

## Full Length Article

## Representation of event boundedness in English and Mandarin speakers

Yue Ji<sup>a,\*</sup>, Anna Papafragou<sup>b</sup><sup>a</sup> Department of English, School of Foreign Languages, Beijing Institute of Technology, No. 8 and 9 Yards, Liangxiang East Road, Fangshan District, Beijing 102488, China<sup>b</sup> Department of Linguistics, University of Pennsylvania, 3401C Walnut St., Suite 300C, Philadelphia, PA 19104, USA

## ARTICLE INFO

## Keywords:

Telicity  
Aspect  
Boundedness  
Event cognition  
Event perception  
English  
Mandarin

## ABSTRACT

Event cognition is sensitive to whether an event is bounded (has a well-defined endpoint, e.g. build a sandcastle) or unbounded (lacks such an endpoint; e.g., play with sand). Boundedness interfaces with telicity in language: telic verb phrases denote events that include an inherent or natural endpoint while atelic verb phrases denote events that lack such an endpoint. Given that languages encode telicity in different ways, could these cross-linguistic differences influence the perception of event boundedness? We address this question by comparing English and Mandarin native speakers. We show that the two groups differ in their use of telicity in event descriptions (Experiment 1) but perform similarly when rating the likelihood of an event having a natural endpoint (Experiment 2) or attending to the temporal structure of bounded vs. unbounded events in a perceptual task (Experiment 3). These findings reveal commonalities in the representation of the temporal profile of events despite cross-linguistic differences.

## 1. Introduction

The world provides us with a dynamic and continuous flow of visual input but we perceive and represent this input in terms of specific units, or *events* (Filip, 1993; Parsons, 1990; Zacks & Tversky, 2001). Language is also a powerful tool for encoding events, including the temporal profile of situations in the world (Parsons, 1990; Partee, 1999; van Hout, 2016; Vendler, 1957). How do non-linguistic event representations connect to the linguistic encoding of events? Do differences between languages affect the way events are perceived and conceptualized? This is a highly debated topic within cognitive science (for different perspectives, see Baier et al., 2023; Flecken et al., 2015; Gerwien & von Stutterheim, 2018; Golinkoff et al., 2025; Konishi et al., 2019; Papafragou et al., 2008; among many others). To address these questions, here we focus on a fundamental aspect of event cognition, namely event boundedness (i.e., the temporal contour of an event), explore its manifestation cross-linguistically, and ask whether it generalizes across members of different linguistic communities or is instead shaped by language-specific patterns.

## 1.1. Boundedness in event cognition

Human minds draw an important distinction between *bounded events*

that are conceptualized as having an inherent, or natural endpoint (e.g., building a sandcastle) and *unbounded events* that are conceptualized as lacking such an endpoint (e.g., playing with sand; Ji & Papafragou, 2020a). A key feature of bounded events is that they can be divided into distinguishable stages. For example, building a sandcastle may include the steps of digging up sand, making the base, building walls, etc. By contrast, unbounded events have a largely undifferentiated internal structure. For example, playing with sand may involve scooping and pouring sand in a repetitive manner. The boundedness distinction in the event domain can be thought of as an analog of the object/substance distinction in the spatial domain (Ji & Papafragou, 2020a). As with objecthood, boundedness is a mental perspective on events, rather than an objective feature of the dynamic input: the same happening can be construed as either a bounded or an unbounded event (e.g., the same situation can be construed as either a bounded event of building a sandcastle, or an unbounded event of playing with sand; Vurgun et al., 2024).

Boundedness is an important formal property that shapes the way events are conceptualized. For example, when presented with a few videos showing everyday events, both adults and 4-to-5-year olds were able to classify events such as drawing a balloon, eating a pretzel, and stacking a deck of cards into one (bounded) category, and events such as drawing circles, eating cheerios, and shuffling a deck of cards into

\* Corresponding author at: Department of English, School of Foreign Languages, Beijing Institute of Technology, No. 8 and 9 Yards, Liangxiang East Road, Fangshan District, Beijing 102488, China.

E-mail addresses: [jiyue@bit.edu.cn](mailto:jiyue@bit.edu.cn) (Y. Ji), [anna4@sas.upenn.edu](mailto:anna4@sas.upenn.edu) (A. Papafragou).

<https://doi.org/10.1016/j.cognition.2026.106443>

Received 2 June 2025; Received in revised form 4 December 2025; Accepted 4 January 2026

Available online 9 January 2026

0010-0277/© 2026 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

another (unbounded) category; they could further extend the bounded vs. unbounded categories to new event items (Ji & Papafragou, 2020a; Ji & Papafragou, 2024). Importantly, people could classify events and distinguish between the two event categories even when their use of language was blocked by a secondary linguistic shadowing task (Ji & Papafragou, 2020a). This finding suggests that boundedness drives people's event representation independently of their use of language during the task.

Further evidence shows that boundedness is spontaneously computed in event perception, even when not relevant in the context. In an interruption-detection task (Ji & Papafragou, 2022), participants watched videos showing bounded events (e.g., stacking a deck of cards) or unbounded events (e.g., shuffling a deck of cards), some of which included very brief interruptions inserted at either event middles or close to event endpoints. The interruptions were external to the event content, and thus their detection was expected to be lower when more processing resources were drawn by the event stimuli themselves (see also Huff et al., 2012; Papenmeier et al., 2019). The results show that viewers exposed to bounded events were more likely to miss interruptions close to the endpoints as their attention was drawn towards moments of culmination. In comparison, viewers of unbounded events detected midpoint and late interruptions equally well since such events are composed of largely undifferentiated temporal slices. In a word, viewers spontaneously track the temporal texture of unfolding events and treat endpoints as salient compared to other time points (such as the midpoints) only in bounded but not unbounded events.

The notion of boundedness, or the availability of an inherent endpoint, is tied to the internal structure of events. Indeed, bounded events behave like individuated entities with well-defined structure, unlike unstructured events (Lee et al., 2024). First, bounded events but not unbounded ones resist restructuring, as the temporal slices within a bounded event are organized and cannot be rearranged. Second, and relatedly, bounded events but not unbounded ones do not tolerate breaks: viewers were more likely to judge interruptions as problematic when these occurred within a bounded event compared to an unbounded one. Third, bounded but not unbounded events have distinct parts: viewers perceived the temporal slices of a bounded event as more distinct from one another than those of an unbounded event. Importantly, these three signatures of individuation in bounded (but not unbounded) events also apply to objects (but not substances) and thus reveal architectural constraints on entities across the temporal and spatial domains (Lee et al., 2024).

In sum, boundedness is an important property of event cognition that shapes the way events are processed and represented. However, research on boundedness (and event cognition more broadly) has so far been conducted almost exclusively with English speakers. Note that event structure – including temporal structure – is encoded very differently across languages (e.g., Botne, 2003; Filip, 2004). To what extent does this variation affect non-linguistic representations of boundedness? In the next section we introduce English and Mandarin data as a test case for addressing this question.

## 1.2. The temporal structure of events cross-linguistically

The linguistic counterpart of the notion of boundedness is *telicity*, which refers to internal temporal event structure (Vendler, 1957; see Filip, 2012 for an overview). In English, telic predicates (e.g., *build a castle*) describe events with different development stages evolving towards a “built-in terminal point” (Comrie, 1976) or “culmination” (Parsons, 1990). Such events have an inherent, or natural endpoint. For example, an event of building a castle comes to an end when a castle came into being. By contrast, atelic predicates (e.g., *play with sand*) describe events with largely undifferentiated stages. Such events lack an inherent endpoint and can terminate arbitrarily. An event of playing with sand does not specify how or when it ends.

The telicity of a verb phrase depends compositionally on the nature

of the action encoded in the lexical semantics of the verb as well as other verbal elements and the nature of the affected object(s) (for reviews, see Filip, 2012; Rothstein, 2004). First, telic predicates require verbs denoting an action leading to a change of state in the affected object (or some other salient change), and thus the endpoint is naturally the resultant state; atelic phrases involve verbs denoting actions that do not affect the object in a perceptible way. The contrast is shown in (1a) and (1b). Second, telic predicates are often formed by a quantified object noun phrase (e.g., “an apple” in 2a; Krifka, 1989, 1998), as the changes in the object delimit or “measure out” the way the event develops (Tenny, 1987). In comparison, atelic phrases have direct objects unspecified for quantity (e.g., the bare plural form “apples” in 2b).

- |     |                      |          |
|-----|----------------------|----------|
| (1) | a. He opened a box.  | (telic)  |
|     | b. He pushed a box.  | (atelic) |
| (2) | a. She ate an apple. | (telic)  |
|     | b. She ate apples.   | (atelic) |

The telicity of a verb phrase can be diagnosed through specific linguistic tests. As shown in (3), a telic predicate is congruent with a delimited temporal phrase (i.e., *in X time*), which denotes that the event described by the predicate has reached the endpoint within a certain amount of time; but telic predicates are incongruent with a durative temporal phrase (i.e., *for X time*). By contrast, an atelic predicate behaves in the opposite way as in (4) (Dowty, 1979; Smith, 1991).

