# Navigating the Event Boundaries: A Multimodal Task-based Exploration Among Chinese Speaking Young Adults

Shisen Yue

School of Foreign Languages, Shanghai Jiao Tong University

Prof. Haiyan Wang

# Abstract

An event boundary refers to the point in time or space where one event ends and another begins. The perception of event boundaries is an essential cognitive ability for humans as they form the building blocks of our understanding of the world. In most cases, a bounded event can be further segmented into many subordinate events, and they are represented with their main events in human cognition in a hierarchical structure. Human capacity in understanding the hierarchical structure of event boundaries and making proper segmentation of events reflects adequate cognitive ability. The present study designed multi-modal tasks to examine the perception of Chinese-speaking young adults to event boundaries. The result reflects the sensitivity of participants to event boundaries. The study also provides cross-linguistic evidence for the universality of event segmentation and boundary perception through comparisons with similar experiments on a different language basis.

*Keywords:* event boundaries, event segmentation, aspects, language and cognition

# Navigating the Event Boundaries: A Multimodal Task-based Exploration Among Chinese Speaking Young Adults

Shisen Yue

**Event Boundary**

Events are a fundamental aspect of human experience, as they form the building blocks of our understanding of the world. The perception of events is not simply the detection of static features in the environment, but rather an active and dynamic process of constructing and updating a representation of the ongoing event (Newtson, 1973). For example, when we perceive the event "He ate an apple", the demonstration of it emerges in our mind as a continuous action with a beginning and an ending point instead of a static picture. The ending of an event detaches the ending of the present event from the beginning of the next event. Such endings are explicitly demonstrated by changes in motion, sound, or object features, which are termed "event boundaries".

Although the term boundary represents the mutual discontinuity of the ending of an event and the beginning of the next event, the research on the ending of events outweighs their beginning in many aspects. For instance, the final destination of a moving ball has been shown to have more psychological weight than other psychological features (e.g. the moving direction of the ball) (He & Arunachalam, 2016; cf. also Sakarias & Flecken, 2019; Strickland & Keil, 2011). Also, the related findings suggest the goal of an event (as in the sentence I ran to school) is more accurately encoded in both language and memory as opposed to the source (as in the

sentence I came from home) (Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007).

From the perspective of human cognition, event boundaries are the product of people processing the perceived experience. This process is tightly associated with language. The presence of a connection between the linguistic and cognitive representation of event boundaries is proved by research. In a recent study (Strickland et al., 2015), individuals without knowledge of sign languages tended to associate signs with a visual boundary to verbs denoting bounded events, such as "decide," and signs lacking a visual boundary to verbs indicating unbounded events, such as "think." This pattern reflects the iconic manner in which sign languages typically represent boundedness (signs for bounded events feature a noticeable visual boundary while signs for unbounded events do not), implying that the iconic correspondence between a sign's meaning (i.e., boundedness) and form (visual boundary) is universally comprehensible to both signers and non-signers.

Moreover, different language expressions describing the same event may lead to significant cognitive differences. Consider the sentences:

(1)A girl played the piano.

(2)A girl played the Moonlight Sonata.

In this pair of sentence, (2) triggers a higher sense of the event boundary than (1) does in human cognition, though they might describe the same event. This phenomenon was observed and used to define the boundedness of an event.

Researchers define sentences like (2) as unbounded events and sentences like (1) as bounded events(Bach, 1986; Harley, 2003; Jackendoff, 1991).

The components of a sentence, such as verbs, objects and aspect markers contribute differently to building the cognitive image of an event. The combination of these different linguistic features influences the boundedness of an event in human cognition. Recent research has highlighted the dichotomy between two fundamental categories of objects: bounded and unbounded. Bounded objects are characterized by clearly defined boundaries that allow for their unambiguous identification and differentiation from other entities. An example of a bounded object is an apple, which possesses a well-defined contour that distinguishes it from other fruits. In contrast, unbounded objects exhibit undefined or flexible boundaries, rendering their exact limits difficult to establish. A prominent example of an unbounded object is time, which lacks a distinct form or boundary, yet can be measured and segmented into discrete units for practical purposes (Shen, 1995). This classification of objects into bounded and unbounded categories represents a valuable conceptual framework for understanding the nature of the world around us.

Numerous studies have investigated the correlation between bounded and unbounded objects and the count/mass noun distinction across a range of languages (Bale & Barner, 2009; Rothstein, 2010; Kulkarni, Shen, 1995; Srinivasan & Barner, 2020). Count nouns are considered bounded objects due to their countability, whereas mass nouns are categorized as unbounded objects. In the English language, for

instance, count nouns can be accompanied by an indefinite article or a plural marker, and can be directly modified by numerals, unlike mass nouns (Li & Thompson, 1989).

