbial clause (e.g., *when John was walking to school*) or a perfective adverbial clause (e.g., *when John walked to school*). The finding was that participants completed the imperfective clause with more actions (e.g., *he tripped and fell*) compared to the perfective adverbial sentence. In Experiment 2, participants answered the question *How many houses?* after reading either the sentence *John was painting houses last* summer or the sentence *John painted houses last*

summer. It was found that participants significantly included more houses in their answers for the imperfective sentence compared to the perfective sentence. The results indicate that participants conceptualized more painting activities for the imperfective sentence rather than the perfective sentence. When more activities of an event are simulated due to imperfective aspect, it is reasonable to assume that the time spent on an event should also be longer when imperfective aspect is used. In Experiment 3, Matlock (2011) targeted the temporal duration of an atelic event "to drive" in combination with imperfective aspect versus perfective aspect. Participants answered the question *How long (number of minutes or hours)?* after reading either *John was driving last weekend* or *John drove last weekend*. The result was that participants considered the driving time significantly longer for the ongoing event than the completed event. This finding demonstrates that, in comparison to perfective aspect, imperfective aspect seems to also create a more extended duration of the event, besides its effect on the number of activities performed in an event (Experiment 2). Matlock (2011) further stated that the observed effect that imperfective aspect exerted on the temporal dynamics of an event was due to the general ability that people can simulate events. Imperfective aspect could trigger the comprehender's mental simulation of an ongoing event, including the motor stimulation of the action that often possesses both quantitative and durative features. With the mental simulation of the action, the features that define the action, including its quantity and duration, could be scrutinized in great detail. In contrast, perfective aspect could deactivate the motor simulation of the action of an event by emphasizing only the result state of the event. As a result, the event depicted with perfective aspect is viewed as a static state rather than a dynamic process in which motor activities play an important part.

Some other studies have employed eye movements as an online measurement when investigating the effect of grammatical aspect on the temporal dynamics of an event during online sentence processing (Huette, Winter, Matlock, & Spivey, 2012; Huette, Winter, Matlock, Ardell, & Spivey, 2014). Eye movements have been adopted in many language comprehension studies before and are treated as signals that reflect the process of language comprehension (Altmann, 2004). As low-level sensorimotor signals, eye movements are also relevant to the simulated motion and action during language processing. Research has shown that when more motion and action are simulated during language comprehension, more eye movements are generated (Richardson & Matlock, 2007). In Huette's studies, participants listened to a story that contained three to four sentences that were either mainly imperfective sentences or mainly perfective aspect sentences[4]. No specific task was given to the participants. While participants were listening to the story, their eye movements were

4 We view both Huette's studies as dealing with sentence-level processes, even though a story instead of independent sentences was used as experimental stimuli. Using a story that included several sentences here was to elicit as much as continuous and additive eye-movement data from each participant rather than to examine the process of discourse comprehension.

tracked and recorded for analysis. The results show that participants' eye movements were more active (shorter fixations and more dispersed saccades across the screen) when they heard the story with mainly imperfective sentences compared to when they heard the story with mainly perfective sentences. In the latter case, participants' eye movements showed more fixations on the same point, as if they were looking at a static picture. These results suggest that participants indeed mentally simulated more dynamic motion and action when they encountered imperfective sentences rather than perfective sentences. This more dynamic representation of an event triggered by imperfective aspect is combined with previously-activated sensorimotor experiences that need not be provoked by any visual or motor stimulation.

So far, we have discussed the role of grammatical aspect and how it affects the mental simulation of event features during sentence comprehension. Generally speaking, there is a clear difference between imperfective and perfective sentences and how events are being perceived by the reader. Imperfective sentences are represented more often as ongoing than perfective sentences that tend to be represented more often as completed (Madden & Zwaan, 2003). Also, when combined with other manipulations, such as differences in accomplishment versus activity verbs (Yap et al., 2009), the instrument used by the protagonist during an event (i.e., Madden et al., 2017), and the semantic relationship of the location with the verb (Ferretti et al., 2007), differences between representations of ongoing events versus completed events emerged. Moreover, grammatical aspect activates powerful representations that determine to a large extent how readers represent the sentence and the actions described in the sentence as a mental model. Imperfective sentences clearly result in representations that are not completed and may take longer, contain more actions (Matlock, 2011), and are more dynamic (Huette et al., 2012; 2014) than perfective sentences.

Language-driven event comprehension, however, is not limited to sentence-level comprehension. Events usually do not stand alone but are interconnected within a certain discourse. Most of the time, we read or hear a story that is embedded in a large context in which often more than one sentence and one event is involved. Understanding event comprehension in a naturalistic way requires more than studying comprehension of a sentence taken out of context. It is important and necessary to study how events are connected and how they develop across time and space within a discourse (Graesser, Millis, & Zwaan, 1997). In fact, many researchers have already studied event comprehension from a discourse perspective, including those who focused on grammatical aspect.

Discourse-level processes

In this section, we will give a detailed review of those studies that have investigated the effect of grammatical aspect on event comprehension that is embedded within a discourse, specifically a narrative. Given the above-discussed differences that imperfective and perfective sentences have in event comprehension at the sentence level, we can expect an

effect of grammatical aspect on event comprehension at the discourse level as well. In the following subsections, we will discuss each specific role that grammatical aspect may play in discourse comprehension according to the studies reviewed.

Protagonist's location and attributes

In a narrative, readers tend to update their representations of a text based on the locations and goals of the protagonist. They follow the protagonist as they are moving through the events from room to room, over time, and when they accomplish a goal or not. This means that the protagonist always operates in dynamic situations in which events unfold because otherwise, nothing would happen in relation to the protagonist that would be worth mentioning in a text. Moreover, linguistic categories that convey the spatial and temporal structure of events that unfold with a protagonist who is in motion can affect the construction and updating process of situation models. Exactly this issue has been studied empirically in several studies.

Morrow (1985) examined the role of grammatical aspect in the context of a change of location from the protagonist from one location to another. In Experiment 3, progressive aspect (was walking versus walked) was examined in combination with prepositions (from versus to) to determine the prominence of a motion event with a protagonist in language comprehension. For example, in the sentence *John walked past the living room into the kitchen*, the kitchen (goal) could be more prominent because John is (arrived) there. In contrast, in the sentence *John was walking past the living room to the kitchen*, the living room (path) could be more prominent because John is still on his way to the kitchen. This is indeed what the results indicated. In progressive sentences, path referents were chosen more often as the location when referent questions were asked after the narrative, making the path more prominent. In contrast, simple past sentences, in combination with 'from' and 'to', made the goal location more prominent for the reader. It appears that progressive aspect adds a form of prominence to path information only.

In subsequent research, Morrow (1990) addressed the issue of grammatical aspect in situation model construction and updating with a protagonist again. Prepositions, such as 'to', 'into', and 'toward', further specified the location the protagonist moved toward or past a possible goal (to, into) or path (toward) location. However, using progressive aspect (was walking) rather than simple past (walked) would change the location indicated with goal to path when 'to' was used. The results indicated that readers constructed situation models around the protagonist's path of motion and that it depended on the sentence (grammatical aspect and preposition). Simple past (walked) in combination with 'into' specified the location of the protagonist in the goal room, whereas use of progressive aspect (walking) in combination with 'to' led to the localization of the protagonist in the path location. It seems that readers rely on prepositions and grammatical aspect to follow and infer the location of the protagonist in a narrative describing dynamic motion events. When the preposition 'toward' is

used, the situation is considered to be in progress, resulting in placing the protagonist on the path in the motion event rather than in a source or goal location. Prepositions, such as 'through' can further emphasize the protagonist being located at the path and not the goal landmark.

The studies discussed above suggest that grammatical aspect helps the construction and updating of situation models with regard to motion events and that they can modulate this effect to a certain extent in combination with prepositions such as 'to', 'into', and 'toward'. These results illustrate how subtle the process of situation model construction and updating unfolds when a reader progresses through sentences in a short narrative. The next study focuses on the role of grammatical aspect on text comprehension and situation model construction and updating as well, but here prominence is determined by the attributes of the protagonist, rather than by the spatial location of the protagonist.

Carreiras and colleagues (Carreiras et al., 1997) examined the role of verb tense and aspect in relation to current and past attributes from a protagonist. The idea behind this study was that verb tense and grammatical aspect can foreground states and events. Information associated with a protagonist and described in the present tense is more activated in the situation model than when this information is described in the past. Experiment 3 assessed whether the use of grammatical aspect differentially affected the speed with which a character was described in a narrative. These narratives contained descriptions of two characters who either performed them simultaneously (past progressive: was finishing) or at different times (past perfect: had finished). The target character probed after the narrative was recognized more quickly in the simultaneous/past progressive condition than in the sequential/past perfect condition. The explanation is that when performing an action simultaneously as indicated by a progressive aspect, character information remains available and foregrounded, whereas this is not the case when one character has already completed this action which makes it less relevant and will not be included in the same situation model as in the simultaneous condition.

This study underscores the importance of grammatical aspect in the form of a temporal marker in the foregrounding of situation model dimension information, such as the character/protagonist. Just the implication that the action of one of two characters has finished, suggested by perfective aspect, is sufficient to remove this character from the foregrounded position in the situation model that then gets updated to a new model that no longer involves the presence of this character. No wonder that probes about the character are responded to faster when both characters remain in the situation.

Coreferential processing

Switching attention from one character to another is also common in a narrative. When reading/hearing a narrative that involves at least two characters, people tend to predict who will be mentioned in the next sentence. The character who is expected to be mentioned next seems to be the most relevant to the current situation model. In many cases, a pronoun (e.g.,

he or she) will be used to refer to this character. Studies have shown that grammatical aspect affects the comprehender's prediction of the character that will be mentioned next and hence also affects the processing of the pronoun that is used accordingly.

For example, Ferretti et al. (2009) have conducted two experiments in studying the role of grammatical aspect in affecting the coreferential processing during narrative reading. In Experiment 1, participants performed a story continuation task. They first read a transfer-of-possession sentence (i.e., depicting one character giving something to another character). The sentence was either in its imperfective form (e.g., *John was handing a book to Mary*) or in its perfective form (e.g., *John handed a book to Mary*). Then they continued the story with their own sentences. The results showed that participants were in general more likely to continue the story with the goal referent (e.g., Mary) than the source referent (e.g., John). However, this goal bias was significantly more prominent in the perfective-aspect condition than in the imperfective-aspect condition. In Experiment 2, an ERP reading task was employed to examine the reader's expectancy of the next character during online narrative comprehension. The sentence after the context sentence (e.g., *John was handing/handed a book to Mary*) always started with either a male pronoun that represents the source referent (i.e., John) or a female pronoun that signals the goal referent (i.e., Mary). They found that, in general, people encountered more difficulty when processing the pronoun that represented the source referent (i.e., he) than the one representing the goal referent (i.e., she). However, in accordance with the results reported in Experiment 1, only in the perfective-aspect condition, an increased P600 (i.e., an event-related potential that signifies expectancy violation in response to grammatical errors) was detected in response to the source referent, which means that participants' expectations of the referent were violated when a source referent was mentioned.

Both experiments conducted by Ferretti et al. (2009) demonstrated that when perfective aspect was used as a context sentence, the end result of the depicted event was more prominent in the current situation model, and hence the goal referent was more likely to be expected to occur in the following context compared to when imperfective aspect was used. Another study conducted by Grüter, Takeda, Rohde, and Schafer (2018) provided further evidence for this finding. They employed the visual world paradigm and presented participants with visual representations of the source referent, the goal referent, and an irrelevant object on the screen. While participants heard a sentence (e.g., *Donald brought Melissa a fancy drink*), their eye movements toward each of the visual representations were tracked for analysis. They found that even before the occurrence of the pronoun in the next sentence, participants already looked at the goal referent areas (e.g., the picture of Melissa) more often than at the source referent area (e.g., the picture of Donald) after hearing a perfective sentence (e.g., *Donald brought Melissa a fancy drink*) than an imperfective sentence (e.g., *Donald was bringing Melissa a fancy drink*).

The availability or integration of an event in a later context

Apart from its influence on tracking a specific event component (e.g., protagonist; Carreiras et al., 1997), studies have provided empirical support for the notion that grammatical aspect can also affect the availability of a whole event in a later context. For example, Magliano and Schleich (2000) discovered that an event that was depicted with imperfective aspect was viewed more often as still continuing and was later retrieved more quickly by the reader as the narrative continues compared to when it was depicted with perfective aspect. Specifically, they presented participants with stories in which the target event was either described with imperfective aspect (e.g., *Betty was delivering their first child*) or with perfective aspect (e.g., *Betty delivered their first child*). The target event was followed by three more sentences that were not relevant to the temporal status of the target event.

In Experiment 1, participants were asked to answer whether the target event was still ongoing or completed either right after the critical sentence or after one of the three sentences. The results revealed that across the four answering positions, participants were more likely to consider an event ongoing when imperfective aspect was used rather than when perfective aspect was used. In experiment 2, both grammatical aspect and the typical duration of the target event (i.e., short duration vs. long duration; e.g., *writing/wrote a check* vs. *writing/wrote a novel*) were manipulated. The same findings reported in Experiment 1 were replicated. That is, for both short- and long-duration events, the likelihood of the target event being viewed as ongoing in a later context was significantly higher when the event was in its in-progress version than when it was in its completed version. The effect of the typical duration of the target event was also detected in Experiment 2. However, its effect only appeared at question position 4 (after three intervention sentences) instead of at position 1 (right after the critical sentence) and only for the imperfective condition. This three-way interaction indicates that when an event is described as ongoing, it is more likely to be viewed as still in process right after its progressive description, regardless of its typical duration. Only after some intervention context is the difference between short- and long-duration events shown. That is, after some moments passing by in the narrative, events that have a longer duration are more likely to be viewed as still ongoing compared to events that are typically short.

In contrast, Mozuraitis, Chambers, and Daneman (2013) argued that when the temporal status of the event was not directly asked but more implicitly measured, an apparent interaction effect between grammatical aspect and typical duration of the depicted event could be found. Moreover, this effect was detected right after the event was described instead of after some intervening sentences. Specifically, they measured both the comprehender's reading times and eye movements when investigating the extent to which an event could be integrated with another event that appears in a later context. They found that participants (both younger and older adults) were slower at reading the second event (e.g., *she wore her new garment*) when the first event was described in an imperfective sentence (e.g., *she was knitting a sweater*) rather than a perfective sentence (e.g., *she knitted a sweater*).

It should be noted that the second event always indicated that the first event should be completed. Moreover, when the sentence depicted the first event as ongoing, upon reading a subsequent event, both younger and older adults looked more frequently back to the verb area of the sentence compared to when the sentence described the event as completed. The results indicate that participants had more difficulty with comprehending the second event and were less sure about the connection between the two events when the first event was still described as ongoing. Most importantly, Mozuraitis and colleagues found that those effects of grammatical aspect were only detected for events that had longer (e.g., *knitting a sweater*) rather than shorter (e.g., *writing a personal check*) durations. And those effects were detected when the second event was presented immediately after the first event without an intervening sentence. The researchers explained that with a more implicit measurement that did not involve direct questions about an event's temporal status, the effect of the typical duration of an event was in fact directly observed right after its description. That is, when the first event was typically short, even though it was just described as ongoing (e.g., *writing a personal check*), the reader seems to have less trouble connecting it with the second event that implied its completion (e.g., *sent the payment*). However, for an event that needs a longer time to complete (e.g., *knitting a sweater*), it was more challenging for the reader to process the information that indicated its completion (e.g., *wore her new garment*) when it was just being described as still in progress.

Despite the differences discussed above, both studies demonstrate that readers do refer to both grammatical information (i.e., grammatical aspect) and word knowledge (i.e., the typical duration of an event) when they need to decide whether an event is still ongoing or completed as the story goes on. However, neither of the studies distinguished the event types between accomplishments and activities (i.e., lexical aspect). Moreover, the intervention sentences used in both studies mainly only depicted some short events or thought processes that could happen within a short period.

Previous research has suggested that lexical aspect and grammatical aspect can interact with each other in affecting the ease of the construction of a mental representation of an event (e.g., Yap et al., 2009). The duration of a time shift (e.g., a moment later vs. an hour later) in a narrative also seems to affect the availability of an event later in the discourse (e.g., Zwaan, 1996). Therefore, it is possible that event types and the duration of intervention events could both interact with grammatical aspect in affecting the availability of an event in a later context. Becker, Ferretti, and Madden (2013) have investigated exactly this possibility. Using the electroencephalogram (EEG) measurement, Becker et al. (2013) demonstrated that the target concepts (e.g., lunch) that were derived from the target event (e.g., *he was packing his lunch*) were more available later in the narrative when the target event was described as ongoing rather than completed (e.g., *he packed his lunch*). However, this effect of grammatical aspect was only detected for accomplishments (e.g., *he was packing/packed the lunch*) not for activities (e.g., *he was exercising/ he exercised*), and only when the intervening events were short

events (e.g., *he grabbed the map*) not long events (e.g., *he studied the map*). They explained that for events that do have a natural endpoint, the difference between whether it is in process or completed should be greater than for events that do not have a natural endpoint. Therefore, regarding the construction of a situation model, the effect of grammatical aspect on tracking event information across discourse is larger for accomplishments than for activities. However, after a long intervening event, regardless of the temporal status and the type of the event, this event is no longer relevant to the current situation model, including features and details that belong to this event. Hence, the effects of grammatical aspect and lexical aspect on the availability of this event in a later context could be annihilated due to a long time shift. This also corresponds to what is proposed by the event-indexing model that a change in the temporal information could lead to the updating of a situation model, and the information that is not relevant to the current situation model would then be unavailable in the current working memory.

Event segmentation

As described earlier, situation models are mental models constructed based on the state of affairs described in a text. When something happens in a situation, the situation model will be updated and incorporate the change. An important question regarding this change would be how the boundaries between dynamic changes in situation models are determined. Is it really necessary for a reader to correctly identify boundaries of episodes in a text in order to update the situation model correctly or not? Research has shown that readers do indeed monitor those changes closely. Reading times increase when readers perceive certain shifts in relevant dimensions of situation models, such as time and location (Therriault, Rinck, & Zwaan, 2006; Zwaan et al., 1995). This suggests that updating occurs and becomes part of the representation that readers have and develop based on the events that unfold.

Grammatical aspect is one of the linguistic features that may signal such a change or not. Imperfective aspect emphasizes the fact that an event is ongoing. The sentence *Yasmin was driving to the store* will not require an update of the situation model and can go on for some time with elements of the situation being elaborated upon while being underway. In contrast, the sentence *Yasmin drove to the store* signals the boundary of the event and readies the reader for the next event, i.e., going inside the store, which would require a shift in location of the situation model. The Event Horizon Model (Radvansky & Zacks, 2014) postulates that readers segment larger events into smaller events as a form of event segmentation. Segmentation judgments tend to correlate with situation model updates (Kurby & Zacks, 2012). A relevant question in this respect would be whether a change in grammatical aspect could initiate the perception of an event boundary among readers.

Feller et al. (2019) tried to answer this question in three experiments in which they examined the role of grammatical aspect in event segmentation and situation model updating. Specifically, they assessed whether the likelihood of event segmentation increased when

events were conveyed in a narrative using perfective aspect compared to events using imperfective aspect. Participants were instructed to identify sentences that in their opinion conveyed a change in events in the narratives they read. The results indicated that sentences with perfective aspect had a greater likelihood to be segmented than sentences with imperfective aspect, regardless of event duration (short versus long). Readers were also less likely to continue a sentence as an ongoing event after they read a perfective target sentence in a narrative than after an imperfective target sentence.

Apparently, grammatical aspect affects event segmentation processes. Readers consider events in narratives as ongoing when imperfective aspect is used and completed but ready to be segmented as an event boundary when perfective aspect is used. Once again, we see a demonstration of comprehenders' sensitivity to grammatical aspect in sentences, which has far-reaching consequences for how they construct and update situation models. This sensitivity is not limited to dimensions of situation models, such as location or protagonist, but extends to the perception of event boundaries as well. Surprisingly, the duration of the event itself (short versus long) did not seem to play a role, which may be explained by the fact that the narratives were always short. In longer narratives, different outcomes might be possible.

Grammatical aspect beyond text-based event comprehension

From the research discussed, converging evidence shows that grammatical aspect, by specifying the current temporal status of an event, modulates how an event or a situation is represented during text comprehension. Whereas imperfective aspect emphasizes the ongoing phase of an event, perfective aspect focuses the attention on the result phase of an event. An event is not only represented as perceptually richer (e.g., simulation of the actual event, possibly including the protagonist, the instrument, and the location information; e.g., Madden & Therriault, 2009; Madden & Zwaan, 2003; Ferretti et al., 2007) but also temporally more dynamic (i.e., more actions simulated and may take longer; e.g., Huette et al., 2014; Matlock, 2011) when it is depicted with imperfective aspect compared to perfective aspect. When an event is described with perfective aspect, only the result state of the event (e.g., the affected object) is targeted and is hence the focus of the current representation, which is possibly more of a static representation than a dynamic one.

Moreover, grammatical aspect contributes to the construction and updating of a situation model during discourse comprehension. According to the event-indexing model (Zwaan et al., 1995), a situation model is constructed based on the following dimensions: characters, location, time, intentionality, and causation. Any changes that happen in these dimensions will lead to the updating of the situation model. The studies we have discussed demonstrate that during narrative comprehension, imperfective aspect highlights the ongoing process of an event. Consequently, the event details that are relevant to this process are included and maintained in the current situation model (e.g., Carreiras et al., 1997; Magliano & Schleich,

2000). In contrast, perfective aspect signifies the completion of the depicted event, which highlights only the result state of the event. Consequently, the old information that is no longer relevant to the current situation does not occupy the current working memory anymore. Moreover, when an event is completed, it often signifies changes, for instance, in the status of the protagonist (Ferretti et al., 2009) or in the location of the protagonist (Morrow, 1985, 1990). In accordance with the event-indexing model, the situation model needs to be updated when a change is detected.

Both sections discussed the role of grammatical aspect within the context of language comprehension. However, language comprehension is also relevant to higher-level cognitive processes, such as problem-solving and decision-making. The framing effect, for example, illustrates that the way information is presented in language may contribute to cognitive bias and has a great influence on people's choices and decisions (Tversky & Kahneman, 1981). For instance, in a vignette describing a dilemma about an important decision, it matters how the options are worded. If a surgery option is phrased in terms of its success rate, such as "55% success", a patient considering this option might be more willing to go for it than when it is phrased in terms of its failure rate like "45% failure", even though it is the same surgery.

As we have discussed earlier, aspectual markers affect how an event or a situation is comprehended. Imperfective aspect directs people's attention to the ongoing process of an event and often leads to detailed processing of the action. On the contrary, perfective aspect presents an event as a whole, which does not target the actual action performed in an event. Even though it is the same event we are talking about (e.g., *he flirted with a girl* vs. *he was flirting with a girl*), the way it is interpreted can be different depending on which aspectual marker is used. And this might also have an influence on how people solve problems or make decisions. Indeed, some studies have investigated the extent to which the effect of grammatical aspect on event comprehension extends to other cognitive processes that are built upon language comprehension.

Grammatical aspect and problem-solving

In order to solve a problem, one needs to find and define the problem first. How the problem is framed in texts can either help or hinder the process of spotting the problem and finding the solution for the problem. Grammatical aspect might be one of the linguistic cues that can affect such problem-solving processes. Salomon, Magliano, and Radvansky (2013) have investigated this issue in more detail. Specifically, they examined whether grammatical aspect affects the process of solving insight problems. Insight problems are often detected and solved through unconventional and alternative understanding. Consider the following example: *a woman was not carrying her driver's license, was not stopping at a railroad crossing, and was ignoring a one-way traffic sign. However, she did not get a ticket even when the policeman saw her behaviors. Why was that?* The answer is that because she was walking. In order to arrive at this answer, one needs to first pay attention to the actions the woman was

performing and then come up with an alternative interpretation, namely that the woman was not driving but that she was walking.

According to Salomon et al. (2013), imperfective aspect directs one's attention to the actions depicted in the situation, which could facilitate the construction of a situation model that is relevant to the correct answer for the above-mentioned situation. When perfective aspect is used (e.g., *a woman did not carry her driver's license, did not stop at a railroad crossing, and ignored a one-way traffic sign*), however, the constructed situation model would have less emphasis on the actions because they were completed and therefore not relevant anymore to the current situation. Consequently, less attention would be paid to the protagonist's action that is actually the cue for the answer for such problems. Indeed, in Experiment 1, participants were more accurate in finding the solution for such problems when imperfective aspect was used compared to when perfective aspect was used.

The use of imperfective aspect does not always help to solve insight problems. In Experiment 2, the authors studied another type of insight problem that does not require special attention to the protagonist's action. An example of such problems is: *our basketball team was winning 72–49, without one man scoring as much as a single point. How is that possible*? The correct answer is that because it was a women's team. To solve such problems, one needs to find a solution that comes from long-term semantic memory, such as gender stereotypes. When imperfective aspect is used, however, most of the attention is allocated to the protagonist's action, which can leave little room for other potentially relevant information, including considering gender roles. Consequently, grammatical aspect might even hinder the process of finding the correct solution due to the attentional bias it created toward the protagonist's action. In contrast, perfective aspect creates more room for considering information from semantic memory, given that the attentional resources are not devoted to detailed processing of the protagonist's action. Indeed, the results of Experiment 2 demonstrated that the chance of finding the correct answer for such problems was significantly higher in the perfective aspect condition than in the imperfective aspect condition.

Salomon et al.'s study demonstrates that grammatical aspect appears to influence the process of problem-solving. When the solution to the problem is relevant to the ongoing event, such as the protagonist's action, the use of imperfective aspect in the description of the situation exhibits a facilitative effect. However, when the solution is not related to the protagonist's action, imperfective aspect can have an adverse effect by misdirecting the problem solver's attention. In this case, perfective aspect demonstrates an advantage over imperfective aspect by leaving more space for considering other information rather than the protagonist's action.

Grammatical aspect and decision-making

Decision-making is another cognitive process that can be built upon language comprehension. It is reasonable to assume that if a decision is made based on textual

descriptions, the manipulation of the texts can result in differences in comprehension and as a result in different decisions (e.g., the framing effect; Tversky & Kahneman, 1981). Studies have found that the use of grammatical aspect can affect the process of decision-making by influencing people's impression of the described protagonist and people's emotions. For example, Fausey and Matlock (2011) investigated whether the use of grammatical aspect in descriptions of political candidates would affect people's attitudes to their electability in a political campaign. They found that when imperfective aspect was used to describe the past negative actions of a senator (e.g., *he was taking hush money*), participants were more confident to vote "no" to this senator than when perfective aspect was used (e.g., *he took hush money*). Moreover, participants also inferred more negative actions when imperfective aspect was used rather than perfective aspect. However, no effect of grammatical aspect was found for positive actions (e.g., he was collecting donation money vs. he collected donation money). One possible explanation they provided was that people might in general pay more attention to negative actions than to positive actions. Consequently, negative information is more highlighted than positive actions. With the additive effect from the use of imperfective aspect that led people's attention toward the details of the described negative action, the intensity of the negative emotion was emphasized and enhanced the confidence of participants to vote "no" to the described politician.

A direct investigation on the extent to which grammatical aspect affects people' emotions during language comprehension was conducted by Havas and Chapp (2016). Using a 5-point Likert-type scale, they examined the valence of different emotions (afraid, angry, anxious, excited, happy, or sad) expressed in either imperfective sentences (e.g., *she was packing up a birthday package for him*) or perfective sentences (e.g., *she packed up a birthday package for him*). They found that, in general, sentences in the imperfective form were rated as conveying more intense emotions compared to perfective sentences. However, a *post hoc* test showed that grammatical aspect only had a significant effect on the valence of the negative emotions (afraid, angry, and sad) instead of the positive emotions (happy and excited). This finding corresponds to the results reported in Fausey and Matlock (2011) that the effect of grammatical aspect was only found for negative actions.

Another series of studies have been conducted regarding the effect of grammatical aspect on intentionality attribution and legal or blame judgment. Understanding a character's intention, as one of the five dimensions proposed by the event-indexing model (Zwaan et al., 1995), is important to the construction and updating of a situation model. Moreover, how a character's goal is perceived and comprehended can also affect people's judgment of the character. For instance, if a person hurt others on purpose, then this person is often considered more guilty when standing trial and should be blamed to a greater extent than a person who did unintentional harm to others (Fontaine, 2007). As discussed in section 3.1, grammatical aspect affects the temporal dynamics of the depicted event. When an event is described with imperfective aspect, this event is often mentally represented as more dynamic, containing

more actions and taking a longer time compared to when an event is described with perfective aspect (e.g., Matlock, 2011). Consequently, the agent might also be considered performing more intentional actions and hence also more guilty when imperfective aspect is used (e.g., *Jack was hitting his wife*) than when perfective aspect is used (e.g., *Jack hit his wife*). This idea has been tested in a series of studies.

This line of research started with Hart and Albarracín (2010). In Experiment 3, Hart and Albarracín (2010) found that when imperfective aspect was used in a story (e.g., *Westmoreland **was firing gunshots**, one of which stuck MacElroy in the back, paralyzing him*), participants considered the protagonist to have higher criminal intentionality (i.e., to knowingly, intentionally, and deliberately harm the other person) and higher intentionality with his behavior (Mind attribution scale on the capabilities of doing things on purpose and of planning things with a goal; Kozak, Marsh, & Wegner, 2006) compared to when perfective aspect was used (e.g., *Westmoreland **fired gunshots**, one of which stuck MacElroy in the back, paralyzing him*). Moreover, when imperfective aspect was used, the actions performed by the protagonist (e.g., to fire gunshots) were processed in greater detail. For instance, the criminal behaviors were imagined in a more detailed way than when perfective aspect was used (i.e., an effect of grammatical aspect on the temporal dynamics of the event). Interestingly, the mediation effects of grammatical aspect (through its effect on the temporal dynamics of the event) on both the perception of criminal intentionality and the intention attribution were significant. Specifically, when a criminal event is represented as temporally more dynamic, the criminal's intention of hurting others on purpose is considered stronger.

