
3.1 Introduction

In face-to-face conversations, people subconsciously make gestures to express how an event unfolds through time, to depict what an agent or patient of an action looked like, and to model what states an action result in, etc. Such gestures bring the thought about an event into existence on a concrete plane (McNeill 2005: 99). The co-occurring speech and gesture are considered as an integrated “ensemble” to convey a “single thought” (Kendon 2004; McNeill, 1992, 2005). As a window to thought, gestures provide access to the speaker’s mental imagery or simulation of a particular experienced event, which can be vividly brought back into the immediate communicative context. However, to what extent the speaker goes into the details of an event is dependent on the temporal perspectives from which s/he chooses to construe it. Thus, aspect plays an important role in modulating the conceptualization of events (e.g., Madden & Zwaan 2003; Bergen & Wheeler 2010; Matlock et al. 2012; Duncan 2002; McNeill 2003; Parrill et al. 2013).

Comrie (1976) defines aspect as “different ways of viewing the internal temporal constituency of a situation” (p.3), which are realized grammatically and/or lexically across languages. In spite of being a tenseless language, Mandarin Chinese is rich in grammatical aspect, which has a series of particles to indicate the temporal contouring of events, such as -le to express the actualization of an event, -guò to express someone’s having experienced the event, and zài on the event’s progressive nature, -zhe on its durativeness. The former two represent perfective aspect, viewing the event as an indecomposable whole from an external perspective, whereas the latter two express
kinds of imperfective aspect, viewing the extension of an event from an internal perspective. That is, the conceptual information encoded by the perfective aspect takes an external viewpoint on an event and focusing its boundaries, while that encoded by the imperfective aspect takes an internal viewpoint on an event and defocuses its boundaries. Previous studies on grammatical aspect and gesture (or even instrumental, actual goal-directed actions in McNeill’s experiment) show that gestures (actions) accompanying imperfective aspect-marked utterances last longer temporally and display more complexity in form (Duncan 2002; McNeill 2003; Parrill et al. 2013).

Most previous studies of aspect and gesture focused on English aspects, used elicited narrative data, and narrowed their lens on physical motion events and concrete iconic gestures. Although Duncan (2002) took Mandarin Chinese into consideration, she conflated the different types of Chinese aspect-marked event construal. Based on these studies, I systematically explore how Mandarin Chinese speakers conceptualize events as expressed in their co-speech gestures accompanying grammatical aspect-marked utterances in natural conversations, which concern not only concrete but also abstract events. The following specific questions are addressed.

1) Are the gestures accompanying speech with imperfective aspect, including the progressive and durative aspect, more complex in form than those with perfective aspect, including the actualized and experiential aspect?

2) Is speech with imperfective aspect, including the progressive and durative aspect, accompanied by more iconic gestures than that with the perfective aspect, including the actualized and experiential aspect?

3.2 Methods

3.2.1 Multimodal corpus

The materials in this study are from Yang Lan One on One. Twenty shows consisting of deep interviews with native Chinese speakers were randomly selected; they were recorded between July 2014 and October 2015. They cover a wide range of topics
(about work, life experiences etc.), which provide relatively rich and natural Chinese conversation data. Each video, without the background introduction and advertisements, lasts for about 20 minutes. This multimodal corpus is about 8.5 hours long in total.

3.2.2 Participants

Thirty people (10 females) were involved in the interview corpus. Their mean age is 49.53 (SD = 10.82), ranging from 24 to 69. Most of them are actors or actresses. All of them are native Chinese speakers.

3. 2. 3 Coding procedures

3.2.3.1 Speech

There were 342 multimodal utterances marked by aspect markers, which were all affirmative and marked by just a single aspect particle. The negative utterances were all excluded. I also excluded utterances marked by the combination of two aspect markers such as chī guò le (eat-EXP-ACTL), as the categorical data can only be designated to one category for Chi-square analysis later. One coder coded the utterance aspect for the whole dataset, and a second coded the grammatical aspect in 50% of the dataset. Grammatical aspect was coded at the macro level for perfective and imperfective, and at the micro level for ACTL-le, EXP-guò, PROG zài and DUR-zhe; the inter-rater agreement was 98%.