Mandarin Chinese, as a classifier language, exhibits a distinct mass-count syntax from number-marking languages. In contrast to languages with an article system or plural morphology, Mandarin lacks these features (Li & Thompson 1981), rendering the use of indefinite articles or plural morphology inadequate for distinguishing between count and mass nouns. Instead, Mandarin uses classifiers to modify numerals, as in "yi bei shui" (a bottle of water), to express the bounded situations of water. Generally speaking, Mandarin Chinese speakers adopt the structure [numerals + classifier + noun] to express a bounded situation of either bounded nouns or unbounded nouns, while the format [xuduo (much) + noun] serves to express the bounded situations.

Furthermore, in the temporal domain, actions can be classified as either bounded or unbounded. A bounded action has a discernible starting point and endpoint, such as the action of "making a decision" which ends when the decision is made, whereas an unbounded event has either no discernible starting point and endpoint or only one starting point without an endpoint, such as the sentence "Mary loves John" which is unknown when Many dislikes John (Shen, 1995). From a linguistic perspective, the verb's telicity is associated with temporal boundedness (Dowty, 2012; Tenny, 1994; Vendler, 1957). Telicity refers to a temporal property of verb meanings: telic verbs refer to events with a natural endpoint, whereas atelic verbs refer to events with no such endpoint. This characteristic is grammatically encoded in natural language. For

6

instance, Mandarin Chinese has telic verbs like dakai(open) and zhaodao(find) and atelic verbs like sikao (ponder), and zhidao (know).

In considering the combinations of verbs and aspect markers, atelic words are compatible in both bounded and unbounded situations, while telic words are not allowed to co-occur with imperfective aspect markers. As demonstrated in the following sentences, (5) and (6) respectively involve a perfective aspect marker and imperfective marker, and the atelic word kan le can combine with them to form bounded and unbounded events. However, the telic du wan is incompatible with the imperfective aspect marker. Thus it can only be included in a bounded event. The compatibility with the following aspect markers manifests an inclination to the bounded event contained in all telic words.

(5) "wo kan le yi ben shu."(I've read a book.) (bounded)

(6) "wo kan zhe yi ben shu." (I'm reading a book) (unbounded)

(7) "*wo du wan zhe yi ben shu." (*I'm finishing reading a book)

The studies on the influence of syntactic components serve as theoretical support to research young adults' perceptions of bounded events. Existing research has proved the sensitivity of young adults to event boundaries. In a recent study (Ji & Papafragou, 2020), the researchers investigated how event boundaries are defined and whether sensitivity to boundedness shapes the way events are processed. The study used a category identification task to show that English-speaking young adult participants are sensitive to boundedness and identify bounded events more easily than unbounded events. The study concludes that event cognition relies on highly abstract properties

of events and their boundaries, which have implications for how events are described, processed, and used to interact with the world.

Another study explores the sensitivity to event boundedness of Chinese-speaking children (Yang et al., 2018). The researchers investigated the perception of aspectual markers among 30-months Chinese-speaking children. The research adopted a preferential-looking paradigm. Specifically, The experiment presented the children with a choice between two video clips, one showing a closed event and the other showing an ongoing event, while playing an auditory stimulus of either the le sentence, the zhe sentence, or a control sentence without any aspect marker. The results indicated that the children showed sensitivity to the aspectual contrast between the le and the zhe, with an increase in looking towards the closed event when hearing the le sentence and an increase in looking towards the ongoing event when hearing the zhe sentence. The study also found that the effect of le on children's looking behavior was immediate, while the effect of zhe was late. The results suggest that Chinese children have an emerging sensitivity to aspectual distinctions early in their language development and prefer event boundaries, laying a solid foundation for the research of the event perception of Chinese people.

The experiment on Chinese children undoubtedly fills in the research gap, as most similar studies are conducted in English-speaking participants. Such a study of event boundary perception in individuals speaking languages with vastly different syntactic and morphological structures is crucial for the progress in finding the universality in human cognition. However, the literature on the perception of event

boundaries currently lacks direct tests of the sensitivity to event endings among Chinese-speaking young adults. Presumed to possess the highest level of cognitive abilities, Chinese-speaking young adults can serve as a baseline for studying event perception in Chinese.

**Event Segmentation**

Event segmentation is the process of identifying the boundaries between events in continuous streams of perception or experience. This process is a critical component of human cognition, as it allows individuals to organize and make sense of their experiences by breaking them down into meaningful units.

The research on event segmentation experienced a transfer from the linguistic field to the cognitive field. Early investigations of event perception have been primarily focused on the perception of speech, where researchers found that individuals segment speech into phonetic units (Bregman, 1978; Cutler & Norris, 1988). However, the understanding of event segmentation has since extended beyond the domain of language and into the realm of perception and cognition.

