Chapter 4. Chinese Grammatical Aspect and Co-speech Gesture in Different Types of Events

4.1 Introduction

In everyday conversations, we often talk about events we can neither witness nor experience. Whether exclaiming about the passage of time, describing changes of emotional states, or discussing changes in the stock market, we have to conceptualize and describe such events that we cannot directly perceive or perform. Then how can we communicate about those events to which we have no direct access without physical experience? The embodiment approach to language and cognition provides an answer to this question. That is, people understand and experience abstract concepts in terms of richer and more experience-based concrete concepts (Boroditsky & Ramscar 2002; Gibbs 1996; Lakoff & Johnson 1980). The way people talk about concrete and abstract events can provide some evidence for this proposal. The same verb can be used literally and metaphorically to describe both concrete and abstract domains. Action verbs in Mandarin Chinese provide a case in point. In their literal uses, they describe the actual action of entities, as in “tā zài páo” (She is running), “zhàdàn zhà le” (The bomb exploded), and “shuǐ miàn zài zhǎng” (The water level is increasing). In their metaphorical uses, which are rich in everyday speech, action verbs are descriptive of time, emotions, change, and other abstract domains, as in “shí jiān pǎo dé hǎo kuài” (Time is running very quickly), “tā yīxià zhà le” (She exploded at once), and “gūpiào zài zhǎng” (The stock is increasing). Obviously, these abstract concepts are grounded metaphorically in embodied and situated knowledge. More and more cognitive linguists and psychologists are providing evidence that metaphor is not only a matter of language but also a matter of thought and action (Boroditsky & Ramscar 2002; Gibbs 2006; Lakoff & Johnson 1980). Consequently, metaphor is realized not only by language but
also by gestures (Cienki 1998a, b; Cienki & Müller 2008a, b). Furthermore, the same event can be construed in different ways and an event’s objective properties are insufficient to predict its construal (Langacker 1991a, 2008a). In viewing and describing a situation, people may choose either the ongoing intermediate phase or the completion of the entire event to focus on, which can be realized in relevant grammatical aspect forms in language (Comrie 1976; Langacker 1987, 2008a).

Previous studies demonstrated that the basic aspectual distinctions between the imperfective aspect and the perfective aspect modulate how people conceptualize events. Events described in imperfective aspect (realized by be...-ing in English; the progressive aspect particle zài and durative aspect particle zhe in Mandarin) stimulate the comprehenders to focus more on the middle portion of a path or dynamic nucleus of motion (Anderson et al. 2008; Anderson et al. 2010; Athanasopoulos & Bylund 2013a; Madden & Zwaan 2003), to yield greater facilitation of the manual ACE or LCE (Bergen & Wheeler 2010; Liu & Bergen 2016), to infer more action in a given period (Fausey & Matlock 2011; Matlock 2011), and to produce more complex and longer co-speech gestures (Duncan 2002; Parrill et al. 2013) than the perfective aspect (realized by the simple past or perfect in English; by the actualized aspect marker -le and the experiential aspect marker -guo in Mandarin). The results can be interpreted as more mental simulation of the internal structure of events involved in the imperfective aspect. All of the above studies compared the concrete (motion) events in the imperfective aspect description and the perfective aspect in artificial tasks. It is unclear how well the findings could be generalized to real-world natural conversation situations, which are replete with metaphoric, abstract, as well as concrete events. Whether, and to what extent, abstract concepts, including grammar constructions like grammatical aspect, are also grounded in action and perception is a critical, but relatively less explored question in embodied cognition theories, particularly from the perspective of language production. Co-speech gestures as visible components of utterances provide the most direct window for us to see how people mentally represent the concrete, metaphoric, and inherently abstract events. According to the GSA framework (Hostetter & Alibali 2008, 2010), gestures arise when speakers activate representations during speech
production that evoke mental simulations of spatial and motor events. The hypothesis was based on physical manual actions without taking metaphorical and inherently abstract actions into consideration. Therefore, this study aims at investigating how Mandarin Chinese speakers conceptualize the concrete, metaphoric and abstract events in speech with grammatical aspect markers and in co-speech gestures in natural conversations.

4.2 Mental simulation in literal, metaphoric and abstract language processing

Studies of the activation of the motor system in literal, metaphoric and abstract action language comprehension using neurophysiological techniques have yielded somewhat inconsistent results due to different techniques and tasks. For instance, Aziz-Zadeh et al. (2006) used functional magnetic resonance imaging (fMRI) to test the activation of areas in the human premotor cortex that is responsible for executing and observing actions when people watch actions and read (literal or metaphorical) phrases relating to three effectors – foot, hand, or mouth actions. They found effector-specific activations in the premotor cortex of the left hemisphere for literal action (e.g., biting the peach, grasping the scissors, pressing the car brake), but not for metaphorical action (e.g., chewing over the details, grasping the idea, kicking off the year). Similarly, through a set of fMRI studies, Raposo et al. (2009) found significant activation in motor and premotor regions when participants listened to action verbs in isolation (e.g., kick) and to a lesser extent in literal sentences (e.g. kick the ball), but not in idiomatic sentences (e.g. kick the bucket). In contrast, Boulenger et al. (2012) used Magnetoencephalograph (MEG) and found somatotopic activation (the activation of motor-related areas in the brain) both in silent reading of literal sentences (Pablo kicked the ball) and idiomatic sentences (Pablo kicked the bucket). But none of the above studies considered inherently abstract events. Desai et al. (2011) filled this gap. They used fMRI studies to compare the neural responses to literal action (e.g. The daughter grasped the flowers),
metaphoric action (e.g. *The public grasped the idea*), and abstract action (e.g. *The public understand the idea*) sentences of varying familiarity. The results showed that the left anterior inferior partial lobe (aIPL) was activated both in literal and metaphoric contexts, that the left superior temporal areas were activated in both metaphoric and abstract contexts, and that the right aIPL regions associated with abstract language was also activated with the metaphoric sentences. Importantly, the more familiar and conventionalized the literal and metaphoric sentences are, the less the primary motor and biological motion perception areas were activated. Thus, understanding metaphoric language involves not only the process of understanding literal language but also that of abstract language. Relatively few studies explored the involvement of the motor system during language production. Morsella & Krauss (2005) employed Electromyography (EMG) to measure electro-muscular activity in participants’ hands and arms when they retrieved difficult lexical items. The results showed that retrieving concrete words led to more activity than retrieving abstract words. They interpreted this as evidence that motor activation is involved in language production.