The events described by each multimodal utterance were also coded as concrete events or abstract events (with metaphorical and abstract use of verbs). Concrete events refer to those that depict what physically happens in our daily life, whereas abstract events depict what happens without concrete action. The abstract event involves the metaphorical use of concrete verbs in construing abstract concepts (such as life attitude, beliefs, changes in quantity etc.), and inherently abstract verbs (such as the inner psychological state predicate xiǎng (think), and the superordinate verb zuò (do)). Example (1a) is a concrete event, in which holding a machine is physically done and is
visible, while (1b) is an abstract event with the metaphoric use of a concrete verb *bàoyòu* (hold), in which the life attitude of indifference and calmness is abstract; “holding it” is an invisible mental state. Finally, (1c) is an abstract event with the use of an inherently abstract predicate *biànhuà* (change).

(1) a. Yìmóu yōngyuǎn bāo zhe jījí.
Yimou forever hold DUR machine

“Yimou was always holding the machine.”

b. tā yōngyuǎn bāo yǒu zhe yí-zhǒng huā kāi huā luò
de tānrán hé cóngróng.
he forever hold have DUR a-CLF flower open flower fall

DE unperturbed and calm.

“He always holds the life attitude of being unperturbed and calm, like flowers blossom and fall naturally.”

C. shìchǎng zài biànhuà.
market PROG change.

“The market is changing.”

3.2.3.2 Gesture

Previous studies reported that gesture forms such as handedness and movement quality (i.e. iteration) were related with aspects (Duncan 2002; Parrill et al. 2013). The handedness of gestures accompanying durative aspect tended to be more complex than those with perfective aspect and progressive aspect. The gestures accompanying the imperfective aspect were more likely to be iterated (or agitated) than those with the perfective aspect. Thus, those forms of gestures accompanying the aspect-marked utterances were coded. In terms of handedness, I considered whether the speaker used a single hand or both hands. In the case of both hands, they were coded as both hands mirroring each other in shape and position, or as both hands contrasting with each other in either one of these respects. Movement qualities here were coded as single when the
movement of a gesture stroke occurred only once, or iteration when the same movement pattern of a gesture stroke was repeated twice or more.

All the aspectual speech synchronous gestures were also coded according to gesture types adapted from McNeill (1992, 2005) and Kendon (2004). They were categorized into *concrete iconic* (hands represent the shape, size and/or action of the concrete events under description), *metaphoric iconic* (hands represent the shape, size and/or action of the abstract events), *deictic* (hands point at some physically present or imagined entities in the immediate context), and *pragmatic* (hands contribute to the pragmatics of the utterance meaning; they may function as performative, modal, or parsing gestures). The first two are iconic gestures, and the other two non-iconic. The iconic gestures focus more on the internal structure of events, such as being iconic for the properties of entities including agents or patients of events, iconic for the manner of action or path left by the action, or iconic for the ongoing or durative process of the events being described. However, the non-iconic gestures are more related to attention alignment via deictic gestures or discourse management via pragmatic gestures.

One coder coded the entire dataset, and a second coded 25% of the dataset. The inter-rater agreement was good: 100% for handedness, 97% for iteration, and 96% for gesture type.

### 3.2. 4 Data analysis

The data in this study are all categorical, so the Chi-square test was used to investigate the significance of differences between two variants. When the interaction effects were calculated, the *Loglinear Model* was used. The alpha level was set at 0.05.

### 3.3 Results

I first report the results of gesture forms of the aspect-marked utterances. Then the usage of gesture types accompanying different grammatical aspects will be thoroughly examined.
3.3.1 Gesture handedness

Generally speaking, a single-handed gesture is less salient than a 2-handed gesture. Furthermore, a 2-handed mirror gesture is also less complex than a gesture where the two hands have contrasting forms or motions. Now, I am going to see whether such complexities were embodied in gestures accompanying different aspect-marked utterances.

3.3.1.1 At the macro-level

As displayed in Table 3.1, the difference in handedness of gestures accompanying speech marked for perfective and imperfective aspectual distinctions is statistically non-significant ($\chi^2(2) = 5.124, p > 0.05$). Both the perfective and imperfective aspect-marked utterances were dominantly accompanied by single-handed gestures, less frequently with gestures where the hands mirrored each other, and the least when the hands were making contrasting gestures, as presented in Fig. 3.1.