The formal notion of event segmentation in the cognitive field was first posited as individuals segmenting ongoing events into meaningful units based on perceptual cues in the environment (Zacks and Tversky, 2001). According to their theory, events are segmented into discrete units at event boundaries, which are characterized by a change in one or more perceptual features, including motion, sound, object features, and others. They suggested that event boundaries provide natural segmentation points,

where individuals can reorganize and consolidate their representations of ongoing

events.

Just as the perception of event boundaries, the connection between the linguistic

and cognitive representations of event segmentation is a valuable issue in identifying

the link between language and cognition. Numerous subsequent studies were

conducted to develop the event segmentation model and the influence of different

factors in people's segmentation((Zacks et al., 2007; Kurby & Zacks, 2008; Silbert et

al., 2014). The studies confirmed their theories with evidence in neural mechanisms.

For instance, it has been shown that the inferior parietal lobule, a region of the brain

involved in attention and spatial processing, is activated during event segmentation

(Kurby & Zacks, 2008). Additionally, the medial prefrontal cortex has been

implicated in encoding event boundaries and integrating information across

segmented events (Silbert et al., 2014). Thus it provides a feasible way to investigate

the event segmentation in the cognitive level by examining the participants'

performance in linguistic tasks.

The study of event segmentation also confirmed some conclusions drawn from

the research on event boundaries. Existing research found event segmentation is

another evidence of people's sensitivity to event boundaries. Specifically, as a whole

events contain many discontinuities reflecting changes in motion and object features,

the viewers of the event perceive these discontinuous points and regard them as the

new boundaries of subordinate events within the main event (Zacks et al., 2007). This

study demonstrated that people's sense of boundedness is hierarchical as they make

further segmentations based on the perception of the boundary of the main event. Their findings were confirmed by another study, which suggested that among all event fragments, individuals more accurately recall information presented at event boundaries compared to information presented within events, and that information presented at event boundaries is more easily incorporated into long-term memory (Kurby and Zacks, 2008).

In specifying the general approach humans segment a long event, researchers designed a video segmentation task to reflect how participants segment the event in their cognition (Zacks, 2007). With a well-designed grading system for participants' performance, the experiment found the hierarchical distribution of the event boundaries in people's cognition. Moreover, the study tested participants' other cognitive abilities such as working memory and learning abilities in subsequent experiments. The completion of these tests is largely associated with the videos they watch in the video segmentation task. For instance, they are asked to complete a jigsaw puzzle after they've watched and segmented the video demonstrating the way to complete the puzzle. The study reflects that participants with higher scores in the video segmentation task perform better in the subsequent tasks of working memory and learning abilities.

In spite of the experiment design and suggestive conclusions, the previous experiments were mostly English based. Also, the studies didn't provide enough characteristics of those segmented events. For example, are people segment the events based on some determined rules? just as the inclination to perceive an event in an

order of beginning-proceeding-ending sequence. It's significant to fill in these gaps, as similar evidence in other languages about event segmentation, especially in a language with different syntactic and morphological structures, can provide convincing evidence of the universality of human cognition.

Furthermore, probing into the rules of event segmentation helps researchers understand the sequence and combining way of events at the cognitive level.

**The Present Study**

The present study answers the below questions:

(1) How the Chinese-speaking young adults perceive the event boundaries?

(2) Is there any rules directing people's event segmentation?

To address the first question, we deliberately selected verbs and collocated noun phrases to form 15 bounded Chinese sentences. The telicity of verbs, aspect markers and object boundedness were carefully combined and tested to guarantee the consistency of boundedness in every sentence. The experiment is divided into two parts. First, we invited young adult participants to perform the actions in the sentences they hear. This step is to examine Chinese-speaking young adults' ability to understand the event boundedness through syntactic and morphological knowledge and present the actions in real person actions. Different from the studies simply involving the judgment of event boundaries, the performance task allows listeners to express their cognitive understanding of the events through body language. The disengagement of language production directly reflects their understanding of events at the cognitive level. Their performance was recorded for use in the second part of

this experiment. we designed a video description task for Chinese participants to obtain and analyze their descriptions of the actions in the videos. Their production of aspect markers is a precise indication of their perception of event boundaries, because of the rules of Chinese sentences expressing boundedness.

We designed an event segmentation task to answer the second question. A set of videos describing bounded events and containing subordinate events were shot. The appropriate segmentation can demonstrate their sharp sense of smaller boundaries within a continuous event. In order to produce more specific rules of the segmented event perception, the number of fragmented a video is fragmented it was fixed. Moreover, each participant was permitted to segment the event infinite times until they confirm their answers.

The two experiments are closely connected as they serve as a continuous and progressive approach toward human cognition from linguistic evidence. In particular, the first experiment confirms the sensitivity of Chinese-speaking young adults to event boundaries, laying the theoretical foundation for the second experiment, since the segmentation of events is a higher level of event boundary perception. Thus the second experiment is able to proceed into examining the specific rules characterized by the event fragments, to approach the more abstract study of human cognition.