Some behavioral studies found that mental simulation happens only in response to literal language but not in metaphoric and abstract language understanding. For example, Bergen et al. (2007) used a visual object categorization task to investigate whether language for concrete, metaphorical and abstract upward or downward motion (e.g. “The mule climbed”, “Oil prices climbed above $ 51 per barrel”, “The price multiplied”) interferes with visual processing in the same area of the visual part on the computer screen. They found that both the up- or down-associated nouns and verbs in concrete motion events, rather than in metaphoric and abstract motion events, activated the mental simulation of visual imagery.

Both the neuropsychological studies and behavioral studies reviewed above demonstrated that mental simulation is affected not simply by words but also by the sentential contexts where the concrete, metaphoric and abstract language are embedded. Grammar, as the other end of a lexico-grammar continuum (Langacker 1987; Talmy 2000), must also have an influence on mental simulation. As one of the most important grammatical devices in the world’s languages, grammatical aspect has recently gained
great attention in this field. For example, Liu & Bergen (2016) carried out a location-sentence compatibility experiment to investigate in what conditions (progressive aspect vs. perfect aspect (i.e., non-progressive in English); concrete events vs. abstract events) language comprehenders mentally simulate the ending location of a motion event. They found that the progressive aspect in both the concrete and abstract events facilitates mental simulation not simply of the central action but also of the ending location of the event, whereas the perfect aspect description of events did not elicit simulation of locations. Furthermore, the concreteness of events led to different simulation effects of locations even within the progressive aspect context: the concrete events showed a Location-sentence compatibility effect, whereas the abstract events showed an incompatibility effect. Thus, they conclude that mental simulation of locations must be different between concrete vs. abstract events.

Based on the findings surveyed above, this study intends to explore the following two questions.

1) Is the internal structure of an event more highlighted by the gestures accompanying imperfective aspect utterances than perfective aspect utterances in reference to concrete, metaphoric, and inherently abstract events? If the metaphoric and inherently abstract description of events simulates visual and motor images in the same way as the description of concrete events does, then imperfective aspect utterances should be accompanied by more iconic gestures than perfective aspect utterances in reference to metaphoric and inherently abstract events, as it does in reference to concrete events.

2) Are gestures accompanying imperfective and perfective aspect speech in concrete, metaphoric and inherently abstract event descriptions iconic for similar or different event semantic elements? If metaphoric and abstract events yield mental simulation similar to that performed in producing concrete events, then this would suggest that grammatical elements such as imperfective aspect and perfective aspect have the same effect on the three sorts of language.
4.3 Methods

4.3.1 Data

The multimodal data used for this study were taken from 20 videos from *Yang Lan One on One*. Each video lasts for about 20 minutes, and 8.5 hours long in total. Our corpus consists of 30 native Chinese speakers (10 females). There were 342 grammatical aspect marked utterances which were accompanied by gestures.

4.3.2 Annotation

I annotated the linguistic features and the gestural features of the 342 multimodal utterances with aspect markers.

4.3.2.1 Speech

In terms of speech, I encoded the utterance as either perfective aspect or imperfective aspect. In Mandarin, the perfective aspect is mainly realized by the particles *le* and *guo*. The particles -*le* attached to verbs usually in the middle, sometimes also in the end of an utterance, is used to express the actuality of an event, whereas -*guo* attached to verbs indicates the experientiality of an event. The imperfective aspect is mainly realized by the particles *zài* and *zhe*. The particle *zài* in front of verbs encodes an ongoing event, and -*zhe* attached to verbs focuses on the durativity of an event.

Then each aspect marked utterance was categorized into three types of events: concrete, metaphoric or abstract. A concrete event used a concrete action verb (e.g., *dā kāi chuāng hù* “open a window”, *bào jīqì* “hold a machine”) to depict a physical action. A metaphoric event used a concrete action verb in a figurative manner (e.g. *dā kāi shìyé* “broaden the vision”, *bào zhe xī wàng* “hold the hope”), so that no physical action was involved. An abstract event used an inherently abstract verb (e.g. *xiǎng* “think”, *dān xīn* “worry”) to describe the inner psychological state or a highly schematic verb (e.g. *zuò* “do”, *gōngzuò* “work”) to over-generalize physical actions.
One coder coded the utterance aspect for the whole dataset, and a second coded 50% of the dataset. The inter-rater agreement for the perfective and imperfective aspect coding was 98%, and for the concrete, metaphoric, and abstract events coding was 96%.