- |     |   |          |
|-----|---|----------|
| (3) | a. A child built a castle in half an hour.    | (telic)  |
|     | b. *A child built a castle for half an hour.  |          |
| (4) | a. A child played with sand for half an hour. | (atelic) |
|     | b. *A child played with sand in half an hour. |          |

In Mandarin, verb phrases also have the telic-atelic distinction.<sup>1</sup> Similar to English, the verb component contributes to expressing telicity. However, to denote an action that leads to a clear state change in the affected object, a resultative verb compound (henceforth RVC) is widely used and forms a telic verb phrase. An RVC (e.g., *da-kai* “hit-open” in 5a)<sup>2</sup> is composed of two verbal elements, with the first one (e.g., *da* “hit”) denoting an action and the second one denoting the result of the action (e.g., *kai* “open”, Li & Thompson, 1981; Sybesma, 1999; Tham, 2014). Importantly, Mandarin lacks accomplishment mono-morphemic verbs that are typical for forming a telic verb phrase in English. The mono-morphemic verbs either express states or activities that are inherently atelic (e.g., *tui* “push” in 5b; Lin, 2004; Peck et al., 2013; Sybesma, 1997; Tai, 1984; but see Smith, 1994; Soh & Kuo, 2005 for an alternative view) or are achievement verbs (e.g., *kai* “open”, the second component of the RVC *da-kai* “hit-open” in 5a).

- |     |                    |          |     |        |        |          |
|-----|--------------------|----------|-----|--------|--------|----------|
| (5) | a. ta              | da-kai   | le  | yi-ge  | he-zi. | (telic)  |
|     | he                 | hit-open | PFV | one-CL | box    |          |
|     | “He opened a box.” |          |     |        |        |          |
|     | b. ta              | tui      | le  | yi-ge  | he-zi. | (atelic) |
|     | he                 | push     | PFV | one-CL | box    |          |
|     | “He pushed a box.” |          |     |        |        |          |

Unlike English, Mandarin lacks an overt count-mass distinction, and all nouns can appear in their bare form (Chierchia, 1998). Bare nouns are “indeterminate”: depending on the syntactic and pragmatic context, they can be interpreted as referential (i.e., referring to entities identifiable in the context), or non-referential (i.e., not pointing to existent

<sup>1</sup> Similar to the literature on English, many studies on Mandarin have analyzed telicity at the verb phrase level. At the sentential level, the *ba*-construction (a sentence structure with S-*ba*-O-V word order) also expresses telic meaning (Sybesma, 1992; Xuan, 2017). This construction behaves as a telic verb phrase when tested with the telicity diagnostics mentioned in the text.

<sup>2</sup> The abbreviations for grammatical constituents throughout this paper are as follows: CL = classifier; PFV = perfective aspect marker.

entities in the discourse), definite or indefinite (Chen, 2014; cf. Soh & Kuo, 2005). The semantics of the verb component may determine the interpretation of a bare noun. Specifically, when a bare noun (e.g., *pingguo* “apple”) appears in a verb phrase that includes an RVC (e.g., *chi-diao* “eat-fall”), the bare noun receives a quantified interpretation and the whole phrase is telic as shown in (6a). In comparison, a bare noun in a verb phrase with a mono-morphemic verb (e.g., *chi* “eat”) has a default unquantified interpretation, and the verb phrase denotes an activity as in (6b). To sum up, since a bare noun does not carry any information about the quantity of its referent, it alone cannot delimit events.<sup>3</sup> The telicity of a verb phrase that includes a bare noun depends on the verb component.

(6)	a. ta	chi-diao	le	pingguo.
	she	eat-fall	PFV	apple
	“She ate up the apple(s).”			
	b. ta	chi	le	pingguo.
	she	eat	PFV	apple
	“She did some apple-eating activity.”			

In Mandarin, the telicity of a predicate can be tested through the possible locations of a temporal noun phrase (i.e., “*X time*”, Duan, 2019; Ernst, 1987; Lin, 2008). If a temporal phrase can appear before the verb but not before the direct object as in (7a-b), then it modifies a telic predicate and expresses similar meaning as the English temporal adverbial *in X time*.<sup>4</sup> If a temporal phrase can appear before the direct object but not before the verb as in (8a-b), then it modifies an atelic predicate and behaves like the English temporal adverbial *for X time*.

(7)	a. xiao pengyou	ban-ge xiaoshi	gai-qi	le	cheng-bao.
	little friend	half-CL hour	build-rise	PFV	castle
	“A child built a castle/some castles in half an hour.”				
	b.* xiao pengyou	gai-qi	le	ban-ge xiaoshi	cheng-bao.
	little friend	build-rise	PFV	half-CL hour	castle
	**“A child built a castle/some castles for half an hour.”				
(8)	a. xiao pengyou	wan	le	ban-ge xiaoshi	sha-zi.
	little friend	play	PFV	half-CL hour	sand
	“A child played with sand for half an hour.”				
	b.* xiao pengyou	ban-ge xiaoshi	wan	le	sha-zi.
	little friend	half-CL hour	play	PFV	sand
	**“A child played with sand in half an hour.”				

In sum, to express whether an action leads to a salient change in state of the affected object, English speakers typically choose between different verbs, while Mandarin speakers choose between RVCs that specify the resultant state and mono-morphemic verbs that denote a process or an activity. Another major difference between English and Mandarin lies in expressing object quantity. When a single object is involved, English speakers have to use a quantified noun phrase since a bare singular form is ungrammatical (e.g., \**build castle*) but Mandarin speakers have the option of not specifying the quantity as bare nouns are

<sup>3</sup> A numeral and a classifier (e.g., *yi-ge pingguo* “an apple”) can precede a noun to overtly specify the quantity of the referent in Mandarin. Numeral-classifier phrases can delimit events and form a telic verb phrase (see further analysis and comparisons with other languages in Koenig & Chief, 2008; Soh & Kuo, 2005; Zhang, 2020, among others). The present study does not elaborate on number-classifier phrases for two main reasons. First, there is similarity between such phrases and the quantified noun phrases in English in terms of forming telic phrases. Second, such phrases are not frequent in event descriptions; when there is no need to specify quantity, Mandarin speakers simply use bare nouns. Our study does not examine potential effects of context and pragmatics on the expression of telicity, even though the role of these factors is well understood.

<sup>4</sup> The English *in/within X time* can be translated into *zai X nei* (*nei* is a spatial term meaning “inside”). This temporal phrase can only appear before the verb and modify a telic verb phrase (Duan, 2019; Zhang, 2018; Zhang & Lin, 2019). In other words, it behaves the same as the temporal noun phrase in (7a).

allowed, as in (6a-b) and (7a-b). If Mandarin speakers use bare noun phrases more often, they would give more atelic descriptions than English speakers. Specifically, the telic-atelic contrast in quantification of the affected object in English would be weakened or even neutralized in Mandarin.

### 1.3. The present study

Could language-specific telicity patterns affect the non-linguistic representations of event temporal structure? The answer to this question bears on efforts to disentangle language-specific from language-independent (potentially universal) components of event structure, and more generally connect language to event cognition. Prior work on event cognition raises the possibility that boundedness is grounded in largely universal conceptual representations (Ji & Papafragou, 2020a, 2020b, 2024; see also Filip, 1993; Folli & Harley, 2006; Kuhn et al., 2021; Strickland et al., 2015; Wagner et al., 2024). On this view, language-specific patterns of encoding telicity would be unlikely to influence the perception of temporal event structure in non-linguistic tasks (see similar arguments in the well-studied field of motion events, as in Gennari et al., 2002; Papafragou et al., 2008; Skordos et al., 2020). An alternative possibility is that the conceptual representation of temporal event structure depends on linguistic encoding and should differ across speakers of different languages (e.g., Flecken et al., 2015; Gerwien & von Stutterheim, 2018; Konishi et al., 2019; Sakarias & Flecken, 2019; Santin et al., 2021, among many others). According to this view, language-specific patterns of expressing telicity would have a global influence on the perception of event structure even when people are not explicitly using their native language.

In the current study, we focus on English and Mandarin speakers to test these two hypotheses. We first compare how English and Mandarin speakers describe events with and without an inherent endpoint (Experiment 1). Turning to non-linguistic event representations, in Experiment 2, we compare how viewers drawn from the two language groups distinguish between the bounded and the unbounded event categories by rating the likelihood that a specific dynamic occurrence in the world has an inherent endpoint. In Experiment 3, we further examine how speakers of the two languages perceive the temporal structure of bounded vs. unbounded events as they unfold.

## 2. Experiment 1

Experiment 1 asked how English and Mandarin native speakers described bounded and unbounded events. As already mentioned in the Introduction, the same episode can be potentially construed from both a bounded and an unbounded perspective. Nevertheless, observers are inclined to conceptualize their experience from one perspective rather than the other. Here we adopted event items that were known to be perceived by English-speaking adults as typical instances of one or the other boundedness category (Ji & Papafragou, 2020a). Of interest was whether Mandarin speakers would be more inclined than English speakers to use atelic verb phrases with bare noun forms for both event types.

### 2.1. Method

#### 2.1.1. Participants

Thirty native speakers of English (age range: 18–23.5;  $M_{age} = 19.2$ ) and 30 native speakers of Mandarin (age range: 18–23;  $M_{age} = 19.1$ ) participated in the experiment. The English-speaking participants were undergraduates at the University of Delaware. These participants reported that English was their only native language (i.e., the language they were exposed to from birth); some participants learned a foreign language such as Spanish at school. The Mandarin-speaking participants were students recruited from Beijing Institute of Technology. These participants were exposed to Mandarin from birth and some of them also

heard a Chinese dialect at home. The participants learned English at school from around age 6 but reported that they used Mandarin Chinese (and/or their Chinese dialect) in daily life and studies the vast majority of time. The sample size was chosen based on similar work on event description (e.g., Sakarias & Flecken, 2019; Santin et al., 2021; Únal et al., 2024). In this and the following experiments, all participants signed an informed consent form approved by the institutional review board of the corresponding universities and received course credit for participation.