The study aimed to confirm the connection between language and cognition by proving the features of a sentence are able to influence our cognition about the event expressed by the sentence. Besides, we aimed to examine the universality of human cognition by observing different language speakers' perceptions of event boundaries.

The cross-linguistic evidence is of much research importance because of the disparate syntactic and morphological rules in English and Chinese.

## Experiment I

### Participants

We invited 28 young adults aged 18 to 30 (Mage = 21.93, SD = 2.73; five males and 23 females) to participate in our description task. They are all intact in cognitive abilities and educated (with an education level of at least high school). Each participant received 5 Yuans reward after they finish all questions in the questionnaire of the description task.

### Materials

### *Sentences*

According to the Chinese syntactic theories and rules in expressing boundedness, we first selected 15 verbs that can be easily performed with actions and understood by viewers. Besides, the 15 sentences were all telic verbs in considering the role telicity plays in the perception of event boundedness. Then we chose an appropriate noun phrase after each telic verb. According to the previous literature, the format of the noun phrases was fixed as [one + classifier + noun]. Furthermore, The aspect marker *le* was added within each sentence to explicitly reflect the boundedness of the sentence. These steps are to unify the boundedness produced by the linguistic expressions. All 15 sentences are listed in the appendix.

### *Videos*

Fifteen event videos (length = 8.6m) were captured and are depicted in Figure 1. The process in which the videos were obtained was in Experiment 1with each video featuring a solo actor, except for the smoking events. The completed and ongoing videos all had the same beginning, where the action of picking up the target object was initiated. A completed video was considered concluded once the action was finished and the actor ceased looking at the camera, with the final frame displaying the completed action. On the other hand, the subject determined the conclusion of an ongoing video by analyzing whether the actor gazed at the camera and ceased the action. The completion of an action, such as washing and placing a bowl on the table, writing on paper, or finishing writing and laying down the pen, confirmed the conclusion of a video. In ongoing videos, the actor maintained continuous action while explicitly presenting the action's main idea, such as drawing a flower. This was done to lessen irrelevant variables and solely examine the subject's use of le and zhe, rather than guessing the action's name. To ensure consistency in the video length, all videos over 10 seconds were edited by removing the beginning, only starting from the action of picking up the object, deleting any redundant endings, and slightly increasing the speed in the middle, with a speed of 1.1-1.2 times the original speed. The increased speed sections did not create an apparent visual difference in speed when compared to the other videos.

***Questionnaires***

All the videos were compiled into an online questionnaire: a video-description questionnaire to examine participants' production. We designed the questionnaires using the website Wen Juan Xing: https://www.wjx.cn/newwjx.

The video-description questionnaire contained three parts: a personal-information registration section, example sentences (practice trail), and a description section (test trial). The first section was identical to the first section of the sentence video matching questionnaire. In the second part, we selected two unambiguous videos -- "ta dao le yi bei niunai" (he has poured a cup of milk) and "ta dao zhe yi bei niunai" (he is pouring a cup of milk)-- as examples, because the color of the object-milk was salient allowing participants to quickly perceive and differentiate the different endings of the two videos. In addition, we provided the corresponding sentences (with the aspect markers -le and -zhe) of these two videos and suggested that the participants could imitate the pattern of the sentences. The testing part asked participants to describe the videos and encouraged them to write sentences as examples.

To gather data for the study, the research team used online platforms to distribute questionnaires to young adults between the ages of 18 and 30. To ensure that the collected data was reliable, the team limited the distribution of the questionnaires to their own older relatives and to parents of children attending elementary school. Participants were compensated with a reward of 5 Yuan for completing any of the questionnaires. Additionally, participants were instructed not to disclose the content of the questionnaires to any third-party individuals.

**Production Coding and Data Analysis**

The participants' verbal responses were classified into six categories: Target Aspect Marker Production, Alternative Aspect Marker Production, Wrong Aspect Marker Production, Bare Verb, Irrelevant Sentences, and No Response. Target Aspect Marker Production indicated that participants correctly produced the aspect marker and placed it in the appropriate position, namely after the verb. Alternative Aspect Marker Production referred to participants describing whether the events were completed without using the target aspect markers. For -le1, alternative productions could be:

1)guo (usually occurs when describing a past experience)

2)-le2 (sentence-final le, reckoned as an informative aspect)

3)resultative adjectives, such as wan(finished), ganjin(clean), kai (open), which denote the completion of the event

4)Verb + -le + Verb (e.g., "xi le xi" (have washed it over and over again))

Bare Verb productions referred to participants using a bare verb without adding the corresponding aspect markers. Irrelevant Sentences were used to categorize responses that were not related to the events, such as "the orange is delicious. No Response was used for participants who did not answer. Data from one adult were excluded as they did not respond to any questions. All data were analyzed using the python library Pandas.