4.3.2.2 Gestures

Gestures were coded as iconic and non-iconic gestures. Iconic gestures referred to hand/arm movements that enacted the semantic content of the accompanying speech, including concrete iconic (i.e. depicting the shape, size and action of concrete events) and metaphoric iconic gestures (i.e. depicting the shape, size and action of abstract events). Non-iconic gestures were gestures that represented no semantic information of the speech itself, including the deictic gestures (i.e. pointing to concrete or abstract entities, locations etc. in the gesture space), and pragmatic gestures (i.e. hand movements contribute to the pragmatics of the utterance meaning, they may function as performative, modal, or parsing gestures).

The iconic gestures were further coded in terms of their iconicity for relevant semantic elements as follows. (1) Iconic for entities: gestures depicted the shape, size or other properties of the agent, patient or settings of an event. (2) Iconic for action: gestures reenacted out the movements of an event. (3) Iconic for path: gestures traced the path left by a motion event. (4) Iconic for process itself: gestures depicted the schematic processing of an event. For example, while uttering “she’s sitting there worrying”, the speaker rotated her two hands around one another alternately to indicate the dynamic duration of this abstract event (Duncan 2002: 202). However, these semantic elements in gestures were not mutually exclusive. Thus, I added the item (5) iconic for combination. For example, when uttering “He was writing”, the speaker mimicked the action of writing by moving the right hand horizontally from the left to right, and the handshape in which the right index finger touched upon the thumb and other fingers curled in looked like the head of a pen. Thus, this gesture is iconic for the action of “writing” and the entity “pen” at the same time.

The author of this thesis coded the entire gesture dataset, and a second coded 25%
of the dataset. The inter-rater agreement was good: 96% for gesture type; 94% for gestural iconicity relating to semantic elements of speech.

4.4 Results

4.4.1 Concrete events

As Table 4.1 shows, the utterances with imperfective aspect were accompanied by more iconic gestures than those in the perfective speech contexts (72.7% vs. 51.6%). When imperfective aspect was used, it co-occurred far more frequently iconic gestures than non-iconic gestures (72.7% vs. 27.3%). The differences were statistically significant enough ($\chi^2(1) = 4.796$, $p < 0.05$). The results suggest that the speakers taking the imperfective aspect to construe events were more prone to mentally simulate the internal structures of concrete events than those taking the perfective aspect.

<table>
<thead>
<tr>
<th>Grammatical aspect</th>
<th>Gesture type</th>
<th>Iconic</th>
<th></th>
<th>Non-iconic</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Perfective</td>
<td></td>
<td>32</td>
<td>51.6</td>
<td>30</td>
<td>48.4</td>
<td>62</td>
<td>58.5</td>
</tr>
<tr>
<td>Imperfective</td>
<td></td>
<td>32</td>
<td>72.7</td>
<td>12</td>
<td>27.3</td>
<td>44</td>
<td>41.5</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>64</td>
<td>60.4</td>
<td>42</td>
<td>39.6</td>
<td>106</td>
<td>100</td>
</tr>
</tbody>
</table>

In terms of the iconic gestures accompanying the aspect-marked utterances in the concrete events, I examined what elements they were iconic for in detail. As Fig. 4.1 reveals, the gestures accompanying the imperfective aspect were predominantly iconic for the combination of semantic elements of events (e.g., the combination of entity and action) that the speech denoted, and twice more so than gestures with the perfective aspect (51.5% vs. 22.9%). However, those accompanying the perfective aspect were more iconic for entity or action than those in the imperfective aspect contexts (31.4% vs. 15.2%, 25.7% vs. 15.2% respectively). These results suggest that the gestures accompanying imperfective aspect speech were more complex than those in the
perfective aspect contexts for concrete events, which is consistent with previous studies on concrete motion events (Duncan 2002; Parrill et al. 2013).

In the perfective aspect speech contexts, the speakers usually depicted the entities involved in the concrete events. As is the case in (1a). When the actress narrated her experience of searching for jobs, she mentioned the effect of the outstanding résumé she designed. Once initiating the topic related to the résumé, she extended her left palm-up open hand, which was iconic for the résumé. The gesture preceded the lexical affiliate jiānli (résumé) and was synchronous with dōu shì kàn wǒ de (all just look at my). It held to the ending of this turn, though there were new gestures added to it in the following utterances. In the second actual aspect-marked utterance, as Fig. 4.2 (1a) depicts, the speaker lifted her left hand a little and moved her gaze from the addressee to her left hand when the predicate kàn (look) was said. In the third negative utterance, she shook her right hand several times. In the fourth actual aspect-marked utterance, her extended right index finger pointed down at the left-hand-as- résumé, which was metonymic for herself. The gesture was synchronous with the adverb zhījì jiū (directly then), and held until its lexical affiliate dìng le wǒ (decide to hire me) appeared. The series of gestures reflected that the speaker simulated the situation where the directors made the decision to hire her due to her excellent résumé. In fact, she didn’t witness the situation, so all of the situation was recreated by her via imagination.

Some gestures accompanying the perfective aspect speech also combined the
semantic elements of concrete events on some occasions. For example, in (1b) and Fig. 4.2 (1b), the position of the thumb and index of both hands was iconic for holding a camera, and the movements of the speaker’s index fingers as if clicking the shutter button was iconic for the action of taking a picture. Such a gesture reflects that he used a camera rather than a mobile phone to take a picture at that time.