### 2.1.2. Stimuli

We adopted the 20 pairs of videos showing bounded and unbounded events in Experiment 1 in Ji and Papafragou (2020a; see Table 1). All videos involved a girl who did an everyday action which began with the girl picking up an object or tool from a desk and ended in her putting down the object or tool. Paired videos had the same duration (4.4–12.0 s,  $M = 7.8$  s,  $SD = 2.4$ ). Inspired by the linguistic literature, we created the contrast between bounded and unbounded events through two factors: the nature of the action and the nature of the affected object. For half of the videos, paired bounded and unbounded events involved the same object(s) but differed in terms of the nature of the action: the bounded event displayed an action that caused a clear change of state in the object(s) (e.g., stack a deck of cards) while its unbounded counterpart did not involve such a change (e.g., shuffle a deck of cards; Fig. 1a-b). Within this subset of event pairs, five pairs included a single object (e.g., dress vs. pat a teddy bear) and the other five pairs included a mass quantity (e.g., stack vs. shuffle a deck of cards). For the other half of the videos, the bounded and unbounded events involved the same action but differed in terms of the nature of the affected object: the bounded event involved a single individual (e.g., draw a balloon) but its unbounded counterpart involved either an unspecified plurality of objects or a mass quantity (e.g., draw circles; Fig. 1c-d).

Two norming studies (with English speakers) showed that the bounded and the unbounded videos did not differ in the degree of intentionality or in the degree of visual similarity (Ji & Papafragou, 2020a). A third norming study (also with English speakers) elicited judgment about the temporal structure of the stimuli. Videos of bounded events were considered as “something with a beginning, midpoint and specific endpoint” 87 % of the time while videos of unbounded events were considered as such only 21.5 % of the time (a significant difference,  $t(39) = 20.33$ ,  $p < .0001$ ; Ji & Papafragou, 2022).

The 20 pairs of events were split into 2 lists, such that each list included only one member of each pair with boundedness and the source of boundedness counterbalanced. In each list, bounded and unbounded

events were intermixed such that items of the same event type could not appear successively more than 3 times. Two new lists were created where the order of the event items were reversed compared to the original 2 lists. In sum, 4 event lists were used in the description task.

### 2.1.3. Procedure

Participants were randomly assigned to one of the 4 event lists. They were instructed (in their native language) to watch each video and then to describe the video in a full sentence. Participants responded by writing down their sentences in an answer sheet.

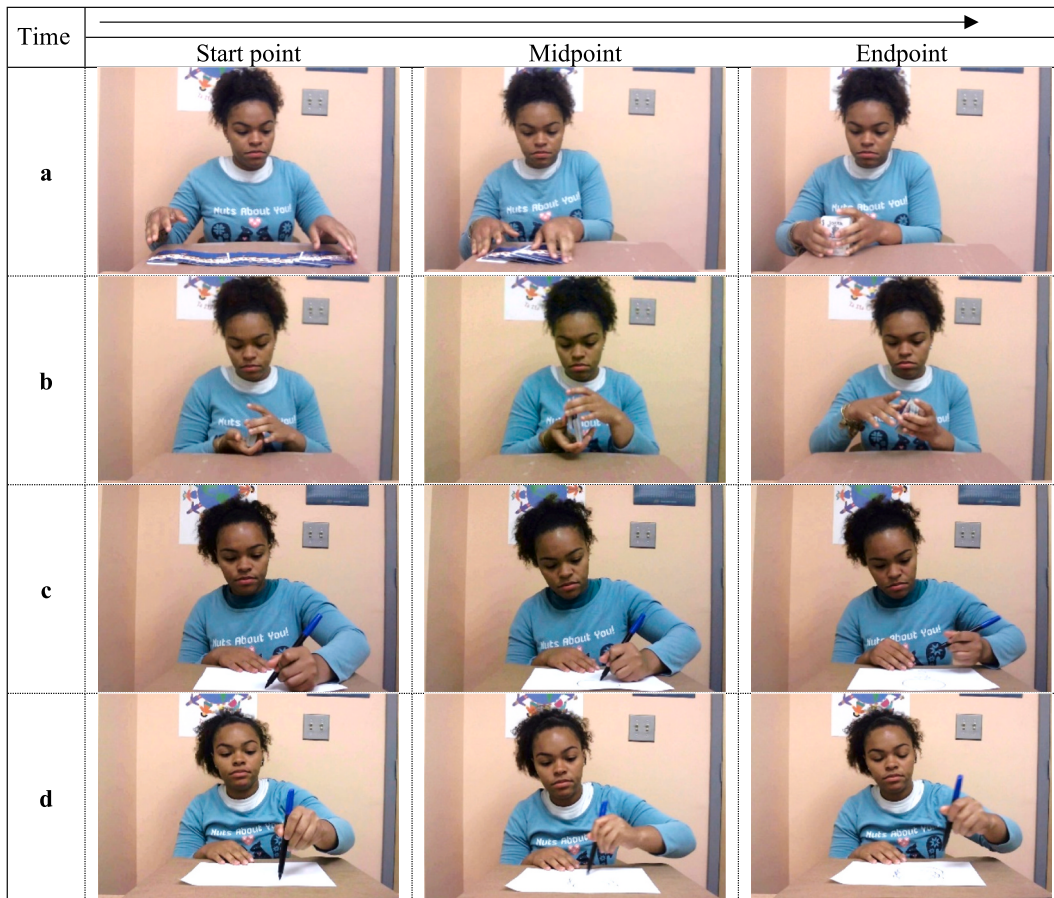
## 2.2. Results

For each language group, two RAs, both native speakers of that language, performed telicity diagnostics on the event descriptions. Specifically, when a verb phrase in an English description could be modified by *in five minutes* but not by *for five minutes*, it was judged as telic; when a verb phrase behaved in the opposite way, it was judged as atelic. As for Mandarin descriptions, a verb phrase was diagnosed as telic when a temporal noun phrase (e.g., *wu fenzhong* “five minutes”) could be added before the verb but not before the direct object in the phrase; by contrast, a verb phrase was diagnosed as atelic when a temporal noun phrase could only appear before the direct object. Overall, 95.7 % of English descriptions and 92.3 % of Mandarin descriptions received consistent telicity judgments. Discrepancies were then resolved by the two RAs and a third native speaker through discussion.

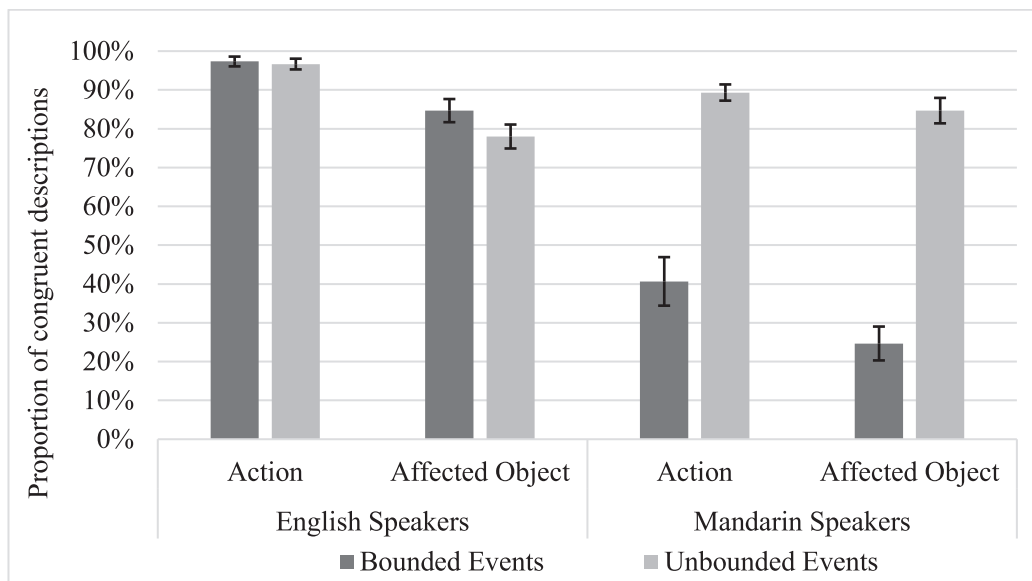
For bounded events, telic predicates were coded as congruent descriptions, i.e., descriptions that are congruent with the expectation that bounded events with an inherent endpoint elicit telic verb phrases; all other descriptions were coded as non-congruent descriptions. Similarly, for unbounded events, atelic predicates were coded as congruent descriptions, all others as non-congruent descriptions. We use the term “congruent” as a shorthand for whether the temporal information we are interested in was grammatically encoded at a certain level of precision within a sentence (obviously, from the perspective of language production, there is no sense in which a sentence has to “match” a scene in terms of the information it contains). The binary data (whether a description was a congruent or a non-congruent one) were analyzed using logit mixed-effects models. Random intercepts were provided for each Subject and each Item (random slopes often did not converge, Baayen et al., 2008; Barr, 2008; Barr et al., 2013). All models were fitted using the *glmer* function of the *lmer4* package in R. We examined the fixed effects of Language (English vs. Mandarin), Event Type (Bounded vs. Unbounded) and the Boundedness Source (Action vs. Affected