The study conducted by Hart and Albarracín (2010) supported the idea that imperfective aspect could enhance both criminal intentionality and intention attribution by highlighting the actions and the details of the event performed by the protagonist. In contrast, perfective aspect does not have this function. However, a large-scale direct replication of this third experiment, initiated by Eerland et al. (2016) involving 11 labs, did not find support for the findings reported by Hart and Albarracín (2010). Eerland et al. (2016) performed a meta-analysis that included the results from these 11 labs which failed to find any effects of grammatical aspect. That is, imperfective aspect did not trigger more detailed processing of event details and also did not enhance the perceptions of intentionality (both criminal intentionality and intention attribution) compared to perfective aspect. Moreover, imperfective aspect did not indirectly enhance the perceptions of intentionality by creating more detailed processing of the criminal behavior.

Eerland and colleagues suggested that possible reasons why they failed to find those effects of grammatical aspect could be due to the "false positive" findings reported by Hart and Albarracín (2010). That is, the detected effects of grammatical aspect could be due to the design and measurement of the study (i.e., a between-subject design that involved only one vignette). Another reason could be that the replication studies changed one sentence stimulus of the original study in the imperfective vignette (i.e., "was pulling out his gun" to

"pulled out his gun"). This reduced one imperfective sentence in the imperfective vignette, which might lead to the vignette used in the replication studies less "imperfective" than in the original study. However, regardless of the possible explanations given by Eerland et al. (2016), the fact that a large-scale replication study could not replicate the original findings in Hart and Albarracín (2010) means that the effects found by Hart and Albarracín (2010) are certainly not robust enough and are not likely to show up beyond a lab study (e.g., in actual legal trials).

Apart from the direct replication study conducted by Eerland et al. (2016), two other studies examined whether the effects of grammatical aspect on the temporal dynamics of an event and on the perceived intentionality would extend to legal decision-making (i.e., first- or second-degree murder judgment; Sherrill et al., 2015) and blame judgment (Eerland, Sherrill, Magliano, & Zwaan, 2017). Both studies based their hypotheses mainly on the findings reported by Hart and Albarracín (2010). For example, Sherrill et al. (2015) conducted four experiments to examine the effect of grammatical aspect on first-degree or second-degree murder judgment of the murder described in a vignette. Following the findings reported by Hart and Albarracín (2010), they expected to find a significant effect of grammatical aspect on the attribution of criminal intentionality, and hence also on the legal decisions made (i.e., first-degree murder or second-degree murder). However, with more than 2.5 times the number of participants tested in Sherrill et al. (2015) compared to Hart and Albarracín (2010), the effect of grammatical aspect on intentionality attribution was not replicated. The effect of grammatical aspect on legal judgment seemed to be relevant to both its effect on the temporal dynamics of the murder action and the provocation action, and the mentioned order and degree of provocation. Sherrill et al. (2015) concluded that the effect of grammatical aspect on temporal dynamics of the depicted event was robust. That is, grammatical aspect did affect the construal of the temporal dynamics of an action (i.e., more iterations of the murder or provocation action indicated by the use of imperfective aspect; see also Matlock, 2011). However, they also concluded that the effect of grammatical aspect on social cognition, such as legal judgment, was limited and therefore not robust. This study demonstrates that a subtle difference in the use of the aspectual markers might indeed influence event comprehension, but the extent to which this influence can go beyond event comprehension and extend to other cognitive processes, such as decision-making, still await more research and exploration. Similar claims were also made by Eerland et al. (2017) concerning the effect of grammatical aspect on blame judgment.

Languages and aspectual pairs investigated

In the previous discussions, we did not deliberately distinguish between the specific types of imperfective aspect and perfective aspect that were investigated in those studies. We also did not specify what languages were studied in those studies. Our reason was that our major focus in the previous sections was on the representations of events that were either grammatically marked as ongoing or completed. The general distinction between

imperfective aspect and perfective aspect sufficed to represent the discussed two temporal statuses of an event. The type of languages in combination of which aspectual markers was also not the main focus of the previous discussions.

Now that we have covered the relevant differences between imperfective and perfective aspect, it is time to discuss what languages and specific aspectual markers were investigated in those studies for the following reasons: First, discussing the specific aspectual markers in a certain language helps to understand grammatical aspect and its role in representing an event's temporal status from a bottom-up manner. Second, we gain a better understating of the researcher's preferred choice in terms of language types and aspectual markers when conducting language comprehension research. Last but not least, knowing which aspectual markers are usually combined with which language has cross-linguistic implications in language comprehension research.

Therefore, in the current section, we provide an overview of what languages and aspectual markers have been examined in the reviewed studies. Particularly, we discuss the preference of choices of languages and aspectual markers when different levels of processes were targeted (i.e., sentence-level processes, discourse-level processes, and higher-level cognitive processes). Table 1 presents the proportions of the languages and aspectual pairs that have been investigated from different research targets in the reviewed studies.

From Table 1, we can see that the combination of English and the aspectual pair *past progressive versus past tense* (e.g., *the man was making a fire* vs. *the man made a fire*; Madden & Zwaan, 2003) was the most common choice by researchers regardless of the level of processes investigated. This is not surprising. As we have mentioned in the beginning, English is the most studied language in language comprehension research. When the research concerns grammatical aspect, this is even more the case. This is because aspect is grammatically marked in English. There is also a clear contrast between progressive aspect and the simple past tense in terms of the temporal development of an event in English. While progressive aspect is the typical instantiation of imperfective aspect that expresses the progression of an event in English, past tense is conventionally a marker of perfective aspect to express the completion of an event in English (e.g., Comrie, 1976). Moreover, when past tense is used to express an event that happened in the past, past progressive is a natural choice to signify a past event as well (but was ongoing in the past).

Besides the most popular aspectual pair *past progressive versus past tense*, researchers have also examined the aspectual pair *progressive aspect versus perfect aspect* (e.g., *she is/was using her umbrella* versus *she has/had used her umbrella*; Liu & Bergen, 2016; Madden & Therriault, 2009). As for this aspectual pair, perfect aspect instead of past tense is adopted to form a contrast with progressive aspect. Even though both perfect aspect and past tense can be used to signify the end state of an event, the former often implies a continuing relevance of the depicted event (e.g., *she has/had used her umbrella*) to a reference time (the present or

Table 1.

The proportions of the languages and aspectual pairs investigated in the reviewed studies.

Research scopes	Languages	Aspectual pairs	Number of studies	Proportions of studies
Sentence-level processes (n = 16)	English	**Past progressive vs. past tense**	**7/16**	**0.44**
		Past progressive vs. past perfect	2/16	0.13
		Past progressive vs. present perfect	1/16	0.06
		Present progressive vs. present perfect	2/16	0.13
	Cantonese	Present progressive vs. present perfect	1/16	0.06
	Mandarin	Present progressive vs. present perfect	1/16	0.06
	Persian	Present progressive vs. present perfect	1/16	0.06
	French	Past progressive vs. past perfect	1/16	0.06
Discourse-level processes (n = 9)	English	**Past progressive vs. past tense**	**7/9**	**0.78**
		Past progressive vs. past perfect	1/9	0.11
		Past progressive vs. past tense vs. present perfect	1/9	0.11
Higher-level cognitive processes (n = 7)	English	**Past progressive vs. past tense**	**7/7**	**1**

the past), whereas the latter does not (Comrie, 1976). Interestingly, when grammatical aspect was investigated during sentence comprehension, the aspectual pair *progressive aspect versus perfect aspect* was chosen more often compared to when discourse comprehension was targeted. One possible reason is that sentence comprehension studies allow more freedom in the choice of aspectual markers, given that no contextual factors have to be considered.

In contrast, for discourse comprehension studies and the studies that target higher-level cognitive processes, a narrative that depicts a complete story is necessary. The choice of aspectual markers is often more limited in a narrative than in one sentence. For instance, the temporal relations between events need to be carefully considered to avoid temporal contradictions in a narrative. Past perfect aspect can only be used to depict an event that happened before a past event (e.g., *He was hungry because he had only eaten an apple*). This might account for why the aspectual pair *progressive aspect versus perfect aspect* was used less when a narrative was needed for the study. On the other hand, the use of both past progressive and the simple past tense was less constrained in terms of the order of the occurrence of an event in relation to the other. Therefore, the combination of English and the aspectual pair *past progressive versus past tense* was almost exclusively chosen when narratives were studied.

From Table 1, we can also see that more languages were investigated when sentence comprehension was studied (i.e., Cantonese, Mandarin Chinese, Persian, and French). All of the studies that investigated a different language other than English chose the aspectual pair *progressive aspect versus perfect aspect*. It is mostly because all those languages have a different aspectual system from English. For instance, the English simple past is pervasively used to describe a completed event that happened in the past (e.g., he ate an apple). However, the French simple past is only used in formal writings, such as literary and historical writings. To describe the completion status of an event in everyday language, the passé compose (similar to the English present perfect) is usually used (Madden et al., 2017).

Moreover, languages, such as Cantonese and Mandarin Chinese, are tenseless languages, which means that whether an event happens before or at the time of speech is not marked by verb tense. To understand whether an event happens in the past or in the present, one often resorts to context or temporal adverbials (e.g., *now*, *last Friday*). Hence, the simple past tense is not an option in those languages. In those languages (e.g., Cantonese and Mandarin Chinese), whether an event is completed or ongoing is marked by grammatical aspect but no tense marker is involved. Examples of the Cantonese grammatical aspectual markers are the perfect marker *zo*2 (e.g., *ngo*5 *sik*6 ***zo2*** *faan*6 'I eat rice already') and the progressive marker *gan*2 (e.g., *ngo*5 *sik*6 ***gan2*** *faan*6 'I eating rice') (Yap et al., 2009).

A cross-linguistic investigation on how an event's temporal development is mentally represented is important and necessary. First of all, as we have just discussed, aspectual systems vary from language to language. Cross-linguistic studies help us understand

the aspectual systems in different languages from a bottom-up manner. That is, we gain an understanding of how specific aspectual markers function in event representations rather than from a general linguistic definition. Secondly, some languages do not even have grammatical aspect to mark an event's temporal development, such as Dutch and German (Flecken, 2011). In a recent study, Liao, Dijkstra, and Zwaan (2021) demonstrated that directional prepositions may also contribute to the construal of an event's temporal development in Dutch when grammatical aspect is absent. In particular, the preposition *richting* 'toward' may be used to emphasize the ongoing status of an event by highlighting the direction of motion rather than the endpoint of motion (e.g., *de brandweerauto rijdt **richting** het brandende gebouw* 'the fire truck drives **toward** the building on fire'). By doing so, we can broaden our horizons when studying language-driven event comprehension and gain a deeper understanding of how events are represented when cross-linguistic differences are taken into account.

Conclusions and future directions

Grammatical morphemes are used to mark the general, invariant, and pervasive properties of objects and actions (e.g., number, space, and time; Morrow, 1986). Temporal development (e.g., ongoing or completed) is one of the general properties that apply to a wide range of events and is hence marked by grammatical aspect in many languages. Grammatical aspect is one of the most studied grammatical morphemes in the field of language comprehension. The current article provides a review of the empirical studies that target the role of grammatical aspect in event comprehension during language processing. How grammatical aspect in fact affects event representations during language comprehension is discussed in terms of both sentence-level and discourse-level processes. In addition, we have also discussed the empirical studies that examined whether the effect of grammatical aspect on event comprehension extends to other cognitive processes, such as problem-solving and decision-making.

From our discussion, we can draw the following conclusions. First, by specifying the current temporal status of an event, grammatical aspect has shown a robust modulation effect on event representations during sentence comprehension. It constrains the perspective under which an event is viewed and modulates what components of an event are simulated in the current event representation (e.g., Madden & Zwaan, 2003). Specifically, it modulates the sensorimotor simulations during event comprehension (e.g., Madden & Therriault, 2009) and the temporal dynamics of an event representation (Matlock, 2011).

Second, robust empirical evidence supports that grammatical aspect also contributes to the construction and updating of situation models during discourse comprehension. Perfective aspect marks the completion status of an event, which often indicates that a change has happened and an event boundary is detected (Feller et al., 2019). When perfective aspect is used in describing an event, a situation model is updated in a way that the depicted event

or certain event components are no longer stored in the working memory and require extra effort to retrieve them in a later context. In contrast, when imperfective aspect is used, the information of the depicted event keeps being relevant to the current situation model and is retrieved easily later in the discourse (e.g., Carreiras et al., 1997; Magliano & Schleich, 2000; Mozuraitis et al., 2013). Moreover, the effect of grammatical aspect extends to the prediction of future events during discourse comprehension (e.g., Ferretti et al., 2009).

Third, apart from the above-mentioned effects grammatical aspect have on event comprehension, we should be aware that other factors, such as event types (i.e., lexical aspect), world knowledge (e.g., the typical duration of an event), and contextual factors (e.g., the duration of a time shift) often interact with grammatical aspect in affecting event comprehension and the availability of the depicted event in a later context (e.g., Becker et al., 2013; Madden & Ferretti, 2009; Magliano & Schleich, 2000; Yap et al., 2009). This is not surprising, given that the properties of an event are not limited to its temporal development, and an event description does not use only grammatical aspect. When conducting future research on event comprehension during language processing, we should not limit our attention to aspectual markers only. We should be careful about the types of the events we are studying and take our world knowledge about those events into account. For narrative comprehension, factors from other linguistic cues and contextual factors should certainly not be neglected.

Fourth, even though there is evidence that grammatical aspect affects event comprehension during language processing, the effect does not necessarily extend to social cognition. When a decision has to be made, grammatical aspect is just one factor that creates a certain impression on the depicted event. In particular, imperfective aspect creates a stronger impression on the action performed by the protagonist than perfective aspect does. However, other factors, such as the linguistic context and contextual factors, certainly play a part in affecting decision processes (Sherrill et al., 2015). We should not neglect the possible effect of grammatical aspect on social cognition but also need to be careful not to overestimate its effect in this aspect. We are also aware that the studies that investigated the effect of grammatical aspect on decision processes only employed self-reports questionnaires, including the studies that aimed to study the effect of grammatical aspect on emotions and decision-making processes (e.g., Fausey & Matlock, 2011; Havas & Chapp, 2016). Future research should use more implicit measures to examine the effect of grammatical aspect in other cognitive processes besides language comprehension.

Last but not least, we discovered that only a few sentence comprehension studies have investigated languages other than English. The narrative comprehension studies and the studies that targeted the effect of aspectual markers on social cognition exclusively targeted only the English language. The finding is in line with what Henrich, Heine, and Norenzayan (2010) have discussed regarding the biased and limited sampling of behavior sciences. Many important findings that claim to have discovered universal patterns about

human behavior were in fact based on observation of a single and atypical group, primarily American undergraduate students. This also seems to be the case in psycholinguistic studies. This may be an important factor in why English has become the most studied language in the event comprehension studies discussed in the current review. However, as we have discussed earlier, English is in fact different from many languages regarding their aspectual and grammatical systems. Focusing only on the English language might bring about findings that are limited to the characteristics of English grammar and are not applicable to people who speak different languages. Therefore, for future research, a more diverse set of languages should be investigated. This does not only make our research findings more general and more applicable across different languages but conducting cross-linguistic empirical studies can also help us to gain a deeper understanding of how an event is comprehended in general across different cultures.

Supplementary materials

Supplementary material 1. ASReview results
Supplementary material 2. The recorded ASReview project during abstract screening
Supplementary material 3. A full list of the classification of the reviewed studies
(All supplementary materials are available at:
https://osf.io/nc7p4/?view_only=af33b300dc90445398cf64f3cc146587)

Chapter 3

A cross-linguistic study of motion event description and conceptualization between Dutch and mandarin Chinese speakers

This chapter has been published as:

Liao, Y., Flecken, M., Dijkstra, K., & Zwaan, R. A. (2020). Going places in Dutch and in mandarin Chinese: conceptualising the path of motion cross-linguistically. *Language, Cognition and Neuroscience*, *35*(4), 498–520. https://doi.org/10.1080/23273798.2019.1676455

Abstract

We study to what extent linguistic differences in grammatical aspect systems and verb lexicalization patterns of Dutch and mandarin Chinese affect how speakers conceptualize the path of motion in motion events, using description and memory tasks. We hypothesized 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 conceptualized 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.

Keywords: Motion events, language production, cross-linguistic analysis, the path of motion, grammatical aspect, verb lexicalization patterns

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 conceptualize 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; Ochsenbauer & Hickmann, 2010; Slobin, 1996, 2004; 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 lexicalization patterns) influence how speakers conceptualize 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 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 prioritized over other types of path information in motion encoding (in particular, the source; Lakusta & Landau, 2005; Papafragou, 2010; Regier & Zheng, 2007; Stefanowitsch & Rohde, 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 & von 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; a particle ***out*** in the English sentence *a boy is* ***walk****ing* ***out*** *of a house*; Talmy, 2000), distribute their attention differently when viewing and describing boundary-crossing events (Papafragou, Hulbert, & Trueswell, 2008), and they also memorize 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 conceptualization 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 conceptualize 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 conceptualized 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 conceptualization 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 conceptualization 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 and colleagues (2019) speculate that verb lexicalization patterns have a stronger impact on speaker's motion conceptualization than grammatical aspect, and that speakers of satellite-framed languages are more likely to conceptualize motion events in terms of their endpoint than speakers of verb-framed languages (see also Slobin, 1996). This study underlines the importance and necessity of combining the two typological features, i.e., verb lexicalization patterns and aspectual systems, in a cross-linguistic comparison of motion event conceptualization. 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 lexicalization patterns as well as grammatical aspect, conceptualize 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 characterized 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; Smith, 1991; 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 conceptualization. We study (1) how the path of motion is ***conceptualized*** in language production, i.e., what element of the path of motion in a visually depicted event do people select for verbalization? That is, do they refer to the (potential) Endpoint, Trajectory or Location, as shown in videos of Endpoint-reached and Endpoint-oriented events? During conceptualization 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 conceptualization, 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-verbalization event memory task, in which participants were tested on their memory representation of the endpoints in Endpoint-oriented events.

We hypothesize, first of all for Endpoint-reached events, that speakers of both languages will most frequently select the endpoint for verbalization, 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 lexicalization patterns: Dutch is hypothesized to conceptualize 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 hypothesize 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 conceptualization 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 the 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 verbalization 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).

Theoretical background

Comparing Dutch and Mandarin Chinese

Verb lexicalization 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 *zuo-guo* in this sentence is "walk-cross" in English. The first element *zou* is a manner verb that means "walking" and the second element *jin* 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 & 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 xiaoche tui-xia shanpo* "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 la zhe yi-liang yingerche guo jie* "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)

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*** *zou xiang tushuguan* "he is walk**ing** towards a library") and a perfective aspectual marker *le* (e.g. *ta zou xiang* ***le*** *tushuguan* "he walk**ed** 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*** *zou xiang tushuguan* "he is walk**ing** towards a library"). The perfective *le* can be used to describe goal-reached motion (as in *ta zou-jin* ***le*** *tushuguan* "he walk**ed** into a library"). However, it should be noted that unlike English, in which tense and aspect is 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 zou xiang tushuguan* "he walks towards a library" is progressive by default; the sentence *ta zou-jin tushuguan* "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 summarize our characterization, 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.

Experiment

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.

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, mean = 14.4 years, *SD* = 2.30 years; Dutch: Range 5-19 years, mean = 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.

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 seconds in length. There was a blue screen with a centered fixation cross in between each item. Participants were instructed that they could start to describe each video clip when they had recognized "what was happening" in the video. Descriptions were recorded with an external voice recorder. The blue screen between each video was shown for 8 seconds, leaving ample time for participants to verbalize 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 second 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).

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

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.

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). The 40 clips were presented in randomized order.

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 randomized across participants. This was done to cancel out potential distance effects.

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, 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 were automatically saved as .wav files on the computer. Each session took about 20 minutes. 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 to 7 minutes. After this part, participants filled in a language background questionnaire. The experimental procedure took around 40 minutes in total. All the instructions were presented in the participants' native language (Dutch or Chinese).

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[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*)[2]. 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 countryside]"). 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

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.

was not considered in the current study. The coding scheme reflects the differences in the viewpoints that people can take during conceptualization: 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".

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[3] coded were those which were used independently as predicates and which encoded path information in a main clause. Examples of Chinese path verbs are *jin* "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[4] 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,

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").

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).

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).

Analysis and results[5]

Conceptualizing the path of motion

To test the effects of Language and Event type on the use of each path element, we set up separate mixed-effect binomial logistic regression models for each Path type in R[6] (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 so that we could get main effects in relation to the use of each path element. 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, *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

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)

below for examples of each path component in each language.

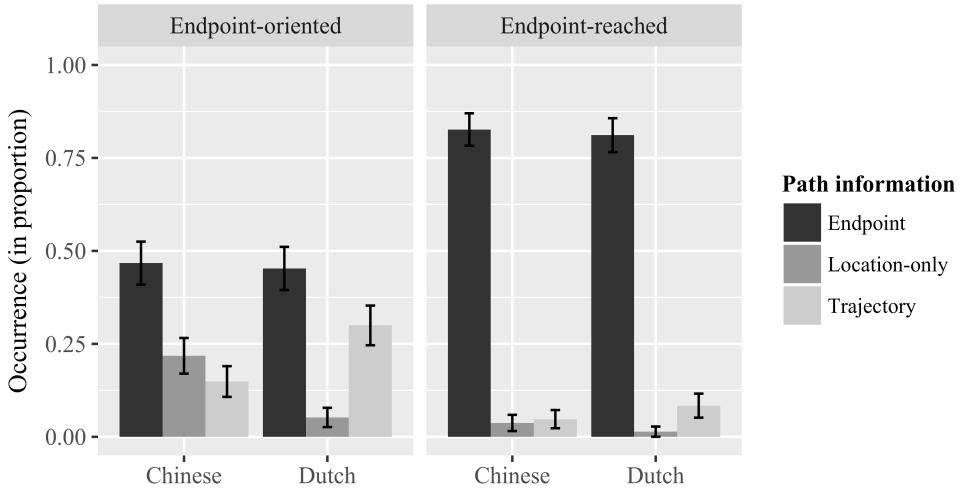
Table 1.

Examples of path component descriptions in Chinese and Dutch.

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

Figure 4.

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



We found that the frequency of Endpoints was significantly different across event types, but not across languages (*Event type* β = -1.212, *SE* = 0.376, *z* = -3.222, *p* = .001; *Language* β = 0.054, *SE* = 0.132, *z* = 0.405, *p* = .686). There was no interaction between the two factors (*Event type*Language* β = 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 hypothesized. In terms of references to **Location-only**, there was a main effect of Language (β = 0.803, *SE* = 0.262, *z* = 3.066, *p* = .002) and a main effect of Event type (β = 1.141, *SE* = 0.452, *z* = 2.522, *p* = .011). No interaction between Language and Event type was found (β = 0.242, *SE* = 0.183, *z* = 1.326, *p* = .185). Dutch participants showed lower frequency of mentioning Location-only information than Chinese participants, as hypothesized. 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* β = -0.468, *SE* = 0.152, *z* = -3.068, *p* = .002; *Event type* β = 0.952, *SE* = 0.365, *z* = 2.610, *p* = .009). The interaction was not significant (β = -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[7]. Event type (fixed factor) was dummy coded and path type was the dependent variable[8]. First, in the **Chinese** data, there was a significant effect of Event type ($\chi 2$ =85.699, $p < .001$): In Endpoint-oriented events[9], Endpoints were mentioned significantly more often than both Location-only and Trajectory information (*Location-only* vs *Endpoint*: *Intercept* β = -0.762, *SE* = 0.153, *z* = -4.995, *p* < .001; *Trajectory* vs *Endpoint*: *Intercept* β = -1.144, *SE* = 0.175, *z* = -6.534, *p* < .001, respectively). In Endpoint-reached events [see footnote 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 *Endpoint*: *Intercept* β = -3.095, *SE* = 0.308, *z* = -10.041, *p* < .001 and *Trajectory* vs *Endpoint*: *Intercept* β = -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

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 1, 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.

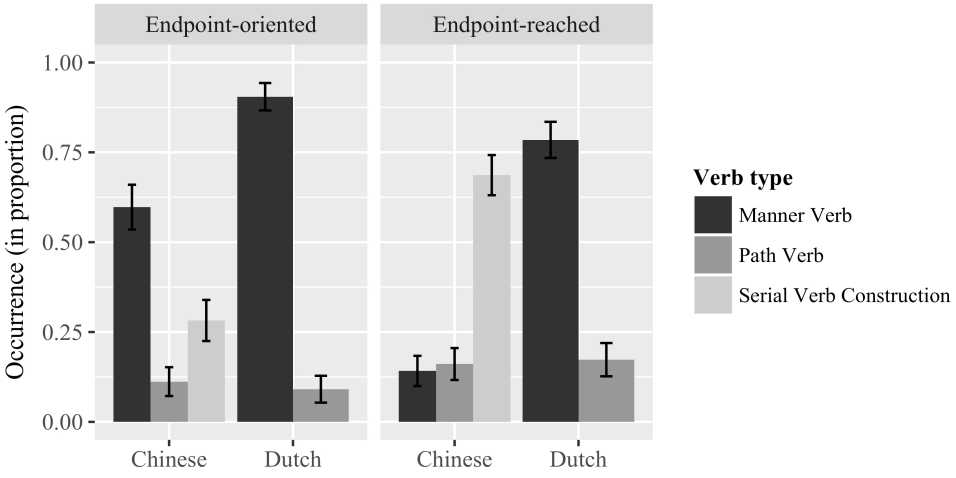
Endpoints significantly more often than Location-only and Trajectory information in both Endpoint-oriented events (*Location-only* vs *Endpoint*: *Intercept* β = -2.160, *SE* = 0.273, *z* = -7.919, *p* < .001; *Trajectory* vs *Endpoint*: *Intercept* β = -0.413, *SE* = 0.139, *z* = -2.973, *p* = .003, respectively) and Endpoint-reached events (*Location-only* vs *Endpoint*: *Intercept* β = -4.060, *SE* = 0.504, *z* = -8.052, *p* < .001 and *Trajectory* vs *Endpoint*: *Intercept* β = -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.

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[10]. 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).

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)))

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.



For **manner verbs**, there was a significant main effect of Language (β = -1.388, *SE* = 0.157, *z* = -8.862, *p* < .001) and Event type (β = 0.725, *SE* = 0.181, *z* = 4.007, *p* < .001). The interaction was also significant (β = 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 (β = 0.084, *SE* = 0.186, *z* = 0.449, *p* = .653), nor of Event type (β = -0.300, *SE* = 0.257, *z* = -1.169, *p* = .243). The interaction was also not significant (β = 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[11]. 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* β = 1.577, *SE* = 0.178, *z* = 8.852, *p* < .001; *serial verb construction* vs *path verb*: *Intercept* β = 1.454, *SE* =

11 mlogit(VerbType~1|EventType, data = VerbC2, reflevel = 1) #reflevel 1 = MV
glmer(VerbType~EventType + (1|PP) + (1|Stimulus), data = VerbD3, family=binomial)

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* β = 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* β = -1.674, *SE* = 0.210, $z = -7.982$, $p < .001$; *serial verb construction* vs *manner verb*: *Intercept* β = -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* β = 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* β = -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* β = 3.639, *SE* = 0.697, $z = 5.220$, $p < .001$; *Endpoint-reached*: *manner verb* vs *path verb Intercept* β = 2.658, *SE* = 0.646, $z = 4.115$, $p < .001$).

Binomial mixed-effect logistic regression models were built for each verb and adjunct combination that existed in the languages[12]. 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. Figure 6 and Figure 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 (located at page 76).

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)

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).

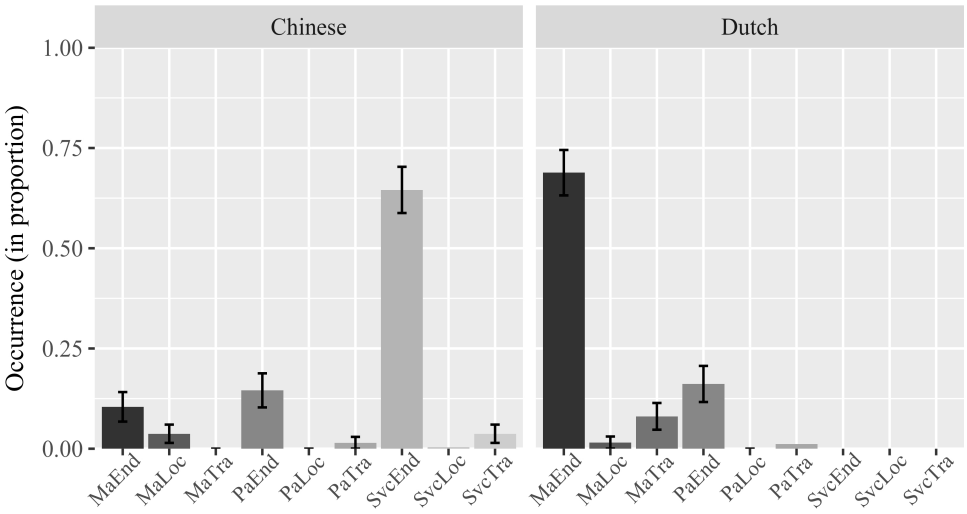
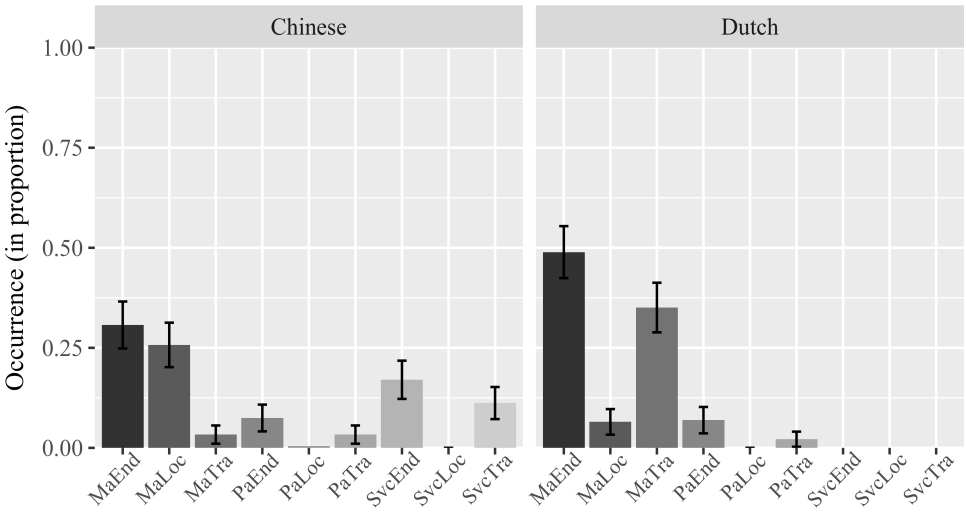


Figure 7.