Table 3.1: Frequencies of gesture handedness across macro grammatical aspect contexts

<table>
<thead>
<tr>
<th>Macro GA</th>
<th>Gesture handedness</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single hand</td>
<td>Hands mirror</td>
<td>Hands contrast</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N  %</td>
<td>N  %</td>
<td>N  %</td>
<td>N  %</td>
<td>N  %</td>
</tr>
<tr>
<td>Perfective</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perfective</td>
<td>121 52.4</td>
<td>62 26.8</td>
<td>48 20.8</td>
<td>231 67.5</td>
<td></td>
</tr>
<tr>
<td>Imperfective</td>
<td>47 42.3</td>
<td>43 38.7</td>
<td>21 18.9</td>
<td>111 32.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>168 49.1</td>
<td>105 30.7</td>
<td>69 20.2</td>
<td>342 100</td>
<td></td>
</tr>
</tbody>
</table>
3.3.1.2 At the micro-level

According to Table 3.2, there was no significant difference of gesture handedness among the micro grammatical aspect-marked utterances ($\chi^2(6) = 6.567, p > 0.05$). The actualized aspect, the experiential aspect, the progressive aspect and the durative aspect marked speeches were all mainly accompanied by single-handed gestures, then by the hands as mirror gestures, and the least by hands as contrast gestures, as displayed in Fig. 3.2. The gestures in durative aspect contexts were even less often two-handed as contrasted both with those in the progressive aspect contexts and with those in the perfective (the actualized and experiential aspects) contexts.

<table>
<thead>
<tr>
<th>Micro GA</th>
<th>Gesture handedness</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTL-le</td>
<td>Single hand</td>
<td>108</td>
<td>51.9</td>
<td>54</td>
<td>26.0</td>
<td>46</td>
<td>22.1</td>
<td>208</td>
<td>60.8</td>
</tr>
<tr>
<td></td>
<td>Hands mirror</td>
<td>11</td>
<td>47.8</td>
<td>8</td>
<td>34.8</td>
<td>4</td>
<td>17.4</td>
<td>23</td>
<td>6.7</td>
</tr>
<tr>
<td>EXP-guo</td>
<td>PROG zai</td>
<td>26</td>
<td>44.1</td>
<td>21</td>
<td>35.6</td>
<td>12</td>
<td>20.3</td>
<td>59</td>
<td>17.8</td>
</tr>
<tr>
<td>DUR-zhe</td>
<td></td>
<td>22</td>
<td>42.3</td>
<td>22</td>
<td>42.3</td>
<td>8</td>
<td>15.4</td>
<td>52</td>
<td>15.2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>167</td>
<td>48.8</td>
<td>105</td>
<td>30.7</td>
<td>70</td>
<td>20.5</td>
<td>342</td>
<td>100</td>
</tr>
</tbody>
</table>

The handedness of gestures produced together with the macro- and micro-
grammatical aspect utterances in this study did not show the preferences revealed by Duncan’s study: the perfective and progressive aspect contexts preferred 2-hand/mirror gestures, whereas the durative aspect contexts favored 2-hand/contrast gestures (Duncan 2002: 197). This may result from the different types of data investigated. Duncan only considered motion events and iconic gestures, but I examined natural conversation which included not only motion events but also other events like causative events and states, and different dimensions of gestures, like iconic, metaphoric, deictic and pragmatic.

![Fig. 3.2 Frequencies of gesture handedness across micro grammatical aspects](image)

### 3.3.2 Movement quality

#### 3.3.2.1 At the macro-level

According to Table 3.3 and Fig. 3.3, although both the perfective and imperfective marked speeches predominantly co-occurred with the single movement gestures, the imperfective aspect-marked utterances were accompanied by twice as many more gestures with iteration than the perfective aspect-marked utterances were. The differences were reliably significant ($\chi^2(1) = 13.159, p<0.05$).
### Table 3.3: Frequencies of movement quality of gestures across macro-grammatical aspects

<table>
<thead>
<tr>
<th>Macro GA</th>
<th>Movement Quality</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single N (%)</td>
<td>Iteration N (%)</td>
<td>Total N (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perfective</td>
<td>197 85.3</td>
<td>34 14.7</td>
<td>231 67.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imperfective</td>
<td>76 68.5</td>
<td>35 31.5</td>
<td>111 32.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>273 79.8</td>
<td>69 20.2</td>
<td>342 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 3.3 Movement quality of gestures across macro-grammatical aspects**

### 3.3.2.2 At the micro-level

As Table 3.4 shows, the differences of gestural iteration among the micro grammatical aspect-marked utterances were highly significant ($\chi^2(3) = 25.563, p < 0.05$). Like the actualized aspect- and experiential aspect-marked speech, the durative aspect-marked speech was mainly accompanied by gestures with a single occurrence, as displayed in Fig. 3.4. Thus, the differences mainly resulted from the gestural behavior of progressive aspect-marked speech, which was far more iterated than those in the other three aspect contexts.