**Table 1**  
Event stimuli in Experiment 1.

Boundedness Source	No.	Bounded Events	Unbounded Events	Duration
Nature of Action	1	fold up a handkerchief	wave a handkerchief	8.00s
	2	put up one's hair	scratch one's hair	8.00s
	3	stack a deck of cards	shuffle a deck of cards	6.33s
	4	group pawns based on color	mix pawns of two colors	7.50s
	5	dress a teddy bear	pat a teddy bear	12.00s
	6	roll up a towel	twist a towel	7.50s
	7	fill a glass with milk	shake a bottle of milk	8.27s
	8	scoop up yogurt	stir yogurt	5.33s
	9	close a fan	use a fan for oneself	4.40s
	10	crack an egg	beat an egg	6.00s
	11	draw a balloon	draw circles	8.00s
	12	tie a knot	tie knots	7.00s
	13	eat a pretzel	eat cheerios	12.00s
	14	flip a postcard	flip pages	4.67s
Nature of Affected Object	15	peel a banana	crack peanuts	11.13s
	16	blow a balloon	blow bubbles	9.00s
	17	tear a paper towel	tear paper towels	8.00s
	18	paint a star	paint stuff	11.33s
	19	cut a piece of paper in half	cut ribbon from a roll	6.40s
	20	stick a sticker	stick stickers	4.67s



**Fig. 1.** Paired event stimuli in Experiment 1: (a) stack a deck of cards (bounded) vs. (b) shuffle a deck of cards (unbounded); and (c) draw a balloon (bounded) vs. (d) draw circles (unbounded).



**Fig. 2.** Proportion of congruent descriptions in Experiment 1. Error bars represent  $\pm$ SEM.

object). All of the factors were coded using centered contrasts (-0.5, 0.5). We built up our logit model in a bottom-up fashion, and only the interactions that significantly contributed to improving the model fit (assessed using chi-square tests on the log-likelihood values of competing models) were included in the final model. The same strategy

of model selection was applied to the analyses in the following experiments.

Results from Experiment 1 are shown in Fig. 2 and Table 2. The two language groups performed differently: overall, English speakers gave more congruent descriptions ( $M = 89.2\%$ ) than Mandarin speakers ( $M$

**Table 2**

Fixed effect estimates for multi-level model of the production of congruent descriptions in Experiment 1.

Effect	Estimate	SE	z value
(Intercept)	1.60	0.15	10.86***
Language (English vs. Mandarin)	-1.96	0.23	-8.46***
Event Type (Bounded vs. Unbounded)	1.19	0.18	6.63***
Source (Action vs. Affected Object)	-1.35	0.28	-4.88***
Language * Event Type	3.22	0.36	8.92***
Language * Source	1.40	0.43	3.28**

Notes. Formula in R: Target ~1 + (1|Subject) + (1|Item) + Language + Event Type + Source + Language: Event Type + Language: Source. Only the interactions that significantly improved the model fit were included in the final model and reported here.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

= 59.8 %) ( $\beta = -1.96$ ,  $z = -8.46$ ,  $p < .001$ ). Moreover, there were fewer congruent descriptions for bounded events ( $M = 61.8\%$ ) compared to unbounded events ( $M = 87.2\%$ ) ( $\beta = 1.19$ ,  $z = -6.63$ ,  $p < .001$ ). Furthermore, there was a significant interaction between event type and language group ( $\beta = 3.22$ ,  $z = 8.92$ ,  $p < .001$ ): Mandarin speakers, but not English speakers, had significantly fewer congruent descriptions for bounded than unbounded events (Mandarin speakers: odds ratio = 0.06,  $SE = 0.01$ ,  $p < .001$ ; English speakers, odds ratio = 1.52,  $SE = 0.42$ ,  $p = .132$ ). As expected, English speakers mostly produced telic predicates for bounded events ( $M = 91\%$ ), and atelic predicates for unbounded events ( $M = 87.3\%$ ). By contrast, Mandarin speakers gave telic descriptions for bounded events 32.7 % of the time, and atelic ones for unbounded events 87 % of the time.

The analysis also revealed a difference between the two sources of boundedness, with fewer congruent descriptions for videos where the bounded-unbounded contrast was instantiated in the affected object than the action ( $\beta = -1.35$ ,  $z = -4.88$ ,  $p < .001$ ). Furthermore, a significant interaction between boundedness source and language group was detected ( $\beta = 1.40$ ,  $z = 3.28$ ,  $p = .001$ ): the difference between the two sources of boundedness was greater among English speakers compared to Mandarin speakers (English speakers: odds ratio = 7.8,  $SE = 3.21$ ,  $p < .001$ ; Mandarin speakers, odds ratio = 1.92,  $SE = 0.53$ ,  $p = .017$ ). English speakers mostly produced congruent descriptions showing the bounded-unbounded contrast in the nature of the action (e.g., stack vs. shuffle a deck of cards,  $M = 97\%$ ); in comparison, fewer congruent descriptions were found in depicting the contrast in the quantification of the affected object (e.g., draw a balloon vs. circles;  $M = 81.3\%$ ). One reason for the latter fact could be that people sometimes specified the quantity of multiple objects involved in an unbounded event (e.g., “The girl cracked five peanuts.” “The girl stuck three stickers.”). Mandarin speakers, compared to English speakers, were less sensitive to both sources of boundedness; they overall used atelic phrases formed by mono-morphemic verbs and bare nouns for both bounded and unbounded events. Specifically, speakers sometimes used a mono-morphemic verb (e.g., *shou* “gather”) rather than an expected RVC (e.g., *shou-qi* “gather up”) to describe actions that led to a clear change of state in the affected object (e.g., the event of stacking a deck of cards;  $M = 59.3\%$ ). Meanwhile, bare nouns were predominant in Mandarin event descriptions ( $M = 92\%$ ). For instance, *tie tie-zhi* (“stick sticker”) was the most frequent (atelic) description for both the bounded event of sticking a sticker and its unbounded counterpart of sticking stickers.

### 2.3. Discussion

Experiment 1 confirmed the presence of cross-linguistic differences in the way event temporality is described. English speakers used telic descriptions for stimuli depicting bounded events and atelic descriptions for stimuli depicting unbounded events. In contrast, Mandarin speakers used atelic phrases for describing unbounded events but also often extended them to bounded events (with context presumably being

expected to resolve the underspecification of endpoints in those cases). Furthermore, this gap was more pronounced when the event endpoint was provided by the object in the videos that underwent a change. A major reason lies in the prevalence of bare nouns in Mandarin: for bounded events that involved a single object such as eating a pretzel, or drawing a balloon, Mandarin speakers simply used atelic phrases such as *chi bing-gan* “eat pretzel”, *hua qi-qiu* “draw balloon” that do not clearly specify the event endpoint. These data provide a starting point for investigating potential effects of cross-linguistic differences on boundedness construals in non-linguistic tasks with the same populations.

## 3. Experiment 2

Experiment 2 explored the extent to which viewers considered an action as a bounded or an unbounded event (i.e., with or without a natural endpoint). Inspired by a method used by Li and colleagues to assess object construals (Li et al., 2009; Experiment 3), we designed a 7-point scale measuring how likely English and Mandarin speakers were to construe the actions in the videos of Experiment 1 as a bounded or an unbounded event. If boundedness is a fundamental property of event construals, the two language groups should give similar ratings to the same event stimuli. By contrast, if the way speakers describe an event (and in particular choose between a telic or an atelic verb phrase) affects or even determines the way people construe the event, the two language groups should have different performance, with Mandarin speakers more likely to consider the events in Experiment 1 as unbounded compared to English speakers, in accordance with their linguistic encoding.

### 3.1. Method

#### 3.1.1. Participants

Forty native English speakers (age range: 18–22.5;  $M_{age} = 18.9$ ) and 40 native Mandarin speakers (age range: 18–19.3;  $M_{age} = 18.5$ ) participated in the experiment. The English-speaking participants were undergraduates from the University of Pennsylvania; the Mandarin-speaking participants were undergraduates recruited from the same university as in Experiment 1. The sample size was chosen based on the sample size in Experiment 3 in Li et al. (2009) and a previous related study (Experiment 1 in Ji & Papafragou, 2020a).

#### 3.1.2. Stimuli

The same 20 pairs of videos showing bounded and unbounded events were used as in Experiment 1. Paired events were arranged into 2 lists, such that each list included only one member of each pair. Boundedness (bounded vs. unbounded) and the source of boundedness (action vs. affected object) were counterbalanced. Within each list, bounded and unbounded events were intermixed such that items of the same event type did not appear successively more than 3 times. Two new lists were created where the event items were arranged in a different order compared to the original 2 lists. Thus a total of 4 event lists were used in the task.

#### 3.1.3. Procedure

Participants were randomly assigned to one of the 4 event lists. They were tested in groups of four to six in a lab room. They were first introduced to the notions of event vs. activity through a pair of examples. We chose to use these terms instead of bounded vs. unbounded event for two main reasons. First, “boundedness” as a term is unfamiliar to non-experts, unlike “event” and “activity” that can be used to capture people’s lay intuitions about the availability of a natural endpoint. Second, although both boundedness and telicity have been discussed in the linguistic literature in Mandarin, researchers have different translations of the two terms. By contrast, the Mandarin equivalents of “event” (i.e., *shi-jian*) and “activity” (i.e., *huo-dong*) are clearer and sound more natural (see OSF page for full materials in Mandarin).