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).



In terms of the combination of **manner verbs with Endpoint adjuncts**, there was a significant main effect of Language (β = -1.171, *SE* = 0.166, *z* = -7.053, *p* < .001), but no effect of Event type (β = 0.052, *SE* = 0.257, *z* = 0.202, *p* = .840). The interaction between Language and Event type was significant (β = 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 (β = 0.064, *SE* = 0.204, *z* = 0.315, *p* = .753), nor Event type (β = -0.613, *SE* = 0.320, *z* = -1.916, *p* = .055) and no interaction either (β = 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 (β = 0.762, *SE* = 0.274, *z* = 2.781, *p* = .005) and Event type (β = 1.142, *SE* = 0.449, *z* = 2.542, *p* = .011). No interaction effect was detected (β = 0.281, *SE* = 0.185, *z* = 1.521, *p* = .128). Chinese speakers used manner verbs plus Location-only adjuncts significantly more often 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 (β = -1.698, *SE* = 0.236, *z* = -7.201, *p* < .001) and also a main effect of Event type (β = 1.098, *SE* = 0.378, *z* = 2.904, *p* = .004). 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[13]. The dependent variable in each model was verb and adjunct combination. There was a significant effect of Event type in the **Chinese** group (χ2 =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 (β = -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 (χ2 =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

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

(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 (β = -0.333, *SE* = 0.146, *z* = -2.287, *p* = .022).

Table 2.

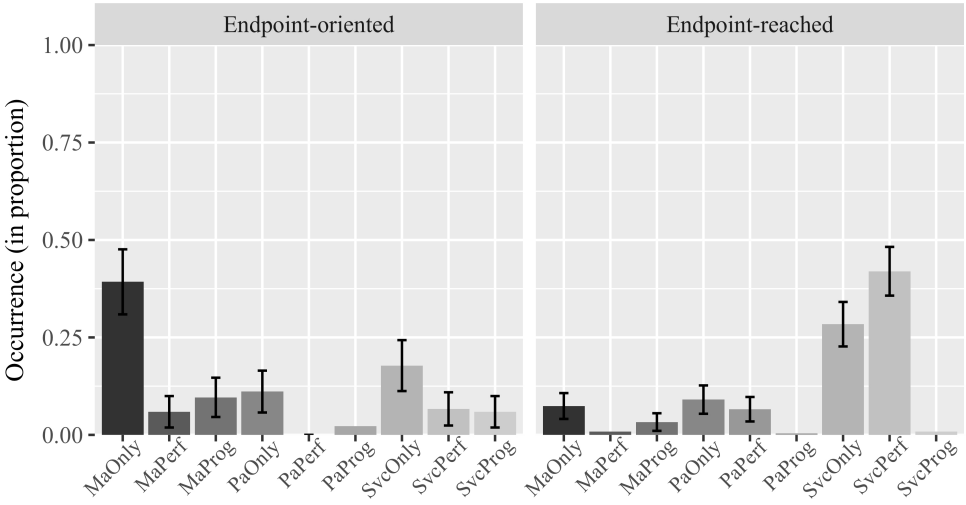
Examples of verb-path combinations in Chinese and Dutch.

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

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) 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) for Endpoint-oriented events. We conducted multinomial logistic regression models[14] 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.

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).



14 mlogit(VerbAspect~1|EventType, data = Csyntax4, reflevel = 1) #reflevel 1 = MV

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

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

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* β = -0.792, SE = 0.246, z = -3.220, p = .001; *path verb* vs *manner verb Intercept* β = -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* β = 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* β = 0.391, SE = 0.156, z = 2.508, p = .012). 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.

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, mean = 2.08, *SD* = 1.23; Dutch: Range 0.00-3.28, mean = 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 3500ms was determined as a reasonable cut-off based on the density of the RT data. RTs that were longer than 3500ms were excluded from analyses. We found that Dutch participants were significantly faster than Chinese participants in giving accurate responses (Chinese: mean = 1757ms, *SD* = 670ms, Range 827.75ms-3216.00ms; Dutch: mean = 1483ms, *SD* = 494ms, Range 850.67ms-1984.00ms; $t(293.67) = 4.223$, $p < .001$). However, the same pattern was also found for filler items (Chinese: mean = 1663ms, *SD* = 634ms, Range 810.67ms-2484ms; Dutch: mean = 1356ms, *SD* = 489ms, Range 676.33ms-1891ms; $t(266.27) = 4.550$, $p < .001$).

Discussion

Our motion event description task set out to discover (1) how the path of motion, the core schema of a motion event, is ***conceptualized*** 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 verbalization. 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).

Conceptualizing the path of motion

We hypothesized 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 conceptualized 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 categorized 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 conceptualization (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 hypothesized 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 lexicalization patterns. However, our results showed similar frequencies regrading endpoint mentioning in both Endpoint-reached and Endpoint-oriented events in Chinese and Dutch. In addition, we hypothesized 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. 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 lexicalization 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. 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 hypothesized 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 conceptualized 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 conceptualization 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 & von 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 conceptualization in Chinese, compared to Dutch.

Verb and adjunct types used

We hypothesized 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 lexicalization 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 hypothesized 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 hypothesized 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.

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 lexicalization 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 lexicalization 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 conceptualization 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.

Implications for typologies of motion expression and theories of event cognition

The present study provides a fine-grained picture of how people conceptualize 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 conceptualization, looking at both grammatical aspect and verb lexicalization 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 conceptualization 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 centered 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 conceptualized during production? What information do people select to form an event "unit" (Gerwien & von Stutterheim, 2018), when the event boundary is *not* salient? We found that other elements of the path of motion became anchoring points for event conceptualization, 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 conceptualize 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 conceptualization 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 localizing 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 conceptualized 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 conceptualization 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 von Stutterheim (2018) were able to show differences in the granularity of event segmentation between French and German, because of differences in the languages' verb lexicalization and motion conceptualization patterns.

It is important to note that the current findings of cross-linguistic differences were obtained in settings with very specific task demands (scene conceptualization during an event description task). One important open question is, how do people perceive and conceptualize 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-verbalization 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 conceptualization (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 conceptualization 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.

Conclusion

The current study sheds light on the extent to which cross-linguistic differences between Dutch and mandarin Chinese affect their speakers' conceptualization of the path of motion in motion events. We first investigated motion event conceptualization 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 lexicalization patterns) into account and analyzed 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 lexicalization 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 (e.g., *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 conceptualization 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 lexicalization patterns and grammatical aspect usage. By reporting detailed comparisons of the linguistic patterns (considering both aspectual systems and verb lexicalization patterns) in both languages, for two event types, we provide an in-depth understanding of the relation between linguistic typology and event conceptualization in language production.

Appendix

Endpoint-oriented events

1. a van driving down a country lane (towards a village/houses)
2. a woman walking across the parking lot (towards a car) event
3. a woman walking down an alley (towards a barrier)
4. a little boy walking along a path (towards a playground)
5. a man climbing up a ladder (towards a balcony)
6. a man crossing a street (towards a car)
7. two girls walking along a path (towards a house)
8. a girl on a horse riding (towards an entrance)
9. a mother and child walking through a park (towards a slide)
10. a car driving down a road (towards a petrol station)

Endpoint-reached events

1. a car driving into a garage
2. a girl entering the station
3. a van turning into a driveway
4. a man on a bicycle turning into a gateway
5. a woman is entering a supermarket
6. a dog running through the door of a building
7. a cat walking into the kitchen
8. a child going through a gate into a playground
9. a man walking into a church
10. a girl on a horse riding into a barn/stable

Chapter 4

Examining motion event endpoint conceptualization: the choice between two Dutch directional prepositions

This chapter has been published as:

Liao, Y., Dijkstra, K., & Zwaan, R. A. (2021). Directional prepositions and event endpoint conceptualization: a study of *naar* and *richting* in Dutch. *Language and Cognition*, *13*(2), 161–190. https://doi.org/10.1017/langcog.2020.31

Abstract

Two Dutch directional prepositions (i.e., *naar* and *richting*) provide a useful paradigm to study endpoint conceptualization. Experiment 1 adopted a sentence comprehension task and confirmed the linguistic proposal that when *naar* was used in motion event descriptions, participants were more certain that the reference object was the goal of the agent than when *richting* was used. Experiment 2 and Experiment 3 used this linguistic pair to test the effect of two factors (i.e., the actor's goal and the interlocutor's status) on endpoint conceptualization via language production tasks. We found significant effects of both factors. First, participants chose *naar* more often, when there was an inference in the referential situation that the reference object was the actor's goal than when was no such inference. Second, participants chose *richting* more often, when they were told to describe the referential scenario to a police officer than to a friend. Participants were more cautious with their statements and were less willing to commit themselves to stating the goal of the agent when talking to a police officer than to a friend. The results are discussed in relation to relevant linguistic theories and event theories.

Keywords: Event endpoint conceptualization; directional prepositions; the actor's goal; the interlocutor's status

Introduction

Events are an important topic in the fields of cognitive psychology and linguistics. Much of the research on event cognition is embedded in the context of language comprehension studies. This is because language is a major way of presenting and talking about events in our daily life. Moreover, it is much easier to construct events and manipulate the spatial and temporal relations among those events in language than in the real world (Radvansky & Zacks, 2014). Researchers have proposed that language comprehension is conceptualized as the construction of a coherent event model or situation model. Situation models are mental representations of a state of affairs described in texts. Linguistic cues on both spatial and temporal relations among events described in texts contribute to the construction of situation models to achieve successful event comprehension (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan, 2016; Zwaan & Radvansky, 1998).

In event cognition studies, there is an increasing focus on event segmentation, given that segmenting meaningful discrete events (e.g., shopping) from a continuous flow of experience (e.g., walking around in a store, taking things from the shelf, bringing them to the counter, etc.) is key to effective memorization and learning. Being able to perceive and detect event boundaries is fundamental to segmenting events (Zacks & Swallow, 2007). Changes that happen in space (e.g., entering a store) or in time (e.g., an egg got cooked) can both be considered as important clues that we use to perceive and define event boundaries. In language, we can easily present such changes by using spatial or temporal related expressions, which function as event-segmentation cues to the comprehender. One of the most studied linguistic cues of this type is grammatical aspect (Anderson, Matlock, & Spivey, 2013; Becker, Ferretti, & Madden, 2013; Madden & Zwaan, 2003; Magliano & Schleich, 2000perfective aspect; Matlock, 2011).

Grammatical aspect is the morphosyntactic marking of verbs that provide different viewpoints on the internal temporal structure of an event (e.g., beginning/ongoing/completed; Comrie, 1976; Declerck, 2007). A broad aspectual distinction is between imperfective and perfective aspect. The use of imperfective aspect in event descriptions directs the comprehender's attention to the internal temporal structure of an event (most often, the progression of an event), while the endpoint is defocused (e.g., *he is cutting a carrot*). In contrast, perfective aspect does not convey information on the internal structure of an event. It gives rise to a completed construct of an event and hence the attention is led to the end-state of the event (e.g., *he cut a carrot*) (Madden & Zwaan, 2003).

Knowing the actor's goal of performing an action is also key to event boundary conceptualization. If the actor's goal is clear, it is then easy for the comprehender to combine relevant actions into a whole event (e.g., Graesser, Singer, & Trabasso, 1994). Take cooking a meal as an example. Cooking a meal is composed of actions such as cutting vegetables and meat, stirring them in a pan, and finally serving them in a bowl. Given that we as a

comprehender know that the actor's goal is to cook a meal, we would not expect the actor to stop after just cutting a single carrot. We would expect more to happen for us to be able to construct a situation model of a person cooking a meal. However, often the actor's goal is not clear. The goal is an internal feature and is therefore not always directly noticeable and obvious. Often, the goal has to be inferred from other aspects of an event, such as actors' movements and changes in location or in time (Radvansky & Zacks, 2014).

Another factor that might affect the conceptualization of an event boundary is the speech context, e.g., the communicative status of the listener. Event conceptualization does not always happen in isolation. Most of the time, we are talking to one another about an event and sometimes we are describing an event with a certain listener or reader in mind. Papafragou and Grigoroglou (2019) argued that to whom we are talking affects how we conceptualize an event endpoint during message planning. Speakers might even define the endpoint of a simple event, for instance making a bed, differently depending on to whom they are talking, their mother or an exacting 5-star hotel manager. When talking to their mother, they are less worried about the end-state of a bed being made than when talking to a 5-star hotel manager, who is expected to have a high standard on how a bed should be made.

In addition to grammatical aspect, there is another linguistic cue that also contributes to event endpoint conceptualization, that is, directional prepositions. We mentioned that many studies have experimentally studied the function of grammatical aspect in event comprehension. Directional prepositions, however, are currently under-studied in experimental research. The current study aims to experimentally investigate the effect of directional prepositions on event endpoint conceptualization when grammatical aspect is not involved. Therefore, we present more details on the linguistic definition of directional prepositions, in comparison to that of grammatical aspect, in the next section, to distinguish between grammatical aspect and directional prepositions concerning their role in event endpoint conceptualization.

Grammatical aspect and directional prepositions

There is an abundance of complex linguistic literature on the distinction between grammatical aspect and lexical aspect (Comrie, 1976; Croft, 2012; Dowty, 1979; Langacker, 1987; Vendler, 1967). Grammatical aspect is known in the literature as being a grammatical way of representing the internal temporal structure of an event (e.g., ongoing or completed; Declerck, 2007). It is also called 'viewpoint aspect'. Imperfective aspect provides a viewpoint on the ongoing phase of an event and enables us to view the internal temporal contour of an event (e.g., *he is writing a letter*). Perfective aspect, however, emphasizes the completion of an event and offers an external viewpoint to inspect the event as a whole (e.g., *he wrote a letter*; Comrie, 1976; Langacker, 1987).

Unlike grammatical aspect, directional prepositions (such as *to* and *towards*) contribute to lexical aspect. As part of a verb phrase (e.g., *walking to/towards the church*), they refer to the inherent structure of an event (the telicity of an event): whether an event has an inherent endpoint or not. A distinction is often made in the literature between telic directional prepositions and atelic directional prepositions (for example, Krifka, 1998).

A telic directional preposition (e.g., *to*) is often labelled as a goal preposition (Bogaert, 2008; Eschenbach, Tschander, Habel, & Kulik, 2000). When a telic directional preposition is used in a motion event description (e.g., *he is walking to the house*), the reference object (e.g., *the house*) is implied to be the goal of the agent. Therefore, this event is considered as having an inherent event endpoint and is thus a telic event. However, when an atelic directional preposition (e.g., *towards*) is used to describe a motion event (e.g., *he is walking towards the house*), there is no such implication. The reference object (e.g., *the house*) only implies the general moving direction of the agent. Therefore, the event is regarded as having no inherent event endpoint and is thus an atelic event.

In English, aspect is grammatically marked. The use of grammatical aspect in event descriptions is frequent and pervasive. The combinational use of grammatical aspect and directional prepositions (lexical aspect) is common, and they both contribute to the endpoint conceptualization. For example, *he is walking to a church* describes a telic event (*church* is implied by *to* as the inherent endpoint of the event of *walking*) that is still in progress and has not reached its inherent endpoint (an internal viewpoint provided by imperfective aspect). In other languages, grammatical aspect is not used or is rarely used. An example is Dutch (Flecken, 2011). As previously mentioned, the current study aims to understand the extent to which directional prepositions purely contribute to endpoint conceptualization when no effect comes from grammatical aspect. We chose Dutch in the current study for this reason (see detailed elaborations in the next section).

Despite the fact that directional prepositions are experimentally understudied, they have been theoretically analyzed by many linguists (see Gruber, 1976; Piñón, 1993; Vandeloise, 2017; Verkuyl & Zwarts, 1992; Zwarts, 2005, 2008; Zwarts & Winter, 2000). Before introducing our experiments, we explain in the next section why Dutch is a suitable choice for our current study, considering its aspectual system. We also conduct some linguistic analyses on the two Dutch directional prepositions (i.e., *naar* and *richting*) based on the existing literature. In doing so, we present a clear picture of the semantic features of the two directional prepositions and provide a solid theoretical background for the hypotheses made in the current study.

Theoretical background and the present study

Two Dutch directional prepositions: *naar* and *richting*

Dutch is often considered as a non-aspectual language. It has progressive constructions, such as the *aan-het* construction and postural verb constructions (e.g., *zitten te* + infinitive), but

they are less frequently used and hence more marked than the simple forms in Dutch (Flecken, 2011). They are often constrained by more contexts and more situation types than the simple forms (see event description studies in Flecken, 2011; Liao, Flecken, Dijkstra, & Zwaan, 2020; von Stutterheim, Carroll, & Klein, 2009). For instance, von Stutterheim et al. (2009) found that, when describing a motion event in which the agent moves in relation to a reference object, Dutch speakers exclusively used simple present tense (e.g., *Hij **loopt** naar/richting het station* 'he **walks** to/towards the station') instead of a progressive construction (e.g., *Hij is naar/richting het station **aan het lopen*** 'he is **walking** to/towards the station'). This fact allows us to use the simple present tense in all our experiments that concern the same type of motion events, to eliminate the possible influence of grammatical aspect, and to focus purely on the role of directional prepositions with regard to endpoint conceptualization.

The two common directional prepositions in Dutch (i.e., *naar* and *richting*), therefore, became our topic of interest. Like *to* in English, *naar* in Dutch is also a goal preposition. It also implies the goal of the agent and hence a telic motion event that owns an inherent endpoint. It sounds odd to say *hij liep urenlang naar de kerk* 'he walked to the church for hours' but *hij liep in twee uur naar de kerk* 'he walked to the church in two hours' sounds perfectly fine.

Richting (meaning *direction*), on the other hand, is an atelic directional preposition in Dutch, which is similar to *towards* in English. It does not imply that the reference object is the goal of the agent and thus does not refer to an inherent endpoint of an event. *Hij liep urenlang richting de kerk* 'he walked towards the church for hours' sounds acceptable, but *hij liep in twee uur richting de kerk* 'he walked towards the church in two hours' sounds odd. *Richting* is originally a noun and is typically used in the phrasal preposition *in de richting van* 'in the direction of'. In fairly recent usage (in 1984; van der Sijs, 2001), it has been denominalized into a monomorphemic directional preposition that can be used independently in a sentence (Bakker & Siewierska, 2002; Broekhuis, 2013; Norde, 2008).

As previously mentioned, directional prepositions have been investigated by many linguists. One common analysis in the linguistic literature is the traditional partitive analysis. According to the traditional partitive analysis, atelic directional PPs (e.g., *richting*-PPs, *towards*-PPs) denote parts of the complete paths denoted by telic directional PPs (e.g., *naar*-PPs, *to*-PPs). Similarly, this analysis is also applied to define imperfective aspect, such that imperfective telic sentences denote parts of complete VP-events (see Jackendoff, 1991; Krifka, 1998; Piñón, 1993; Verkuyl & Zwarts, 1992).

Zwarts (2005), however, argued that the traditional partitive analysis works well with straight paths but not with curved paths. He took a new perspective and proposed that atelic directional PPs are comparatives, whereas telic directional PPs are superlatives. That is, if atelic directional PPs refer to paths that get 'nearer' to the reference object, then telic directional PPs refer to paths that get 'nearest' to the reference object (pp. 766). This proposal captures the cumulative feature of atelic directional PPs (i.e., nearer and nearer to the reference object)

and the non-cumulative feature of telic directional PPs (i.e., 'nearest' means that one cannot get any nearer to the reference object) (see Zwarts, 2005, 2008 for a detailed explanation of *cumulativity*). However, it fails to distinguish between, for example, *to* and *towards* in the sense that the former implies that the reference object is the goal of the agent, whereas the latter does not have this implication.

It is mentioned in Zwarts (2005) that, *to* is an "informationally stronger" preposition than *towards* (pp. 765). This idea captures the main semantic characterization of the two types of prepositions, but it is not further elaborated nor studied in Zwarts. In the current study, we elaborate upon this idea by incorporating the existing literature on implicature studies (e.g., Buccola & Haida, 2019; Grice, 1975; Zhan, 2018), and we apply it to distinguish between *naar* and *richting*. Two types of implicature are of relevance: scalar implicature and ignorance inference.

A typical example of scalar implicature comes from the use of quantifiers such as 'some' and 'all': the use of 'some' (e.g., *Some students participated in this exam*) leads to a scalar implicature that a stronger meaning 'all' is not satisfied (that is, not all students participated in this exam). An ignorance inference occurs when the speaker is ignorant about contextually based propositions. For example, when a speaker says *she lives in Rotterdam or Amsterdam*. The implicature is that the speaker is ignorant about both the proposition 'she lives in Rotterdam' and the proposition 'she lives in Amsterdam'. An ignorance inference also occurs when the speaker is uncertain about something stronger than what is said.

As previously mentioned, *naar* is a goal preposition and it implicates that the reference object is the goal of the agent, whereas *richting* does not have such implication. Based on this definition, we further propose that *naar* is a semantically stronger expression than *richting*. This is because the interpreter could adopt the above-mentioned Gricean reasoning and draw either of the following two inferences from the use of *richting*:

(1) *a scalar implicature*: the speaker knows that *naar* is not satisfied; therefore, the speaker believes that the reference object is not the goal of the agent;

(2) *an ignorance inference*: the speaker does not know whether *naar* is satisfied; therefore, the speaker does not know whether the reference object is the goal of the agent.

Following the above line of thought, we conducted a language comprehension task in Experiment 1. We asked participants to read sentences and drag a slider on a bar to indicate the distance between a moving agent and a reference object mentioned in these sentences. We hypothesized that when *richting* is used in a motion event description (e.g., *hij loopt richting een kerk* 'he walks towards a church', participants should locate the agent further from the reference object than when *naar* is used (e.g., *hij loopt naar een kerk* 'he walks to a church'). This is because participants could conclude from inference (1) or inference (2) triggered by a *richting*-PP that the agent has or may have another goal instead of the reference object.

Accordingly, they should locate the agent further from the reference object to leave room for the possibility that the agent deviates from the path to the reference object.

We furthermore investigated possible factors that might influence the choice between the two prepositions in event description tasks (Experiment 2 and Experiment 3). We have mentioned in the 'Introduction' two factors that might affect event endpoint conceptualization: the actor's goal and the interlocutor's status. We hypothesized that when the inferred goal of the actor is clear to be the reference object, speakers should use *naar* more often than when it is not; This is also consistent with the idea in Experiment 1. Moreover, speakers should use *naar* more often when the speech context does not require cautious and accurate statements than when it does. The use of *richting* should exhibit opposite patterns from the use of *naar* under the influence of these two factors. It is important to note that the simple present tense was used in the sentence stimuli of all three experiments. Reasons for this have been elaborated above (Flecken, 2011; von Stutterheim et al., 2009).

Pre-registration

Hypotheses, sample size, materials, design, exclusion criteria, and analyses of all experiments reported in this paper were pre-registered on the Open Science Framework in advance of data collection and analysis. (See details at:

[1]https://osf.io/9ncdv/?view_only=b4751c303fbb4cd68a130f02b0d758cf; https://osf.io/k52tx/?view_only=bddc5fe4c23b499e800795a816ab340d; https://osf.io/7f2be/?view_only=4283aa91b93b4618891b4be6efac8440.)

Frick's COAST method

We adopted Frick's COAST method (Frick, 1998) to conduct flexible sequential testing during data collection for all three experiments in the current study. We adopted the sequential stopping rule that was proposed by Frick:

> *The researcher can perform a statistical test at any time. If the outcome of this statistical test is $p < .01$, the researcher stops testing participants and rejects the null hypothesis; if $p > .36$, the researcher stops testing participants and does not reject the null hypothesis; and if $.01 < p < .36$, more participants are tested. (Frick, 1998: 691)*

Frick conducted computer simulations to show that the conventional alpha level of .05 is preserved in this procedure.

There are two main reasons to adopt sequential analysis for our study. First of all, there were no similar experiments that had been done before. We could not infer an estimated effect size from an already existing and highly related study to perform an a priori power analysis. Second, as we mentioned in our pre-registration of Experiment 1, we did a pilot

1 Experiment 2 mentioned in this pre-registration is not included in the current study.

study for Experiment 1, but the effect size inferred from this pilot study was not reliable. This was because we used a small sample size in the pilot study. The power of the pilot study was not high enough to provide a reliable effect size to perform an a priori power analysis for the current study. A sequential analysis allowed us to test participants without having to determine a fixed sample size in advance of data collection. Therefore, we decided that sequential analysis was the best way to follow for all our experiments.

Experiment 1

Experiment 1 asked participants to read sentences and drag a slider on a bar to indicate the distance between a moving agent and a reference object mentioned in these sentences.

We had three conditions in Experiment 1: *richting*, *naar,* and goal-reached conditions. To make sure that participants understood the task, we added the goal-reached condition as a control condition (using sentences like *de man loopt de kerk in* 'the man walks into the church'). If participants understood our task correctly, they should always drag the slider to the very right side of the bar to represent that the agent had arrived at the reference object when they read a goal-reached sentence. We hypothesized that the distance between an agent and a reference object should be the shortest in the goal-reached condition, compared to the *naar* and *richting* conditions.

Furthermore, we hypothesized that when *naar* is used in the description of a motion event, participants should drag the slider, representing a moving agent, closer to a reference object than when *richting* is used; when *richting* is adopted, the location of an agent on the trajectory is less constrained by a reference object, given that the use of *richting* implies that the agent has or may have another goal instead of the reference object (i.e., participants should drag the slider further from a reference object to leave room for the possibility that the agent deviates from the path to the reference object).

Method

Participants

Participants were recruited from the Department of Psychology, Education, and Child Studies at Erasmus University Rotterdam. They all received course credits for their participation.

We adopted Frick's COAST method (Frick, 1998) to conduct flexible sequential testing during data collection while preserving an overall alpha level of .05. We decided to recruit participants in batches of 40 (as 40 was the minimum number of participants we decided to test). We would conduct a paired-samples *t*-test between the *naar* and *richting* conditions. According to the stopping rule proposed in Frick (1998), if the planned paired-samples *t*-test between the *naar* and *richting* conditions showed $p > .36$ or $p < .01$ after 40 participants (after data exclusion), we would stop testing participants. If p was within these boundaries, we would test another 40, determine p, and then decide if we needed another 40. We would stop

at $N = 120$, regardless of the p-value of the paired-samples t-test at that time. Following this line of thought, we stopped data collection when the number of participants reached 56 (7 males, mean age 20.36, range: 18-28 years). This warranted that we had 40 participants after data exclusion for the first batch. This was also our last batch, given that the lower boundary of .01 was reached at this point.

Materials

Experiment 1 was programmed using the Qualtrics Survey Software and used a slider answer option to represent a path (trajectory) between a moving agent (e.g., a vehicle, a person) and a reference object (e.g., a gas station, a church).

Figure 1.
An example of the pictures used in Experiment 1.

The moving agent was always presented on the left side of the bar with a value of 0, whereas the reference object was presented on the right side with a value of 100 (see an example of the stimuli in Figure 1). Participants were asked to move the slider on the bar to a specific point to indicate the location of the moving entity based on the sentence they read (e.g., *het meisje loopt naar de speeltuin* 'the girl walks to the playground'). Values were not shown on the bar to prevent participants from making their decisions by remembering the displayed values instead of by reading the sentences carefully. Sixteen pictures with 8 different moving entities (man, woman, girl, boy, car, truck, van and bus) and 16 different reference locations were created. Each moving entity was used twice in combination with the 16 locations to form the 16 pictures. Three sentences representing each condition were created for each picture. Thus, 48 sentences were formed (see 'Appendix I')[2].

Design and procedure

Experiment 1 was a within-participants design in which 2 blocks (each with 24 sentences) were created. Each block contained 3 sentence types (*richting*, *naar* and goal-reached

2 No filler items were added, given that it did not matter whether participants were making a conscious comparison between the conditions or not. It was even preferable if they knew they were comparing different prepositions because their responses were then more consistent.

sentences). The same reference object did not combine with the same directional preposition (i.e., *richting* or *naar*) in the same block. Thus, if a reference object was combined with a *richting* sentence in Block 1, then it was combined with a *naar* sentence in Block 2 (for example, when the sentence *de man loopt naar de kerk* 'the man walks to the church' appeared in Block 1, the sentence *de man loopt richting de kerk* 'the man walks toward the church' would only appear in Block 2). The order of blocks was randomized and the sentences within each block were also randomly presented.