To sum up, there was a significant difference between the frequencies for perfective and imperfective utterances in terms of gestural iteration, but almost half of the progressive aspect-marked utterances co-occurred with iterated gestures in contrast with the situation with the durative aspect-marked utterances.
Table 3.4: Frequencies of movement quality of gestures across micro grammatical aspect

<table>
<thead>
<tr>
<th>Micro GA</th>
<th>Movement Quality</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
<td>Iteration</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>ACTL-le</td>
<td>178</td>
<td>85.6</td>
<td>30</td>
<td>14.4</td>
</tr>
<tr>
<td>EXP-guo</td>
<td>19</td>
<td>82.6</td>
<td>4</td>
<td>17.4</td>
</tr>
<tr>
<td>PROG zai</td>
<td>33</td>
<td>55.9</td>
<td>26</td>
<td>44.1</td>
</tr>
<tr>
<td>DUR-zhe</td>
<td>43</td>
<td>82.7</td>
<td>9</td>
<td>17.3</td>
</tr>
<tr>
<td>Total</td>
<td>273</td>
<td>79.8</td>
<td>69</td>
<td>20.2</td>
</tr>
</tbody>
</table>

Fig. 3.4 Movement quality of gestures across micro grammatical aspect

3.3.3 Gesture types/functions across grammatical aspects

3.3.3.1 At the macro-level

Table 3.5 shows that the multimodal utterances marked by the perfective aspect were twice as frequent as those with the imperfective aspect (231 vs. 111). This result is consistent with the case in English, in which the imperfective aspect (the progressive forms) is used less frequently than the perfective aspect (the non-progressive forms like the simple past tense) (e.g., Biber & Reppen 2002). This makes sense for natural conversational discourse, in which the speakers mostly narrate past events in a snapshot way.
According to Table 3.5 and Fig. 3.5, we can see that the utterances marked by the imperfective aspect were accompanied by significantly more iconic gestures than those in the perfective aspect contexts (69.4% vs. 51.1%). In the imperfective aspect speech contexts, twice as many iconic gestures were produced than non-iconic ones (69.4% vs. 30.6%), whereas in the perfective aspect-marked speech contexts, there was almost equal use of iconic and non-iconic gestures (51.1% vs. 48.9%). The Chi-square test of significance showed a reliable effect ($\chi^2(1) = 10.231, p < 0.05$).

Table 3.5: Frequencies of iconic gesture across macro grammatical aspect

<table>
<thead>
<tr>
<th>Macro GA</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>118</td>
<td>51.1</td>
<td>113</td>
<td>48.9</td>
</tr>
<tr>
<td>Imperfective</td>
<td>77</td>
<td>69.4</td>
<td>34</td>
<td>30.6</td>
</tr>
<tr>
<td>Total</td>
<td>195</td>
<td>57.0</td>
<td>147</td>
<td>43.0</td>
</tr>
</tbody>
</table>

Fig.3.5 Frequency of iconic gesture across macro grammatical aspect

The results suggest that the imperfective event descriptions were accompanied by more complex gestures, which is consistent with the results of the previous studies (Duncan 2002; Parrill et al. 2013). The visible gestural evidence further proves that the internal structures of situations in linguistic imperfective contexts are more likely to be mentally stimulated.
3.3.3.2 At the micro-level

As Table 3.6 shows, for the multimodal events, the actualized aspect was dominantly employed by the Mandarin Chinese speakers, whereas the experiential aspect-marked multimodal utterances were the least frequent in natural conversations. The progressive aspect and durative aspect were almost equally used, both of which were twice more frequent than the experiential aspect and almost four times less frequent than the actualized aspect.