**Results**

According to the statistics, the average number of correct responses (the response that is classified as Target Aspect Marker production) among the selected adult subjects is 12.61 with the standard deviation of 4.03 and the average accuracy is 84%.

**Discussion**

The accuracy participants attained in the video description task demonstrates their sensitivity to event boundaries. This finding is consistent with previous research on event boundary perception, which suggests that people have a natural tendency to perceive and segment events in the world around them (Newtson, 1976; Zacks & Tversky, 2001). The present study provides further evidence that Chinese-speaking healthy young adults have the expected sensitivity to the event boundaries and can be involved in future studies of Chinese people's cognition as control groups possessing adequate cognitive abilities.

In addition, the result of the present study also shed light on the influence of syntactic expressions on sentence boundedness. This confirms the previous study(Tamly, 2000), which exhaustively discusses the different sentence components' contribution to the expression of event boundedness. The finding contributes to bridging the linguistic representation and people's cognition, demonstrating the close relationship and mutual impacts among human beings. Moreover, by comparing the results with that of the previous study (Kurby and Zacks, 2008), we found universality in the model of event perception across languages, especially between speakers of languages with distinguished grammatical systems (English and Chinese). This

finding makes further progress in understanding the universality of humans' cognitive abilities.

## Experiment II

**Participants**

We first invited 5 participants to perform the actions in the sentences they hear. We then recruited 10 participants from campus to segment the recorded videos. They were instructed to watch each video and make segmentations according to their cognition. Each participant received 5 Yuans as a reward after they finish the experiment.

**Materials**

we made 5 sentences that contain bounded events for the performance task and video segmentation task. Similar to the sentences prepared in Experiment 1, we selected actions according to their difficulty to act and explicitness to be perceived, and we balanced the influence of telicity and noun phrases to ensure each sentence expresses the same extent of boundedness. All 5 sentences are listed in the appendix.

**Procedure**

For each sentence in our preparation material, the experimenters read it to the subjects. Then after the signal of the beginning, the subjects began their performance. They were asked to perform the actions in the sentences they hear with the objects on the table. The videos ended when the actions are completed. The performer could give the experimenters a signal of completion through hints such as modifying their sitting

posture. The experimenters and participants repeated the above procedures until the performance for the last sentence was finished.

**Results**

We examined the recorded videos with the professional movie editor iMovie and found the actors in the videos demonstrated explicitly the end of the event with body language, such as modifying the sitting posture and raising the head to see the camera. We also observed that the actors slightly paused when the goal of the current motion changed. For example, in performing the action of drawing a flower, the actor first picked up the pen, then after a tiny pause he began drawing, and when he finished drawing paused again to put down the pen. We collected the pause positions and listed them in the appendix.

The 10 participants segmented the actions at almost the same points when pauses happened. The average segmenting point was recorded in the same file as the pause positions.

**Discussion**

According to the previous theory of event boundary perception (Tamly, 2000), The results that actors displayed apparent signals of completion manifests that they perceived the changes in the properties of action they were conducted. Specifically, they were able to perceive the state changes from an incomplete state to a completed state, the motion stopped and the goal turned off. Thus their reaction of displaying a signal of completion demonstrates their capacity of perceiving the event boundaries. This result indicates that Chinese-speaking young adults are sensitive to event

boundaries. Their cognitive abilities are adequate and are able to be involved as a control group in future studies with respect to event boundaries segmentation.

Moreover, The pauses in their actions also indicate their perception of sudden changes in physical attributes in the midst of continuous events. For example, in drawing a flower, their goals changed from picking up a pen to starting drawing, and their motion changed subsequently. It reflects their capacity in recognizing the hierarchical event distribution and the subordinate events under the main events. The pauses reflect the sensitivity to the boundaries of the subordinate events in their cognition, as also indicated in the previous studies (Zacks & Tversky, 2001; Zacks et al., 2007).

The results of the video segmentation task provide direct evidence of Chinese-speaking young adults' event perception. In particular, they observed the physical changes within the continuous events that were performed by actors and perceived the boundaries of subordinate events. This finding is consistent with the previous study(Ji & Papafragou, 2020). Besides, in analyzing the segmenting points drawn by participants, we found the viewers' divided each event in the format of "beginning--proceeding--ending". For example, picking up the pen and putting the pen on the table can be the beginning and ending parts of drawing a flower. In these processes, the actors prepare for the main events or clear up the materials used for the main events. Similarly, the process of cutting the meat also was also divided as beginning--picking up the kitchen knife, proceeding--cutting the meat, and ending--putting down the kitchen knife.