(1) a. [[[tāmen dōu shì kàn wǒ de jiānli], [kàn le jiānli yǐhòu]],
   they all be look I DE résumé, look ACTL résumé after,
   [bù yòng jiàn wǒ rén], [zhījiē jiù dìng le wǒ.]]
   no use meet I people, directly then hire ACTL me.
   “All of them just had a look at my résumé. After looking at it, they directly decided to hire me without needing to give me an interview.”

b. [ránhòu wǒ pāi le yī-gè zhàopiān].
   then I take ACTL a-CLF picture.
   “Then I took a picture.”

c. [wǒmen jiù ná zhe gē zhī bēi ]
   we then hold DUR CLF paper cup.
   “Then we held a paper cup.”

d. [[tā hái zài pá], [cái qī-gè duǒ yuè], [tā zài pá]]
   he still PROG crawl, still seven-CLF more month, he PROG crawl.
   “He was still in the phase of crawling, as he was only more than seven months old, and he was crawling.”
Gestural iconicity for the combination of semantic elements of concrete events was dominant in the imperfective aspect speech contexts. However, it occurred mainly in the durative aspect speech contexts, and the progressive aspect-marked utterances were usually accompanied by gestures iconic for action only. For example, in (1c) and Fig. 4.2 (1c), the speaker’s left thumb and index finger touched each other to be iconic for the rim of the paper cup, the other fingers curled around as if holding it. He simulated the entity and action carried out by the actress beside him in the past with details. The progressive aspect speech such as in (1d) went together with the speakers’ focus on the action. As Fig. 4.2 (1d) depicts, the actress’s two fists moved forward and backward in alternate ways three times, each movement being synchronous with the predicate of each utterance. Although the second utterance did not use the predicate zài pá (was crawling), the lexical affiliate of the gestures, she also made the gesture to complement her explanation. Such a series of gestures demonstrates the characteristics of a crawling phase, as the speech described the state of the child as a whole.

4.4.2 Metaphoric events

It is widely acknowledged that people understand abstract concepts metaphorically in terms of concrete notions (Lakoff 1987; Lakoff & Johnson 1980). Metaphorical experiences of abstract concepts are expressed not only in language but also in gestures.
We saw in the previous section that speech encoding the literal meaning of concrete verbs to construe events was more likely to be accompanied by iconic gestures in imperfective aspect than the perfective aspect contexts. Will this effect be extended to the utterances that include verbs denoting metaphorical meanings? As Table 4.2 reveals, there was no significant difference in the gesture use between the two grammatical aspect situations with metaphorical events ($\chi^2(1) = 0.325, p > 0.05$). Different from the pattern in the concrete event contexts, the gestures co-occurring with not only imperfective but also perfective speech contexts were highly iconic (78.3% vs. 72.1%), and with both aspects they occurred three times more frequently with iconic than with non-iconic gestures. This result suggests that metaphoricity is a much more important predictor for recalling the details of events when they are metaphoric events, without regard to aspectual distinctions.

Table 4.2: Frequencies of gestures across grammatical aspects in metaphoric events

<table>
<thead>
<tr>
<th>Grammatical aspect</th>
<th>Gesture type</th>
<th>Iconic</th>
<th>Non-iconic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Perfective</td>
<td>44</td>
<td>72.1</td>
<td>17</td>
<td>27.9</td>
</tr>
<tr>
<td>Imperfective</td>
<td>18</td>
<td>78.3</td>
<td>5</td>
<td>21.7</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>73.8</td>
<td>42</td>
<td>39.6</td>
</tr>
</tbody>
</table>

Taking into consideration the semantic elements of the iconic gestures referred to, both the imperfective and perfective aspect-marked utterances were accompanied by gestures which were also dominantly iconic for the action they depicted (50.0% and 42.9% respectively) (see Fig. 4.3). The gestures co-occurring with the perfective aspect were more iconic for combinations (e.g. entity and action) with the path than those with the imperfective aspect, which were more iconic for process and entity.
Unlike in the concrete events, not so many gestures were iconic for the entities of the metaphoric events, as in (2a). The speaker’s palm-up open right hand was iconic for the track of the playground, and this stroke hold was synchronous with 小孩什么 (the child what). When uttering the topic 小孩什么 (the child what), his left index finger drew a straight line on his right palm, and he held it through the comment of the whole sentence 失在了起点 (lose at the starting line), as shown in Fig. 4.4 (2a). The concrete domain RUNNING MATCH was mapped onto the abstract domain LIFE in speech, consisting in the metaphor “LIFE IS A RUNNING MATCH”. The competitors, the track, the starting point and the winning or losing results were all projected onto the abstract concept LIFE. His gestures highlighted the track and starting line.

There were also a number of motion events in metaphoric contexts. Their co-speech gestures were more likely to be iconic for the path in the perfective aspect speech contexts than in the imperfective speech contexts. For example, as in (2b), the speaker was narrating that he liked digging holes (i.e., to direct some TV series or films whose topics were very disputable and dangerous) as he progressed smoothly and successfully in his career. His right index finger traced the path of walking around the round hole,
as presented in Fig. 4.4 (2b). We can infer from the gesture that he mentally simulated the abstract event of overcoming difficulties in work or life as the concrete event of walking around a dangerous hole.

(2) a. jiù shì [tā //pā zhègè xiǎo hái shénme shū zài le qí pāo xiàn.]  
then be she fear this-CLF little child what lose at ACTL starting line.

“That she feared that the little child lost at the starting point (in life).”

b. dánhshì [[[hòulái shì kěng de biàn[biàn shèng zhuǎn le yī quān],]]
but later be at hole DE side side on walk ACTL a round,

[méi diào xiàqù.]]
not fall down

“But later I walked around the hole and did not fall into it.”