Table 3.6: Frequencies of micro gesture type across micro grammatical aspect

<table>
<thead>
<tr>
<th>Micro Grammatical aspect</th>
<th>Deictic N</th>
<th>Deictic %</th>
<th>ConIconic N</th>
<th>ConIconic %</th>
<th>MetaIconic N</th>
<th>MetaIconic %</th>
<th>Pragmatic N</th>
<th>Pragmatic %</th>
<th>Total N</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTL-le</td>
<td>30</td>
<td>14.4</td>
<td>30</td>
<td>14.4</td>
<td>80</td>
<td>38.5</td>
<td>68</td>
<td>32.7</td>
<td>208</td>
<td>60.8</td>
</tr>
<tr>
<td>EXP-guo</td>
<td>4</td>
<td>17.4</td>
<td>1</td>
<td>4.3</td>
<td>7</td>
<td>30.4</td>
<td>11</td>
<td>47.8</td>
<td>23</td>
<td>6.7</td>
</tr>
<tr>
<td>PROG zai</td>
<td>3</td>
<td>5.1</td>
<td>8</td>
<td>13.6</td>
<td>35</td>
<td>59.3</td>
<td>13</td>
<td>22.0</td>
<td>59</td>
<td>17.3</td>
</tr>
<tr>
<td>DUR-zhe</td>
<td>11</td>
<td>21.2</td>
<td>23</td>
<td>44.2</td>
<td>11</td>
<td>21.2</td>
<td>7</td>
<td>13.5</td>
<td>52</td>
<td>15.2</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>14.0</td>
<td>62</td>
<td>18.1</td>
<td>133</td>
<td>38.9</td>
<td>99</td>
<td>28.9</td>
<td>342</td>
<td>100</td>
</tr>
</tbody>
</table>

Notes: ConIconic=Concrete Iconic, MetaIconic=Metaphoric Iconic

![Fig. 3.6 Frequencies of micro gesture type across micro grammatical aspect](image)

As Fig. 3.6 and Table 3.6 show, the actualized aspect utterances were mainly accompanied by metaphoric iconic gestures and pragmatic gestures, the experiential aspect utterances were predominantly accompanied by pragmatic gestures, then by
metaphoric iconic gestures, and the least by concrete iconic gesture. Although the actualized aspect and experiential aspect all construe the events perfectly as an entirety, the gestural behaviors with them display some differences. If I conflate the concrete iconic and metaphoric iconic gestures into iconic gestures, and the deictic and pragmatic gestures into non-iconic gestures, you can see that, like the progressive and durative aspect-marked utterances, the actualized aspect-marked utterances were accompanied slightly more by iconic gestures than by non-iconic gestures (52.9% vs 47.1%), whereas the experiential aspect-marked utterances were accompanied more by non-iconic gestures than by iconic gestures (65.2% vs 34.7%). This result suggests that the speakers were also more prone to simulate the internal structure of events while using the actualized aspect. When the experiential aspect was used, pragmatic gestures were produced most frequently. The progressive aspect-marked utterances predominantly co-occur with metaphoric gestures, at a level three times that with durative aspect-marked utterances and twice as much as with those of actualized and experiential aspects-marked utterances (59.3% vs. 21.2% vs. 38.5% vs. 30.4%). However, the durative aspect-marked utterances were mainly accompanied by concrete iconic gestures, almost ten times more than utterances with the experiential aspect, and four times more than those with actualized aspect and progressive aspect (44.2% vs. 4.3% vs. 13.6% vs. 14.4%). The differences between them are statistically significant as the Chi-square test shows ($\chi^2(9) = 49.753$, $p < 0.05$). The results reveal an interesting phenomenon, namely, that the progressive aspect may be preferentially chosen to construe abstract events, but the progressive aspect may be more frequently employed to conceptualize concrete events. The next section will examine this phenomenon in detail.

To our knowledge, these results have been discovered here for the first time, which is of great significance for us to understand the micro differences of profiled focus in conceptualizing events between different aspect-framed contexts.
3.3.4 Aspect and gesture in concrete vs. abstract events

3.3.4.1 At the macro-level

As is shown in Table 3.7, there is no significant difference between the perfective and imperfective aspects ($\chi^2_{(1)} = 1.497, p > 0.05$) when it comes to their use in characterizing concrete and abstract events. Both of them were employed more frequently with the abstract events than with the concrete events in conversational discourse, as indicated in Fig. 3.7. Therefore, when studying aspect and gesture use by speakers of Mandarin Chinese, we need to pay attention not only to concrete motion events as previous studies did, but also to abstract events, which are more frequently framed by the aspect particles.