Participants signed their name on an informed consent form first. They then filled out a questionnaire about their demographic information (i.e., age, gender, education) and linguistic background (i.e., native language). After that, participants were instructed to read a sentence and look at a scenario presented below the sentence. They were informed that the picture on the left side of the bar represented the moving person/vehicle they read in the sentence, the picture on the right side of the bar represented the location mentioned in the sentence, and their task was to drag the slider to indicate where the person/vehicle was in relation to the location picture shown in the scenario on the basis of the sentence. They were told to imagine that the slider represented the moving entity when doing so. They did three practice trials first and then moved on to the experimental trials. This experiment took around 10 minutes. The language used in this experiment was Dutch.

Exclusion criteria

Participants who indicated that their first language was not Dutch and those who mentioned several languages as their mother tongue were excluded (12 out of 56 participants). This was to ensure that all participants were monolingual native Dutch speakers. We also removed Participant*Condition groups fulfilling at least one of the three criteria: (a) the standard deviation of the value on the slider was > 25 (i.e., more than a fourth of the total range); (b) the standard deviation was < 1 and the mean was > 95 (i.e., where the slider value was just maxed out); (c) the mean was < 5 (i.e., essentially no directionality in sentences that were clearly directional). Based on these three criteria, we further excluded 3 participants. In order to strictly follow our pre-registered sample size, we excluded the last participant to ensure we had exactly 40 participants for the first batch (which later on proved to be the only batch of participants we needed). Therefore, we excluded in total 16 out of the 56 participants we recruited.

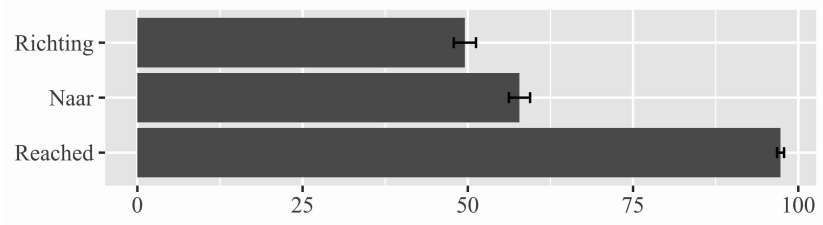
Results and discussion

A paired-samples *t*-test was conducted in R (R Core Team, 2016) between the *naar* and the *richting* conditions. Results showed that participants dragged the slider significantly closer to the reference object in the *naar* condition ($M = 57.8$, $SD = 20.6$) than in the *richting* condition ($M = 49.6$, $SD = 21.6$); $t(39) = 3.604$, $p < .001$, $d = 0.39$). To also include the control condition, we built a multilevel linear model in R using the function *lme* in package *nlme* (Pinheiro,

Bates, DebRoy, Sarkar, & R Core Team, 2019). We adopted the model (Value ~ Condition, random = ~1|Participant/Condition), given that Experiment 1 was a one-way repeated measures design and this model considered dependency in our data. The random effect we included in this model (random = ~1|Participant/Condition) specified that data of different levels of Condition were from the same participant. Besides that, we included in this model our predictor Condition. To assess the overall effect of Condition, we built a baseline model that did not include the predictor Condition but did include the intercept and the random effect (Value ~ 1, random = ~1|Participant/Condition). A model comparison using the ANOVA function showed that our predictor Condition had a significant effect on the values shown on the bar ($\chi^2(2) = 147.8803$, $p < .001$). Additional Tukey Post Hoc analyses using the function *glht* from the package *multcomp* in R (Hothorn, Bretz, & Westfall, 2008) also showed that there was a significant difference between the *naar* condition and the *richting* condition ($\beta = 8.178$, $SE = 2.640$, $z = 3.098$, $p = .006$). Moreover, there was a significant difference between the goal-reached condition and the *naar* condition ($\beta = 39.600$, $SE = 2.640$, $z = 15.001$, $p < .001$), and between the goal-reached condition and the *richting* condition ($\beta = -47.778$, $SE = 2.640$, $z = -18.009$, $p < .001$). Figure 2 presents the mean values of the position of a moving agent in relation to a reference object (a potential destination) on the bar (the larger the value, the closer an agent is to a reference object).

Figure 2.

Mean values of the position of a moving agent in relation to a reference object on the bar (from 0 to 100).



Our findings support our hypothesis. Participants indeed located the agent significantly further from the reference object when *richting* was used than when *naar* was used. By doing so, they assumed the possibility that when *richting* is used the moving agent has another goal instead of the reference object, and deviates from the path to the reference object. In other words, they were more certain about the reference object being the goal of the agent when *naar* was used than when *richting* was used.

This leads to the following question. What are the factors that affect a speaker's certainty about the reference object being the goal of the agent, and subsequently the choice between *naar* and *richting*? We used event description tasks to answer this question in Experiment 2. We investigated two factors (i.e., the actor's goal and the interlocutor's status) in Experiment

2. Experiment 3 was an extension of Experiment 2. We manipulated the strength of the actor's goal in Experiment 3 and tested whether a stronger intention/goal had a larger effect.

Experiment 2

We aimed to investigate two possible factors that might influence the use of *naar* and *richting* in real-life situations: the actor's goal (Intention) and the interlocutor's status (Context). We hypothesized that *naar* should be used more often when the actor's goal could be inferred from the referential situation. For example, a man is walking with a trash bag in his hand and there is a trash bin in the near distance. The trash bin could be used as an inference that the man's goal is to throw the trash bag into the trash bin. On the other hand, *richting* should be chosen more frequently when the reference object could not be inferred as the goal of the agent in the referential situation. For example, a man is walking with nothing in his hand and there is a trash bin in the near distance. There is no extra information in the referential situation that can be used to infer that the trash bin is the goal of the man.

We furthermore hypothesized that speech context should also affect the choice between the two prepositions. For example, when there is a (perceived) requirement of an accurate description of an event, such as talking to a police officer, people should be more conservative and cautious with their linguistic choice than normally (Ainsworth, 1993). This means that they should use *richting* more often than they normally do, for *richting* does not come with the implication that the reference object is the goal of the agent.

Specifically, Experiment 2 tested the choice between *naar* and *richting* in two different referential situations (an intention shown vs. no intention shown) and in two distinct contexts (a description to a police officer vs. to a friend). To manipulate intention, we created two versions of a picture. In both versions, a person was shown and the same reference object was placed in its near distance. The person was either holding something that was semantically related to the reference object (e.g., a trash bag vs. a trash bin) or not holding anything. In total we created two different pictures, each with two versions (See 'Appendix II').

To manipulate context, we created two versions of the instruction for each version of each picture. The instruction was either "*You describe the following scenario to a police officer as a witness. Please choose a word that you think fits better in the sentence*" (e.g., *hij loopt (?) de container*) or "*You describe the following scenario to a friend. Please choose a word that you think fits better in the sentence*" (literally translated from the original Dutch versions). *Naar* and *richting* were options shown below each scenario.

We hypothesized that (1) *naar* should be chosen more often when there is an intention of arriving at the reference object than when there is no such intention shown in the referential situation; *richting* should exhibit the opposite pattern; and (2) *richting* should be adopted more often when participants are required to describe a scenario to a police officer than to a friend; *naar* should show an opposite pattern. Overall, we hypothesized a main effect of

Intention and a main effect of Context on the choice between *richting* and *naar*.

Method

Participants

As in Experiment 1, we adopted Frick's COAST method to conduct flexible sequential testing. We decided to recruit participants in batches of 160 (the minimum number of participants we planned to test), 10 in each cell (that is, 80 per intention and 80 per context). If $p < .01$ or $p > .36$ for each main effect we tested (Intention and Context), we would stop testing participants. If p was within these boundaries for any one of the two main effects predicted, we would test another 160 participants. We would stop at $N = 480$ (240 per intention and 240 per context), regardless of the p-values at that time. In the end, we collected valid data from 480 participants (203 males, mean age 22.34, range: 16-71 years) after excluding 17 participants whose mother tongue was not Dutch. Participants were recruited in various ways, such as by posting the survey link on social media, sending emails, or asking people on campus, etc.

Materials

Experiment 2 was programmed using the Qualtrics Survey Software. Two scenarios were created and each scenario was created with two versions (with intention and without intention (see 'Appendix II'): a person dragging a trash bag with a posture of walking and a trash bin in the near distance (with intention) /the same person without dragging a trash bag with a posture of walking and the trash bin in the near distance (without intention); a person holding a bike with a posture of walking and a bike repair shop in the near distance (with intention)/the same person without holding a bike with a posture of walking and the bike repair shop in the near distance (without intention).

Figure 3.
An example of the stimuli used in Experiment 2 (with an intention shown in the referential scenario).

The person with/without an object on hand was always on the left side of the display and the referred object was always on the right. A sentence was shown below each scenario,

for example, *hij loopt (?) de container* 'he walks (?) the trash bin'. A choice between *naar* and *richting* was shown below the sentence (see an example of the stimuli in Figure 3).

Design and procedure

Sixteen cells were created (2 Scenario × 2 Intention × 2 Context × 2 Option Order). Instructions were manipulated between 'a description to a police officer' and 'a description to a friend' (Context). The order of *naar* and *richting* as options was counterbalanced (Option Order). This experiment was a between-participants design. Each cell presented only one item. Each participant was randomly assigned to any one of the 16 cells. Participants first read an informed consent form attached to the survey. After they agreed to continue with the survey, they then answered online questions about their demographic information (i.e., age, gender) and linguistic background (i.e., native language). After that, they read the instruction that appeared on the screen, made a choice between *naar* and *richting* based on the scenario they saw by completing the incomplete sentence shown below the scenario (e.g., *hij loopt (?) de container* '*he walks (?) the trash bin*').

Results and discussion

Confirmatory analyses

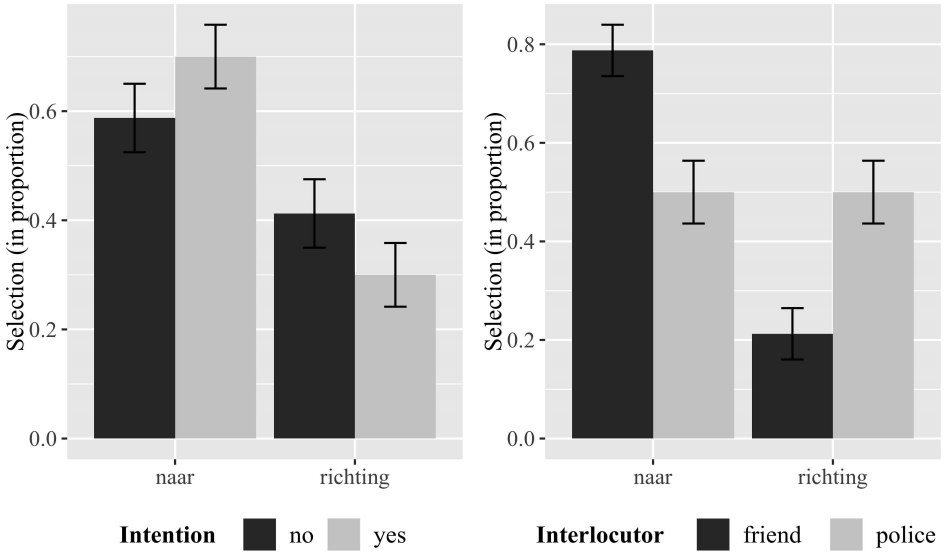
A binomial logistic regression model that included the main effect of Intention and the main effect of Context was conducted in R using the *glm* function implemented in the package *lme4* (Bates, Mächler, Bolker, & Walker, 2015)[3]. Both factors were dummy coded. In the first batch of collected data (N = 160), we did not find a significant main effect of Intention but the p-value was within the boundary from .01 to .36 (β = -0.35, SE = 0.34, z = -1.024, p = .31, odds ratio: 0.70, 95% CI: 0.35-1.38). We did find a significant main effect of Context (β = 1.515, SE = 0.35, z = 4.374, p < .001, odds ratio: 4.55, 95% CI: 2.34-9.13). Given that the p-value found for the effect of Intention was within the boundary from .01 and .36, we continued data collection until we reached 480 participants. The p-value of the main effect of Intention reached lower than .01 when N = 480 (β = -0.544, SE = 0.20, z = -2.686, p = .007, odds ratio: 0.58, 95% CI: 0.39-0.86): based on the standards of the sequential analysis we pre-registered, it was considered as a significant effect. A significant main effect of Context remained when N = 480

3 In the pre-registration, we stated that we would include the random effect of items in our binomial logistic regression model (formula used in R: preposition ~ intention + context + (1|item)). However, we obtained a singular fit using this model. This often indicates that the model is overfitted. Moreover, the AIC value of a more parsimonious model (with only the fixed effects of intention and context) was lower than that of the model that included the random effect of items (202.5 vs. 204.6). This means that the parsimonious model is better. Therefore, we decided to adopt a binomial logistic regression model that only contained the two fixed effects (i.e., intention and context; formula used in R: preposition ~ intention + context). This model was used for the analysis of the later on added data as well (i.e., the second and the third batch of data).

(β = 1.331, *SE* = 0.21, *z* = 6.462, *p* < .001, odds ratio: 3.79, 95% CI: 2.54-5.70). Figure 4 shows the mean proportions of the selection of each preposition in each Intention condition and in each Interlocutor condition (i.e., Context).

Figure 4.

Mean proportions of the selection of the two prepositions (i.e., *naar* and *richting*) in each Intention condition (left) and in each Interlocutor condition (right).



Exploratory analyses

Besides the main effects of Intention and Context on the choice between *naar* and *richting*, we were interested in whether the scenario type (a trash bin in the near distance or a bike shop in the near distance) showed a main effect or an interaction effect with Intention or with Context. Moreover, we were also interested in whether there was a main effect of Option Order (i.e., *naar*/*richting* or *riching*/*naar*). Therefore, we built a binomial logistic regression model that included the main effect of Intention, Context, Option Order, Scenario, the interaction between Scenario and Intention, and the interaction between Scenario and Context (the formula used in R: preposition ~ intention*scenario + context*scenario + option order). All the factors were sum coded except for the factor Option Order (dummy coded), as we were interested in only the main effect of this factor. Significant main effects of Intention (intercept β = -1.024, *SE* = 0.15, *z* = -6.657, *p* < .001; β = 0.278, *SE* = 0.10, *z* = 2.708, *p* = .007, odds ratio: 1.32, 95% CI: 1.08-1.62) and Context (β = -0.685, *SE* = 0.11, *z* = -6.526, *p* < .001, odds ratio: 0.50, 95% CI: 0.41-0.62) were remained in this model. There was no significant main effect of Scenario (β = 0.096, *SE* = 0.10, *z* = 0.919, *p* = .36, odds ratio: 1.10, 95% CI: 0.88-1.35).

No interaction effect was found between Scenario and Intention (β = 0.045, *SE* = 0.10, *z* = 0.434, *p* = .67, odds ratio: 1.05, 95% CI: 0.86-1.28), nor between Scenario and Context (β = 0.026, *SE* = 0.11, *z* = 0.249, *p* = .80, odds ratio: 1.03, 95% CI: 0.84-1.26). There was a significant main effect of Option Order (β = 0.680, *SE* = 0.21, *z* = 3.305, *p* < .001, odds ratio: 1.97, 95% CI: 1.33-2.97). Participants used *richting* more often when the option order was shown as *richting*/*naar* than when it was *naar*/*richting*.

In sum, our findings support our hypotheses on the main effect of Context and the main effect of Intention. We found a significant effect of Context on the use of *naar* and *richting*. That is, when participants were required to describe a scenario to a police officer, they were more cautious and more conservative with their linguistic choice, and they used *richting* 3.79 times more often than when they were asked to describe the scenario to a friend. As expected, the opposite pattern occurred for the use of *naar*. *Naar* was used 3.79 times more often in the context of talking to a friend than in the context of talking to a police officer. We take this to mean that, when the speech context was talking to a friend, participants were less careful in making a certain statement and resorted to the default. Given that *naar* is very common in daily conversations, participants showed a more frequent use of *naar* in the context of talking to a friend than that of talking to a police officer. We also found a significant effect of Intention on the use of *naar* and *richting*. Specifically, when there was an intention shown going to a certain place (e.g., a trash bag and trash bin, a broken bike and a bike repair shop), participants used *naar* 1.72 times (1/0.58) more often than when no intention of going to a certain place was presented. The opposite pattern appeared for the use of *richting*. However, this effect is smaller than the effect of Context (odds ratios: 3.79 vs. 1.72).

Why is the effect of Intention on the use of *naar* and *richting* smaller than that of Context? Perhaps our manipulation of intention was not powerful enough. We used one picture with a person dragging a trash bag with a trash bin in the near distance, and another picture with a person holding a bike and a bike shop in the near distance. Both pictures only showed a relational combination of two objects that belong to the same semantic category (e.g., trash bag and trash bin; bike and bike shop). Their relation might not be strong enough to imply that a trash bag would necessarily be thrown into a trash bin and a bike would definitely be taken to a bike shop. When the conversational context implies a more conservative and cautious answer, the uncertainty that one object (e.g., the trash bag) does not necessarily end at the other (e.g., the trash bin) becomes even stronger. Therefore, we hypothesized that the effect of Intention should get stronger if we could increase the strength of intention, which is what we tried to accomplish in Experiment 3.

Experiment 3

Following the above line of thought, Experiment 3 investigated the effect of Intention on the use of *naar* and *richting* by increasing the strength of intention. We hypothesized that, when an intention could be more readily inferred from the picture, speakers should use

naar more often than when the intention is not as salient. We manipulated the degrees of intention from *strong*, via *moderate* to *weak*. Compared to Experiment 2, the addition was the *strong* condition (see details in the 'Materials' section). Furthermore, we only retained the context of talking to a police officer from that experiment. As Experiment 2 showed, speakers used *richting* as often as *naar* (*naar* vs. *richting*: 120 vs. 120) when the speech context was a description to a police officer, whereas *naar* was dominantly used (*naar* vs. *richting*: 189 vs. 51) when it directed at a friend. This suggests that there is more room to detect an increase of the use *naar* when the context is talking to a police officer instead of to a friend. There might be a ceiling effect on the use *naar* when the interlocutor is a friend. Therefore, our hypothesis was that in the context of talking to a police officer, with the increase of the strength of intention of motion from *weak*, *moderate*, to *strong*, the use of *naar* would, accordingly, increase.

Method

Participants

As in Experiment 2, we adopted Frick's COAST method to conduct flexible sequential testing. We decided to recruit participants in batches of 120 (the minimum number of participants we planned to test), 20 in each cell (that is, 40 per scenario: *strong* vs. *moderate* vs. *weak*). If $p > .36$ or $p < .01$ for the main effect of Scenario and each comparison we tested, we would stop testing participants. If p was within these boundaries for the main effect of Scenario and any one of the three comparisons, we would test another 120 ($N = 320$; 160 per intention and 160 per context). We stopped when the number of participants reached 540 (251 males, mean age 21.34, range: 17-38 years)[4] after excluding 13 participants whose mother tongue was not Dutch. Participants were recruited mainly by asking people on campus.

Materials

This experiment was programmed using the Qualtrics Survey Software. Three scenarios were created. The strong condition showed a fire truck with its siren on and a building on fire in the near distance. In this condition, there is a causal connection between the firetruck and the building that is on fire. That is, there is a big fire nearby, the fire truck has its siren on and is therefore on its way to combat the fire.

The moderate condition showed a fire truck with its siren off and a fire station in the near distance. In this condition, there is a semantic connection between the firetruck and the building but not necessarily a causal one. A fire truck with its siren off is not necessarily going to a fire station. It can be that it is still going to a building that is on fire (since the road is not busy, the siren can be off) or going to another fire station.

4 In the pre-registration, we stated that we would stop data collection when the number of participants reached 360. However, we decided to continue data collection because the *p*-value of the main effect of Scenario was still within the boundary of .01 to .36 when $N = 360$.

Finally, the weak condition showed a fire truck with its siren off and a church in the near distance. In this case, there is no particular connection between the firetruck and the building. They do not even belong to the same semantic category. When people talk about a fire truck, they normally do not think of a church (see all the stimuli in 'Appendix III').

A sentence was shown below each scenario, for example*, de brandweerauto rijdt (?) het brandende gebouw* 'the fire truck drives (?) the building on fire'. Participants had to choose between *naar* and *richting* to complete the sentence based on the scenario they saw. Participants' age, gender, and mother tongue were also recorded.

Design and procedure

Six cells were created (3 Scenario × 2 Option Order). The order of *naar* and *richting* as options was counterbalanced. As in Experiment 2, Experiment 3 was also a between-participants design. Each cell had only one item. Each participant was randomly assigned to any one of the 6 cells. Participants first read an informed consent form attached to the survey. After they agreed to continue with the survey, they then filled in their demographic information (i.e., age, gender) and linguistic background (i.e., native language). After that, they followed the instruction that appeared on the screen (*Please describe the scenario to a police officer as a witness*), made a choice between *naar* and *richting*, based on the scenario they saw, to complete the incomplete sentence shown below the scenario (e.g., *de brandweerauto rijdt (?) het brandende gebouw* 'the fire truck drives (?) the building on fire').

Norming study

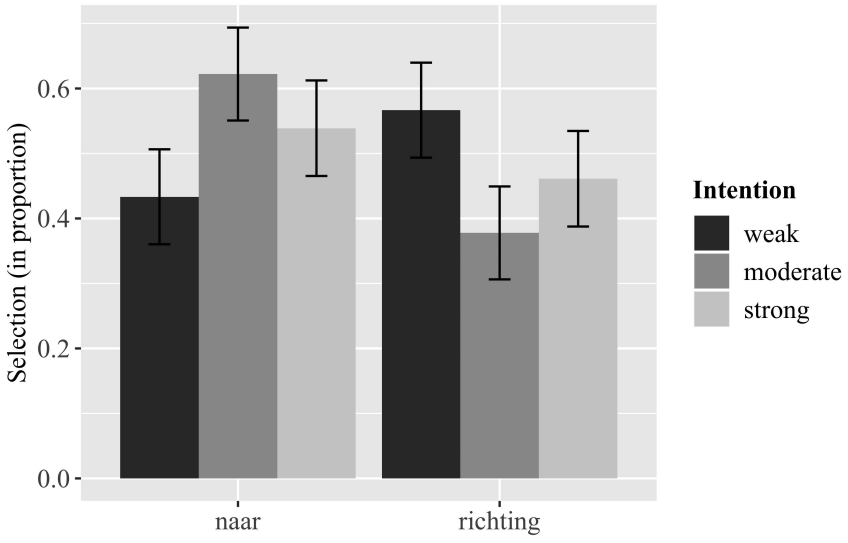
To ensure that there was indeed a difference in the degree of connection between the two objects shown in the three scenarios, we performed a norming study. We asked 120 Dutch speakers (40 per scenario) to rate the possibility of the fire truck's destination being at a building on fire/a fire station/a church using a 7-point Likert scale (very likely, moderately likely, a bit likely, neither likely nor unlikely, moderately unlikely, a bit unlikely, very unlikely: from 1 to 7). Each participant was assigned to only one of the three scenarios to avoid them forming conscious comparisons between scenarios and from over-thinking (e.g., the same fire truck with three destinations). We found that 40% (16 out 40) and 30% (12 out of 40) participants chose 'very likely' and 'moderately likely', respectively, for the scenario with a strong intention (i.e., a fire truck with its siren on and a building on fire; possibility rating from 1 to 7: M = 2.58, 95% CI = 1.95-3.21). 12.5% (5 out of 40) and 32.5% (13 out of 40) participants chose 'very likely' and 'moderately likely', respectively, for the scenario with a moderate intention (i.e., a fire truck with its siren off and a fire station; possibility rating from 1 to 7: M = 3.15, 95% CI = 2.65-3.66). Only 2.5% (1 out 40) and 7.5% (3 out 40) participants chose 'very likely' and 'moderately likely', respectively, for the scenario with a weak intention (i.e., a fire truck with its siren off and a church; possibility rating from 1 to 7: M = 4.55, 95% CI = 4.06-5.04). We assume that the extent to which speakers choose 'very likely', compared to other

possibility options, actually determines whether they use *naar* or *richting* in their language production. Any bit of uncertainty in this respect might lead to the use of *richting* in real speech. Only when speakers are very certain about the destination (when they choose 'very likely') will they use *naar* in the context that requires speech accuracy (e.g., talking to a police officer). Therefore, we did find a difference between these three scenarios, in the sense that people chose 'very likely' much more often for the building on fire scenario than the other two scenarios (building on fire vs. fire station vs. church: 40% vs. 12.5% vs. 2.5%).

Results and discussion

A binomial logistic regression model that included the main effect of Scenario was conducted in R. This factor was dummy coded. We found a significant difference between IntentionModerate and IntentionWeak (*intercept* β = -0.499, *SE* = 0.1537, *z* = -3.246, *p* = .001; β = 0.7673, *SE* = 0.2151, *z* = 3.567, *p* < .001, odds ratio: 2.15, 95% CI: 1.42-3.29). There was no significant difference between IntentionStrong and IntentionWeak (β = 0.4241, *SE* = 0.2121, *z* = 2.000, *p* = .046, odds ratio: 1.53, 95% CI: 1.009-2.320), nor was there a significant difference between IntentionStrong and IntentionModerate (β = 0.3431, *SE* = 0.2145, *z* = 1.600, *p* = .110, odds ratio: 1.41, 95% CI: 0.93-2.15). An extra Wald test showed that there was a significant main effect of Scenario ($\chi^2(2) = 12.8$, *p* = .002). Figure 5 exhibits the mean proportions of the selection of each preposition in each Intention condition.

Figure 5.
Mean proportions of the selection of the two prepositions (i.e., *naar* and *richting*) in each Intention condition (i.e., *weak*, *moderate*, *strong*).



Experiment 3 supports what we have found in Experiment 2; that is, *naar* was used significantly more often when there was an intention of arriving at the reference object than when there was no such intention (IntentionModerate vs. IntentionWeak: $p < .001$). However, contrary to our prediction we did not find an increased effect of intention even though we increased the salience of the intention. Instead, the expected moderate degree of intention had, if anything, a stronger effect than the expected strong degree of intention. Possible reasons will be discussed in the general discussion section.

General discussion

We conducted three experiments to examine event endpoint conceptualization. Two Dutch directional prepositions (i.e., *naar* and *richting*) provide a useful paradigm to ask precise questions about endpoint conceptualization. Experiment 1 adopted a sentence comprehension task and confirmed the linguistic proposal that the two directional prepositions differ in their implication on the goal of the agent. Experiment 2 and Experiment 3 used this linguistic pair to test the effect of two factors (i.e., the actor's goal and the interlocutor's status) on event endpoint conceptualization via language production tasks.

Experiment 1 used a slider dragging task and showed that when *richting* was used to describe a motion event, participants placed the moving agent further from the reference object than when *naar* was adopted; moreover, when a goal-reached expression was used, participants placed the moving agent nearest to the reference object.

It is proposed that comprehenders adopted Gricean reasoning when deciding on the distance between the moving entity and the reference object. The three expressions (i.e., *naar*, *richting*, and goal-reached expressions) differ in their strength of semantics. Distances were divided up into three corresponding sections based on their semantics (i.e., scalar implicature): nearest/inside, near, and not near. A goal-reached expression means that the agent arrives at the reference object (that is, nearest to the reference object/inside of the reference object). It is semantically the strongest, compared to *naar* and *richting*. *Naar* implies the reference object is the goal of the agent but the goal is not yet reached. Therefore, on a semantic scale, it is not as strong as a goal-reached expression. Accordingly, the agent was placed not as close to the reference object as a goal-reached expression, but near it. *Richting* has the weakest semantics, given that it triggers the inferences that the agent has or may have another goal instead of the reference object. Therefore, participants located the agent the furthest from the reference object, compared to the other two expressions. By doing so, they left room for the possibility that the agent deviates from the path to the reference object and goes to other places.

Having shown that *richting* and *naar* indeed differ semantically, we set out to investigate factors that might affect the use of the two words and hence endpoint conceptualization in Experiment 2 and Experiment 3. In both experiments, we asked participants to make a very specific linguistic choice between *naar* and *richting* based on a motion scenario they were

presented with. This task is simple and ecological. By giving participants a well-defined task, we obtained a linguistic report that can be taken as an indicator of the thought process that occurs during decision-making (Ericsson & Simon, 1993).

In Experiment 2, as hypothesized, we found a very strong main effect of Interlocutor (odds ratio: 3.79). Specifically, when participants were required to describe a scenario of a person moving *naar* or *richting*, they used *richting* more often when they were asked to describe the scenario to a police officer than to a friend. This presumably occurred because they wanted to be more cautious and more conservative with their linguistic choice when the recipient of their description was an officer of the law. Whereas *naar* implies that the goal of the person is the reference object, *richting* merely indicates the direction. In other words, by using *naar*, participants were committing themselves to stating the goal of the protagonist, something they were less willing to do when talking to a police officer. When the context was talking to a friend, participants showed a more frequent use of *naar*, compared to that of talking to a police officer. This finding provides empirical support for the view that contextual factors affect endpoint conceptualization ((Papafragou & Grigoroglou, 2019).

One might argue that the effect of Interlocutor might be due to the formality of the situation (i.e., register). Describing an event to a police officer normally requires a more formal register than to a friend. *Richting*, as a denominalized preposition, might appear more often in a more formal register than *naar*. Hence, *richting* was used more often when the interlocutor was a police officer instead of a friend. This line of argument is plausible but does not provide the main reason for the effect. The chief reason for a person to choose between *naar* and *richting* still has a semantic basis (i.e., a goal or a direction), a semantic basis that is also evident in the effect of the type of stimuli, which we will discuss next. Moreover, in formal occasions, people usually feel the need to be cautious with their statements. Talking to a police officer is very likely to be an example of this. Therefore, we propose that register is a possible factor that can affect the choice between *naar* and *richting*, but is definitely not the only or main reason, and it can go hand in hand with other reasons (i.e., being cautious, accurate and specific).