The process of segmenting the events was elaborately recorded and analyzed in the study of Kurby and Zacks (2008), and he reached his conclusion as all events have a hierarchical distribution and are naturally segmented into subordinate events in human cognition. The present study further explores the characteristics of the cognitive representations of event segmentation and specifies the possible paradigm for humans' processing of a continuous event.

### General Discussion

The present study proves that Chinese-speaking young adults are sensitive to event boundaries, which is represented not only in their perception of bounded events but also in the boundaries of subordinate events.

The study indicates the inextricable link between language and cognition, revealing their interdependence. In Experiment 1, the processing of the visual information and selecting the appropriate aspect marker le to express the aspectual features of the events demonstrate the conversion between cognitive signals and linguistic expressions. Specifically, the current study provides valuable cross-linguistic data that supports the continuity view of the relationship between language and cognition, which posits that language and cognition are closely intertwined and mutually influential (Bowerman & Choi, 2001).

Moreover, The present study supports that Chinese speakers share the same way to recognize and process bounded events as English speakers. Both of them perceive the events in a hierarchical structure and make segmentations in a similar approach. In

particular, the findings suggest that the ability to perceive event boundaries is a universal aspect of human cognition, regardless of linguistic and cultural differences.

## Conclusion

Our study captured video footage of bounded events, analyzed it, and invited participants to segment them. we explored the process of event segmentation, a cognitive mechanism that operates naturally and constantly in our minds. Also, We asked participants to describe the actions involved in the videos. We proved that Chinese-speaking young adults are able to discern the event boundaries and produce the correct aspect marker. Our findings contribute to the growing body of research on event perception in a cross-linguistic context, highlighting the importance of investigating how language and culture shape cognitive processes.

# References

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy, 9*, 5–16.

Baddeley, A. D. (1983) Working Memory. *Philosophical Transactions of the Royal Society of London. 302*(1), 311–324

Bale, A. C., & Barner, D. (2009). The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics, 26*(3), 217–252.

Barner, D., Wagner, L., & Snedeker, J. (2008). Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition, 106*, 805–832.

Caplan, D., & Waters, G. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory (Hove), 13*(3-4), 403-413.

Chen, L., & Durrleman, S. (2022). Comprehension of Mandarin Aspect Markers by Preschool Children With and Without Developmental Language Disorder. *Frontiers in Psychology*, *13*, 839951.

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese, 174*, 99–149.

Cho, S., Nevler, N., Shellikeri, S., Parjane, N., Irwin, D., Ryant, N., . . . Grossman, M. (2021). Lexical and Acoustic Characteristics of Young and Older Healthy Adults. *Journal of Speech, Language, and Hearing Research, 64*(2), 302-314.

Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: Reidel.

Erbaugh, M. (1982). Coming to Order: Natural Selection and the Origin of Syntax in the Mandarin-Speaking Child. Unpublished Ph.D. dissertation, The University of California, Berkeley, CA, USA.

Erbaugh, M. (1992). The acquisition of Mandarin. *The Crosslinguistic Study of Language Acquisition*. Edited by Dan Isaac Slobin. Hillsdale: Lawrence Erlbaum Associates, vol. 3, pp. 373–455.

Fyndanis, V., & Themistocleous, C. (2019). Are there prototypical associations between time frames and aspectual values? Evidence from Greek aphasia and healthy aging. *Clinical Linguistics & Phonetics, 33*(1-2), 191-217.

Fyndanis, V., Arcara, G., Christidou, P., & Caplan, D. (2018). Morphosyntactic Production and Verbal Working Memory: Evidence From Greek Aphasia and Healthy Aging. *Journal of Speech, Language, and Hearing Research, 61*(5), 1171-1187.

Grant, B. J., & Suddendorf, T. (2010). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly. 26*. 87-95.

Jackendoff, R. (1991). Parts and boundaries. *Cognition, 41*, 9–45.

Jin, L., & Hendriks, H. (2005). The development of aspect marking in L1 and L2 Chinese. *Working Papers in English and Applied Linguistics 9*: 69–99.

Kemper, S. (1986). Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics, 7*(3), 277-287.

Kemper, S. (1987). Syntactic complexity and elderly adults' prose recall. *Experimental Aging Research, 13*(1), 47-52.

Kemper, S., Herman, R., & Lian, C. (2003). Age Differences in Sentence Production. The Journals of Gerontology. Series B, *Psychological Sciences and Social Sciences, 58*(5), P260-268.

Kemper, S., Herman, R., & Liu, C. (2004). Sentence Production by Young and Older Adults in Controlled Contexts. The Journals of Gerontology. Series B, *Psychological Sciences and Social Sciences, 59*(5), P220-P224.

Kemper, S., Kynette, D., Rash, S., O'Brien, K., & Sprott, R. (1989). Life-span changes to adults' language: Effects of memory and genre. *Applied Psycholinguistics, 10*(1), 49-66.