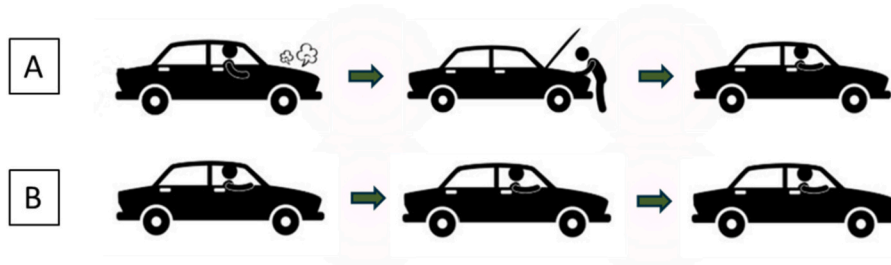


Fig. 3. Introductory images in Experiment 2: (A) “fix a car” and (B) “drive a car”.

Participants saw the examples in Fig. 3 along with a brief explanation of the task terminology in their native language. In English, the explanation was as follows: “The world is dynamic. We tend to think of our dynamic experience as either an event or an activity. For example, when we see the situation in A, we generally think of it as an event, fixing a car, because it has a natural endpoint. The event comes to an end when the car gets fixed and starts to work again. But when we see the situation in B, we generally think of it as an activity, driving a car, because it does not have a natural endpoint. The driver may stop driving at any moment (to take a break, to refuel the car, etc.). When the activity comes to an end is not specified.”

Participants were then introduced to the rating task as follows (English version): “Watch each video clip and rate your likelihood of seeing what is going on in the clip as an event or an activity on a 7-point scale. Circle 1 if you would definitely think of it as an event and 7 if you would definitely think of it as an activity.”

Each time, participants watched a video clip in the center of the screen. The 7-point scale with Event (value of 1) at one end and Activity (value of 7) at the other was below the video and stayed on screen; the same scale could be found on the answer sheet. After watching the video, participants responded by circling a number.

### 3.2. Results

Participants' rating scores of the videos were analyzed using multi-level mixed modeling with Language (English vs. Mandarin), Event Type (Bounded vs. Unbounded) and Boundedness Source (Action vs. Affected object) as the fixed factors, and crossed random intercepts for Subjects and Items. All the two-level categorical predictors were coded with centering contrasts (-0.5, 0.5).

Results are shown in Fig. 4 and Table 3. As expected, bounded events received significantly lower rating scores ( $M = 2.48$ ) than unbounded events ( $M = 5.00$ ) ( $\beta = 2.52, t = 26.95, p < .001$ ): bounded stimuli were

Table 3

Fixed effect estimates for multi-level model of the ratings of the videos in Experiment 2.

Effect	Estimate	SE	t value
(Intercept)	2.48	0.13	19.55***
Language (English vs. Mandarin)	-0.07	0.13	-0.58
Event Type (Bounded vs. Unbounded)	2.52	0.09	26.95***
Source (Action vs. Affected Object)	0.54	0.14	3.89***
Event Type * Source	-0.90	0.19	-4.79***

Note. Formula in R: Target  $\sim 1 + (1|Subject) + (1|Item) + Language + Event Type + Source + Event Type: Source$ .

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

more likely to be considered as events with a natural endpoint while unbounded stimuli were more likely to be considered as activities without a natural endpoint. Importantly, the two language groups did not differ in their rating scores (English speakers:  $M = 3.78$ ; Mandarin speakers:  $M = 3.71$ ;  $\beta = -0.07, t = -0.58, p > .250$ ). Overall, bounded-unbounded pairs showing the contrast in the nature of the action received lower ratings ( $M = 3.63$ ) compared to the pairs showing the contrast in the nature of the affected object ( $M = 3.86$ ) ( $\beta = 0.54, t = 3.89, p < .001$ ). A significant interaction between Event Type and Boundedness Source was detected ( $\beta = -0.90, t = -4.79, p < .001$ ): the difference between bounded and unbounded events was larger when they contrasted in the nature of the action (Bounded events:  $M = 2.14$ ; Unbounded events:  $M = 5.12$ ; odds ratio =  $-22.44, SE = 0.13, p < .001$ ) than the nature of the affected object (Bounded events:  $M = 2.82$ ; Unbounded events:  $M = 4.89$ ; odds ratio =  $-15.67, SE = 0.13, p < .001$ ).

### 3.3. Discussion

Despite the differences in encoding endpoints in their event

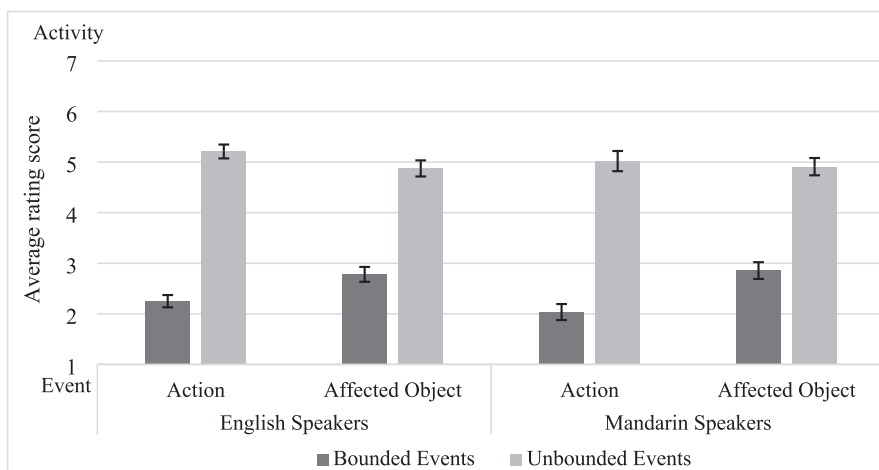


Fig. 4. Average rating scores in Experiment 2. Error bars represent  $\pm SEM$ .

descriptions (Experiment 1), English and Mandarin speakers in Experiment 2 did not differ in their categorization of temporal units as entities with or without inherent endpoints. The results suggest that boundedness is a shared event property, such that viewers speaking different languages arrive at similar judgments about how an event comes to an end. Closer inspection of the data reveals more specific similarities across members of the two language communities: people's rating for bounded vs. unbounded events differed to a greater extent when these events contrasted in the nature of the action (e.g., stack vs. shuffle a deck of cards) than in the nature of the affected object (e.g., draw a balloon vs. circles). One possible explanation is that, in the latter cases, the videos of unbounded events might have more readily afforded an alternative, bounded interpretation. In those videos (e.g., sticking stickers, eating cheerios, etc.), although a large number of objects were present (e.g., a roll of stickers, a large bag of cheerios), the number or the amount of objects involved in the event was often limited and thus the event might have been construed in some instances as bounded (e.g., the girl stuck three stickers; the girl ate a handful of cheerios). The fact that this pattern was similar in both language groups points to shared sensitivity to cues that contribute to the construal of (un)boundedness in the human mind.

Experiment 2 used explicit judgments to provide a first demonstration of how English and Mandarin speakers conceptualize events as temporal entities. An important question is whether the similarities observed would generalize to implicit measures of bounded vs. unbounded event perception. Experiment 3 addressed this question.

#### 4. Experiment 3

In Experiment 3, we examined how English and Mandarin native speakers tracked the temporal structure of events online. We adopted the break-detection task in Ji and Papafragou (2022). Participants watched short video clips, some of which included brief interruptions inserted either at event middles or close to event endings. Participants had to report whether or not they had seen an interruption after watching each clip. The task of detecting interruptions was irrelevant to the process of perceiving each event; nevertheless, previous results with English speakers (Ji & Papafragou, 2022) show that viewers exposed to bounded events were more likely to miss interruptions close to the endpoints compared to midpoints but viewers of unbounded events were equally likely to detect midpoint and late interruptions. These results show that viewers spontaneously track the temporal texture of unfolding events, even when not relevant for the task. In addition, because bounded events truly culminate, viewers' attention is drawn to the changes occurring towards the end of these events, making viewers more likely to miss an external disruption to the unfolding stimulus (e.g., a visual interruption) compared to other time points (such as the midpoints). This attentional asymmetry is not present in unbounded events, since such events are composed of less distinct and differentiated temporal slices.

If boundedness is part of foundational and largely universal event representations, Mandarin speakers should perform similarly to English speakers on the same task in Experiment 3. But if language-specific patterns of encoding telicity shape the way people compute boundedness, the two language groups should perform differently. In the limit, Mandarin speakers' accuracy in detecting interruptions should be similar for midpoints and endpoints regardless of event boundedness, since the lack of detailed telicity marking should blunt their attention to temporal differences within boundedness classes.

#### 4.1. Method

##### 4.1.1. Participants

Sixty-four native English speakers (age range: 18–24;  $M_{\text{age}} = 19.3$ ) and 64 native Mandarin speakers (age range: 18–23;  $M_{\text{age}} = 19.2$ ) participated in the experiment. Both language groups were recruited

from the same populations as in Experiment 2. The sample size was decided based on the calculated power of Experiment 2 in Ji and Papafragou (2022) that adopted the same design. The power analysis was based on the reported mixed-effects model (reported in Table 3 in Ji & Papafragou, 2022, p. 7) with the interaction between Event Type (Bounded vs. Unbounded) and Break Placement (Mid vs. Late) as the fixed predictor of interest, using the *simr* package in R (Green & MacLeod, 2016). The estimated effect size for the interaction was 1.99 and the power of this predictor was 92%, suggesting that 64 participants (per language group) were adequate to achieve a power of 0.80 at  $\alpha = 0.05$ .