In addition to the effect of the social status of the interlocutor, we found that the choice between *naar* and *richting* was also influenced by the characteristic of the stimulus, namely the degree to which the intention of the protagonist going to the reference object could be inferred from the stimulus. This also supports our hypothesis. Specifically, when it could be inferred that the reference object in the display was the goal of the agent, participants used *naar* more often than when no clear inference was presented (i.e., no trash bag or bike was shown), and the opposite patterns occurred for *richting*. This finding clearly confirms the idea that knowing the actor's goal of performing an action is important in event endpoint conceptualization. However, this effect was smaller than the effect of Interlocutor (odds ratios: 1.72 vs. 3.79).

In Experiment 3 we tried to increase the salience of the intention shown in the referential

situation and investigated whether this would increase its effect on the use of *naar* and *richting*. We argued that there is a semantic relation between, for instance, a trash bag and a trash bin. That is, these two concepts belong to the same semantic category and they do co-occur in the same text frequently[5]. We proposed that such a relation, however, might not be strong enough to exhibit a strong effect of Intention. The mere presence of a trash bin does not necessarily cause a trash bag to be thrown in it. Hence, we created a causal relation in Experiment 3 that we argued could be a stronger semantic relation (i.e., a firetruck and a burning building). However, we did not find a significant difference between the causal scenario and the relational scenario, not even between the causal scenario and the irrelevant scenario, in the use of *naar* and *richting*. Participants did not use less *richting*/more *naar* in the causal scenario than in both the relational scenario and the irrelevant scenario.

This unpredicted finding prompts further thoughts. One possible reason is that, although a burning building is indeed a more salient goal for the fire truck driver than a fire station when the siren of the fire truck is on, the inference that the fire truck driver's goal is the burning building shown in the referential situation is still a defeasible inference. That is, people can still think of other places being the fire truck driver's goal instead of the one shown in the referential situation; for example, another burning building. Therefore, it is not evident that we would find a stronger effect of Intention if we choose a burning building as the reference object rather than a fire station. Another possibility is that we ignored a possible 'ongoingness' implication in the strong condition. That the building is on fire implies that this situation is happening right now. This could create a sense of urgency, which makes participants want to emphasize that the firetruck is on its way and is getting nearer and nearer to the reference object. *Richting* then becomes a more proper choice than *naar* in this case, given that *richting* emphasizes the direction and the trajectory of motion.

As mentioned earlier, Zwarts (2005, 2008) proposed that atelic directional PPs have a cumulative nature, whereas telic directional PPs do not. Atelic directional PPs denote paths of a progression that are connected and add up cumulatively. This proposal might point out the progressive and dynamic nature of atelic directional PPs, such as *richting*-PPs, and indicate that the use of *richting*-PPs in event descriptions creates a richer mental representation with more details on what is happening along the way than the use of *naar*-PPs. This kind of use of *richting* is reminiscent of the role progressive aspect plays in event descriptions. There is empirical support for the claim that the use of progressive marker makes the details on the happening of an event more available, for example, actions of the character, location, etc., than the use of perfective aspect (Ferretti, Kutas, & McRae, 2007; Matlock, 2011). For future research, it would therefore be meaningful to test experimentally whether the use of *richting*

5 We searched collocations of the word 'trash' via Sketch Engine and found that there was a high association strength between the word 'trash' and the word 'bin', 'dumpster', and 'container': with a logDice of 8.80, 8.31, 7.75, respectively (for more explanations about logDice, see Rychlý, 2008 and Gablasova, Brezina, & McEnery, 2017).

indeed evokes a more detailed representation of a depicted event, compared to the use of *naar*.

Future research should also explore more languages that differ in their aspectual systems. For instance, English differs from Dutch in its aspectual system, given that aspectual markers are more grammaticalized in English than in Dutch. It would then be useful to study whether *to* and *towards* in English work the same way as *naar* and *richting* work in Dutch, or whether the use of progressive aspect in English creates a different understanding and use of *to* and *towards*. It is possible that the function of *towards* in English is less salient than that of *richting* in Dutch in representing an event, as progressive aspect in English might direct the comprehender's attention to other aspects of the event, for instance, the actor's manner. By investigating more languages, we are able to build a systematic understanding of how grammatical aspect and directional prepositions interplay in building situation models of an event.

Conclusion

Our results contribute to a deeper understanding of event cognition and event conceptualization. Spatial relations between one entity and another entity offer important information in building an event model (Radvansky & Zacks, 2014; Zwaan & Radvansky, 1998). By employing experimentally less studied linguistic features (i.e., two directional prepositions in Dutch: *naar* and *richting*), we are able to assess the extent to which the two directional prepositions differ in representing the spatial relations in an event through a sentence comprehension task. Our study shows that people are sensitive to such spatial relations, whether the reference object is the goal or just an indicator of the direction of a moving agent when building an event model.

Moreover, we used an event description task to explore the effect of two factors (i.e., the actor's goal and the interlocutor's status) on event conceptualization, targeting endpoint conceptualization. We find that both factors affect how people conceptualize an event endpoint. Our study manifests the importance of considering contextual factors, such as the social status of the interlocutor, in event description studies.

Appendix

I. Sentences used in Experiment 1

(1) GOAL-reached condition:

1.	de auto rijdt de tunnel in	'the car drives into the tunnel'
2.	de auto rijdt de wasstraat in	'the car drives into the carwash'
3.	de bus rijdt de poort in	'the bus drives into the gate'
4.	de bus rijdt de stad in	'the bus drives into the city'
5.	de jongen loopt de dierentuin in	'the boy walks into the zoo'
6.	de jongen loopt het huis in	'the boy walks into the house'
7.	de man loopt de kerk in	'the man walks into the church'
8.	de man loopt het museum in	'the man walks into the museum'
9.	de vrachtwagen rijdt de parkeergarage in	'the truck drives into the parking lot'
10.	de vrachtwagen rijdt het bos in	'the truck drives into the wood'
11.	de vrouw loopt de winkel in	'the woman walks into the store'
12.	de vrouw loopt het ziekenhuis in	'the woman walks into the hospital'
13.	het busje rijdt de garage in	'the van drives into the garage'
14.	het busje rijdt het dorp in	'the van drives into the village'
15.	het meisje loopt de flat in	'the girl walks into the apartment'
16.	het meisje loopt de speeltuin in	'the girl walks into the playground'

(2) *Naar* and *richting* conditions:

1.	de auto rijdt naar/richting de tunnel	'the car drives to/towards the tunnel'
2.	de auto rijdt naar/richting de wasstraat	'the car drives to/towards the carwash'
3.	de bus rijdt naar/richting de poort	'the bus drives to/towards the gate'
4.	de bus rijdt naar/richting de stad	'the bus drives to/towards the city'
5.	de jongen loopt naar/richting de dierentuin	'the boy walks to/towards the zoo'
6.	de jongen loopt naar/richting het huis	'the boy walks to/towards the house'
7.	de man loopt naar/richting de kerk	'the man walks to/towards the church'
8.	de man loopt naar/richting het museum	'the man walks to/towards the museum'
9.	de vrachtwagen rijdt naar/richting de parkeergarage	'the truck drives to/towards the parking lot'
10.	de vrachtwagen rijdt naar/richting het bos	'the truck drives to/towards the wood'
11.	de vrouw loopt naar/richting de winkel	'the woman walks to/towards the store'
12.	de vrouw loopt naar/richting het ziekenhuis	'the woman walks to/towards the hospital'
13.	het busje rijdt naar/richting de garage	'the van drives to/towards the garage'
14.	het busje rijdt naar/richting het dorp	'the van drives to/towards the village'
15.	het meisje loopt naar/richting de flat	'the girl walks to/towards the apartment'
16.	het meisje loopt naar/richting de speeltuin	'the girl walks to/towards the playground'

II. Stimuli used in Experiment 2

Hij loopt (?) de container.

naar richting

Hij loopt (?) de container.

naar richting

Hij loopt (?) de fietsenmaker.

naar richting

Hij loopt (?) de fietsenmaker.

naar richting

III. Stimuli used in Experiment 3

De brandweerauto rijdt (?) het brandende gebouw.

naar

richting

De brandweerauto rijdt (?) de brandweerkazerne.

naar

richting

De brandweerauto rijdt (?) de kerk.

naar

richting

Chapter 5

Linguistic and non-linguistic cues in motion event endpoint conceptualization: the choice between two English directional prepositions

This chapter has been submitted for publication:

Liao, Y., Dijkstra, K., & Zwaan, R.A. (submitted). Linguistic and non-linguistic cues in motion event endpoint conceptualization: The selection between English *to* and *towards*

Abstract

The current study investigates the role of both linguistic (i.e., grammatical aspect) and non-linguistic factors (i.e., the actor's goal and the interlocutor's social status) in motion event endpoint conceptualization. In language production experiments on the use of two Dutch directional prepositions ((i.e., *naar* and *richting*), a previous study found that both the actor's goal (Intention) and the interlocutor's social status (Interlocutor) affect motion event endpoint conceptualization (Liao, Dijkstra, & Zwaan, 2021). The current study aims to extend these findings by examining a similar pair of directional prepositions in English (i.e., *to* and *towards*). Moreover, we aim to study whether grammatical aspect (i.e., the English simple present and the English progressive aspect) affects the sensitivity to the two non-linguistic factors and consequently affects endpoint conceptualization. In Experiment 1, we used the English simple present for all sentence stimuli (e.g., *he walks (?) the trash bin*). We found a significant effect of Interlocutor on preposition choice, but no significant effect of Intention. In Experiment 2, we replaced the English simple present with the English progressive aspect (e.g., *he is walking (?) the trash bin*). We found significant main effects of both Interlocutor and Intention on preposition choice. These findings extend those reported in Liao et al. (2021) in that protagonist intention and interlocutor status were found to indeed affect motion event endpoint conceptualization. The current findings furthermore show that grammatical aspect affects people's sensitivity to these factors, thus also affecting event endpoint conceptualization.

Keywords: The actor's goal; the interlocutor's social status; grammatical aspect; directional prepositions; motion event endpoint conceptualization

Introduction

To describe an event in words, we first need to process all the relevant information about the event and then decide what message we want to convey to achieve a certain communicative purpose. These processes of message planning before the formation of utterances are called the conceptualization phase (Levelt, 1989, 1999). In motion event conceptualization studies, converging evidence shows that the endpoint of a motion event is more salient than the source of a motion event during the event conceptualization phase (e.g., Lakusta & Carey, 2015; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007). Endpoints are mentioned more often (e.g., *the bird flew **to a tree***) than sources (e.g., *the bird flew **from a signpost***) in motion event descriptions (e.g., Papafragou, 2010). Moreover, endpoints are remembered better than sources after the description task (e.g., Lakusta & Landau, 2012). Plenty of studies have discussed this endpoint-bias phenomenon and the reasons why it occurs (e.g., Johanson, Selimis, & Papafragou, 2019; Papafragou, 2010). Hence, we do not elaborate on this matter any further. In the current study, what we are particularly interested in is the factors that might affect motion event endpoint conceptualization, given the salient status of an event endpoint in motion event conceptualization.

The first factor that we find relevant and important to the identification of a motion event endpoint during motion event conceptualization is the intention of the moving agent. Anticipating upcoming information and predicting the near future is a fundamental part of our daily life. Given that human actions are often goal-directed, knowing the goal of the actor can greatly help the observer to understand an ongoing activity by predicting its possible endpoint (Zacks, 2004). For instance, if you see that your sister is drawing and you know that your sister wants to draw a house, you will then not expect her to stop after she just finished drawing a roof. A predicted endpoint of her drawing activity will be a complete picture of a house. Studies have shown that the ability to infer an event endpoint from the actor's goal has already been found in infants (Baldwin, Baird, Saylor, & Clark, 2001).

In motion event studies, researchers have also discovered a strong connection between the animacy/intentionality of the agent and the memory of the event endpoint during event cognition (Lakusta & Carey, 2015; Lakusta & Landau, 2012). However, most of these studies focus on comparing animate agents with inanimate agents. Not much attention has been paid to comparing between animate agents, such as the extent to which the presence/absence of a clear goal of an animate agent affects the construal of a motion event endpoint. Therefore, in the current study, we would like to investigate this aspect. We expect to find a strong connection between the actor's goal and the identification of a motion event endpoint during event conceptualization.

Another factor that we consider important to motion event endpoint conceptualization is the formality of the speech context. As proposed by Heylighen and Dewaele (2002), there are four parameters that determine the degree of formality in speech context, namely the

speech topic, the setting, the speech modality (written vs spoken), and the interlocutor. In the current study, we are especially interested in examining the effect of the social status of the interlocutor (i.e., the social distance between the speaker and the interlocutor) on motion event endpoint conceptualization. As social animals, we care about the social status of our audience/our interlocutor. Many studies have provided evidence that speakers accommodate their speech in correspondence to the knowledge of their interlocutor to achieve successful communication (Accommodation Theory; Giles & Powesland, 1975; Giles & Smith, 1979). However, in event conceptualization studies, including motion event conceptualization studies, the effect of the interlocutor is often overlooked and is certainly understudied (Papafragou & Grigoroglou, 2019; Ünal, Ji, & Papafragou, 2021).

Event conceptualization rarely happens in isolation. An audience or an interlocutor is usually involved. Regarding event endpoint construal, it is probable that an event endpoint will be defined differently depending on the social distance between the speaker and the interlocutor. Just take a simple daily event as an example: cleaning dishes. Depending on who is listening: their mom or a strict manager of a three-star Michelin restaurant, speakers might even define the endpoint of cleaning dishes differently. In the latter case, the standard of speaking of completing cleaning dishes should be much higher. Therefore, in the current study, we are interested in whether the effect of the interlocutor's social status will also affect motion event endpoint conceptualization.

Rarely any studies have studied the effects of both the actor's goal and the interlocutor's status on motion event endpoint conceptualization, except for a recent study conducted by Liao, Dijkstra, and Zwaan (2021). In their study, two Dutch directional prepositions (i.e., *naar* 'to' and *richting* 'towards/direction') have provided a nice paradigm to study motion event endpoint conceptualization. Specifically, they examined the extent to which the above-mentioned two factors affected the choice between the two Dutch directional prepositions in a motion event description task. Before we discuss their study in more detail, we would like to first explain what directional prepositions are and why they make a nice paradigm for studying motion event endpoint conceptualization.

Directional prepositions, as part of a verb phrase, contribute to the telicity of an event, specifically whether a motion event has an inherent endpoint or not. There are two types of directional prepositions: telic directional prepositions and atelic directional prepositions (Krifka, 1998; Zwarts, 2005). The use of telic directional prepositions, such as *naar* in Dutch and *to* in English, implies that a motion event is telic and has an inherent endpoint (e.g., *he is walking* ***to the bus station***). On the contrary, the use of atelic directional prepositions, such as *richting* in Dutch and *towards* in English, only implies the direction of a motion event but not its endpoint. Therefore, a motion event that is described with an atelic directional preposition is considered lacking an inherent endpoint and is, therefore, atelic (e.g., *he is walking* ***towards the bus station***). Given the definition of directional prepositions and their classification, studying the choice between the two types of direction prepositions in a motion event

description task does provide a nice paradigm for studying motion event endpoint conceptualization.

The current study is an extension of Liao et al. (2021). We aim to investigate the effects of the actor's goal and the interlocutor's status on motion event endpoint conceptualization by examining a different pair of directional prepositions in a different language, that is, two English directional prepositions (i.e., *to* and *towards*). Furthermore, the current study also goes beyond Liao et al. (2021). Before we explain why this is the case, we would like to first provide a brief summary of Liao et al. (2021) and highlight the limitation of their study that we are going to address in the current paper.

A brief summary of Liao et al. (2021) and the limitation

Liao and colleagues (2021) adopted a motion event description task when examining the effects of the actor's goal and the interlocutor's status on motion event endpoint conceptualization. Specifically, they investigated the effects of the two non-linguistic factors on the use of two Dutch directional prepositions (i.e., *naar* and *richting*).

They found that if the actor's goal can be clearly inferred from the referential scenario, Dutch speakers use the telic preposition *naar* more often, compared to if no clear actor's goal is presented in the scenario. For example, if the referential scenario depicts a man carrying a trash bag and a trash bin in the near distance, then his goal can be easily inferred from the scenario, that is, to go to the trash bin to dispose of the trash bag. However, if the referential scenario depicts a man carrying nothing and a trash bin in the near distance, the goal of the person is then not as clear as in the previous scenario. Consequently, Dutch speakers use *naar* more often (e.g., *hij loopt naar de container* 'he walks to the trash bin') when describing the first scenario to indicate the endpoint of the motion event is the trash bin, compared to when describing the second scenario. The opposite pattern is found for the use of the atelic preposition *richting*. That is, Dutch speakers use *richting* more often (e.g., *hij loopt richting de container* 'he walks towards the trash bin') for the second scenario than for the first scenario. This is because the use of *richting* does not indicate the reference object is the endpoint but just the actor's moving direction.

Furthermore, they found that when the interlocutor is a police officer, Dutch speakers are more likely to use *richting* than when the interlocutor is a friend. This is because the social distance is larger between the speaker and the interlocutor if the interlocutor is a police officer, compared to if the interlocutor is a friend. When the social distance is larger, the speech context is more formal (Koppen, Ernestus, & van Mulken, 2019). When the speech context is more formal, people also tend to be more specific and cautious with their statements. Given that *richting* only refers to the moving direction, not the endpoint of a motion event, the use of *richting* in a motion event description is considered a more conservative and more cautious expression compared to the use of *naar* because the speaker

does not commit to a destination. Importantly, the effect of Interlocutor is also found larger than the effect of Intention (odds ratios: 1.72 vs. 3.79) in Liao et al. (2021), which highlights the importance of considering contextual factors in event conceptualization studies.

In Liao et al. (2021), only one verb form is used for all sentence stimuli, that is, the simple present tense (e.g., *Hij loopt naar/richting de container* 'he walks to/towards the trash bin'). A possible effect of grammatical aspect is ruled out in their study, given that there is no grammaticalized progressive marker in Dutch (Flecken, 2011). The unmarked simple present is the major way to express ongoing events in Dutch. This is especially so when ongoing directional motion events are described (e.g., *Hij loopt naar het station* 'he walks to the station'; Liao, Flecken, Dijkstra, & Zwaan, 2020; von Stutterheim, Carroll, & Klein, 2009).

However, the possible effect of grammatical aspect cannot be ruled out if an aspectual language is studied, such as English. In English, the simple present is not the only way to express ongoing events. In fact, English has a grammaticalized progressive marker (i.e., *-ing*) that plays the main role in expressing ongoing events, including ongoing directional motion events (e.g., *he is walking to/towards the church*). As shown in many previous language-based event comprehension studies (e.g., Anderson, Matlock, Fausey, & Spivey, 2008; Ferretti, Kutas, & McRae, 2007; Madden & Zwaan, 2003; Matlock, 2011), grammatical aspect provides individuals with different viewpoints on the internal temporal structure of an event. When progressive aspect is used to express an event, more details about the event, especially the details relevant to the ongoing phase of the event, are activated during event comprehension, compared to when perfective aspect is used (e.g., Madden & Zwaan, 2003).

It is possible that the use of the English progressive aspect also creates a different event representation, compared to the use of the English simple present (e.g., *he walks to the trash bin* vs *he is walking to the trash bin*). In the current study, we are specifically interested in whether the use of the simple present and the use of progressive aspect in English would result in people's different sensitivity to the two factors that we are investigating, namely the actor's goal and the interlocutor's status. The more sensitive people are to these two factors, the larger effects they might create on people's event endpoint conceptualization, which can be shown based on their choice between *to* and *towards*. Liao et al. (2021) did not investigate this possibility, which is a limitation of their study. Our current study, therefore, goes beyond their study by exploring the potential effect of grammatical aspect on event comprehension, and consequently on event endpoint conceptualization.

In the next section, we provide a theoretical comparison between the English simple present and the English progressive aspect in their differences in representing ongoing events. A special focus will be put on how they can possibly affect the sensitivity to the two non-linguistic factors that we manipulate in the experiments.

The simple present and progressive aspect in English

The simple present in English is often introduced in a dictionary as representing habits (e.g., *I smoke*), general truth (e.g., *he has long hair*), or even future events (e.g., *our meeting starts at 10:00 am*), and so on (see, for example, https://www.ef.com/wwen/english-resources/english-grammar/simple-present-tense/). In these cases, the English simple present is "timeless" (e.g., Vraciu, 2015). It takes a specific time reference from its linguistic environments, such as from the adverbial phrases (e.g., *at 10:00 am*), or it is used with a state (e.g., *he has long hair*), or it functions as a "stativizer" (Vraciu, 2015, p. 294) on dynamic predicates and creates a habitual reading of the dynamic event (e.g., *I smoke*).

The English simple present can also be used to describe dynamic ongoing events, such as in sports commentaries and narrative commentaries, without creating a habitual interpretation of these events. However, progressive aspect is still the major way of expressing ongoing events in English. Therefore, the question is what the differences are between the simple English present and the English progressive aspect when they are both used to describe ongoing events.

Many linguists have already provided some answers. One main difference is that they might be different in representing event details. There is already much empirical evidence that the use of the English progressive aspect directs the comprehender's attention to more event details, such as the location of the event (Ferretti et al., 2007), the intention of the actor (Sherrill, Eerland, Zwaan, & Magliano, 2015), and the manner of action (Madden, Dominey, & Ventre-Dominey, 2017), compared to the use of perfective aspect.

The English simple present, however, often leads to a perfective interpretation of an event (see Cowper, 1998; Vraciu, 2015). When an event is viewed with a perfective viewpoint, this event is considered complete and "an unanalyzable whole" (Vraciu, 2015, p. 295). This perfective viewpoint provided by the English simple present is thus in contrast with the internal viewpoint provided by the English progressive aspect and is similar to the use of the English perfective aspect (i.e., *-ed*) (Williams, 2002).

When the English simple present is used, for example, in sports commentaries (e.g., *Inzaghi passes the ball to Totti*; Williams, 2002, p. 1239), it represents the event (e.g., 'to pass the ball') in a holistic way, even though this event can still be happening at the moment of speech. By doing so, the speaker aims to push the whole story forward. Hence, sports commentaries in which the English simple present is used always consist of a series of events, for example, *Inzaghi passes the ball to Totti, he shoots, and the ball bounces off the goalpost* (Williams, 2002, p. 1241). It is odd to use the English progressive aspect here (i.e., **Inzaghi is passing the ball to Totti. He is shooting. The ball is bouncing off the goalpost*). This is because the English progressive aspect emphasizes the progression of each event and places a spotlight on the ongoing phase of each event, which is part of an event instead of a whole event. This would

affect the fluency of the development of the whole scenario/story.

Therefore, it is relevant to test whether the use of the English progressive aspect indeed creates a more salient representation of the internal temporal structure of an event, such as the actor's manner, compared to when the English simple present is used. In the current study, the actor's goal is indicated by the actor's walking manner (e.g., holding a trash bag or not) in the referential scenario, which is relevant to the ongoing phase of the walking event. It is possible that the use of progressive aspect makes the participants more aware of the information relevant to the ongoing phase of the event (e.g., the actor's goal/ the actor's manner of walking), than when the English simple present is used. When the English simple present is used, the event is viewed as a whole, which might defocus the ongoing phase of the event and consequently, decrease the salience of our manipulation on the actor's manner of walking (e.g., holding a trash bag or not) in the description task.

There is another difference between the English simple present and the English progressive aspect, when they are both used to describe ongoing events. When the English progressive aspect is used to report a scenario (e.g., *They're all coming out of the front door. Two of them are wearing masks. One of them's opening the car door and the others are getting in. They're driving off now towards Friar Lane*; Williams, 2002, p. 1247), the unpredictable nature of the event is highlighted. The speaker is probably not sure about what is going to happen next and, therefore, chooses only to focus on describing the progression of the event. On the contrary, the English simple present is often used to report situations that are "rule-based", "complete", and "self-contained (Williams, 2002, p. 1248), such as sports, ceremonies, or demonstrations. Under these circumstances, the speaker is not focusing on reporting an unpredictable scenario whose progression is important to the hearer. Instead, the speaker is trying to create an "eventful" scenario in which proper developments of complete events are presented.

Therefore, it is also important to test whether the use of the English progressive aspect will lead to a larger effect from the interlocutor's status, compared to the use of the English simple present. It is possible that the use of the English progressive aspect makes the participants more aware of the speech context (i.e., to whom they are describing the scenario) than when the English simple present is used, given that the former highlights the unpredictable nature of an event, whereas the latter does not. For instance, when the speaker is aware that an event is still happening and its development is unpredictable, to whom they are asked to report this event could create a stronger effect on the cautiousness they have with their statements than when the speaker is not aware of the unpredictable nature of the event they are describing.

The current study

The first goal of the current study is to investigate the effects of two non-linguistic factors (i.e., the actor's goal and the interlocutor's status) on motion event endpoint conceptualization,

by targeting the choice between two English directional prepositions (i.e., *to* and *towards*). Specifically, we are interested in whether the observed effects of the actor's goal and the interlocutor's status on event endpoint conceptualization in Liao et al. (2021) are upheld when another language is studied. English is chosen because it is different from Dutch in the aspectual system. This relates to the second goal of our study, which is to investigate the effect of one linguistic cue, grammatical aspect, on the sensitivity to the two non-linguistic factors, and consequently on motion event endpoint conceptualization. This second goal of our study goes beyond Liao et al. (2021), in that in their study, the effect of grammatical aspect was not considered. They adopted only one verb form, that is, the simple present because it is exclusively used to describe ongoing directional events in Dutch (von Stutterheim et al., 2009).

We combine both comprehension and production tasks in the current study. The comprehension task targets the comprehension of the simple present and progressive aspect used in the sentence stimuli (e.g., *he walks/is walking (?) the trash bin*). At the same time, participants would need to perform a description task by choosing between *to* and *towards* to complete the sentence stimuli (e.g., *he walks* (? *to/towards) the trash bin*).

In Experiment 1, we would use the simple present for all our sentence stimuli (e.g., *he walks (?) the trash bin*). This is a reasonable start, given that our first goal is to extend the findings in Liao et al. (2021) (i.e., the main effects of the actor's goal and the interlocutor's status on motion event endpoint conceptualization). By using the English simple present in Experiment 1, we manage to keep the surface form of the verbs of our sentence stimuli (e.g., *he walks (?) the trash bin*) the same as those in Liao et al. (2021). Our hypothesis for Experiment 1 is that we would find the main effects of both the actor's goal (Intention) and the interlocutor's status (Interlocutor) on the choice between *to* and *towards* in the event description task. At the same time, we are aware of the possibility that the use of the English simple present might weaken the effects of the two factors.

In Experiment 2, we would replace the English simple present with the English progressive aspect for all the sentence stimuli (e.g., *he is walking (?) the trash bin*). Even though we have discussed the possible differences between the English simple present and the English progressive aspect in affecting the sensitivity to the two factors (i.e., Intention and the Interlocutor) that we manipulate in the description task, these lines of thoughts are still speculative. To our knowledge, in addition to the above-discussed linguistic analyses, no experimental studies have yet been published on comparing the English simple present to the English progressive aspect regarding their role in representing event details, let alone from the perspectives of studying the actor's goal and the interlocutor's status. Hence, we decided to form two hypotheses for Experiment 2:

Hypothesis 1: The English progressive aspect does not differ from the English simple present in affecting people's sensitivity to the two factors that we manipulate in the description task (i.e., Intention and Interlocutor). The effects of the two factors (on the use of *to* and *towards*) that we obtain in Experiment 2 would be the same as those we found in Experiment 1.

Hypothesis 2: The English progressive aspect does differ from the English simple present in affecting people's sensitivity to the two factors (i.e., Intention and Interlocutor). The English progressive aspect would make people more sensitive to the two factors in the description task, compared to the English simple present. We would find the main effects of both Interlocutor and Intention (on the use of *to* and *towards*), and the effects of both factors would be larger than those found in Experiment 1.

Frick's COAST method and sequential testing

As in Liao et al. (2021), we adopted Frick's COAST method for both experiments conducted in the current study. This method preserves an overall alpha level of .05 while allowing for sequential testing (Frick, 1998). There are many advantages for researchers to choose sequential analyses over a conventional fixed-sample testing method (see Lakens, 2014). One major advantage is that sequential analyses can greatly help researchers to run sufficiently powered studies without running an inefficiently large number of participants. Determining a fixed sample for a high-powered study is not easy and often faces much uncertainty. If the determined sample size is too small, the study has the risk of being underpowered and the obtained effect size is often inaccurate. On the other hand, if the sample size is too big than actually needed, it is a waste of time, resources, and energy. Sequential analyses can be used for testing larger samples and at the same time allow for earlier termination of data collection. This increases the statistical power of the study and also prevents researchers from wasting participants. Moreover, given the adjusted lower alpha level, if data collection is stopped earlier with a relatively small sample size, the estimated effect size is still more reliable than a traditional small-scale study (Lakens, 2014, p. 703).

Experiment 1

Experiment 1 was designed to extend the findings reported in Liao et al. (2021). Based on their findings, we hypothesized the main effects of both the actor's goal and the interlocutor's status in this experiment. Except that we replaced the scenarios that contained a man and a bike repair shop with a more distinct version (see detailed clarification in the *Material*s section and in the preregistration https://osf.io/7c5zh/?view_only=54cdbbb89cfb4f58a952edf8bd7331ab), the design, the data collection plan, and the data analysis plan were all kept the same as those of the second experiment in Liao et al. (2021). As in Liao et al. (2021), we also used the simple present in Experiment 1 for all the sentence stimuli (e.g., *he walks (?) the trash bin*). The whole experiment was in English, including the instructions and the sentence stimuli.