**Fig. 4.4 (2a)** Gesture with “lose ACTL”

**Fig. 4.4 (2b)** Gesture with “walk ACTL”

c. [dǎkāi le yī-gè xīn de zhègè rénmen wán jiū]  
open ACTL a-CLF new DE this people study

zhè-ɡè shì jiè de yī-gè] fāngshì huòzhě shuō dào.lù.
this-CLF world DE a-CLF way or say road.

“(It) opened a new way or road for people to study the world.”

d. [[yǒu hěnduō yě yǒu. zhènɡfù…(0.6s)] [zài tuǐdònɡ nèixī diàn])zì shānɡwù.
have many also have government PROG push those electronic commerce.

“There were many. also governments which were pushing those e-commerces.”
In metaphoric events, gestures accompanying both perfective and imperfective aspect speech were more likely to be iconic for the action embodied in the concrete events the speakers experience here and there. Take (2c) for instance; the speaker mapped the concrete action of opening doors, windows, cupboards etc. in everyday life onto the abstract action of opening a new way or road to study the world. As Fig. 4.4 (2c) depicts, the speaker moved his two palms towards the center, open hands moving forward and upward, separating from each other when uttering dākāi le (opened), and he held them this way for a while. Similarly, in example (2d) and Fig. 4.4 (2d), the speaker moved his palm downwards open right hand up and held at the height of his right shoulder before the stroke when uttering hěnduō yě yǒu (many also have), and then moved it downwards and outwards quickly when uttering the topic of this utterance zhèngfǔ (government). When the progressive aspect-marked comment zài tuōdòng (is pushing) occurred, the same gesture was repeated twice in a diminishing manner. Thus, the speaker simulated the abstract concept of facilitating e-commerce development via the concrete action of pushing a button or something in everyday life.

In most metaphoric events, the agents or patients are abstract, but the actions have an embodied basis. Therefore, the actions were more likely to be activated and represented by gestures.
4.4.3 Inherently abstract events

Abstract utterances, like the metaphoric utterances in the last section, represented abstract events but did so using verbs that did not have a concrete meaning denoting the physical action either (verbs such as biàn “change” and xiǎng “think”). The strong version of the embodied approach to language proposes that representation and comprehension of all concepts, be they concrete or abstract, are all embodied in our experience in, and interaction with, the world. Therefore, it is of great importance to investigate whether the gestures accompanying them will behave similarly to those in the concrete contexts when the inherently abstract events are framed by grammatical aspects.

As Table 4.3 shows, the gestures accompanying the imperfective aspect speech were more likely to be iconic than those in the perfective aspect speech contexts (61.4% vs. 38.9%). This was also the case with the concrete events. The imperfective aspect-marked utterances were accompanied by more iconic gestures than non-iconic gestures (61.4% vs. 38.6%), and the situation was the opposite in the perfective speech contexts (38.9% vs. 61.1%). The differences between them were significant via the Chi-square test ($\chi^2_{(1)} = 6.371$, $p < 0.05$). The results suggest that the speakers choosing the imperfective aspect to describe an event with inherently abstract verbs were more prone to mentally simulate the internal structure of the event than they were when choosing the perfective aspect.

**Table 4.3:** Frequencies of gestures across grammatical aspects in inherently abstract events

<table>
<thead>
<tr>
<th>Grammatical aspect</th>
<th>Gesture type</th>
<th>Iconic</th>
<th>Non-iconic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Perfective</td>
<td>42</td>
<td>38.9</td>
<td>66</td>
<td>61.1</td>
</tr>
<tr>
<td>Imperfective</td>
<td>27</td>
<td>61.4</td>
<td>17</td>
<td>38.6</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>45.4</td>
<td>83</td>
<td>54.6</td>
</tr>
</tbody>
</table>

As Fig. 4.5 shows, the iconic gestures accompanying the imperfective aspect were dominantly iconic for the actual process itself of an event, to a much greater extent than
that in the perfective aspect context (59.3% vs. 33.3%). Furthermore, the gestures accompanying the imperfective aspect speech were two times more likely to be iconic for the combination (mainly combination of metaphoric entity and process) than those in the perfective aspect speech contexts (11.1% vs. 4.8%). The iconic gestures with the perfective aspect were more iconic for the entity and action referred to in an inherent abstract event than those with the imperfective aspect (38.1% vs. 14.8%, and 21.4% vs. 14.8% respectively).

![Fig. 4.5 Gestural iconicity for different semantic elements in abstract verb contexts](image)

Previous studies reported that the “cyclic” gesture (i.e. a continuous circling motion of the loose hand) (Ladewig 2011) often accompanied the English progressive construction (be v-ing) to express progressivity (Harrison 2009). Taking a close look at the examples given by Harrison (2009), I found that the cyclic gestures encoding progressivity all occurred in inherently abstract events, such as “it’s turning out to be quite expensive”, “there’s something going on in the city that...”, “Because I never know what’s gonna change”. I also found some, but not so many, examples in our corpus, such as in example (3a). The actress made a two-handed (mirroring) cyclic gesture with the whole predicate by moving the hands downward-outward-inward vertically and then ended it with a palm-up open hand gesture with a hold for 10 seconds before the next utterance, as shown in Fig. 4.6 (1a). Such cyclic gestures “directly symbolize duration and dynamism” of the processing of an inherently abstract event.
(Harrison 2009: 234).