<table>
<thead>
<tr>
<th>Macro GA</th>
<th>Event types</th>
<th>Concrete</th>
<th>Abstract</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>74</td>
<td>32.0</td>
<td>157</td>
<td>68.0</td>
</tr>
<tr>
<td>Imperfective</td>
<td>43</td>
<td>38.7</td>
<td>68</td>
<td>61.3</td>
</tr>
<tr>
<td>Total</td>
<td>117</td>
<td>34.2</td>
<td>225</td>
<td>65.8</td>
</tr>
</tbody>
</table>

Table 3.7: Distribution of macro grammatical aspects across event types

Fig. 3.7 Macro grammatical aspect in concrete and abstract events
In order to know the relationship among the macro grammatical aspect, the event type, and the macro gesture types in multimodal contexts, I carried out a LogLinear Selection Analysis. The results showed that there was no interaction effect between macro grammatical aspect, event type and macro gesture type ($\chi^2_{(1)} = 1.199$, $p > 0.05$). But the interaction between macro grammatical aspect and macro gesture type was significant ($\chi^2_{(1)} = 8.578$, $p < 0.05$). It revealed a main effect of event type ($\chi^2_{(1)} = 34.696$, $p < 0.05$). Such results imply that event type influences not only the use of grammatical aspect, but also the native Chinese speakers’ gestural behaviors when they take aspectual views on events.

### 3.3.4.2 At the micro-level

According to Table 3.8, the actualized aspect, experiential aspect, and progressive aspect in particular were used more frequently with abstract events than with concrete events, whereas the durative aspect was employed more frequently with concrete events than to characterize abstract events in conversational discourse, which is clearly displayed in Fig. 3.8. The differences are also shown to be significant via the Chi-square test ($\chi^2_{(3)} = 12.978$, $p < 0.05$).

**Table 3.8: Distribution of micro grammatical aspects across event types**

<table>
<thead>
<tr>
<th>Event Type</th>
<th>ACTL-le</th>
<th>EXP-guo</th>
<th>PROG-zai</th>
<th>DUR-zhe</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Concrete</td>
<td>67</td>
<td>32.2</td>
<td>8</td>
<td>34.8</td>
<td>15</td>
</tr>
<tr>
<td>Abstract</td>
<td>141</td>
<td>67.8</td>
<td>15</td>
<td>65.2</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>208</td>
<td>60.8</td>
<td>23</td>
<td>6.7</td>
<td>59</td>
</tr>
</tbody>
</table>
In order to know the interaction effects concerning the micro grammatical aspect, the event type, and the micro gesture types in multimodal contexts, I also carried out a LogLinear Selection Analysis. The results showed that the interaction effect of micro grammatical aspect, event type and micro gesture type was not significant ($\chi^2_{(9)} = 8.677$, $p > 0.05$). However, the interaction between micro grammatical aspect and micro gesture type was significant ($\chi^2_{(9)} = 34.450$, $p < 0.05$), as it was between event type and micro gesture type ($\chi^2_{(3)} = 111.308$, $p < 0.05$). Such results suggest that both micro grammatical aspect and event type influence the speakers’ co-speech gestural functions. Thus, it furthered our awareness of the influence of event type on the grammatical aspect usage and the speakers’ gestural performance.

3.4 Discussion

3.4.1 Grammatical aspect and gesture forms

In terms of the physiological form of gestures accompanying aspect-marked utterances, the study revealed that there is no significant relation between aspect and gesture handedness, but aspects at both macro and micro- levels and gesture iteration were found to be interrelated.
When Mandarin Chinese speakers construe events using grammatical aspects in speech, they did not differ in the use of handedness between the imperfective and perfective speech contexts, and among micro-grammatical aspect contexts. The use of one-handed or two-handed (symmetrical or asymmetrical) gestures may not be framed by the linguistic aspects but the contexts of their use. Maybe due to the principle of least effort (Zipf 1949), single-handed gestures dominated the utterances with different grammatical aspect markers. Two-handed mirror (i.e. symmetrical) gestures were mainly used in the contexts where (1) the original action was done by two hands simultaneously, as in example (2a) below; (2) where two events with the same patient occur at the same time, as in (2b), in which the first event was mostly marked by the durative aspect; (3) where a metaphorical event with implied two-handed action was involved, as in (2c).