##### 4.1.2. Stimuli

The video stimuli in Experiment 1 were edited in Corel VideoStudio to introduce a “break” of 0.03 s (i.e., 1 editing frame, with a display rate of 30 frames per second). The break consisted of a blurry picture created by applying an Iris Blur Effect in Adobe Photoshop to portions of the original video (see the examples in Fig. 5). Each video was edited twice. In the mid-break version, the break replaced the frame of the temporal midpoint of the event (e.g., in the video of stacking a deck of cards with 190 frames, the mid-break replaced the 96th frame). In the late-break version, the break began at the point that corresponded to 80% of the event (e.g., in the same video of stacking a deck of cards, the late-break replaced the 153rd frame). Edited videos with either a mid-break or a late break were used as test items, and their original versions were used as fillers.

The video stimuli of bounded events were arranged into 4 lists. The position of the break (mid vs. late) and the source of boundedness (action vs. affected object) in videos were counterbalanced across the lists. Each list began with a practice session including 4 videos. For this session, the first and third videos always had a mid-break and a late-break respectively and the other two videos did not include any break. The same 4 events were used as practice items for all 4 lists but each event appeared in the mid-break version in one list, in the late-break version in a second list, and as a filler without any break in the remaining two lists. The testing session of each list was composed of 8 test videos (4 with a mid-break, 4 with a late-break) and 8 fillers. Whether an event appeared as a test item or a filler was rotated across the lists. Unlike the practice session, the events were presented in the same order across the 4 lists. Therefore, the order between test items and fillers differed among the lists. In each list, test items and fillers were intermixed such that items of the same type could not appear successively more than 3 times. The stimuli of unbounded events were also arranged into 4 lists in the same way.

##### 4.1.3. Procedure

Participants were randomly assigned to one of two conditions depending on the event type that they were exposed to throughout the experiment (Bounded or Unbounded). Within each condition, they were randomly assigned to one of the 4 lists. Participants were tested in groups of four to six in a lab room. They were told (in their native language) to watch each video carefully and decide whether they saw a break in the video. They responded by circling either “Break”, or “No break” on an answer sheet. The text in the answer sheet was in their native language.

Participants were first given a practice session meant to illustrate what a break was. After each practice trial, participants noted their answer, and then the experimenter gave the correct answer. If participants were wrong, the video was played a second time. In the testing session, no feedback was given.

#### 4.2. Results

We coded “Break” responses to test items and “No break” responses to fillers as correct. The binary accuracy data were analyzed using logit mixed-effects models. Random intercepts were provided for each



**Fig. 5.** Examples of two versions of a bounded event (stack a deck of cards) and an unbounded event (shuffle a deck of cards) in Experiment 3: (a) mid-break and (b) late-break in the bounded event; (c) mid-break and (d) late-break in the unbounded event.

Subject and each Item. All the two-level categorical predictors were coded with centering contrasts  $(-0.5, 0.5)$ .

Performance on the filler items did not significantly differ between event types (Bounded events:  $M = 94.1\%$ ; Unbounded events:  $M = 93.8\%$ ;  $\beta = -0.05$ ,  $z = -0.19$ ,  $p > .250$ ), or between the two language groups (English speakers:  $M = 92.8\%$ ; Mandarin speakers:  $M = 95.1\%$ ;  $\beta = 0.42$ ,  $z = 1.44$ ,  $p = .151$ ). Turning to test items, Language (English vs. Mandarin), Event Type (Bounded vs. Unbounded), and Break Placement (Mid vs. Late) were included as predictors.<sup>5</sup> The interactions between the three variables were also examined and only the interactions that significantly improved the model fit were included in the final model.

As shown in Fig. 6 and Table 4, the two language groups did not differ (English speakers:  $M = 90.4\%$ ; Mandarin speakers:  $M = 93.2\%$ ;  $\beta = 0.37$ ,  $z = 1.39$ ,  $p = .192$ ). The difference between Bounded ( $M = 89.7\%$ ) and Unbounded event types ( $M = 94.0\%$ ) was also not significant ( $\beta = 0.45$ ,  $z = 1.51$ ,  $p = .130$ ). Overall, Break Placement affected results: participants were better at detecting mid-breaks ( $M = 94.1\%$ ) than late-breaks ( $M = 89.5\%$ ) ( $\beta = -0.53$ ,  $z = 2.06$ ,  $p = .039$ ). In addition, a significant interaction between Event Type and Break Placement was detected ( $\beta = 1.78$ ,  $z = 3.44$ ,  $p < .001$ ). Participants watching videos of bounded events were better at detecting mid-breaks ( $M = 95.3\%$ ) than late-breaks ( $M = 84.0\%$ ) (odds ratio = 4.16,  $SE = 1.46$ ,  $p < .001$ ); by contrast, participants watching videos of unbounded events did not differ in their detection of mid-breaks ( $M = 93.0\%$ ) and late-breaks ( $M = 94.9\%$ ) (odds ratio = 0.70,  $SE = 0.27$ ,  $p > .250$ ). Most importantly for present purposes, there was no additional interaction with Language.

#### 4.3. Discussion

In Experiment 3, no difference was found between English and Mandarin speakers' break-detection performance. In both language groups, people perceiving bounded events were more likely to miss a visual interruption when it occurred towards the end of an event, suggesting that more attention was drawn to the event culmination; however, people exposed to unbounded events could detect interruptions at

<sup>5</sup> Adding List, Gender, Boundedness Source (Action vs. Affected Object), or any interaction between Boundedness Source and other predictors to the model did not reliably improve model fit so we excluded these factors from further analysis.

midpoints and close to endpoints equally accurately as these events had largely similar temporal slices. The patterns replicated the results from Ji and Papafragou (2022) obtained from English speakers, and extend them to Mandarin speakers even though the latter tend to depict both event types using atelic verb phrases. Our results support the hypothesis that boundedness is a basic conceptual distinction that affects event perception in similar ways cross-linguistically.

## 5. General discussion

The human mind rapidly and seamlessly processes events, and yet human languages differ substantially in how they encode events and their boundaries. Do cross-linguistic differences affect event perception and conceptualization? This paper addressed this question in the domain of boundedness, a formal property concerning the presence of an inherent endpoint in the mental representation of an event that interfaces with the linguistic distinction of telicity (Ji & Papafragou, 2020). We began with evidence that English and Mandarin speakers differ in their use of telic vs. atelic phrases when describing events: compared with English speakers, Mandarin speakers tend to use more atelic phrases for bounded events as Mandarin allows bare noun forms (Experiment 1). Despite these differences, both groups make similar judgments about whether different dynamic occurrences in the world have clear endpoints or not (Experiment 2). Furthermore, the two language groups perceive bounded vs. unbounded events in similar ways: they track the internal temporal contour of these events and prioritize endpoints only for bounded events (that culminate) and not unbounded events (that simply stop; Experiment 3). Taken together, these findings provide evidence that event viewers from different language backgrounds are likely to share similar tendencies to compute event boundedness. This picture suggests that boundedness is a foundational property in the conceptual representation of events, and is independent of language-specific aspectual patterns in encoding event temporality.

Our data suggest that boundedness may serve as a conceptual basis for the linguistic notion of telicity. They also raise interesting questions about how shared representations of (un)boundedness interface with the need to acquire and use different lexical and/or grammatical means to encode telicity – not only in the two languages examined in the current work, but across diverse languages (Bar-El et al., 2005; Botne, 2003; Filip, 2004; Kardos, 2016; Singh, 1998; Soh & Kuo, 2005; Zhang, 2020). Although infants and children are sensitive to event endings (e.g., He &

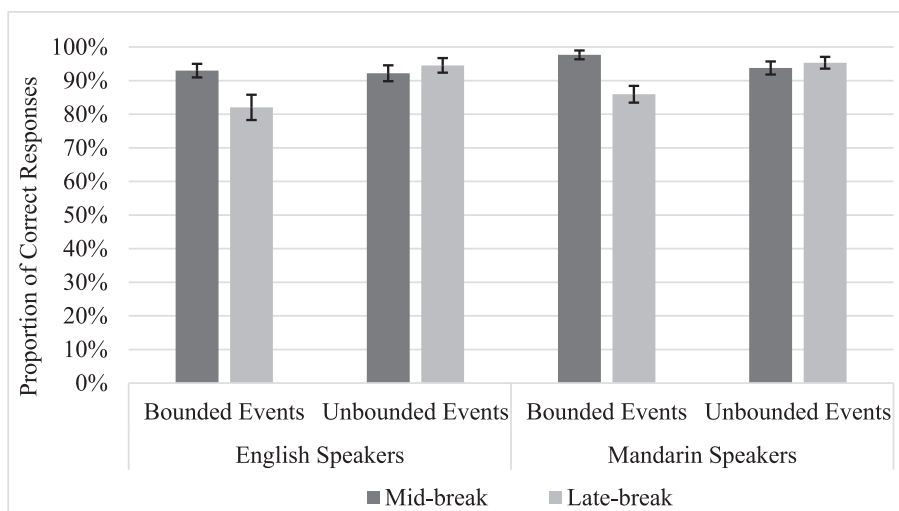


Fig. 6. Proportion of correct responses in Experiment 3. Error bars represent  $\pm$ SEM.

Table 4

Fixed effect estimates for the multi-level model of the accuracy of break detection in Experiment 3.