Kuhn, J., Geraci, C., Schlenker, P., & Strickland, B. (2021). Boundaries in space and time: Iconic biases across modalities. *Cognition, 210,* 104596.

Kulkarni, R., Rothsein, S., & Treves, A. (2013). Statistical investigation into the cross- linguistic distribution of mass and count nouns: Morphosyntactic and semantic perspectives. *Biolinguistics, 7*, 132–168.

Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. University of California Press.

Li, H., Liang, L., & Wu, D. (2022). Predicting Chinese Preschoolers' Acquisition of Aspect Markers: A Corpus-Based Study. *Languages, 7*(2), 133. MDPI AG.

Li, P. (1990). Aspect and Aktionsart in Child Mandarin. Unpublished doctoral dissertation, University of Leiden.

Li, P., & Bowerman, M. (1998). The acquisition of lexical and grammatical aspect in Chinese. *First Language, 18*(54), 311–350.

Li, P., & Shirai, Y. (2000). *The Acquisition of Lexical and Grammatical Aspect*. New York: Mouton de Gruyter. 10.1515/9783110800715.

Lu Jianming. (2014). About "bounded and unbounded" theory and its application. *Linguistics Review* (02), 29-46 + 360-361.

Maguire, M. J., Brumberg, J., Ennis, M., & Shipley, T. F. (2011). Similarities in object and event segmentation: A geometric approach to event path segmentation. *Spatial Cognition and Computation, 11,* 254–279

McCormack, T. and Hoerl, C. (2008). Temporal Decentering and the Development of Temporal Concepts. *Language Learning, 58*: 89-113.

Peng, X., & Li, J. (2018). Case study on Chinese children's acquisition of sentences led by Le and Guo. *Journal of Xingtai University,* 33: 166–68.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition, 83,* 141–165.

Rothstein, S. (2010). Counting and the mass/count distinction. *Journal of Semantics, 27* (3), 343–397.

Saeed, J. (2016). Semantics (Fourth ed., Introducing linguistics).

Samuelson, L., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition, 73*, 1–33.

Sharon, T., & Wynn, K. (1998). Individuation of actions from continuous motion. *Psychological Science, 9*, 357–362

Srinivasan, M., & Barner, D. (2020). Lexical, syntactic, and pragmatic sources of countability: An experimental exploration of the mass-count distinction. To appear in. *Mass-Count in Linguistics, Philosophy, and Cognitive Science*. Amsterdam: Benjamins.

Tenny, C. (1994). *Aspectual roles and the syntax-semantic interface*. Dordrecht: Kluwer Academic.

van Hout, A. (2008). Acquiring perfectivity and telicity in Dutch, Italian and Polish. *Lingua, 118*, 1740–1765.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review, 66*(2), 143–160.

Vendler, Z. (1967). *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press

Wagner, L. (2001). Aspectual influences on early tense comprehension. *Journal of Child Language, 28*, 661–681.

Wagner, L., & Carey, S. (2003). Individuation of objects and events: A developmental study. *Cognition, 90,* 163–191.

Wellwood, A., Hespos, S. J., & Rips, L. (2018). The object:substance::event:process analogy. In T. Lombrozo, J. Knobe, & S. Nichols (Vol. Eds.). *Oxford Studies in Experimental Philosophy, 2,* 183–212.

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science, 8*, 173–181.

World Health Assembly, 69. (1 October 2022). Decade of Healthy Ageing: Plan of Action. World Health Organization. https://www.who.int/initiatives/decade-of-healthy-ageing

Yang, X., Shi, R., & Xu, K. (2018). Grammatical aspect in early child Mandarin: Evidence from a preferential looking experiment. *Journal of Psycholinguistic Research, 47*(6), 1301–1320.

Zhou, P., Crain, S., & Zhan, L. (2014). Grammatical aspect and event recognition in children's online sentence comprehension. *Cognition, 133*(1), 262–276.

龚千炎. (1995). 汉语的时相时制时态. 北京：商务印书馆.

刘一之. (2001). 北京话中的"着(zhe)"字新探. 北京：北京大学出版社.

沈家煊. (1995). "有界"与"无界". *中国语文, (5)*: 367-38.

戴耀晶. (1997). 现代汉语时体系统研究. 杭州：浙江教育出版社.

尚新. (2007). 英汉体范畴对比研究. 上海：上海人民出版社.