In (2a), a famous actress was narrating how she practiced throwing paper fans simultaneously before performing this scene in a film. As the left picture from the original film segment in Fig. 3.9 (2a). shows, the actress grabbed two paper fans downward and threw them upward above her head, and grabbed them again when they fell down at the same time. Simulating the original action in the film, the actress made a grabbing downward gesture in the preparation phrase while saying nà-gè (that), then rotated and threw her two hands upward while uttering the topic shànzi (paper fan). The symmetrical hands as mirror gesture (an upward thrust) that accompanied the spoken reference to the shànzi (paper fan) was expressively the same as the verb liàn (practice), so it was held after the predicate liàn (practice) appeared.

(2) a. [nà-gè shànzi wǒ liàn le wǔ-gè yuè].
That-CLF fan I practice PERF five-CLF month.

“I practiced (throwing) those paper fans simultaneously for five months.”
b. [wǒ ná zhe máojīn níng]
I hold DUR towel wring.
“I held the towel and wrung it.”

c. [[zuìhòu tā yāngāi le] [zhēnshí de shichāng de xiānxìăng]]
finally it cover ACTL real GEN market GEN phenomenon.
“Finally, it covered the real phenomenon of markets.”

In (2b), the actor made two fists as if grabbing a towel and twisted his two hands simultaneously and repeatedly when speaking ná zhe máojīn níng (hold the towel and wring it). The event of holding the towel and the event of twisting the towel were integrated in the symmetrical hands as a mirror gesture, as Fig.3.9 (2b) shows.

In (2c) and Fig. 3.9 (2c), the speaker held his two palm-down hands up to the center in front of his neck as a preparation, then his two similar hands moved apart from the center to either side while uttering yāngāi le (covered), and then repeated it in a
diminishing way when speaking (zhēn)shi de (real); it was held through when the abstract patient (object of the verb) shichăng de xiànxìàng (the phenomenon of market) was mentioned. In this situation, we can see the metaphor REALITY OR TRUTH IS OBJECT. Covering it was acted out as if covering a table or something like that. Such symmetrical hands as mirror gestures occurred very frequently when metaphorical events were depicted.

The two-handed contrast (i.e., asymmetrical) gestures mainly occurred in the contexts where (1) the agent and goal of a motion event were profiled as in (3a), and (2) two different events with different patients happened simultaneously, one as the backgrounding event with durative aspect marker, and the other as a foregrounding one like (3b), and (3) when a reference point of the place where an event happened was profiled as in (3c).

In (3a), there was a long pre-stroke hold during the speaker’s self-correction as the left part in Fig. 3.10 (3a) shows: the left palm-towards-center open hand standing for the “gate of hell” was held statively on the left side as the goal of the motion event, and the right palm-down open hand represented the actors, which moved laterally to the left hand while uttering the aborted dì (earth). Then the fingers of the right hand were inserted between the fingers of the left hand one by one when speaking jinrù le (entered) as displayed in the right part in Fig. 3.10 (3a), expressing the path of the motion in speech; however, the gesture did not only reveal this, but also a very special manner of entering. Finally, the fingers of crossed hands curled down intensely associated with diyù zhī (hell’s) and were held to mén (gate), and this gesture complemented the speech by adding the information of “the immediate closure of the gate of hell after you enter it”. The semantic components of such a metaphorical motion event were integrated vividly and comprehensively into this asymmetrical two-hands as contrast gestures.

(3) a. [[nǐ jiù huì dào jìn rù* hǎoxiàng di*] [jìn rù le d] [yì zu zhī mén]].
you then can reach enter seem earth enter ACTL hell GEN gate

“It seemed that you entered the gate of hell.”
In (3b) and Fig. 3.10 (3b), two simultaneous events with different patients were involved. When introducing the new topic by uttering *ránhòu wōmen jiù* (then we), the left unbounded cupped hand as if holding a cup moved away from his body to the center, and the right palm-down clenched hand as if holding a round object moved downward
towards the left hand. Such an abrupt gesture preceded all the gesture-linked speech. He repeated and enlarged this gesture when the relevant information occurred: the thumb and the index came into contact, like the shape of the round mouth of a cup, and the other fingers stayed slightly curled when uttering *nà zhe gè* (held a) and were held until the *zhī bēi* (paper cup) appeared, which was accompanied by an additional gesture of the right palm-down clenched hand moving rapidly and directly towards the left hand. Then when the predicate of *dào le hěnduō xiāngbīn* (poured a lot of champagne) was spoken, as a reference point, the posture of the left hand was maintained, and the right hand moved cyclically towards the left hand. These asymmetrical hands, contrasting in the gesture, clearly demonstrated that the two hands were involved in different actions at the same time in an event. The movement pattern distinguished pouring a lot of cups of champagne rather than pouring much champagne into the same cup.