Effect	Estimate	SE	z value
(Intercept)	2.80	0.19	14.66***
Language (English vs. Mandarin)	0.37	0.28	1.39
Event Type (Bounded vs. Unbounded)	0.45	0.30	1.51
Break Placement (Mid vs. Late)	-0.53	0.26	-2.06*
Event Type * Break Placement	1.78	0.52	3.44***

Note. Formula in R: Accuracy  $\sim 1 + (1|\text{Subject}) + (1|\text{Item}) + \text{Language} + \text{Event Type} + \text{Break Placement} + \text{Event Type} : \text{Break Placement}$ .

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Arunachalam, 2023; Lakusta et al., 2017), and event boundedness (Ji & Papafragou, 2024), the literature on the acquisition of telicity reveals different developmental patterns triggered by linguistic factors such as the complexity of aspectual marking in a given language (see Martin et al., 2020 for a review). In other words, telicity involves mapping between language-specific properties and conceptualization of happenings in the world. Our work raises the question of how this mapping is implemented cross-linguistically.

Our findings contribute to efforts to disentangle language-specific from language-independent (potentially universal) components of event structure, and more generally connect language to event cognition (for different perspectives see Flecken et al., 2015; Gerwien & von Stutterheim, 2018; Konishi et al., 2019; Sakarias & Flecken, 2019, among many others). It is important to extend the current comparison between two languages to a broader range of language types to assess the generalizability of the current findings. We note that our event perception task was purposefully constructed to provide an implicit test of sensitivity to boundedness. As the task is in principle irrelevant to event processing per se, it makes it possible to measure spontaneous event processing (as opposed to strategic responding). As such, it lends itself to testing a variety of linguistic populations beyond the English-speaking participants typically included in past event cognition studies. While the current evidence is drawn from a comparison of two languages, the task itself provides a language-neutral tool for future studies to systematically investigate whether the observed similarities in boundedness perception hold across a wider sample of languages.

Our work raises several possibilities for future research. Recall that, inspired by the literature on compositional telicity, the current work implemented the bounded-unbounded contrast via either the nature of the action (e.g., open vs. push a box) or the nature of the affected object (e.g., eat an apple vs. apples). In both Experiments 1 and 2, a significant

interaction that involved Source of Boundedness (the nature of the action vs. the affected object) was detected, suggesting that the action may play a greater role in determining boundedness compared to the way an affected object delimits an event. In the context of the description task of Experiment 1, this finding is reminiscent of traditional views that regard telicity as a property of verbs (e.g., Vendler, 1957, see Filip, 2012 for a review) but treat the effects of object noun phrases on specifying event endpoints as malleable and dependent on subtle quantification choices (Ogiela, 2007; Ogiela et al., 2014; Xu & Schmitt, 2025). Similarly, within the non-linguistic rating task of Experiment 2, the effect of the affected object on event perception might be more flexible compared with that of the action itself. For instance, actions that lead to a natural or a self-evident endpoint (e.g., dress a teddy bear) or lack such an endpoint altogether (e.g., pat a teddy bear) might be good cues to (un) boundedness, even when the affected object is the same. However, affected objects by themselves might provide more ambiguous cues: for example, when a plurality of such objects is involved in the event as in some of our stimuli, whether the event is construed as bounded or unbounded depends on whether the plurality is construed as composed of discreet and countable entities (e.g., sticking three stickers, eating a handful of cheerios) or not (e.g., sticking stickers, eating cheerios). Future experiments could test the role of sources of boundedness more directly in the way events are described and categorized.

Finally, future work could ask whether the linguistic differences observed in Experiment 1 might shape other aspects of event cognition, including how events are remembered and reconstructed. Currently, the role of language in event memory is under active discussion (e.g., Chen et al., 2024; Engemann et al., 2015; Flecken et al., 2015; Santin et al., 2021; Trueswell & Papafragou, 2010). If language does modulate event memory, the predominant use of atelic phrases by Mandarin speakers might lead them to store more events as unbounded in memory. Consequently, compared to English speakers, who typically have more accurate memory for the culmination of bounded events over other event time points (Santin et al., 2021), Mandarin speakers might not show such a memory advantage (or might only show an attenuated advantage for the endpoints of bounded events). An alternative possibility is that, despite cross-linguistic encoding differences, event memory remains unaffected across English and Mandarin speakers.

## 6. Conclusion

The present study points to a dissociation between linguistic encoding and conceptual representation in event cognition. English and Mandarin speakers exhibit marked differences in how they aspectually

encode events, but their non-linguistic perception of event boundedness remains strikingly similar. The results align with the idea that boundedness is a core aspect of event cognition that interfaces with (but is independent from) language.

### CRedit authorship contribution statement

**Yue Ji:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Anna Papafragou:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

### Funding

This work was supported by the National Social Science Fund of China grant #21CYY046 (Y.J.) and National Science Foundation grant #2041171 (A.P.).

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

The authors thank Raevyn Johnson for helping construct the videos, and Xiaoto Liu for helping with collecting and entering data.

### Data availability

The materials, data and analysis code for the present study are available at the OSF platform and can be accessed at <https://osf.io/79k35/overview>.

### References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Baier, D., Choi, S., Goller, F., Nam, Y., & Ansorge, U. (2023). Does language rule perception? Testing a radical view of linguistic relativity. *Journal of Experimental Psychology: General*, 152(3), 794–824. <https://doi.org/10.1037/xge0001296>
- Bar-El, L., Davis, H., & Matthewson, L. (2005). On non-culminating accomplishments. In L. Bateman, & C. Ussery (Eds.), *Vol. 1. Proceedings of the 35th Annual Meeting of the North Eastern Linguistics Society* (pp. 87–102). GLSA.
- Barr, D. J. (2008). Analyzing ‘visual world’ eye-tracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474. <https://doi.org/10.1016/j.jml.2007.09.002>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Botne, R. (2003). To die across languages: Towards a typology of achievement verbs. *Linguistic Typology*, 7(2), 233–278. <https://doi.org/10.1515/lity.2003.016>
- Chen, P. (2014). Referentiality and definiteness in Chinese. In C.-T. Huang, Y.-H. Li, & A. Simpson (Eds.), *The handbook of Chinese linguistics* (pp. 404–413). Wiley Blackwell. <https://doi.org/10.1093/oxfordhb/9780199856336.013.0021>
- Chen, Y., Trueswell, J., & Papafragou, A. (2024). Sources and goals in memory and language: Fragility and robustness in event representation. *Journal of Memory and Language*, 135, Article 104475. <https://doi.org/10.1016/j.jml.2023.104475>
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, 339–405. <https://doi.org/10.1023/A:1008324218506>
- Comrie, B. (1976). *Aspect: An introduction to the study of verb aspect and related problems*. Cambridge University Press.
- Dowty, D. R. (1979). *Word meaning and Montague grammar*. Kluwer.
- Duan, W. (2019). *On telicity of eventualities in Mandarin Chinese* [Unpublished doctoral dissertation]. Zhejiang University.
- Engemann, H., Hendriks, H., Hickmann, M., Soroli, E., & Vincent, C. (2015). How language impacts memory of motion events in English and French. *Cognitive Processing*, 16(S1), 209–213. <https://doi.org/10.1007/s10339-015-0696-7>
- Ernst, T. (1987). Duration adverbials and Chinese phrase structure. *Journal of the Chinese Language Teachers Association*, 22, 1–11.
- Filip, H. (1993). *Aspect, situation types and nominal reference* [Unpublished doctoral dissertation]. University of California at Berkeley.
- Filip, H. (2004). The telicity parameter revisited. In R. Young (Ed.), *Proceedings of the 14th Semantics and Linguistic Theory Conference* (pp. 92–109). LSA.
- Filip, H. (2012). Lexical aspect. In R. I. Binnich (Ed.), *The Oxford handbook of tense and aspect* (pp. 721–751). Oxford University Press.
- Flecken, M., Gerwien, J., Carroll, M., & von Stutterheim, C. (2015). Analyzing gaze allocation during language planning: A cross-linguistic study on dynamic events. *Language and Cognition*, 7(1), 138–166. <https://doi.org/10.1017/langcog.2014.20>
- Folli, R., & Harley, H. (2006). What language says about the psychology of events. *Trends in Cognitive Science*, 10, 91–92. <https://doi.org/10.1016/j.tics.2006.01.002>
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79. [https://doi.org/10.1016/S0010-0277\(01\)00166-4](https://doi.org/10.1016/S0010-0277(01)00166-4)
- Gerwien, J., & von Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, 180, 225–237. <https://doi.org/10.1016/j.cognition.2018.07.008>
- Golinkoff, R. M., Katz, S., Jo, J., Singh, L., Collins, M. A., & Hirsh-Pasek, K. (2025). How the perception of events in children is influenced by language. *Cognition*, 259, Article 106123. <https://doi.org/10.1016/j.cognition.2025.106123>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493–498. <https://doi.org/10.1111/2041-210X.12504>
- He, X., & Arunachalam, S. (2023). Event end-state encoding in 13-month-olds – Completed and non-completed events are different. *Language and Cognition*, 1–15. Advance online publication <https://doi.org/10.1017/langcog.2023.54>
- van Hout, A. (2016). Lexical and grammatical aspect. In J. Lidz, W. Synder, & J. Pater (Eds.), *The Oxford handbook of developmental linguistics* (pp. 587–610). Oxford University Press.
- Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition*, 20(7), 848–864. <https://doi.org/10.1080/13506285.2012.705359>
- Ji, Y., & Papafragou, A. (2020a). Is there an end in sight? Viewers’ sensitivity to abstract event structure. *Cognition*, 197, Article 104197. <https://doi.org/10.1016/j.cognition.2020.104197>
- Ji, Y., & Papafragou, A. (2020b). Midpoints, endpoints and the cognitive structure of events. *Language, Cognition and Neuroscience*, 35(10), 1465–1479. <https://doi.org/10.1080/23273798.2020.1797839>
- Ji, Y., & Papafragou, A. (2022). Boundedness in event cognition: Viewers spontaneously represent the temporal texture of events. *Journal of Memory and Language*, 127, Article 104353. <https://doi.org/10.1016/j.jml.2022.104353>
- Ji, Y., & Papafragou, A. (2024). Boundedness supports children’s event representations. *Developmental Psychology*, 60(8), 1457–1473. <https://doi.org/10.1037/dev0001782>
- Kardos, É. (2016). Telicity marking in Hungarian. *Glossa: A Journal of General Linguistics*, 1(1), 1–37. <https://doi.org/10.5334/gjgl.52>
- Koenig, J.-P., & Chief, L.-C. (2008). Scalarly and state-changes in Mandarin (and other languages). In O. Bonami, & P. C. Hofherr (Eds.), *7. Empirical issues in syntax and semantics* (pp. 115–241). <http://www.cssp.cnrs.fr/eiss7/>
- Konishi, H., Brezack, N., & Golinkoff, R. M. (2019). Crossing to the other side: Language influences children’s perception of event components. *Cognition*, 190, Article 104020. <https://doi.org/10.1016/j.cognition.2019.104020>
- Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expression* (pp. 75–115). Foris Publications.
- Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar* (pp. 197–235). Kluwer.
- Kuhn, J., Geraci, C., Schlenker, P., & Strickland, B. (2021). Boundaries in space and time: Iconic biases across modalities. *Cognition*, 210, Article 104596. <https://doi.org/10.1016/j.cognition.2021.104596>
- Lakusta, L., Spinelli, D., & Garcia, K. (2017). The relationship between pre-verbal event representations and semantic structures: The case of goal and source paths. *Cognition*, 164, 174–187. <https://doi.org/10.1016/j.cognition.2017.04.003>
- Lee, S., Ji, Y., & Papafragou, A. (2024). Signatures of individuation across objects and events. *Journal of Experimental Psychology: General*, 153(8), 1997–2012. <https://doi.org/10.1037/xge0001581>
- Li, C. N., & Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology*, 58(4), 487–524. <https://doi.org/10.1016/j.cogpsych.2008.12.001>
- Lin, J. (2004). *Event structure and the encoding of arguments: The syntax of the mandarin and English verb phrase* [Unpublished doctoral dissertation]. MIT.
- Lin, J.-W. (2008). Event decomposition and the syntax and semantics of durative phrases in Chinese. In J. Döling, T. Heyde-Zybatow, & M. Schäfer (Eds.), *Event structures in linguistic form and interpretation* (pp. 31–54). Mouton de Gruyter.
- Martin, F., Demirdache, H., García del Real, I., van Hout, A., & Kazanina, N. (2020). Children’s non-adultlike interpretations of telic predicates across languages. *Linguistics*, 58(5), 1447–1500. <https://doi.org/10.1515/ling-2020-0182>
- Ogiela, D. A. (2007). *Development of telicity interpretation: Sensitivity to verb-type and determiner-type*. Michigan State University. Ph.D. Dissertation.
- Ogiela, D. A., Schmitt, C., & Casby, M. W. (2014). Interpretation of verb phrase telicity: Sensitivity to verb type and determiner type. *Journal of Speech, Language, and Hearing Research*, 57, 865–875. [https://doi.org/10.1044/2013\\_JSLHR-L12-0271](https://doi.org/10.1044/2013_JSLHR-L12-0271)
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184. <https://doi.org/10.1016/j.cognition.2008.02.007>