(3) a. [wǒ jué dé shì zài jīnxíng yī-gè xīllī.]
   I think be PROG doing one-CLF baptism.
   “I think we were going through the baptism (of a tough life)”

b. [shichāng làobāixíng yě zài tiáozhēng]
   market ordinary people too PROG adjust
   “The ordinary people are adjusting themselves to markets too”

Fig. 4.6 (3a) Cyclic gesture for PROG
Fig. 4.6 (3b) Back and forth gesture for PROG

However, the prototypical gesture accompanying the imperfective aspect in Mandarin Chinese conversations was the “back and forth” gesture rather than the cyclic gesture. The prototypical “back and forth” gesture was made with an intense or loose hand(s) alternating away and towards the speaker’s body. Take (3b) for instance; in Fig. 4.6 (3b), the speaker makes a “back and forth” gesture by alternating the tense open right hand away and towards his body while uttering the imperfective aspect-marked utterance làobāixíng yě zài tiáozhēng (ordinary people are also adjusting).

Due to the affordances of gestures and influence of context, inherently abstract events are usually accompanied by gestures iconic for the implicit agents, patients, actions or circumstances located in the events. Take (4a) as an example; before this utterance, the speaker recalled the preparation his son was making for directing a film. He located the center gesture space as the preparation period, the left gesture space as
the ending period and the middle period of film making by making a metaphorical gesture like Fig. 4.7 (4a). His two spread flat palms facing towards center consisted of a bounded container to represent the abstract concept TIME. When he said his son spent much more time in the preparation period than in the postproduction phase or even the shooting period, he explained that because “he was thinking and doing” during it. As verbs like think and do are abstract, their actions are not available for realizing as gestures. Therefore, the speaker referred back to the conduit metaphor gesture (McNeill 1992) in the center. This implies that the abstract events of “his thinking and doing” occurred in the preparation period. Similarly, in example (4b), the actress made a gesture in which her right index and thumb horizontally paralleled, and the other fingers curled onto the palm while uttering wǒ céngjīng (I once), it held statively when the abstract mental predicate hèn guo (hated) was mentioned, then she moved the tensed index-thumb upward quickly in front of her right eye when uttering yī xiǎo xiàzǐ (for a little while), as Fig.4.7 (4b) presents. Therefore, I can infer that the tensed index-thumb was metaphorically iconic for the length of TIME yī xiǎo xiàzǐ that her hate lasted for. However, when we look at the following utterance, we see that the tensed index-thumb gesture was simulated as “hate” at the same time. When she said zìjǐ yòu bù jiàn le (itself dissappeared), the tensed index finger and thumb slowly came together to touch each other and meanwhile moved from the center to the extreme right, which metaphorically mimicked the action of bù jiàn le (disappeared). It is not hard to see that what the gestures represent was very flexible and dynamic, strongly depending on the context.

(4) a. [[ yīnwèi tā zài xiǎng, tā zài zuò.]
Because he PROG think, he PROG do.

“Because he was thinking and doing.”

b. [[[dànshí wǒ céngjīng hèn guo] [//yī xiǎo xiàzǐ,]
but I once hate EXP a small CLF,

[dàn hòulái e zìjǐ yòu bù jiàn le.]]]
but later PRT itself again not see ACTL.
"But I once hated (my ex-boyfriend) for a while, but later it disappeared again."

c. [qíshí wǒmen zuò guò yī gé diào yī gé diáoyán.] actually we make EXP a-CLF invest. a-CLF investigation

   “Actually, we made an investigation.”

d. [nǐ kàn zhè fángdiàn lǐkè jiù jiù nánshòu le.]

   you see the real estate immediately then then unwell ACTL.

   “You see that the real estate immediately felt unwell.”

In the perfective speech contexts, the speakers’ gestures were prone to focus on the entities in the inherently abstract events, such as in example (4c). In speech, the speaker used a super-ordinate verb zuò (make or do) and chose the experiential aspect
to construe the experience of making an investigation of the excellent resources like education, healthcare, science and technology, the job opportunities, and environment, which the ordinary people were concerned about while choosing a city to live in. As Fig. 4.7 (4c) depicts, in the co-speech gesture, his palm-up open left hand and the stretched forearm modeled the questionnaire sheet (using the embodying mode of representation, introduced in Chapter 2), and his open right hand with the palm towards the body moved from the top to the elbow of the left hand, modeling the arrangement of the excellent resources items on the questionnaire sheet. The gesture was iconic for the entity “questionnaire sheet”, which was not explicit in speech but complementary with it to clarify how they made the investigation.

In some inherently abstract events, the speakers sometimes used more schematic verbs to describe an event, but the co-speech gestures represented the action beyond words. Although the implied verbs were seldom employed explicitly before or after what the multimodal utterance said, they could be inferred from the context. In example (4d), in speech, the speaker used the mental state verb nánshòu (unwell, uncomfortable) to depict the situation of market once the broad money (i.e. M2 in economy) was adjusted. Its implied meaning was “The housing prices became unstable, fluctuating up and down”, which was elaborated by the co-speech gesture. The speaker moved his right open hand up and down twice while uttering lìkè jiù... jiù nánshòu le (felt unwell immediately) as in Fig. 4.7 (4d).