Method

Participants

Frick's COAST (Frick, 1998) method was adopted for data collection and for conducting sequential analyses. Specifically, we recruited participants in batches of 160 participants (i.e., the minimum number of participants to be tested; 80 per Intention and 80 per Interlocutor). We predicted the main effects of Intention (the actor's goal) and Interlocutor (the interlocutor's status). If $p < .01$ or $p > .36$ for both the main effects we predicted, data collection would be terminated. If p was within these boundaries for any one of the two main effects predicted, we would test another 160 participants. Data collection would be terminated if the number of participants reached 480, regardless of the p-values at that time. We recruited 480 participants (315 males, 164 females, 1 others; mean age 34.77 years old, range: 20-74 years old) eventually. All participants reported their native language as English. Participants were recruited via Mechanical Turk. The experiment took around 1 to 2 minutes per participant and each participant was paid $0.5 as a reward.

Materials

The experiment was programmed using the Qualtrics Survey Software. As in Liao et al. (2021), we used two different scenarios (one was with a person and a bike repair shop and the other one was with a person and a trash bin). For each scenario, there were two versions. In the first scenario, the person either carried a broken bike or not; in the second scenario, there was a person who either carried a trash bag or not. Therefore, there were in total 4 scenarios as the experimental stimuli (see Appendix). The trash bin scenarios were exactly the same as those used in Liao et al. (2021). We replaced the bike scenarios in Liao et al. (2021) with their more distinct versions (see an example of the ones used in their study and its replacement that was used in the current study in Figure 1 and Figure 2, respectively). We did so because we wanted to have a homogeneous layout for both the bike repair shop and the trash bin scenarios, for example, no background in the picture of the person. Moreover, we thought that if there was no background in the picture of the person, it would be easier for participants to combine the two pictures, for instance, the person and the bike repair shop, into one holistic scenario. The bike repair shop was also replaced with one with signs on it (e.g., repairs, rental) to ensure that participants knew it was a bike repair shop. Below each picture, a sentence was shown (e.g., *he walks (?) the bike repair shop*) and a choice between *to* and *towards* was shown below the sentence (see Appendix).

Figure 1.
An example of the bike scenarios used in Liao et al. (2021).

Figure 2.
An example of the bike scenarios used in the current study.

Design and procedure

Sixteen cells were designed (4 Scenario [Intention]*2 Interlocutor*2 OptionOrder) and each participant was randomly assigned to only one cell. We manipulated the instructions as either "*You are describing a scenario to a police officer as a witness*" or "*You are describing a scenario to a friend*" (Interlocutor). Progressive aspect was used in the instructions because we wanted the participants to imagine their interlocutor as vividly as possible while performing the description task. The choice option order was counterbalanced (*to* at the right side of *towards* or at the left side of *towards*). An informed consent form was attached at the beginning of the survey. Participants could choose whether to continue the survey or to leave freely at any time. Then they answered questions about their demographic information (i.e., gender and age) and their mother tongue. They subsequently read the instructions of the experiment on the screen and chose between *to* and *towards* to complete the sentence stimuli, based on the scenario they were viewing.

Results and discussion

Confirmatory analyses

A binomial logistic regression model that included the main effect of Intention and the main effect of Interlocutor was conducted in R (R Core Team, 2016) using the *glm* function implemented in the package *lme4* (Bates, Mächler, Bolker, & Walker, 2015) (the formula used in R was *preposition ~ intention + interlocutor*). Both factors were dummy coded. No interaction effect was included in the model. The dependent variable was the directional preposition choice, which was a binary outcome. The choice of the directional preposition *to* was coded as "0" and the choice of the directional preposition *towards* was coded as "1" in R (in alphabetic order). In the first batch of collected data (N = 160), we found a significant main

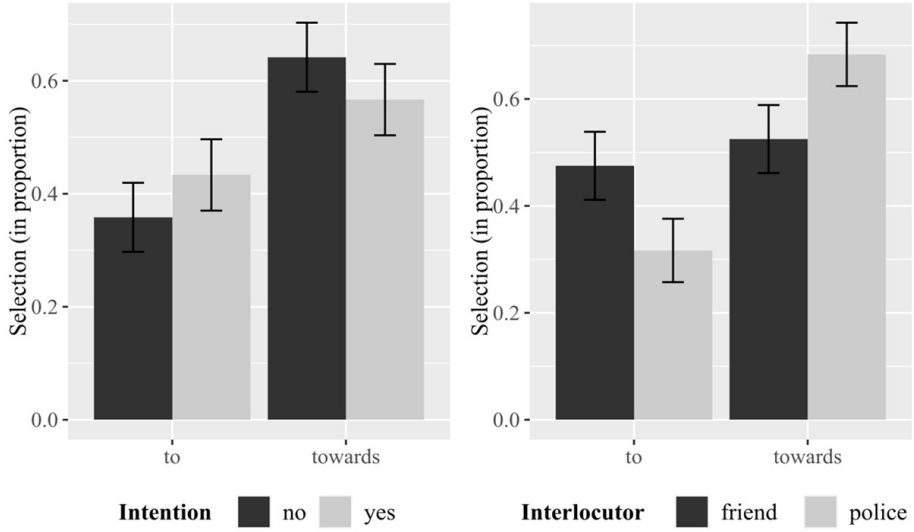
effect of Intention (β = -0.995, *SE* = 0.34, *z* = -2.919, *p* = .004, odds ratio: 0.370, 95% CI: 0.19-0.72). We did not find a significant main effect of Interlocutor but the *p*-value was within the boundary from .01 to .36 (β = 0.676, *SE* = 0.34, *z* = 1.989, *p* = .047, odds ratio: 1.965, 95% CI: 1.02-3.86).

Given that the *p*-value found for the effect of Interlocutor was within the boundary from .01 to .36, we continued data collection. After the second round of data collection (*N* = 320), we performed the same analysis. We found that both the main effect of Intention and the main effect of Interlocutor were not significant (*Intention* β = -0.482, *SE* = 0.23, *z* = -2.071, *p* = .038, odds ratio: 0.62, 95% CI: 0.39-0.97; *Interlocutor* β = 0.534, *SE* = 0.23, *z* = 2.296, *p* = .022, odds ratio: 1.71, 95% CI: 1.08-2.70). The *p*-values for both factors were within the boundary from .01 to .36. Therefore, we continued data collection until we reached 480 participants. Based on our preregistration, this was our last round of data collection.

We did the same analysis again for the last data batch (*N* = 480). We did not find a significant effect of Intention (β = -0.323, *SE* = 0.19, *z* = -1.700, *p* =.089, odds ratio: 0.72, 95% CI: 0.50-1.05). The *p*-value of the main effect of Interlocutor, however, was smaller than .01 (β = 0.673, *SE* = 0.19, *z* = 3.538, *p* < .001, odds ratio: 1.96, 95% CI: 1.35-2.85): based on the standards of the sequential analysis we pre-registered, it was considered a significant effect. Thus, we found support for the hypothesis that addressee status affects preposition choice but not for the hypothesis that protagonist intention does the same. Figure 3 shows the mean proportions of the selection of *to* and the selection of *towards* in each Intention condition and in each Interlocutor condition.

Figure 3.

Mean proportions of the selection of *to* and the selection of *towards* in each Intention condition (left) and in each Interlocutor condition (right).



Exploratory analyses

The interaction effect between Intention and Interlocutor was not considered when forming our hypotheses in the confirmatory analyses. It was possible that different interlocutors might affect the sensitivity to the actor's goal. However, we were not sure in what direction Interlocutor might affect the sensitivity to the two levels of Intention (i.e., a clear goal vs an unclear goal). One possibility was that talking to a police officer might lower participants' willingness to commit themselves to an event endpoint no matter whether the actor's goal was clear or unclear, compared to when talking to a friend. Another possibility was that talking to a police officer would only lower the certainty about an event endpoint when the actor's goal was not clear but would not make much difference when the actor's goal was clear, compared to when talking to a friend. In such circumstances, we would like to explore these possibilities in the exploratory analyses.

Apart from the interaction between Intention and Interlocutor, we also decided to add the scenario type as a fixed effect in the exploratory analyses (including its main effect and its interaction effect with Intention and with Interlocutor). Scenario type was supposed to be taken as a random effect in the confirmatory analyses. However, the inclusion of it as a random effect in the confirmatory analyses brought overfitted warnings. Moreover, the scenario type had only two levels (i.e., trash bin vs bike repair shop). When a random factor has too few levels (normally below 5 levels; Bolker, 2015), the estimated variance can be very imprecise and unstable, especially when a singular fit warning occurs (Oberpriller, de Souza Leite, & Pichler, 2021). Under such circumstances, researchers recommend fitting such a random effect in the statistical model as a fixed effect (Bolker, 2015; Crawley, 2002; Gelman, 2005; Gelman & Hill, 2007).

Moreover, we were also interested in whether there was a main effect of the order of the two prepositions as an option. We counterbalanced this factor in our experimental design to avoid a possible primacy or recency bias in the participants' answers. That is, participants might prefer to choose the first option they encountered (a primacy bias) or on the contrary, participants might be more likely to choose the latest option they saw (a recency bias). In the explanatory analyses, we were interested in whether such biases indeed existed in our experiments or not.

Therefore, we built a binomial logistic regression model that included the main effects of Intention, Interlocutor, Option Order, Scenario, the interaction between Intention and Interlocutor, the interaction between Scenario and Intention, and the interaction between Scenario and Interlocutor (the formula used in R was *preposition ~ intention*interlocutor + intention*scenario + interlocutor*scenario + option order*). All the factors were sum coded except for the factor Option Order (dummy coded), for we were only interested in the main effect of this factor. The dependent variable was the choice between the two directional prepositions: *to* and *towards*. The choice of *to* was coded as "0", and the choice of *towards* was

coded as "1".

The main effect of Intention was still not significant (β = 0.164, *SE* = 0.10, *z* = 1.666, *p* =.096, odds ratio: 1.18, 95% CI: 0.97-1.43). The significant main effect of Interlocutor (β = -0.374, *SE* = 0.10, *z* = -3.759, *p* <.001, odds ratio: 0.69, 95% CI: 0.57-0.84) remained in this model. No interaction effect was found between Intention and Interlocutor (β = 0.106, *SE* = 0.10, *z* = 1.082, *p* = .28, odds ratio: 1.11, 95% CI: 0.92-1.35). We found a significant main effect of Scenario (β = -0.377, *SE* = 0.10, *z* = -3.786, *p* <.001, odds ratio: 0.69, 95% CI: 0.56-0.83). A significant interaction effect between Scenario and Interlocutor (β = 0.273, *SE* = 0.10, *z* = 2.748, *p* =.006, odds ratio: 1.31, 95% CI: 1.08-1.60) was also detected. Specifically, *towards* was used more often when the addressee was a police officer than when it was a friend, only for the scenarios of trash bins. No interaction effect was found between Scenario and Intention (β = -0.064, *SE* = 0.10, *z* = -0.647, *p* =.52, odds ratio: 0.94, 95% CI: 0.77-1.14). There was no significant main effect of Option Order (β = 0.151, *SE* = 0.19, *z* = 0.776, *p* =.438, odds ratio: 1.16, 95% CI: 0.80-1.70).

In sum, we predicted the main effects of Intention and of Interlocutor on the use of *to* and *towards* in Experiment 1. In the confirmatory analyses, we did find a main effect of Interlocutor (*p* < 0.001, odds ratio: 1.96), but not a main effect of Intention (*p* =.089, odds ratio: 0.72) (this was also confirmed in the explanatory analyses). Moreover, the effect sizes of both factors (indicated by their odds ratios: Intention vs Interlocutor: 1.39 vs. 1.96) were lower than those found in Liao et al. (2021) (Intention vs Interlocutor: 1.72 vs. 3.79). According to our previous discussions (see the section *The simple present and progressive aspect in English*), a possible reason is that the use of the English simple present might have weakened the salience of both factors, Intention and Interlocutor, in Experiment 1. The use of the English progressive aspect might help to increase the salience of both factors and hence also strengthen their effects on preposition choice. Therefore, in Experiment 2, we replaced the English simple present with the English progressive aspect for all the sentence stimuli to examine this possibility.

Experiment 2

Experiment 2 was identical to Experiment 1, except that we replaced the English simple present used in the sentence stimuli (e.g., *he* ***walks*** *(?) the trash bin*) with the English progressive aspect (e.g., *he* ***is walking*** *(?) the trash bin*). We formed alternative hypotheses for Experiment 2 (see *The current study* section).

Method

Participants

As in Experiment 1, we adopted Frick's COAST method for data collection and for conducting sequential analyses. We also recruited participants in batches of 160 participants. If *p* < .01 or *p* > .36 for both factors (Intention and Interlocutor), data collection would be terminated.

If *p* was within these boundaries for any one of the two factors, we would test another 160 participants. Data collection would be terminated if the number of participants reached 480, regardless of the *p*-values at that time. In the end, we ended data collection at our first data batch (N = 160, 83 males, 73 females, 4 others; mean age 33.675 years old, range: 17-74 years old), given that the *p*-values for both factors were below .01 when N = 160. All participants were native English speakers and were recruited via the Prolific platform. The experiment took around 2 minutes per participant and each participant received £0.35 as a reward.

Materials

The same materials as in Experiment 1 were used, except that the simple present used in Experiment 1 was replaced with progressive aspect for all the sentence stimuli.

Design and procedure

The design and procedure were identical to Experiment 1, except that participants were not asked to fill in information about their mother tongue. This is because we already excluded people whose first language was not English through the Prolific platform, and it was not allowed to collect information about people's linguistic backgrounds on this platform.

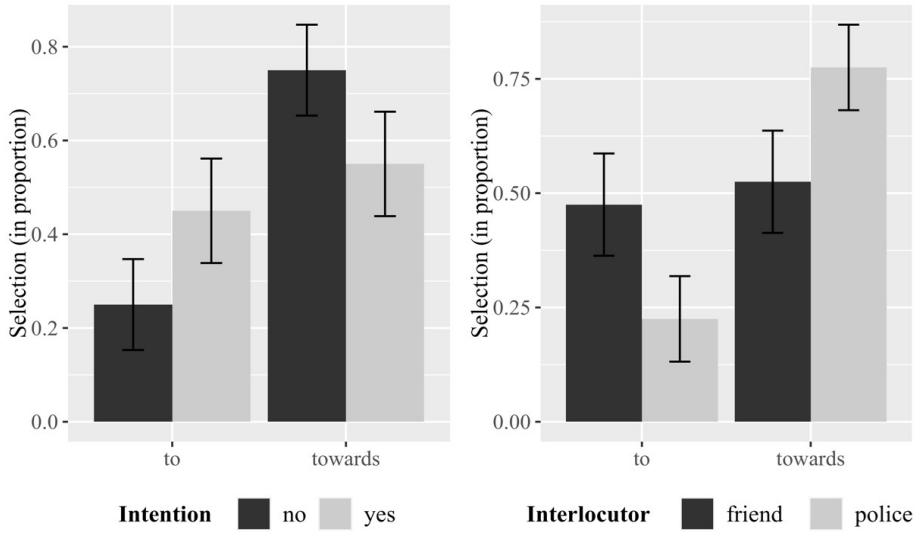
Results and discussion

Confirmatory analyses

The same binomial logistic regression model that included the main effect of Intention and the main effect of Interlocutor was conducted in R using the *glm* function implemented in the package *lme4* (Bates, et al., 2015) (the formula used in R was *preposition ~ intention + interlocutor*). Both factors were dummy coded. No interaction effect was included in the model. The dependent variable was the directional preposition choice, which was a binary outcome. The choice of *to* was coded as "0" and the choice of *towards* was coded as "1" in R (in alphabetic order). We found significant main effects of both Intention (N = 160; β = -0.969, SE = 0.36, z = -2.711, p = .007, odds ratio: 0.379, 95% CI: 0.19-0.76) and Interlocutor (β = 1.195, SE = 0.36, z = 3.322, p < .001, odds ratio: 3.31, 95% CI: 1.66-6.82). Figure 4 shows the mean proportions of the selection of *to* and the selection of *towards* in each Intention condition and in each Interlocutor condition.

Figure 4.

Mean proportions of the selection of *to* and the selection of *towards* in each Intention condition (left) and in each Interlocutor condition (right).



Exploratory analyses

As in Experiment 1, we built the same binomial logistic regression model that included the main effects of Intention, Interlocutor, Option Order, Scenario, the interaction between Intention and Interlocutor, the interaction between Scenario and Intention, and the interaction between Scenario and Interlocutor (the formula used in R was *preposition ~ intention*interlocutor + intention*scenario + interlocutor *scenario + option order*). All the factors were also sum coded except for the factor Option Order (dummy coded), for we were only interested in the main effect of this factor. The reasons why we included these factors in the exploratory analyses were already provided in the *Exploratory analyses* section in Experiment 1.

Both the main effects of Intention and Interlocutor remained significant in this model (*Intention* β = 0.623, *SE* = 0.21, *z* = 2.969, *p* = .003, odds ratio: 1.86, 95% CI: 1.26-2.92; *Interlocutor* β = -0.732, *SE* = 0.21, *z* = -3.471, *p* < .001, odds ratio: 0.48, 95% CI: 0.31-0.71). The interaction between Intention and Interlocutor was found to be marginally significant if we adopted the conventional rule of determining a significant effect when $p < 0.05$ (β = -0.419, *SE* = 0.21, *z* = -1.998, *p* = .046, odds ratio: 0.66, 95% CI: 0.42-0.97). Specifically, *to* was used less frequently when the actor's goal could not be clearly inferred from the scenario, compared to when it could, particularly so when the interlocutor was a police officer, compared to when the interlocutor was a friend. The main effect of Scenario was not significant (β = 0.089, *SE* = 0.20, *z* = 0.455, *p* = .649, odds ratio: 1.09, 95% CI: 0.75-1.62), and neither was its interaction

with Intention (β = 0.079, *SE* = 0.19, *z* = 0.411, *p* = .681, odds ratio: 1.08, 95% CI: 0.75-1.59) and with Interlocutor (β = -0.292, *SE* = 0.20, *z* = -1.501, *p* = .133, odds ratio: 0.75, 95% CI: 0.50-1.09). The main effect of Option Order was also found insignificant (β = 0.255, *SE* = 0.36, *z* = 0.713, *p* = .476, odds ratio: 1.29, 95% CI: 0.64-2.62).

These findings support the second hypothesis formed in Experiment 2. We found the main effects of both Intention and Interlocutor. Their effect sizes (indicated by the odds ratios: Intention vs Interlocutor: 2.64 vs. 3.31) were also larger than those found in Experiment 1 (Intention vs Interlocutor: 1.39 vs. 1.96) (from the confirmatory analyses). Therefore, we conclude that the use of the English progressive aspect indeed leads people to pay more attention to event details and to speech context, compared to when the English simple present is used.

General discussion

In the current study, two experiments were conducted to examine the effects of two non-linguistic factors, namely the actor's goal and the interlocutor's social status, on motion event endpoint conceptualization. Moreover, the effect of one linguistic cue, grammatical aspect, on the sensitivity to the two non-linguistic factors was also investigated. By examining the choice between two Dutch directional prepositions (i.e., *naar* and *richting*), a study conducted by Liao et al. (2021) demonstrates that both the actor's goal and the interlocutor's status affect motion event endpoint conceptualization. Our first goal here was to extend these findings by investigating the English equivalents of the two Dutch prepositions, namely *to* and *towards*. Our second goal was to go beyond their study by studying whether different grammatical aspect (i.e., the English simple present and the English progressive aspect) would affect the salience of the two non-linguistic factors (i.e., Intention and Interlocutor) in the description task and consequently also affect people's motion event endpoint conceptualization.

In Experiment 1, we used the English simple present for all the sentence stimuli (e.g., *he walks to/towards the bike repair shop*). We predicted the same significant main effects of both Intention and Interlocutor as in Liao et al. (2021). We did find a significant effect of Interlocutor ($p < 0.001$, odds ratio: 1.96). Its effect size, however, was almost twice as small as that found in Liao et al. (2021) (odds ratio: 3.79). Moreover, we did not find a significant effect of Intention in Experiment 1 ($p = .089$, odds ratio: 0.72).

We assume that the insignificant effect of Intention and the smaller effect of Interlocutor (compared to the one in Liao et al. (2021)) should be attributed to the use of the English simple present in Experiment 1. As previously mentioned, although the English simple present can be used to express ongoing events, it does not emphasize the progressive phase of the event. Instead, it tends to present an event as a whole and defocuses the internal temporal structure of the event. In our experiments, the actor's goal was indicated by the actor's walking manner (e.g., whether the actor was carrying a trash bag or not), which was

relevant to the progressive stage of the actor's action. For instance, if the actor was carrying a trash bag, it was then clearer that the actor was to go to the trash bin to dispose of the trash bag, than when the actor was carrying nothing. However, when an event was viewed as a unified whole (an external viewpoint provided by the use of the English simple present: *walks*), the ongoing phase of an event, including the actor's walking manner, was then defocused. Consequently, participants in Experiment 1 became less sensitive to the factor Intention when performing the description task. In Liao et al. (2021), the effect of Intention was significant but not particularly strong (odds ratio: 1.72). Hence, it is reasonable that this effect disappeared in Experiment 1 in the current study.

The use of the English simple present also led to a decreased sensitivity to the speech context. As we have discussed previously, the use of the English simple present imbues an event with a more predictable and a more pre-planned sense, than when the English progressive aspect is used. This is because the English simple present does not emphasize the progression of an event and what is exactly happening at the moment of speech, but instead focuses on a more complete and holistic presentation of the event. When an event is more predictable, the speech context could then be comprehended as less important, than when an event is presented as less predictable. This, however, does not necessarily cause the disappearance of the effect of speech context if its effect is strong enough. The effect size of Interlocutor was indeed fairly large in Liao et al. (2021). Therefore, we found that the effect of Interlocutor remained in Experiment 1 but was smaller than the effect found in Liao et al. (2021).

In Experiment 2, we replaced the English simple present with the English progressive aspect for all the sentence stimuli, and we found the main effects of both Intention and Interlocutor on the use of *to* and *towards*. Moreover, the effect sizes of both factors found in Experiment 2 were larger than those found in Experiment 1. These findings support our second hypothesis formed in Experiment 2. The English progressive aspect indeed brings about a more careful reading of both the actor's goal and the interlocutor's status during event comprehension, compared to the English simple present.

Therefore, an important conclusion we can draw based on the results of these experiments is that the effects of the actor's goal and the interlocutor's status on event endpoint conceptualization that were found in Liao et al. (2021) are indeed stable, given that those effects were also detected among English native speakers when *to* and *towards* were investigated. Our study confirms the idea that during the process of event conceptualization, knowing the actor's goal is essential to the identification of an event endpoint. This is not surprising, given that humans are intentional agents whose behaviors are normally goal-directed (see also Zacks & Swallow, 2007). Moreover, a goal is set to be achieved. Where there is a goal, there is an expected endpoint.

Our finding of the role of the actor's goal in identifying an event endpoint during event conceptualization is similar to what Zacks has proposed about the role of the actor's goal

in event segmentation (see Zacks, 2004). As was put by Zacks (2004), the actor's goal is one of the defining features of the "knowledge structures for events" (p. 980). It works as a cue for detecting an event boundary in a top-down manner during the process of ongoing-activity segmentation. Similarly, we assume that how the actor's goal affects event endpoint conceptualization is also a sort of top-down processing.

Moreover, the actor's goal is an internal feature. Most of the time, it needs to be explicitly expressed or inferred from the movements of the actor (Radvansky & Zacks, 2014; Zacks, 2004). Our study demonstrates that people indeed make use of the information provided in the referential scenarios, including the actor's walking manner, to get access to the intention of the actor for predicting an event endpoint during event conceptualization. This is bottom-up processing since the actor's goal is inferred from the sensory information presented in the referential scenarios. This bottom-up processing is incorporated with the later-on top-down processing (i.e., inferring an event endpoint from the actor's goal) during the whole conceptualization phase.

Our study also confirms that speech context, such as the social distance between the speaker and the interlocutor, plays an important role in event endpoint conceptualization. Previous studies have demonstrated that people adapt their speech behavior depending on the formality of the speech context. They found that people use nouns, prepositions, and adjectives more frequently in a more formal speech context, compared to in a less formal speech context (e.g., Heylighen & Dewaele, 2002; Koppen et al., 2019). Our study focuses on the social status of the interlocutor, which is one important parameter that determines the formality of the speech context. We found that even the choice between specific prepositions differs depending on the social distance between the speaker and the interlocutor. Specifically, people use a more specific preposition (e.g., *towards*) to define an event more often if they talk to a formal interlocutor, such as a police officer, compared to if they talk to an informal interlocutor, such as a friend. Importantly, the effect of Interlocutor was found larger than the effect of Intention on motion event endpoint conceptualization. Hence, our study also highlights the urgency of considering contextual factors in event conceptualization studies, which are currently understudied in this field.

The confirmed crucial role of the interlocutor's social status in event endpoint conceptualization also sheds light on eyewitness testimony studies. For example, police officers should be aware that their identity might create an unconscious effect on how their witnesses describe a crime. Witnesses might become more careful and more conservative with their language use than they normally do, which is not always helpful for solving cases, especially if they do not dare to commit themselves to any certain statements.

Furthermore, our study demonstrates that grammatical aspect influences people's mental representations of event details and even people's sensitivity to speech context. This is an important message for both event comprehension and event production studies, given

that even a small difference in the use of grammatical aspect (i.e., the simple present vs progressive aspect) can lead to different representations of the depicted event. For event production studies, it is, hence, important to take into account any possible effects that might come from the verb forms used in the instructions or in any other experimental materials.

Limitations of the current study and future research

For both experiments in the current study, we used only two types of scenarios (i.e., the trash bin scenarios and the bike shop scenarios). This is mainly because for our study, it was very important to ensure that participants did not know the goal of the experiments. Their preposition choice should be based on their linguistic intuition without any awareness of our manipulations. Otherwise, their responses would become useless. To avoid participants from knowing the goal of our experiments, we adopted a between-subjects design. In this design, we assigned one participant only one cell out of the total 16 cells (2 Scenario*2 Intention*2 Interlocutor*2 OptionOrder). Each cell was assigned to 10 participants for the first data batch, based on our preregistration plan. If we had added one more scenario, we would have to create 8 more cells (3 Scenario*2 Intention*2 Interlocutor*2 OptionOrder: 24 cells). Consequently, we would need to recruit at least 80 more participants, which we consider a waste of resources and was not worthwhile to do so.

However, we are aware of the generalizability issue due to the limited number of scenarios used in the current study. We are positive that our findings can be generalized to other scenarios. One main reason is that the two chosen types of scenarios represent two common motion events in daily life. They are not special regarding the nature of the motion events they represent but they still represent two unrelated motion events. Given that the effects of the protagonist intention and the interlocutor's status have already been generalized across these two common but unrelated scenarios in two different languages, we are confident that these effects can be generalized to other types of scenarios. Moreover, the concepts of the two studied factors (i.e., Intention and Interlocutor) and the use of grammatical aspect are not limited to the characteristics of the two chosen scenarios. It is, therefore, feasible to consider other scenarios to manipulate these factors.

What should be noted is that possible differences between scenarios might affect the strength of the effects of these factors on endpoint conceptualization. As we can infer from the explanatory analyses of our Experiment 1, when the sensitivity to the interlocutor's status was weakened by the use of the English simple present, its effect was only detected in the trash bin scenarios, not in the bike shop scenarios. This indicates a possible difference between the trash bin scenarios and the bike shop scenarios in relation to the effect of Interlocutor, even though this difference eventually disappeared in Experiment 2 when the English progressive aspect was used. Hence, we do not formulate strong claims here regarding the specific scenarios that our findings can generalize to but leave that for future research. Future research should explore more types of motion events and that may even go

beyond the scope of motion events. If possible, a more naturalistic depiction of events, for instance, using videos of events, is also recommended.

Another limitation of the current study is that we did not take a possible interaction effect between the actor's goal and the interlocutor's status into account in the preregistered analyses. As shown in the explanatory analyses of Experiment 2, however, there was a marginally significant interaction between the two factors. Specifically, when talking to a police officer, participants were especially less willing to commit themselves to an event endpoint when the actor's goal was unclear (compared to when the actor's goal was unclear), compared to when talking to a friend. However, the effect size of this interaction was relatively small (odds ratio: 1.52). Moreover, detecting a reliable interaction effect often requires a larger sample size than detecting a main effect. Given that we did not plan our sample size for finding an interaction, we are uncertain whether the detected marginally significant interaction between the two factors is a true effect or is just a positive false. Theoretically speaking, it is indeed possible that people show different sensitivity to the actor's goal depending on to whom they are talking. Therefore, future research should consider this interaction effect when conducting event endpoint conceptualization studies.

Conclusion

The current study extends the findings reported by Liao et al. (2021). Our findings support the idea that both the actor's goal and the interlocutor's status affect motion event endpoint conceptualization, even when speakers with a different native language were tested. Our study contributes to motion event conceptualization studies by providing evidence that the absence/presence of a clear intention of the actor is an important factor in event endpoint conceptualization. Moreover, our study highlights the importance of considering contextual factors, such as the social status of the interlocutor, in event conceptualization studies.

Importantly, our study provides further evidence that grammatical aspect (i.e., the English simple present and the English progressive aspect) also affects event endpoint conceptualization, via their influence on event details representation and the perception of speech context. Unlike most event representation studies that focus on the difference between the English progressive aspect and the English perfective aspect, the current study provides a novel perspective in event representation studies, that is, including the contrast between the English simple present and the English progressive aspect. Many linguists have theoretically analyzed their difference in representing eventualities. However, to our knowledge, no studies have experimentally tested this difference. The current study provides experimental evidence for their different role in event representation. A take-home message here is that subtle differences in language use, such as the use of different verb forms, can result in a substantial change in meaning.