**Appendix**

**1. Language Material Samples**

| Verb | Bounded Event |
|------|---------------|
| *Bo* (peel) | *Ta bo le yi gen xiangjiao*<br>(He peeled a banana) |
| *Ca* (wipe) | *Ta ca le yi ge beizi*<br>(He wiped a cup) |
| *Chi* (eat) | *Ta chi le yi ban juzi*<br>(He ate a clove of orange) |
| *Chou*(smoke) | *Ta chou le yi gen yan*<br>(He smoked a cigarette) |
| *Da* (build) | *Ta da zhe yi zu ji mu*<br>(He built a set of blocks) |
| *Da* (pour) | *Ta dao zhe yi bei niunai*<br>(He poured a glass of milk) |
| *Feng* (sew) | *Ta feng le yi mei kouzi*<br>(He sewed a button) |
| *He* (drink) | *Ta he le yi bei cha*<br>(He drank a cup of tea) |
| *Hua* (draw) | *Ta hua le yi duo hua*<br>(He drew a flower) |
| *Jian* (cut) | *Ta jian le yi ge sanjiaoxing*<br>(He cut a triangle) |
| *Pin* (put) | *Ta pin le yi fu pintu*<br>(He played a puzzle) |
| *Qie* (cut) | *Ta qie le yi kuai rou*<br>(He cut a piece of meat) |
| *Xi* (wash) | *Ta xie le yi zhi wan*<br>(He washed a bowl) |
| *Xie* (write) | *Ta xie le yi ye zi*<br>(He has written in one page) |
| *Die* (fold) | *Ta die le yi jia zhifeiji*<br>(He folded a paper airplane) |

## 2. Video Samples


3.mov


5.mov


2.mp4


4.mov


1.mov


6.mov

## 3. Results for the event segmentation task

| 被试者序号 | 性别（M/F) | 年龄 | 受教育年龄 | 洗碗 | | 切肉 | | 写字 | | 画画 | | 切菜 | | 倒可乐 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 切分点1 | 切分点2 | 切分点1 | 切分点2 | 切分点1 | 切分点2 | 切分点1 | 切分点2 | 切分点1 | 切分点2 | 切分点1 | 切分点2 |
| 1 | M | 20 | 13 | 2190 | 6230 | 2270 | 8130 | 3000 | 6290 | 2260 | 7020 | 3000 | 6290 | 2200 | 6240 |
| 2 | M | 20 | 13 | 3010 | 7170 | 3250 | 9070 | 1160 | 7210 | 3030 | 8100 | 3250 | 8130 | 1240 | 7090 |
| 3 | M | 20 | 13 | 1280 | 7250 | 2280 | 8030 | 1260 | 7170 | 2260 | 8240 | 1250 | 7090 | 3010 | 7090 |
| 4 | M | 20 | 13 | 3000 | 7210 | 3070 | 8140 | 1150 | 7030 | 3020 | 8060 | 2190 | 6270 | 2200 | 9050 |
| 5 | M | 20 | 13 | 2190 | 6230 | 3070 | 7220 | 1160 | 5280 | 3120 | 8120 | 2130 | 4180 | 3050 | 6220 |
| 6 | M | 20 | 13 | 2190 | 5280 | 2050 | 8280 | 280 | 6150 | 2200 | 8030 | 1200 | 6190 | 3140 | 7030 |
| 7 | M | 20 | 13 | 1160 | 6040 | 3240 | 7240 | 2290 | 5270 | 4000 | 7280 | 2280 | 6270 | 270 | 7060 |
| 8 | M | 20 | 13 | 3030 | 9150 | 3030 | 9150 | 1240 | 6090 | 6190 | 7130 | 3210 | 5100 | 3040 | 8230 |
| 9 | M | 20 | 13 | 2020 | 6040 | 1190 | 7260 | 1150 | 6020 | 2090 | 7230 | 1290 | 6270 | 2150 | 7080 |
| 10 | M | 20 | 13 | 3000 | 5110 | 3010 | 6080 | 1200 | 5270 | 3020 | 7180 | 3200 | 5060 | 3040 | 6280 |
| 11 | | | | | | | | | | | | | | | |
| 12 | | | | | | | | | | | | | | | |
| 13 | | | | | | | | | | | | | | | |
| 14 | | | | | | | | | | | | | | | |
| 15 | | | | | | | | | | | | | | | |
| 16 | | | | | | | | | | | | | | | |
| 17 | | | | | | | | | | | | | | | |
| 18 | | | | | | | | | | | | | | | |
| 19 | | | | | | | | | | | | | | | |
| 20 | | | | | | | | | | | | | | | |
| 平均时点 | | | | 2307 | 6571 | 2646 | 7860 | 1389 | 6178 | 3119 | 7639 | 2300 | 6085 | 2334 | 7137 |
| 标准差 | | | | 704.841 | 1177.43 | 675.577 | 931.379 | 740.172 | 765.678 | 1229.9 | 503.796 | 843.577 | 1105.42 | 951.189 | 898.097 |
| 停顿位置 | | | | 2050 | 6100 | 2300 | 7450 | 1300 | 6000 | 3150 | 7400 | 2200 | 6000 | 2000 | 7150 |

## 4. Questionnaire Link: https://www.wjx.cn/vm/eypc4Kx.aspx#