Like in the metaphoric events, there were few gestures iconic for the combination of entity and action or process in the inherently abstract events. For instance, in example (5a) the speaker was narrating how his colleague constructed a data model to predict the Academy Award in 2014. He moved his two hands to the center from the armchair in rest position in the preparation phase, then slightly shook his two palms towards center with open hands while uttering yǐzhí zài biàn (was changing all the time) as Fig. 4.8 (5a) displays. In the following utterance, he clarified what the deictic zhège (this) in the last utterance referred to was móxing (model); thus we can infer that his two palm-towards-center open hands were iconic for the entity, i.e., the “data model”. Furthermore, he moved the “data model” from the center to the right side with a slight
shake at the same time while elaborating “you see that his model was changing all the
time”. The slight shake was to be iconic for the ongoing process of the abstract event
(CHANGE IS MOTION [Grady 1997]), and the movements of change in SPACE can
be interpreted as metaphorically iconic for change in QUANTITY of the data model.
Similarly, in example (5b) and Fig. 4.8 (5b), the speaker’s right index finger and thumb
were bent, and the other fingers curled onto the palm. Such a tensed index-thumb
arrangement, as if modeling the length of three or four days, moved laterally from the
center to the extreme right quickly while starting the topic tā yīzhì zài (he was all the
time). After a short pause, he completed the utterance while moving the tensed index-
thumb back parallel to his right shoulder. In speech, the abstract verb nèijiù (guilty) was
used to depict his son’s mental activity after seeing the quarrel between him and his
wife. The holding and movement of the tensed index-thumb from the beginning to the
end of the utterance can be seen as metaphorically iconic for the combination of entity
and process to represent the duration and progressivity of the abstract mental action.

(5) a. [[tā zhègè yīzhì zài biàn],

   he this all the time PROG change,

   [ni kàn tā de móxing yīzhì zài biàn.]]

   you see he DE model all the time PROG change.

   “This was changing all the time. You see that his model was changing all the time.”

b. [[tā yīzhì zài (.0.2s)] [zhē săn sì tiān tā zài nèijiù]].

   he all the time PROG this three four day he PROG guilty.

   “He was guilty all the time during the three or four days.”
To investigate the interaction effects between grammatical aspects, verb types and gesture type, I carried out the *LogLinear Model Selection* process. The results showed that there was no interaction effect between grammatical aspects, event type and gesture type ($\chi^2_{(2)} = 0.739, p > 0.05$). However, there were significant interaction effects between grammatical aspect and event type ($\chi^2_{(2)} = 5.490, p < 0.05$), between grammatical aspects and gesture type ($\chi^2_{(1)} = 8.419, p < 0.05$), and between the event type and gesture type ($\chi^2_{(2)} = 20.326, p < 0.05$). The results suggest that the functions of the co-speech gestures accompanying the grammatical aspectual utterances are influenced by the grammatical aspect and event type, respectively.

4.5 Discussion

The study reported in this chapter investigated the gestures accompanying the grammatical aspect-marked utterances in concrete, metaphoric and inherently abstract events in Mandarin Chinese conversations. In this section, I will first summarize the results so that I can answer the two research questions, and then I will interpret them respectively.

I examined the frequencies of iconic gestures co-occurring with grammatical aspect speech to answer the first question I put forward: is the event internal structure more highlighted by the gestures accompanying imperfective aspect utterances than
perfective aspect utterances in concrete, metaphoric, and abstract events? The results showed that gestures accompanying imperfective aspect speech were more prone to be (concrete or metaphoric) iconic than gestures accompanying perfective aspect speech in the concrete and inherently abstract events, whereas the gestures that speakers produced in both imperfective and perfective aspect-marked speech contexts were equally frequently iconic in the metaphoric events. The more iconic gestures represent the events, the more internal structures were highlighted. Thus, the results suggest that the event-internal structures were more highlighted by the gestures accompanying the imperfective aspect-marked utterances than the perfective aspect-marked utterances in the concrete and inherently abstract events, and were equally highlighted in the metaphoric events.

According to GSA framework, representational gestures emerge from the perceptual and motor simulations that underlie embodied language and motor imagery (Hostetter & Alibali 2008). The speakers who produced more iconic gestures in event descriptions should activate more perceptual and motor information about the internal structures of the events. Our results further suggest that describing metaphoric and inherently abstract events often stimulates the motor system as when describing the physically experienced or perceived concrete events, as evidenced by the motoric movements of gesturing. They were consistent with the results from some neurophysiological studies reviewed in section 4.2 that understanding literal sentences (i.e., describing concrete events), metaphoric sentences (i.e., describing metaphoric events) and abstract sentences (i.e., describing inherently abstract events) could produce somatotopic activation (Boulenger et al. 2012; Desai et al. 2011). The embodied cognition approach proposes that many abstract concepts are understood and experienced via analogies to sensation and action (Bergen et al. 2007; Gibbs 2006; Lakoff & Johnson 1980, 1999). This may be useful for explaining why the metaphoric events were more likely to be accompanied by iconic gestures to highlight their internal structures regardless of the temporal perspectives the speakers took to construe the events. However, in the concrete and inherently abstract events, the imperfective aspect viewpoint appears to correlate with the speakers’ focusing more on the internal
structures of events in contrast with the perfective aspect via their co-speech gesture use.