Ethics statement

This study was approved by the Ethics Review Committee of the Department of Psychology, Education, and Child Studies at the Erasmus University Rotterdam, the Netherlands.

Open practices statement

The data of both Experiment 1 and Experiment 2 are available on Open Science Framework (link: https://osf.io/hy3f9/?view_only=2a05d791ed0d42c5ace6e54179d5343b). Hypotheses, materials, designs, exclusion criteria, data collection plans, and data analysis plans of all experiments were pre-registered on Open Science Framework in advance of data collection and analysis (see Experiment 1: https://osf.io/7c5zh/?view_only=54cdbbb89cfb4f58a952edf8bd7331ab; Experiment 2: https://osf.io/hwupz/?view_only=2e0fea9fbec74f01aa1defb4115cc275).

Appendix

Stimuli used in Experiment 1

He walks (?) the bike repair shop.

○ to ○ towards

He walks (?) the bike repair shop.

○ to ○ towards

He walks (?) the trash bin.

○ to ○ towards

He walks (?) the trash bin.

○ to ○ towards

Chapter 6

General Discussion

In this dissertation, I have looked at language-based event conceptualization from a cross-linguistic perspective by comparing the way mandarin Chinese speakers and Dutch speakers conceptualize and describe the path information of a motion event (**Chapter 3**). Moreover, I have examined language-based event comprehension by focusing on the function of two grammatical morphemes in the construction of event representations. The two grammatical morphemes are ***grammatical aspect*** (**Chapter 2**, **Chapter 5**) and ***directional prepositions*** (**Chapter 4**). Furthermore, I have investigated the effects of two non-linguistic factors (i.e., ***the actor's goal*** and ***the interlocutor's social status***) on the use of two pairs of directional prepositions in two different languages, respectively: the Dutch pair in **Chapter 4** (i.e., *naar* and *richting*) and the English pair in **Chapter 5** (i.e., *to* and *towards*), targeting motion event endpoint conceptualization. In this chapter, I summarize and discuss the major findings from all of the studies presented in this dissertation. I also discuss the limitations of the current dissertation and provide directions for future research.

Grammatical morphemes and event comprehension/representations

For language-based event comprehension, not only lexical morphemes (e.g., nouns, verbs, adjectives, adverbs) play an important role, but grammatical morphemes (e.g., tense, grammatical, aspect, prepositions) are equally important (e.g., Morrow, 1986). When lexical morphemes provide information about the components of an event (e.g., entities, actions), grammatical morphemes add properties to these components, including both their temporal (tense and aspect) and spatial properties (prepositions), to construct a coherent event model.

One such grammatical morpheme is grammatical aspect. **In Chapter 2**, we conducted a review of the studies that investigated the role of grammatical aspect in and beyond event comprehension. Specifically, we have reviewed studies from three perspectives: sentence comprehension, discourse comprehension, and higher level processes such as problem-solving and decision-making. Our major findings are summarized as follows:

During sentence processing, grammatical aspect has shown a robust effect on the mental simulation of the temporal status of an event (i.e., the ongoing versus the completed status of an event). Properties that are related to the ongoing process of an event, such as the instrument information, location information, and the action of the event, are activated when progressive aspect is used in a sentence (e.g., *he is drawing a picture*) (e.g., Ferretti et al., 2007; Madden & Therriault, 2009; Madden & Zwaan, 2003). Event elements that belong to the resultant stage of an event, such as the final stage of an affected object, are mainly present in the mental simulation of an event when perfective aspect is used (e.g., *he drew a picture*) (e.g., Kaplan et al., 2021; Madden & Zwaan, 2003).

During discourse comprehension, grammatical aspect has also shown an reliable effect on successful construction of a situation model. For instance, grammatical aspect affects the availability of an event, including the features of the engaged protagonists, in a later text

(e.g., Carreiras et al., 1997; Magliano & Schleich, 2000). It also contributes to the prediction of future events, such as who will be mentioned in the following text (e.g., Ferretti et al., 2009). Furthermore, grammatical aspect plays a role in event segmentation during discourse comprehension (Feller et al., 2019).

However, studies have shown a mixed picture regarding the effect of grammatical aspect on social cognition which uses event representations (e.g., problem-solving or decision-making which is based on language comprehension). A safe conclusion in this aspect would be that the effect of grammatical aspect on social cognition is limited, small, and unstable (e.g., Eerland et al., 2017).

The effect of grammatical aspect on event representations is further supported by the study on motion event endpoint conceptualization described in **Chapter 5**. We found that when an event is described with the English simple present, this event is viewed as a holistic whole with its ongoing stage defocused. This function of the English simple present is similar to the function of perfective aspect. On the contrary, when an event is described with the English progressive aspect, the progressive stage of the event, and thereby its internal structure, comes into focus, including the properties of the protagonist (e.g., the protagonist's walking manner). This consequently affects how the depicted event is further conceptualized , as was shown in **Chapter 5**. Interestingly, the use of the English progressive aspect also highlights the unpredictable nature of the depicted ongoing event, which draws attention to the speech context during event conceptualization/description.

Directional prepositions are another grammatical morpheme that we have investigated in this dissertation. Specifically, in Experiment 1 that was presented in **Chapter 4**, we examined how directional prepositions contribute to the construction of event representations. We found that when the goal preposition *naar* 'to' was used to describe a directional motion event (e.g., *de man loopt naar de kerk* 'the man walks to the church'), participants located the figure of motion (e.g., *the man*) significantly closer to the referential object (e.g., *the church*), compared to when the directional preposition *richting* 'towards' was used (e.g., *de man loopt richting de kerk* 'the man walks towards the church'). These findings suggest that people are sensitive to the semantic differences between the two types of directional prepositions (i.e., telic and atelic directional prepositions) used during language-based event comprehension. The use of a goal preposition (a telic directional preposition) makes people more certain about the referential object being the endpoint of the motion event, compared to the use of an atelic directional preposition that only indicates the direction of a motion event but not its endpoint.

Taken together, the findings regarding the function of grammatical aspect and directional prepositions in language-based event comprehension further suggest that both the temporal and spatial properties of an event are important information for event comprehension in general.

Cross-linguistic comparison of motion event conceptualization

In **Chapter 3**, we examined the function of grammatical aspect in event conceptualization from a cross-linguistic perspective. Researchers have discovered a typological difference among languages in the use of progressive aspect in describing motion events. They found that speakers of a language that lacks progressive aspect tend to conceptualize an event as a whole and mention an event endpoint more often, compared to speakers of a language that typically uses progressive aspect to describe an ongoing event (e.g., von Stutterheim et al., 2012). Following the lead of these researchers, we compared mandarin Chinese speakers and Dutch speakers in their frequency of encoding an event endpoint during motion event description.

However, we did not find a significant difference between the speakers of the two languages in their frequency of mentioning motion event endpoints in **Chapter 3**. Our explanation is that even though mandarin Chinese language indeed has a progressive marker (i.e., *-zai*), it is not used so often to describe a directional ongoing event (this has been proved by our further investigation on the use of aspectual markers in motion event descriptions in mandarin Chinese in the same chapter). Therefore, mandarin Chinese speakers and Dutch speakers, in fact, do not differ much in their real-life usage of progressive markers when describing ongoing directional motion events. The study described in **Chapter 3** underscores the importance of a close examination of how events are encoded by speakers of different languages. A simple typological classification sometimes is not sufficient to predict the actual differences that exist among speakers of different languages.

In **Chapter 3**, besides the effect of grammatical aspect on motion event conceptualization, we also investigated the effect of lexicalization patterns on motion event conceptualization. As we have discussed in **Chapter 3**, mandarin Chinese differs from Dutch in the way they encode motion events. Specifically, mandarin Chinese exhibits mixed features of both satellite-framed and verb-framed languages, whereas Dutch is a typical satellite-framed language. In other words, the path of motion (i.e., endpoint, trajectory of motion, or location of motion) is typically encoded outside the verb root in Dutch, such as in verb affixes or in particles (e.g., ***in****stappen* 'get **in**'), whereas the path of motion is normally encoded in serial verb constructions (i.e., two or more verbs combined as a verbal construction; e.g., *zou-jin* 'walk-enter'), single path verbs (e.g., *jin* 'enter'), or outside the verb root (e.g., *zou-xiang* 'walk-towards) in Chinese.

We found that for endpoint-oriented motion events, mandarin Chinese speakers encoded the location-only information (e.g., *on the road*) more often than they did the trajectory information (e.g., *over the road*), whereas Dutch speakers displayed the opposite behavior (the trajectory information was encoded more often than the location-only information). We assumed that the more frequent mention of the location-only information over the trajectory information was due to the verb-framed features that Chinese possesses. As we

have observed from the language production data, the location-only information often followed manner verbs in Chinese (e.g., *drive on the road*). Research has provide evidence that for verb-framed languages, manner verbs are often used to focus the attention on the figure of motion, whereas paths verbs are usually employed to express spatial changes (e.g., Flecken et al., 2015). Hence, when manner verbs are used, the location-only information is more preferred to just locate the figure of motion instead of the trajectory information, which concerns the change of space.

In sum, speakers of mandarin Chinese and speakers of Dutch indeed conceptualize the path information of a motion event differently, but this is mainly due to the differences in the lexicalization patterns between the two languages. Moreover, our study shows that typological differences do not necessarily predict differences in event conceptualization patterns (this was indicated by the finding that mandarin Chinese speakers and Dutch speakers showed insignificant difference in their frequency of mentioning an event endpoint during motion event description). Hence, it is important to investigate how languages differ in their specific ways of encoding events by analyzing and parsing language production data.

Motion event endpoint conceptualization

In **Chapter 4** and **Chapter 5**, we investigated the effects on motion event endpoint conceptualization of two non-linguistic factors. The first factor concerns an aspect of the referential scenarios, namely whether a clear goal of the actor is present in the referential scenario (the actor's goal). The second factor is the social status of the interlocutor, specifically the social distance between the speaker and the interlocutor (the interlocutor's social status). We discovered that when a clear goal was presented in the referential scenario, participants tended to use a goal preposition (i.e., *naar* in Dutch and *to* in English) significantly more often, compared to when no clear goal was presented in the referential scenario. We also found that when the social distance between the speaker and the interlocutor was large (e.g., between a witness and a police officer), participants were less willing to commit to an event endpoint (shown by a more frequent use of *richting* in **Chapter 4** or *towards* in **Chapter 5**), compared to when the social distance between the speaker and the interlocutor was small (e.g., between friends).

Both **Chapter 4** and **Chapter 5** contribute to our understanding of motion event conceptualization by highlighting the importance of the intention of the actor in event endpoint conceptualization. Human behavior is often goal-directed. To make sense of human behavior and predict future actions, knowing and tracking the intention behind the behavior is the most natural and important thing to do. Researchers found that infants are already sensitive to the intentionality of the agent (Lakusta & Carey, 2015) and can infer the endpoint of an event from the actor's goal (Baldwin, et al., 2001).

Speech context, on the other hand, is often neglected in event conceptualization studies. Our

studies have provided evidence that it plays an important role in event conceptualization as well. Humans are social animals and therefore, event conceptualization does not always happen in isolation. To whom speech is directed has a great impact on how the speaker describes a situation. The social distance between the speaker and the interlocutor is an important factor that affects event construal. As is shown in both **Chapter 4** and **Chapter 5**, the effect of the interlocutor's social status is even larger than that of the actor's goal on motion event endpoint conceptualization.

It should still be noted that, as shown in **Chapter 5**, linguistic factors also have an impact on people's sensitivity to the above-mentioned two non-linguistic factors (i.e., the actor's goal and the interlocutor's social status). Specifically, the use of the English simple present tense in a motion event description can bring a less sensitive reading of the actor's goal as well as the speech context, compared to the use of English progressive aspect, which further influences motion event endpoint conceptualization.

The finding concerning the different functions that the English simple tense and the English progressive aspect have in event construal is also in line with the ideas presented in **Chapter 2**. That is, during sentence comprehension, grammatical aspect affects the mental representation of an event by directing readers' attention to different aspects of the depicted event. Furthermore, we have contributed to grammatical aspect studies by investigating a less-studied pair of grammatical aspect (i.e., the English simple present and the English progressive aspect) and their roles in affecting readers' sensitivity to the protagonist goal and to the speech context.

The importance of studying speakers of different languages

One important contribution of this dissertation is that we have conducted our studies in different languages, i.e., mandarin Chinese, Dutch, and English. We found that speakers of different languages conceptualize the path information of a motion event differently (i.e., mandarin Chinese speakers versus Dutch speakers; **Chapter 3**). Moreover, mandarin Chinese and Dutch differ in the way they encode motion events regarding the use of verb types, adjunct types, and aspectual markers. Some typological difference does not necessarily cause different conceptualizing patterns (i.e., different aspectual systems), whereas the other does to some extent (i.e., lexicalization patterns). This sheds light on the idea that for cross-linguistic studies, it is important to look at real-time language production data, analyze the frequency of the use of specific linguistic patterns in real-life situations, and understand cross-linguistic differences in a fine-grained way.

When comparing **Chapter 4** and **Chapter 5**, we can also easily see the importance of studying people with different linguistic background. Each language has its unique way of encoding an event, and this could have consequences on the results of our studies. Nowadays, most of the research in the fields of cognitive science and psycholinguistics has

been conducted in the English/American context where only English speakers have been investigated (see also in Henrich et al., 2010). Our review in **Chapter 2** has also mentioned this problem. Without considering speakers of other languages, research findings are likely to be limited and biased by the characteristics of one specific language. This does not mean that all research findings should be generalizable across speakers of different languages, but that we need to be aware of the consequences of focusing on only one language group. Moreover, when we are trying to generalize research findings among speakers of different languages, we need to have a comprehensive understanding of how those languages differ in their way of encoding specific events and the potential consequences of those differences.

Limitations and suggestions for future research

There are several limitations regarding the studies I have presented in this dissertation. First, we have only investigated language-based event conceptualization. Specifically, we only examined the effect of language on event conceptualization when language use is required. This is what Slobin (1996) calls "thinking for speaking". What remains unclear is whether speakers of different languages conceptualize events differently when they are not thinking for speaking. Especially in **Chapter 3**, we discovered that mandarin Chinese speakers and Dutch speakers differ in the frequency with which they select location-only information and trajectory information. Would this finding still be valid if a non-verbal task were to be used? For instance, would speakers of Dutch pay more attention to the change in the trajectory information, compared to mandarin Chinese speakers, if they were only asked to view (not to describe) these events for a later-on memory test? An even more strict non-verbal test would be adding a verbal intervening task during the viewing process (e.g., ask the participants to repeat strings of numbers or words while viewing the events to prevent any possibility of language use during event conceptualization). I believe that this surely deserves further investigation.

Second, considering the studies presented in **Chapter 4** and in **Chapter 5**, we used a limited number of stimuli/scenarios. We cannot deny a generalizability issue in these studies. That is, to what extent can our findings in these studies be generalized to other scenarios? We are positive that our findings are robust enough, given that we have generalized the effects of both the actor's goal and the interlocutor's status across two unrelated motion events in two different languages (i.e., Dutch and English). Moreover, we have also ensured the power of our study by recruiting enough participants using Frick's COAST method (Frick, 1998). However, for future studies, I encourage researchers to test other types of motion events and may even go beyond the scope of motion events. A more natural presentation of events is also recommended, such as using videos instead of static pictures to depict events.

Last but not least, we only tested the main effects of the actor's goal and the interlocutor's social status on motion event endpoint conceptualization. As indicated by the explanatory analyses of Experiment 2 in **Chapter 5**, there is a possible interaction between the actor's goal

and the interlocutor's status in affecting motion event endpoint conceptualization. That is, talking to different interlocutors might lead to different sensitivity to the characteristics of the referential scenario, including the actor's goal. Some interesting hypotheses can be tested in future studies. For instance, will talking to a police officer make people less certain with their guessing of the actor's goal no matter whether a clear goal is represented in the referential scenario, compared to talking to a friend? Or will the certainty about the actor's goal also make a difference? That is, if a clear goal is presented in the referential scenario, there should be no effect from the interlocutor's social status. However, if no clear goal is shown in the referential scenario, talking to a police officer might make people even less certain with the actor's goal, compared to talking to a friend. Such hypotheses deserve further investigation.

We employ many ways when interacting with the world. Language is one of the most amazing products of the history of human evolution. We use language to talk about our experiences, share our feelings, and express opinions. Whether we want it or not, language is affecting our life in many different ways. One intriguing fact about language is that there are so many different languages spoken in the world. There are surprising similarities as well as differences among these various languages. With all these similarities and differences, how does language really play a role in the way we experience with world, then? Moreover, to what extent does our experience with language use shed light on our general understanding of the world? This dissertation provides some compelling insights and answers, but as is the case with all research, there is room left for future exploration.

Summary

Events are commonly conveyed and understood through language. Language plays an important role in event studies. In this dissertation, I have focused on investigating the function of language in event conceptualization and event comprehension. Two specific linguistic cues play a central role in my dissertation, namely ***grammatical aspect*** and ***directional prepositions***. ***Grammatical aspect*** is a morphosyntactic marker of the verb that specifies the temporal development of an event (e.g., progressive or completed***). Directional prepositions*** are used to show some sort of movement (e.g., ***towards***/***to*** *the village*). The majority of the studies in the current dissertation is concentrating on examining how ***motion events*** (i.e., an agent moves in relation to a location/another object) are conceptualized and understood through language (**Chapter 3**, **Chapter 4**, and **Chapter 5**).

As I have just mentioned, grammatical aspect marks the temporal status of an event. Many language comprehension studies have experimentally tested the effect of grammatical aspect on the representations of an event during both sentence comprehension and discourse comprehension. Moreover, some studies have also examined the effect of grammatical aspect on social cognition, such as problem-solving and decision-making. In **Chapter 2**, we provided a review of those studies with the aim of summarizing how robust the effect of grammatical aspect is on event representations and on social cognition which is built upon event representations. Regarding the effect of grammatical aspect on event representations, most studies have provided evidence that this effect is robust and stable (during both sentence comprehension and discourse comprehension). However, regarding social cognition which relies on event representations, the effect from grammatical aspect is limited, unstable, and not robust.

Despite the importance of grammatical aspect on event representations, not all languages have this grammatical marker that specifies the temporal status of an event. Dutch is one of those languages that do not grammatically mark the progression of an event. Mandarin Chinese, on the other hand, does have a grammaticalized progressive marker to specify the ongoing stage of an event (i.e., *-zai*). **Chapter 3** examined whether this typological difference would affect how speakers of the two languages conceptualize the path of motion during a motion event description task. Previous research has shown that speakers of languages that do have a progressive marker mention an event endpoint more frequently than speakers of languages that do not have a progressive marker. Besides the effect of grammatical aspect, we have also considered another typological difference between the two languages that could also affect the conceptualization of the path of motion: how the path of motion is typically encoded in these languages (i.e., lexicalization patterns). We found that Dutch speakers did not differ from mandarin Chinese speakers in the frequency of mentioning an event endpoint. However, we found that Dutch speakers tended to mention the trajectory of motion (e.g., *de auto rijdt over de weg* 'the car drives over the road') more frequently than the location-only information (e.g., *de auto rijdt op de weg* 'the car drives on the road') in a motion event description. Mandarin Chinese speakers presented an opposite pattern, that is, more

frequent mentioning of the location-only information than the trajectory information. We attribute this difference to their different lexicalization patterns in encoding the information about the path of motion. Mandarin Chinese exhibits features of a verb-framed language (i.e., the path of motion is encoded in the verb root). In verb-framed languages, when manner verbs are used (e.g., *drive*, *walk*), the focus is on the figure in motion. Hence, often only the location information is needed to locate the figure in motion. Mentioning the trajectory information (e.g., *across/over the road*) indicates spatial changes and directs the attention to the trajectory instead of to the figure in motion. One important conclusion of this study is that typological differences do not necessarily lead to different conceptualization patterns (e.g., the effect of whether languages have or have not progressive markers on event endpoint mentioning). A more detailed analysis of language production data is needed. For instance, we found that even though mandarin Chinese has a grammaticalized progressive marker that can be used to describe directional motion events, it is not used often in directional motion event descriptions, as shown in our linguistic data. The difference between Dutch and mandarin Chinese in their aspectual systems is not as large as we expected, and the real difference can only be shown in linguistic data that are obtained through real-time language production.

When analyzing the motion event descriptions produced by Dutch speakers in the study described in **Chapter 3**, we found that two directional prepositions (i.e., *naar* 'to' and *richting* 'towards') both occurred frequently in these descriptions. The two directional prepositions differ in their implication concerning the certainty of a motion event endpoint, which provides a nice paradigm for studying motion event endpoint conceptualization. In **Chapter 4**, we, therefore, first tested their semantic differences through a language comprehension task (Experiment 1). We found that the use of the two directional prepositions in motion event descriptions indeed created distinct event representations. We then investigated the factors that could affect the use of the two directional prepositions in motion event descriptions, which sheds light on event endpoint conceptualization (Experiment 2). Two non-linguistic factors that were considered important to event endpoint construal were investigated: the actor's goal and the interlocutor's social status. We found that both the two non-linguistic factors showed significant effects on the choice of the two directional prepositions. This study has shown that both the characteristics of the referential scenario (e.g., whether a clear actor's goal in shown) and the speech context (e.g., the interlocutor's social status) are important factors that affect motion event endpoint conceptualization.

Chapter 5 is an extension of **Chapter 4**. We adopted almost the same experimental design as in **Chapter 4**, except that we examined the effects of the two non-linguistic factors on the choice of another pair of directional prepositions: two English directional prepositions (i.e., *to* and *towards*). The other difference is that we included the effect of grammatical aspect on event representations (i.e., a comparison between the English simple present and the English progressive aspect) in **Chapter 5**. The first goal of **Chapter 5** was to test whether the obtained

effects of the two non-linguistic factors on motion event endpoint conceptualization in **Chapter 4** were still upheld when another language was chosen. The second goal of **Chapter 5** was to go beyond **Chapter 4** by investigating whether grammatical aspect would affect the sensitivity to the two non-linguistic factors during event comprehension and whether this effect would consequently affect event endpoint conceptualization. The results show that the effects of the actor's goal and the interlocutor's social status on motion event endpoint conceptualization are indeed stable enough when another language is chosen. However, the use of different grammatical aspect has an effect on the sensitivity to the actor's goal and to speech context during event comprehension, which also affects motion event endpoint conceptualization.

In sum, all these chapters from this dissertation contribute to our understanding of how language plays a role in event studies and how we convey and understand events through language. They also provide important implications on how we apprehend events in general.

References

Ainsworth, J. E. (1993). In a different Register: the pragmatics of powerlessness in police interrogation. *The Yale Law Journal*, *103*(2), 259–322. https://doi.org/10.2307/797097

Altmann, G. T. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition*, *93*(2), B79–B87. https://doi.org/10.1016/j.cognition.2004.02.005

Anderson, S. E., Matlock, T., Fausey, C., & Spivey, M. J. (2008). On the path to understanding on-line processing of grammatical aspect. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2253-2258). Cognitive Science Society.

Anderson, S. E., Matlock, T., & Spivey, M. J. (2010). On-line interactions of context and grammatical aspect. In S. Ohlsson & R. Catrambone (Eds.), *Cognition in Flux* (pp. 1198–1203).

Anderson, S. E., Matlock, T., & Spivey, M. J. (2013). Grammatical aspect and temporal distance in motion descriptions. *Frontiers in Psychology*, *4*, 337. https://doi.org/10.3389/fpsyg.2013.00337

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.

Bakker, D., & Siewierska, A. (2002). Adpositions, the lexicon and expression rules. In R. M. Usón & M. J. Pérez Quintero (Eds.), *New perspectives on argument structure in functional grammar* (pp. 125–178). Berlin, Germany: Mouton de Gruyter. Retrieved from https://eprints.lancs.ac.uk/id/eprint/1345/

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*(3), 708–717. https://doi.org/10.1111/1467-8624.00310

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, *22*(4), 637–660. https://doi.org/10.1017/S0140525X99532147

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal* of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beavers, J., Levin, B., & Tham, S. W. (2010). The typology of motion expressions revisited. *Journal of Linguistics*, *46*, 331-377.

Becker, R. B., Ferretti, T. R., & Madden, C. J. (2013). Grammatical aspect, lexical aspect, and event duration constrain the availability of events in narratives. *Cognition*, *129*(2), 212–220. https://doi.org/10.1016/j.cognition.2013.06.014

Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and Language*, *112*(3), 150–158. https://doi.org/10.1016/j.bandl.2009.07.002

Bogaert, P. (2008). *A qualitative calculus for moving point objects constrained by networks* (Doctoral dissertation, Ghent University, Ghent, Belgium). Retrived from http://hdl.handle.net/1854/LU-469203

Bohnemeyer, J., & Swift, M. (2004). Event realization and default aspect. *Linguistics and Philosophy*, *27*, 263-296.

Bolker, B. M. (2015). Linear and generalized linear mixed models. In G. A. Fox, S. Negrete-Yankelevich, & V. J. Sosa (Eds.), *Ecological statistics: Contemporary theory and application* (pp. 309-333). New York: Oxford University Press.

Borghi, A. M., & Scorolli, C. (2009). Language comprehension and dominant hand motion simulation. *Human Movement Science*, *28*(1), 12–27. https://doi.org/10.1016/j.humov.2008.07.002

Broekhuis, H. (2013). *Syntax of Dutch: adpositions and adpositional phrases*. Amsterdam, the Netherlands: Amsterdam University Press. https://doi.org/10.26530/OAPEN_462289

Buccola, B., Haida, A. (2019). Obligatory Irrelevance and the Computation of Ignorance Inferences, *Journal of Semantics*, *36*(4), 583–616. https://doi org.eur.idm.oclc.org/10.1093/jos/ffz013

Carreiras, M., Carriedo, N., Alonso, M. A., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory & Cognition*, *25*(4), 438–446. https://doi.org/10.3758/BF03201120

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.

Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review*, *16*(3), 573–577. https://doi.org/10.3758/PBR.16.3.573

Cowper, E. (1998). The simple present tense in English: a unified treatment. *Studia Linguistica*, *52*(1), 1–18. https://doi.org/10.1111/1467-9582.00027

Crawley, M. J. (2002). *Statistical computing: an introduction to data analysis using S-PLUS*. New York: John Wiley & Sons.

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

de Knop, S., & Gallez, F. (2011). Manner of motion: A privileged dimension of German expressions. *International Journal of Cognitive Linguistics*, *2*, 25-40.

de Koning, B. B., Wassenburg, S. I., Bos, L. T., & van der Schoot, M. (2017). Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse Processes*, *54*(7), 493–503. https://doi.org/10.1080/0163853X.2015.1119604

Declerck, R. (2007). *Distinguishing between the aspectual categories "(a)telic", "(im)perfective" and "(non)bounded"* [Working Paper]. University of Kansas. Linguistics Graduate Student Association. https://doi.org/10.17161/KWPL.1808.1787

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.

Dowty, D. R. (1979). *Word meaning and Montague grammar: the semantics of verbs and times in generative semantics and in Montague's Ptq*. Berlin, Germany: Springer Science & Business Media.

Eerland, A., Sherrill, A. M., Magliano, J. P., & Zwaan, R. A. (2017). The blame game: An investigation of grammatical aspect and blame judgments. *Collabra-Psychology*, *3*(1), 29. https://doi.org/10.1525/collabra.113

Eerland, A., Sherrill, A. M., Magliano, J. P., Zwaan, R. A., Arnal, J. D., Aucoin, P., … Prenoveau, J. M. (2016). Registered replication report: Hart & Albarracín (2011). *Perspectives on Psychological Science*, *11*(1), 158–171. https://doi.org/10.1177/1745691615605826

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: verbal reports as data, Rev. ed* (pp. liii, 443). Cambridge, MA, US: The MIT Press.

Eschenbach, C., Tschander, L., Habel, C., & Kulik, L. (2000). Lexical specifications of paths. In C. Freksa, C. Habel, W. Brauer, & K. F. Wender (Eds.), *Spatial cognition II: Integrating abstract theories, empirical studies, formal methods, and practical applications* (pp. 127–144). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-45460-8_10

Fausey, C. M., & Matlock, T. (2011). Can grammar win elections? *Political Psychology*, *32*(4), 563–574.

Feller, D. P., Eerland, A., Ferretti, T. R., Magliano, J. P., Zwaan, R., & Radvansky, G. (2019). Aspect and narrative event segmentation. *Collabra: Psychology*, *5*(1). https://doi.org/10.1525/collabra.182

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(1), 182–196. Scopus. https://doi.org/10.1037/0278-7393.33.1.182

Ferretti, T. R., Rohde, H., Kehler, A., & Crutchley, M. (2009). Verb aspect, event structure, and coreferential processing. *Journal of Memory and Language*, *61*(2), 191–205. https://doi.org/10.1016/j.jml.2009.04.001

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*(3), 479–524. https://doi.org/10.1515/ling.2011.015

Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere: Brain potentials reflect language effects on motion event perception. *Cognition*, *141*, 41-51.

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.