- Papenmeier, F., Brockhoff, A., & Huff, M. (2019). Filling the gap despite full attention: The role of fast backward inferences for event completion. *Cognitive Research: Principles and Implications*, 4, 1–17. <https://doi.org/10.1186/s41235-018-0151-2>
- Parsons, T. (1990). *Events in the semantics of English: A study in subatomic semantics*. MIT Press.
- Partee, B. H. (1999). Semantics. In R. A. Wilson, & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 739–742). MIT Press.
- Peck, J., Lin, J., & Sun, C. (2013). Aspectual classification of mandarin Chinese verbs: A perspective of scale structure. *Language and Linguistics*, 14(4), 663–700.
- Rothstein, S. (2004). *Structuring events*. Blackwell Publishing.
- Sakariäs, M., & Flecken, M. (2019). Keeping the result in sight and mind: General cognitive principles and language-specific influences in the perception and memory of resultative events. *Cognitive Science*, 43(1), 1–30. <https://doi.org/10.1111/cogs.12708>
- Santin, M., van Hout, A., & Flecken, M. (2021). Event endings in memory and language. *Language, Cognition and Neuroscience*, 36(5), 625–648. <https://doi.org/10.1080/23273798.2020.1868542>
- Singh, M. (1998). On the semantics of the perfective aspect. *Natural Language Semantics*, 6(2), 171–199.
- Skordos, D., Bunger, A., Richards, C., Selimis, S., Trueswell, J., & Papafragou, A. (2020). Motion verbs and memory for motion events. *Cognitive Neuropsychology*, 37(5–6), 254–270. <https://doi.org/10.1080/02643294.2019.1685480>
- Smith, C. S. (1991). *The parameter of aspect*. Kluwer.
- Smith, C. S. (1994). Aspectual viewpoint and situation type in Mandarin Chinese. *Journal of East Asian Linguistics*, 3, 107–146. <https://doi.org/10.1007/BF01736124>
- Soh, H. L., & Kuo, J. Y.-C. (2005). Perfective aspect and accomplishment situations in Mandarin Chinese. In H. Verkuyl, H. Swart, & A. van Hout (Eds.), *Perspectives on aspect* (pp. 199–216). Springer.
- Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences of the United States of America*, 112(19), 5968–5973. <https://doi.org/10.1073/pnas.1423080112>
- Sybesma, R. (1992). *Causatives and accomplishments: The case of Chinese Ba* [Unpublished doctoral dissertation]. HIL/Leiden University.
- Sybesma, R. (1997). Why Chinese verb-le is a resultative predicate. *Journal of East Asian Linguistics*, 6, 215–261. <https://doi.org/10.1023/A:1008238319773>
- Sybesma, R. (1999). *The mandarin VP*. Kluwer.
- Tai, J. H.-Y. (1984). Verbs and times in Chinese: Vendler's four categories. In D. Testen, V. Mishra, & J. Drogo (Eds.), *Papers from the parasession on lexical semantics* (pp. 289–296). Chicago Linguistics Society.
- Tenny, C. (1987). *Grammaticalizing aspect and affectedness* [unpublished doctoral dissertation]. MIT.
- Tham, S. W. (2014). Resultative verb compounds in Mandarin. In W. S.-Y. Wang, & C. Sun (Eds.), *The Oxford handbook of Chinese linguistics* (pp. 306–321). Oxford University Press.
- Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, 63(1), 64–82. <https://doi.org/10.1016/j.jml.2010.02.006>
- Ünal, E., Wilson, F., Trueswell, J., & Papafragou, A. (2024). Asymmetries in encoding event roles: Evidence from language and cognition. *Cognition*, 250, Article 105868. <https://doi.org/10.1016/j.cognition.2024.105868>
- Vendler, Z. (1957). Verbs and times. *Philosophical Review*, 66(2), 143–160. <https://doi.org/10.2307/2182371>
- Vurgun, U., Ji, Y., & Papafragou, A. (2024). Aspectual processing shifts visual event apprehension. *Cognitive Science*, 48(6), Article e13476. <https://doi.org/10.1111/cogs.13476>
- Wagner, L., Geraci, C., Kuhn, J., Davidson, K., & Strickland, B. (2024). Non-signing children's assessment of telicity in sign language. *Cognition*, 249, Article 105811. <https://doi.org/10.1016/j.cognition.2024.105811>
- Xu, J., & Schmitt, C. (2025). Demonstratives but not verbs cause non-culmination in mandarin incremental-theme predicates: Evidence from children and adults. In *Proceedings of the 42nd West Coast Conference on Formal Linguistics (WCCFL 42)*. Cascadia Press.
- Xuan, Y. (2017). Baoliu binyu lei baziju yu wanjie duanyu lilun [Ba-construction with retained-object and the telic phrase hypothesis theory]. *Yuyan Jiaoxue yu Yanjiu [Language Teaching and Linguistic Studies]*, 185, 28–39.
- Zacks, J., & 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>
- Zhang, A. (2018). *On non-culminating accomplishments in Mandarin* [Unpublished doctoral dissertation]. University of Chicago.
- Zhang, A. (2020). Referentiality, individuation and incomplete readings. *Journal of East Asian Linguistics*, 29, 435–468. <https://doi.org/10.1007/s10831-020-09217-5>
- Zhang, M., & Lin, Y. (2019). Xiandai hanyu feizhongjie wancheng qingzhuang yanjiu [on non-culminating accomplishments in Mandarin Chinese]. *Xiandai Waiyu [Modern Foreign Languages]*, 42(5), 586–596.