Metaphoric and inherently abstract events had the same potential to be represented iconically as the concrete events did in grammatical aspect speech contexts. But metaphorical and inherently abstract events have a different nature and can activate different regions in the brain in contrast with concrete events (Desai et al. 2011). Let us now turn to the second questions: Are gestures accompanying the imperfective and perfective aspect speech in concrete, metaphorical and inherently abstract events descriptions iconic for similar or different event semantic elements? In this regard, I found that co-speech gestures have “preferences” in terms of iconicity for different semantic elements with different types of events. Gestures accompanying the imperfective aspect-marked utterances were more iconic for the combination of event semantic elements, whereas those co-occurring with the perfective aspect speech were more iconic for entity and for action of the concrete events. For the metaphorical events, gestures accompanying the imperfective and perfective aspect speech were all predominantly iconic for action. Gestures accompanying the imperfective aspect speech were more iconic for the process, whereas those accompanying the perfective aspect were more iconic for entity in the inherently abstract events. Our results extended the previous studies (Duncan 2002; Parrill et al. 2013), as gestures accompanying imperfective-aspect marked speech were shown to be more complex (representing more semantic elements) than those with the perfective-aspect speech not only in concrete motion events but also in various kinds of concrete events. However, this effect did not arise in the metaphorical and inherent abstract events. Behavioral and electrophysiological studies also showed that progressive aspect descriptions of concrete events activate more richly detailed information, such as details about location and the entities of an event (Carreiras et al. 1997; Ferretti et al. 2007). Furthermore, concrete events can provide more detailed information in mental imagery, and the imperfective aspect facilitated greater fine-grained simulation in concrete events. Thus, the entities, locations, actions etc. were more available to be combined in gestural representations of the events. However, in the metaphorical and inherently abstract events,
either the entities or actions involved were tangible without direct access being possible to their mental images of the (abstract) target referents.

Previous studies showed that use of the imperfective aspect focused on the dynamic ongoing actions of an event, whereas the perfective aspect emphasized the completed end-state or affected entities of an event. As our study showed, this only occurred in the concrete and inherently abstract events rather than in the metaphoric events. In the metaphoric event descriptions, the verbs were concrete, embodied in the physical actions we carry out in everyday life. No matter what grammatical aspect the speaker employed to construe the events, motion or action may be simulated first to bridge the mapping between the concrete domain and the abstract domain.

Grammatical aspect is closely related to the process of events and cognitively motivated by our ability to simulate action (Barsalou 2009; Langacker 2008a). With abstract events, the verbs were inherently abstract without direct access to a physical action we can perceive or do. As the semantic poles of their symbolic structures were both highly schematic, their relevant cognitive iconicity depended on a correspondingly schematic phonological structure (Wilcox 2004). The qualitative results about inherently abstract events show that the entities represented by the co-speech gestures in perfective aspect were mainly schematic as conduit metaphors of time to describe the temporal locations of the abstract events. The movements of the gestures did not have some clear and concrete action forms like the concrete and metaphoric use of verbs, but rather the schematic “back and forth” or “cyclic” forms, which were the schematic phonological form of the imperfective aspect. Our results were inconsistent with some psychological or behavioral studies which reported that understanding abstract events did not activate motor systems of the brain or stimulate mental simulation of visual imagery (Aziz-Zadeh et al. 2006; Bergen et al. 2007; Raposo et al. 2009). This may result from the differing contextual scopes provided by psychological experimental materials and by natural conversations. The psychological studies all used one-sentence stimuli, whereas the natural conversations consist of bigger discourse units. The discourse contexts influence the growth points of gestures accompanying speech (McNeill 2005). Thus, the metaphoric entities mentioned in the perfective-
aspect utterances, and the implicit action implied by the grammatical aspect-marked utterances, were also represented in gestures to keep the conversations coherent.

4.6 Conclusions

A comparison of gestures accompanying grammatical aspect-marked utterances in concrete, metaphoric and abstract events revealed that grammatical aspect modulated the conceptualization of concrete and inherently abstract events rather than that of metaphoric events. Gestures co-occurring with imperfective-marked speech were more likely to be iconic than those with perfective-marked speech for the concrete and inherently abstract events, whereas both types of speech were equally accompanied by gestures which were highly iconic for the metaphoric events. Furthermore, the gestures in the imperfective speech contexts were more iconic for a combination of semantic elements such as entity and action for concrete events, and for process for inherently abstract events, whereas those in the perfective speech contexts were more iconic for the (concrete or metaphoric) entities in the two kinds of events. However, the gestures accompanying the imperfective and perfective aspect speech were not different, and for both of them were mainly iconic for actions in metaphoric events. The speakers made more complex gestures when producing imperfective aspect utterances than when producing perfective aspect ones only in the concrete events.

These findings have a number of clear implications. First, metaphoric events and inherently abstract events can strongly activate mental simulation of events as the concrete events did when they were marked by grammatical aspects in speech production. Gestures as simulated action have been argued to be visible manifestations of the speakers' mental simulation of visual or motor images related to events. While (re)constructing situation models, the speakers made more iconic gestures than non-iconic gestures in the concrete and abstract events rather than in the metaphoric events when they are marked by imperfective aspect rather than perfective aspect.

Second, what elements of events the speakers focus on in event conceptualization
is modulated by the grammatical aspectual categories, which rely on the event types with which they occur. The imperfective aspect facilitated the speakers’ focus on the dynamic nucleus, whether it was concrete or schematic, in contrast with the perfective aspect, where the focus was on the entities, in the concrete and abstract events. Metaphoricity based on physical actions overrides the grammatical aspect to mainly simulate the action in both imperfective and perfective aspects speech contexts.

There are some limitations of this study. First, the corpus is relatively small, particularly the samples of metaphoric events. Second, the advantages of natural conversations, which can provide various kinds of events, larger discourse contexts, and show what people actually say and do in event construal, also entail disadvantages at the same time. There are too many factors influencing the gestural behaviors. I cannot definitely say what features are uniquely inherent to the grammatical aspect in different event types. Therefore, the results of this study need to be verified in future studies by strictly controlling other factors such as the nature of agents, patients, and use of the same stimuli to elicit speech and gestures.