Fontaine, R. G. (2007). Disentangling the psychology and law of instrumental and reactive subtypes of aggression. *Psychology, Public Policy, and Law*, *13*(2), 143–165. https://doi.org/10.1037/1076-8971.13.2.143

Frick, R. W. (1998). A better stopping rule for conventional statistical tests. *Behavior Research Methods, Instruments, & Computers*, *30*(4), 690–697. https://doi.org/10.3758/BF03209488

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: identifying, comparing, and interpreting the evidence. *Language Learning*, *67*(S1), 155–179. https://doi.org/10.1111/lang.12225

García, A. M., Moguilner, S., Torquati, K., García-Marco, E., Herrera, E., Muñoz, E., … Ibáñez, A. (2019). How meaning unfolds in neural time: Embodied reactivations can precede multimodal semantic effects during language processing. *NeuroImage*, *197*, 439–449. https://doi.org/10.1016/j.neuroimage.2019.05.002

Gelman, A. (2005). Analysis of variance: why it is more important than ever. *The Annals of Statistics*, *33*(1), 1–31.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

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

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research*, *133*(4), 468–490. https://doi.org/10.1007/s002210000431

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., & von Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, *180*, 225-237.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation*. Oxford: Academic Press.

Giles, H., & Smith, P. (1979). Accommodation theory: optimal levels of convergence. In H. Giles & R. St. Clair (Eds.), *Language and social psychology* (pp. 45–65). Basil Blackwell.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565. https://doi.org/10.3758/BF03196313

Golshaie, R., & Incera, S. (2020). Grammatical aspect and mental activation of implied instruments: A mouse-tracking study in Persian. *Journal of Psycholinguistic Research*. https://doi.org/10.1007/s10936-020-09742-3

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*(1), 163–189. https://doi.org/10.1146/annurev.psych.48.1.163

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and Semantics* (Vol. 3, pp. 41-58). New York: Academic Press.

Gruber, J. S. (1976). *Lexical structures in syntax and semantics*. North-Holland Pub. Co.; American Elsevier Pub. Co. https://catalog.hathitrust.org/Record/000716164

Grüter, T., Takeda, A., Rohde, H., & Schafer, A. J. (2018). Intersentential coreference expectations reflect mental models of events. *Cognition*, *177*, 172–176. https://doi.org/10.1016/j.cognition.2018.04.015

Hart, W., & Albarracín, D. (2010). Learning about what others were doing: Verb aspect and attributions of mundane and criminal intent for past actions. *Psychological Science*. (Sage CA: Los Angeles, CA). https://doi.org/10.1177/0956797610395393

Havas, D. A., & Chapp, C. B. (2016). Language for winning hearts and minds: Verb aspect in US presidential campaign speeches for engaging emotion. *Frontiers in Psychology*, *7*, 899. https://doi.org/10.3389/fpsyg.2016.00899

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heylighen, F., & Dewaele, J.-M. (2002). Variation in the contextuality of language: an empirical measure. *Foundations of Science*, *7*(3), 293–340. https://doi.org/10.1023/A:1019661126744

Hickmann, M., & Hendriks, H. (2010). Typological constraints on the acquisition of spatial language in French and English. *Cognitive Linguistics*, *21*,189-215.

Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation? *Memory & Cognition*, *45*(6), 974–982. https://doi.org/10.3758/s13421-017-0708-1

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Huette, S., Winter, B., Matlock, T., Ardell, D. H., & Spivey, M. (2014). Eye movements during listening reveal spontaneous grammatical processing. *Frontiers in psychology*, *5*, 410. https://doi.org/10.3389/fpsyg.2014.00410

Huette, S., Winter, B., Matlock, T., & Spivey, M. (2012). Processing motion implied in language: Eye-movement differences during aspect comprehension. *Cognitive Processing*, *13*(S1), 193–197. https://doi.org/10.1007/s10339-012-0476-6

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.

Jackendoff, R. (1991). Parts and boundaries. *Cognition*, *41*(1), 9–45. https://doi.org/10.1016/0010-0277(91)90031-X

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*, *195*, 57-71.

Johanson, M., Selimis, S., & Papafragou, A. (2019). The source-goal asymmetry in spatial language: language-general vs. language-specific aspects. *Language, Cognition and Neuroscience, 34*(7), 826–840. https://doi.org/10.1080/23273798.2019.1584323

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kan, Z. (2010). Motion-event typology in mandarin Chinese revisited. *Contemporary Linguistics*, *2*, 126-135.

Kaplan, M. J., Raju, A., & Arunachalam, S. (2021). Real-time processing of event descriptions for partially- and fully-completed events: Evidence from the visual world paradigm. *Proceedings of the Linguistic Society of America*, *6*(1), 118–132. https://doi.org/10.3765/plsa.v6i1.4954

Kintsch, W. (1998). *Comprehension: A paradigm for cognition* (pp. xvi, 461). New York, NY, US: Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

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.

Koppen, K., Ernestus, M., & van Mulken, M. (2019). The influence of social distance on speech behavior: formality variation in casual speech. *Corpus Linguistics and Linguistic Theory*, *15*(1), 139–165. https://doi.org/10.1515/cllt-2016-0056

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*(4), 543–555. https://doi.org/10.1037/0022-3514.90.4.543

Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and Grammar* (pp. 197–235). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-3969-4_9

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition*, *40*(5), 812–826. https://doi.org/10.3758/s13421-011-0179-8

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*, *44*(7), 701–710. https://doi.org/10.1002/ejsp.2023

Lakusta, L., & Carey, S. (2015). Twelve-month-old infants' encoding of goal and source paths in agentive and non-agentive motion events. *Language Learning and Development: The Official Journal of the Society for Language Development*, *11*(2), 152–157. https://doi.org/10.1080/15475441.2014.896168

Lakusta, L., & Landau, B. (2005). Starting at the end: the importance of goals in spatial language. *Cognition*, *96*(1), 1–33. https://doi.org/10.1016/j.cognition.2004.03.009

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: origins of the asymmetry between source and goal paths. *Cognitive Science*, *36*(3), 517–544. https://doi.org/10.1111/j.1551-6709.2011.01220.x

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*, 3-33. COE-ECS Working Paper.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar*. Stanford, CA: Stanford University Press.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: The MIT Press.

Levelt, W. J. M. (1999). Producing spoken language: a blueprint of the speaker. In C. M. Brown & In P. Hagoort (Eds.), *The neurocognition of language* (pp. 83–122). New York: Oxford University Press. Retrieved from http://hdl.handle.net/11858/00-001M-0000-0013-44B7-4

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.

Liao, Y., Dijkstra, K., & Zwaan, R. A. (2021). Directional prepositions and event endpoint conceptualization: a study of naar and richting in Dutch. *Language and Cognition*, *13*(2), 161–190. https://doi.org/10.1017/langcog.2020.31

Liao, Y., Flecken, M., Dijkstra, K., & Zwaan, R. A. (2020). Going places in Dutch and mandarin Chinese: conceptualising the path of motion cross-linguistically. *Language, Cognition and Neuroscience*, *35*(4), 498–520. https://doi.org/10.1080/23273798.2019.1676455

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.

Liu, N., & Bergen, B. (2016). When do language comprehenders mentally simulate locations? *Cognitive Linguistics*, *27*(2), 181–203. https://doi.org/10.1515/cog-2015-0123

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.

Madden, C. J., Dominey, P. F., & Ventre-Dominey, J. (2017). Grammatical verb aspect and event roles in sentence processing. *PLOS ONE*, *12*(12), e0189919. https://doi.org/10.1371/journal.pone.0189919

Madden, C. J., & Therriault, D. J. (2009). Short article: Verb aspect and perceptual simulations: *Quarterly Journal of Experimental Psychology*. (Sage UK: London, England). https://doi.org/10.1080/17470210802696088

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, *31*(5), 663–672. https://doi.org/10.3758/BF03196106

Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes*, *29*(2), 83–112. https://doi.org/10.1207/S15326950dp2902_1

Matlock, T. (2011). The conceptual motivation of aspect. In K.-U. Panther & G. Radden (Eds.), *Motivation in grammar and the lexicon* (pp. 133–147). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/hcp.27.09mat

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, *14*(2), 15–28.

Morey, R. D., Kaschak, M. P., Díez-Álamo, A. M., Glenberg, A. M., Zwaan, R. A., Lakens, D., … Ziv-Crispel, N. (2021). A pre-registered, multi-lab non-replication of the Action-sentence Compatibility Effect (ACE). *Psychonomic Bulletin & Review*. Retrieved from https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_3309331

Morrow, D. G. (1985). Prepositions and verb aspect in narrative understanding. *Journal of Memory and Language*, *24*(4), 390–404. https://doi.org/10.1016/0749-596X(85)90036-1

Morrow, D. G. (1986). Grammatical morphemes and conceptual structure in discourse processing. *Cognitive Science*, *10*(4), 423–455. https://doi.org/10.1016/S0364-0213(86)80012-X

Morrow, D. G. (1990). Spatial models, prepositions, and verb-aspect markers. *Discourse Processes*, *13*(4), 441–469. Scopus. https://doi.org/10.1080/01638539009544769

Mozuraitis, M., Chambers, C. G., & Daneman, M. (2013). Younger and older adults' use of verb aspect and world knowledge in the online interpretation of discourse. *Discourse Processes*, *50*(1), 1–22. https://doi.org/10.1080/0163853X.2012.726184

Norde, M. (2008). East Scandinavian på; a preposition of the 'third kind'. In K. Dekker, A. MacDonald & H. Niebaum (Eds.), *Northern voices. Essays on Old Germanic and related topics offered to Professor Tette Hofstra (= Germania Latina VI)* (pp. 383-398). Leuven, Belgium: Peeters. https://www.rug.nl/research/portal/publications/east-scandinavian-paa-a-preposition-of-the-third-kind(d55b1c0c-43fe-4f36-b17b-b9651b8f7ac0)/export.html

Oberpriller, J., de Souza Leite, M., & Pichler, M. (2021). *Fixed or random? On the reliability of mixed-effects models for a small number of levels in grouping variables* [Preprint]. Ecology. https://doi.org/10.1101/2021.05.03.442487

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*(6), 1064–1092. https://doi.org/10.1111/j.1551-6709.2010.01107.x

Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: evidence from event encoding. *Language, Cognition and Neuroscience*, *34*(9), 1117–1128. https://doi.org/10.1080/23273798.2019.1589540

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.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2019). nlme: linear and nonlinear mixed effects models. R package version 3.1-142, https://CRAN.R project.org/package=nlme.

Piñón, C. J. (1993). Paths and their names. In K. Beals, G. Cooke, D. Katman, S. Kita, K.-E. McCullough, & D. Testen, (Eds.), *Chicago Linguistics Society 29* (Vol. 2, pp. 287–303). Chicago Linguistics Society.

Psychology Software Tools. (2012). E-Prime (Version 2.0). Retrieved from https://www.pstnet.com

Quine, W. V. (1985). Events and reification. In E. Lepore, & B. McLaughlin, (Eds.), *Actions and Events: Perspectives on the Philosophy of Davidson* (pp. 162–171). Oxford, England: Blackwell.

R Core Team. (2016). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. NY, US: Oxford University Press.

Regier, T., & Zheng, M. (2007). Attention to endpoints: a cross-linguistic constraint on spatial meaning. *Cognitive Science*, *31*(4), 705–719. https://doi.org/10.1080/15326900701399954

Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, *102*(1), 129–138. https://doi.org/10.1016/j.cognition.2005.12.004

Rychlý, P. (2008). *A lexicographer-friendly association score*. Masarykova Univerzita. https://www.muni.cz/en/research/publications/937193

Salomon, M. M., Magliano, J. P., & Radvansky, G. A. (2013). Verb aspect and problem solving. *Cognition*, *128*(2), 134–139. https://doi.org/10.1016/j.cognition.2013.03.012

Shen, J. (2003). The resultative construction in Chinese: a typological perspective. *Chinese Teaching in the World*, *3*, 17-23.

Sherrill, A. M., Eerland, A., Zwaan, R. A., & Magliano, J. P. (2015). Understanding how grammatical aspect influences legal judgment. *PLOS ONE*, *10*(10), e0141181. https://doi.org/10.1371/journal.pone.0141181

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. Strömqvist & 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.

Smith, C. S. (1991). *The parameter of aspect*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Smith, C. S. (1997). *The parameter of aspect*. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-5606-6

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Stefanowitsch, A., & 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.

Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Tropical Medicine and Health*, *47*(1), 46. https://doi.org/10.1186/s41182-019-0165-6

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & Cognition*, *34*(1), 78–89. https://doi.org/10.3758/BF03193388

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Ünal, E., Ji, Y., & Papafragou, A. (2021). From event representation to linguistic meaning. *Topics in Cognitive Science*, *13*(1), 224–242. https://doi.org/10.1111/tops.12475

van de Schoot, Rens, De Bruin, Jonathan, Schram, Raoul, Zahedi, Parisa, De Boer, Jan, Weijdema, Felix, … Oberski, Daniel. (2021, January 18). ASReview: Active learning for systematic reviews (Version v0.16). Zenodo. http://doi.org/10.5281/zenodo.4446746

van der Sijs, N. (2001). *Chronological dictionary. The age and origin of our words and meanings (2nd ed)*. Amsterdam/ Antwerpen: Veen.

van Dijk, T. A. V., & Kintsch, W. (1983). *Strategies of discourse comprehension*. NY, US: Academic Press.

Vandeloise, C. (2017). Three basic prepositions in French and in English: a comparison. *Corela. Cognition, Représentation, Langage*, *HS-23*, Article HS-23. https://doi.org/10.4000/corela.5033

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press. https://doi.org/10.7591/9781501743726

Verkuyl, H., & Zwarts, J. (1992). Time and space in conceptual and logical semantics: the notion of path. *Linguistics*, *30*(3), 483–512. https://doi.org/10.1515/ling.1992.30.3.483

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.

von Stutterheim, C., Carroll, M., & Klein, W. (2009). New perspectives in analyzing aspectual distinctions across languages. In W. Klein & P. Li (Eds.), *The expression of time* (pp. 195–216). Berlin: Mouton de Gruyter.

Vraciu, A. (2015). The simple present and the expression of temporality in L1 English and L2 English oral narratives. In D. Ayoun (Ed.), *The Acquisition of the Present* (pp. 289-334). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.196.10vra

Williams, C. (2002). Non-progressive aspect in English in commentaries and demonstrations using the present tense. *Journal of Pragmatics*, *34*(9), 1235–1256. https://doi.org/10.1016/S0378-2166(01)00052-2

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.

Yap, F. H., Chu, P. C. K., Yiu, E. S. M., Wong, S. F., Kwan, S. W. M., Matthews, S., … Shirai, Y. (2009). Aspectual asymmetries in the mental representation of events: Role of lexical and grammatical aspect. *Memory & Cognition*, *37*(5), 587–595. https://doi.org/10.3758/MC.37.5.587

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*(6), 979–1008. https://doi.org/10.1207/s15516709cog2806_5

Zacks, J. M. (2020). Event Perception and Memory. *Annual Review of Psychology*, *71*(1), 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., Speer, N., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*, 307-327.

Zacks, J. M., Speer, N., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin*, *133*, 273-293.

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zhan, L. (2018). Scalar and ignorance inferences are both computed immediately upon encountering the sentential connective: The online processing of sentences with disjunction using the visual world paradigm. *Frontiers in Psychology*, *9*, 61. https://doi-org.eur.idm.oclc.org/10.3389/fpsyg.2018.00061

Zhou, P., Crain, S., & Zhan, L. (2014). Grammatical aspect and event recognition in children's online sentence comprehension. *Cognition*, *133*(1), 262–276. https://doi.org/10.1016/j.cognition.2014.06.018

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1196–1207. https://doi.org/10.1037/0278-7393.22.5.1196

Zwaan, R. A. (2008). Time in language, situation models, and mental simulations. *Language Learning*, *58*, 13–26. https://doi.org/10.1111/j.1467-9922.2008.00458.x

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*, 1028–1034. https://doi.org/10.3758/s13423-015-0864-x

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 224–245). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511499968.010

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0051382

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1–11. https://doi.org/10.1037/0096-3445.135.1.1

Zwarts, J. (2005). Prepositional aspect and the algebra of paths. *Linguistics and Philosophy*, *28*(6), 739–779. Scopus. https://doi.org/10.1007/s10988-005-2466-y

Zwarts, J. (2008). Aspects of a typology of direction. In S. D. Rothstein (Ed.), *Linguistik Aktuell/Linguistics Today* (Vol. 110, pp. 79–105). Amsterdam, the Netherlands: John Benjamins Publishing Company. https://doi.org/10.1075/la.110.05

Nederlandse Samenvatting

(Summary in Dutch)

Gebeurtenissen worden gewoonlijk overgebracht en begrepen door middel van taal. Taal speelt een belangrijke rol in de studie van gebeurtenissen. In dit proefschrift heb ik mij gericht op het onderzoeken van de functie van taal in de conceptualisatie en het begrijpen van gebeurtenissen. Twee specifieke linguïstische categorieën spelen een centrale rol in mijn proefschrift, namelijk ***grammaticaal aspect*** en ***directionele voorzetsels***. ***Grammaticaal aspect*** is een morfosyntactische markeerder van het werkwoord die de temporele ontwikkeling van een gebeurtenis specificeert (bv. progressief of voltooid). ***Directionele voorzetsels*** worden gebruikt om een beweging aan te geven (bv. ***richting/naar*** *het dorp*). Het merendeel van de studies in het huidige proefschrift concentreert zich op het onderzoeken van hoe ***bewegingsgebeurtenissen*** (d.w.z. een agent beweegt in relatie tot een locatie/een ander object) worden geconceptualiseerd en begrepen door middel van taal (**Hoofdstuk 3**, **Hoofdstuk 4**, en **Hoofdstuk 5**).

Zoals ik zojuist heb gezegd, markeert grammaticaal aspect de temporele status van een gebeurtenis. Veel taalbegripstudies hebben experimenteel getest wat het effect is van grammaticaal aspecten op de mentale representaties van een gebeurtenis tijdens zowel het begrijpen van zinnen als van discourse. Bovendien hebben enkele studies ook het effect van grammaticaal aspect op sociale cognitie bestudeerd, zoals het oplossen van problemen en het nemen van beslissingen. In **Hoofdstuk 2** hebben we een overzicht gegeven van deze studies met als doel weer te geven hoe robuust het effect van grammaticaal aspect is op gebeurtenisrepresentaties en op sociale cognitie en hoe die gedeeltelijk beïnvloed wordt door gebeurtenisrepresentaties. Wat betreft het effect van grammaticaal aspect op gebeurtenisrepresentaties, hebben de meeste studies bewijs geleverd dat dit effect robuust en stabiel is (zowel tijdens zinsbegrip als tijdens het begrijpen van discourse). Wat betreft de sociale cognitie die slechts gedeeltelijk gebaseerd is op gebeurtenisrepresentaties, is het effect van grammaticaal aspect echter beperkt, instabiel, en niet robuust.

Ondanks het belang van grammaticaal aspect op gebeurtenis-representaties, hebben niet alle talen deze grammaticale markeerder die de temporele status van een gebeurtenis specificeert. Het Nederlands is één van de talen die de voortgang van een gebeurtenis niet grammaticaal markeert. Het Mandarijn Chinees heeft echter wel een grammaticaal progressieve markering om de lopende fase van een gebeurtenis aan te geven (d.w.z. *-zai*). **In hoofdstuk 3** werd onderzocht of dit typologische verschil van invloed is op hoe sprekers van de twee talen het bewegingspad conceptualiseren tijdens een beschrijvingstaak van bewegingsgebeurtenissen. Eerder onderzoek heeft aangetoond dat sprekers van talen die wel een progressieve markeerder hebben, vaker een eindpunt van een gebeurtenis noemen dan sprekers van talen die geen progressieve markeerders hebben. Naast het effect van het grammaticale aspect hebben we ook gekeken naar een ander typologisch verschil tussen de twee talen dat ook van invloed zou kunnen zijn op de conceptualisatie van het bewegingspad: hoe het bewegingspad typisch wordt gecodeerd in deze talen (d.w.z. lexicalisatiepatronen). We vonden dat Nederlandstaligen niet verschilden van

Mandarijn-Chineestaligen in de frequentie van het noemen van een eindpunt van een gebeurtenis. Wel bleek dat Nederlandstaligen vaker het traject van de beweging noemden (bv. *de auto rijdt over de weg*) dan alleen de locatie-informatie (bv. *de auto rijdt op de weg*) in een beschrijving van een bewegingsgebeurtenis. Chinese sprekers in het Mandarijn vertoonden een tegenovergesteld patroon, dat wil zeggen dat ze de informatie over de locatie vaker vermeldden dan de informatie over het traject. Wij schrijven dit verschil toe aan hun verschillende lexicaliseringspatronen bij het encoderen van de informatie over het traject van de beweging. Mandarijn Chinees vertoont kenmerken van een werkwoord-gekaderde taal (d.w.z. het bewegingspad is gecodeerd in de werkwoordstam). In werkwoord-gekaderde talen, wanneer werkwoorden die de manier van bewegen uitdrukken worden gebruikt (bv. *rijden*, *lopen*), ligt de nadruk op de figuur in beweging. Daarom is vaak alleen de locatie-informatie nodig om de figuur in beweging te lokaliseren. Het vermelden van de trajectinformatie (bv. *over de weg*) geeft ruimtelijke veranderingen aan en richt de aandacht op het traject in plaats van op de figuur in beweging. Een belangrijke conclusie van deze studie is dat typologische verschillen niet noodzakelijk leiden tot verschillende conceptualisatiepatronen (bv. het effect van het al dan niet hebben van progressieve markeerders in talen op het noemen van eindpunten van gebeurtenissen). Een meer gedetailleerde analyse van taalproductiegegevens is nodig. Zo ontdekten we bijvoorbeeld dat hoewel het Mandarijn-Chinees een grammaticaal verantwoorde progressieve markeerder heeft die gebruikt kan worden om richtinggevende bewegingsgebeurtenissen te beschrijven, deze markeerder helemaal niet vaak gebruikt wordt in de beschrijvingen van richtinggevende bewegingsgebeurtenissen, zoals blijkt uit onze linguïstische data. Het verschil tussen het Nederlands en het Mandarijn Chinees in hun aspectuele systemen is niet zo groot als we verwachtten. Het echte verschil kan alleen worden aangetoond in taalkundige gegevens die zijn verkregen door middel van real-time taalproductie.

Bij het analyseren van de beschrijvingen van bewegingsgebeurtenissen door Nederlandse sprekers in de studie beschreven in **Hoofdstuk 3**, vonden we dat twee directionele voorzetsels (d.w.z. *naar* en *richting*) beide vaak voorkwamen in deze beschrijvingen. De twee directionele voorzetsels verschillen in hun implicatie met betrekking tot de zekerheid van een eindpunt van een bewegingsgebeurtenis, wat een mooi paradigma oplevert voor het bestuderen van de conceptualisatie van het eindpunt van een bewegingsgebeurtenis. **In Hoofdstuk 4** testten we daarom eerst hun semantische verschillen door middel van een taalbegripstaak (Experiment 1). We ontdekten dat het gebruik van de twee richtingsvoorzetsels in beschrijvingen van bewegingsgebeurtenissen inderdaad verschillende representaties van gebeurtenissen creëerde. Vervolgens onderzochten we de factoren die het gebruik van de twee directionele voorzetsels in beschrijvingen van bewegingsgebeurtenissen zouden kunnen beïnvloeden, wat licht werpt op het conceptualiseren van het eindpunt van gebeurtenissen (Experiment 2). Twee niet-linguïstische factoren die belangrijk werden geacht voor de constructie van het eindpunt van een gebeurtenis werden onderzocht: het doel van de acteur en de sociale status van de gesprekspartner. We vonden dat de twee niet-

linguïstische factoren significante effecten vertoonden op de keuze van de twee directionele voorzetsels. Deze studie heeft aangetoond dat zowel de kenmerken van het referentiële scenario (bv. of er een duidelijk doel van de acteur wordt getoond) als de spraakcontext (bv. de sociale status van de gesprekspartner) belangrijke factoren zijn die van invloed zijn op de conceptualisatie van het eindpunt van een bewegingsgebeurtenis.

Hoofdstuk 5 is een uitbreiding van **Hoofdstuk 4**. We hebben bijna hetzelfde experimentele design gebruikt als in **Hoofdstuk 4**, met als verschil dat we de effecten van de twee niet-linguïstische factoren op de keuze van een ander paar directionele voorzetsels hebben onderzocht: twee Engelse directionele voorzetsels (d.w.z. *to* en *towards*). Het andere verschil is dat we het effect van grammaticaal aspect op gebeurtenisrepresentaties (d.w.z. een vergelijking tussen het Engelse simple present en het Engelse progressive aspect) in **Hoofdstuk 5** hebben opgenomen. Het eerste doel van **Hoofdstuk 5** was om te testen of de verkregen effecten van de twee niet-linguïstische factoren op de conceptualisatie van bewegingsgebeurtenissen in **Hoofdstuk 4** standhouden wanneer een andere taal wordt gekozen. Het tweede doel van **Hoofdstuk 5** was om verder te gaan dan **Hoofdstuk 4** door te onderzoeken of het grammaticale aspect de gevoeligheid voor de twee niet-linguïstische factoren tijdens het begrijpen van gebeurtenissen beïnvloedt en of dit effect vervolgens van invloed zou zijn op het conceptualiseren van het eindpunt van de gebeurtenis. De resultaten tonen aan dat de effecten van het doel van de actor en de sociale status van de gesprekspartner op de conceptualisering van het eindpunt van de gebeurtenis inderdaad stabiel genoeg zijn wanneer een andere taal wordt gekozen. Het gebruik van verschillende grammaticale aspecten heeft echter een effect op de gevoeligheid voor het doel van de acteur en voor de spraakcontext tijdens het begrijpen van gebeurtenissen, wat ook van invloed is op het conceptualiseren van het eindpunt van een bewegingsgebeurtenis.

Samenvattend, het onderzoek gerapporteerd in dit proefschrift draagt bij aan ons wetenschappelijke inzicht in de rol van taal met betrekking tot het begrijpen van gebeurtenissen en hoe we gebeurtenissen overbrengen en begrijpen door middel van taal. Het biedt ook inzicht in hoe we gebeurtenissen in het algemeen begrijpen.

Portfolio

Curriculum Vitae

Yiyun Liao was born in Hunan, China, on the 25th of March in 1993. From 2009 to 2013, she studied English Language and Literature and obtained her Bachelor's degree at Yunnan University, China. During this period, she also studied at the English Language Department at Kohn Kaen University, Thailand, as an exchange student for half a year. From 2013 to 2016, she studied Linguistics and Applied Linguistics and obtained her Master's degree at Beihang University, China. In September 2016, she started her PhD research at the Department of Psychology, Education and Child Studies at Erasmus University Rotterdam, the Netherlands. Her PhD project focused on the role of language in event comprehension and event conceptualization. The studies she conducted for her PhD project are presented in this dissertation. While conducting research as a PhD candidate, Yiyun presented her work at various national and international conferences and completed various courses. In 2020, she earned the award of Best Presentation at the Graduate Research Day.

Publications

Liao, Y., Dijkstra, K., & Zwaan, R. A. (2021). Directional prepositions and event endpoint conceptualization: a study of *naar* and *richting* in Dutch. *Language and Cognition*, *13*(2), 161-190.

Liao, Y., Flecken, M., Dijkstra, K., & Zwaan, R. A. (2020). Going places in Dutch and mandarin Chinese: conceptualising the path of motion cross-linguistically. *Language, Cognition and Neuroscience*, *35*(4), 498–520.

Manuscripts submitted for publication

Liao. Y., Dijkstra, K., & Zwaan, R. A. Linguistic and non-linguistic cues in motion event endpoint conceptualization: the selection between English *to* and *towards*

Liao. Y., Dijkstra, K., & Zwaan, R. A. The role of grammatical aspect in and beyond event comprehension

Presentations and workshops

2020. *An empirical study of the two Dutch words: naar and richting*. Oral presentation at the Graduate Research Day, Erasmus University Rotterdam, the Netherlands, March 12. (*Winner Best Presentation*)

2019. *Directional prepositions and event boundary conceptualization*. Oral presentation at the 1st International Conference for Young Researchers in Cognitive Linguistics (YRCL1) held at the Universidad Complutense de Madrid, Spain, November 28-29.

2019. *Go to or towards a certain place? A study of naar and richting in Dutch*. Oral presentation at the 15th International Cognitive Linguistics Conference held at Kwansei Gakuin University, Nishinomiya, Japan, August 6-11.

2018. *Going (toward) places: motion encoding in Dutch and Mandarin Chinese*. Poster presentation at the 3rd International Meeting of the Psychonomic Society, Amsterdam, the Netherlands, May 10-12.

2017. *Grammatical aspect and motion event endpoints in Dutch and Chinese languages.* Poster presentation at the 16th NVP Winter Conference, Egmond aan Zee, the Netherlands, December 14-16.

2017. *Grammatical aspect and motion event cognition: a cross-linguistic comparison between Dutch and Chinese*. Oral presentation at the 7th biennial conference of BeNeCLA

(CogLing7) held at Radboud University, Nijmegen, the Netherlands, January 5-6.

2017. *Grammatical aspect and motion event endpoints in Dutch and Chinese*. Poster presentation at Event Representations in Brain, Language and Development Workshop held at Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, October 27-28.

Guest lectures

2019. Speaker. Guest lecture for the Psychology bachelor course 2.1 *Thinking and Remembering*, Erasmus University Rotterdam, the Netherlands, September 23.

2019. Speaker. Guest lecture for the master program Brain and Cognition on *Linguistic Relativity*, Erasmus University Rotterdam, the Netherlands, October 11.

Courses

Data visualisation, web scraping, and text analysis in R (2020)	2.5 EC
English academic writing for PhD candidates (2019)	2 EC
Doing the (systematic) literature review (2018)	2 EC
Shut up and write (2018)	1 EC
Digital research methods for textual data (2018)	2.5 EC
Data analysis with R (2018)	1 EC
Introduction to coding with ATLAS.ti. (2017)	1 EC
Professionalism and integrity in research (2017)	1.5 EC
Brush up your SPSS skills (2016)	1 EC
How to survive your PhD (2016)	2.5 EC