In (3c) and Fig. 3.10 (3c), the actress narrated her experience of performing a scene for a film on a sea by driving a boat with the other antagonist. When they started, the sea was normal. However, when they reached the center of the sea, a sudden rising tide occurred. This was located by her left palm down hand as a reference point, where nearby they turned around abruptly as displayed in Fig. 3.10 (3c). The right palm-down open hand representing the boat moved laterally in a smooth way to the center, parallel to her neck, with the speech *chābùduō dōu kuài* (almost) and held through the words *dào zhè’ér de shíhòu* (when reached here), then moved semi-cyclically and very fast to her left side, co-occurring with the speech *jiù guāiwăn le* (turned around).

The results also showed that the gestures accompanying imperfective aspect speech were more likely to be iterated than those in perfective aspect speech, but actually only the progressive aspect-marked utterances were significantly more frequently accompanied by gestures with iteration. This further proves that when taking the progressive aspect perspective to describe an event, the speakers were more prone to focus on the “ongoingness” and event-internal details of that event. As the previous behavioral studies showed, the progressive aspect resulted in more actions and longer duration of actions in a given period of time than the perfective aspect did (Fausey & Matlock 2011; Madden & Zwann 2003; Matlock 2011). The uses of progressive aspect
were more likely to lead to more gestural action with language production here than the perfective aspect and even the durative aspect. Events described with the progressive aspect were homogeneous, with each portion having the same features and occurring successively (Langacker 2008a). The above conceptual features can be argued to have motivated the repeated gestures with the progressive-aspect speech, such as in (4a) and (4b).

In (4a), the speaker narrated that his son worked very hard until 2 or 3 a.m. every night writing and drawing. You can see in Fig. 3.11 (4a) that his index finger touching upon the thumb of his right hand like the head of a pen moved horizontally from the center to the right as if writing three times while uttering dōu zài xiě a huà a (was writing and drawing).

In example (4b), when the speaker mentioned zhè-gè liàng-gè shì’ér (those two things), which referred to the industrial structure transformation and the aging problem, he held up his two palm-down open hands with extended fingers to the center from his legs, then he moved the two hands, one forward, and the other backward alternatively three times while uttering zài pāobù (are running), as Fig. 3.11 (4b) presents. The nucleus of such a metaphorical event was simulated as the concrete event of running.

(4) a. [tā dōu zài <...0.5s> dōu zài xiě a huà a].
   he all PROG all PROG write PRT draw PRT
   “He was writing or drawing all the time.”

b. [zhè-gè liàng-gè shì’ér zài pāobù]
   this-CLF two-CLF things PROG run.
   “These two things are running.”
The durative aspect also focused on the intermediate phase of an event, emphasizing the durativity of an ongoing event. Therefore, in most cases, where the durative aspect-marker co-occurs with a verb to describe an event, the speakers usually hold the gesture stroke rather than repeat it, as in (5a). The gesture accompanying this durative event preceded the speech. During the 0.5s pause before the speech, the gesture of “holding a rope” (the left palm-up fist in Fig. 3.12 (5a) came into being, and it was held until the ending of this utterance. Even when another subsequent event (looking down) was added, the “holding a rope” gesture was maintained rather than repeated. He made another new gesture by pointing downward at the ground under the desk before him with a downward gaze upon uttering kàn zhe (was looking at). The deictic gesture was also held to the ending of this utterance. However, the gestures accompanying two simultaneous events in durative aspect speech contexts were always repeated as in the progressive aspect speech contexts like example (2b). On closer examination, we can find that the gesture for ná zhe máojīn (was holding the towel) (the two grabbing hands) held through the utterance after it came into existence, and the iteration for three times of the gesture resulted from the forgrounding event níng (máojīn) (wring the towel). This is also the case in example (3b). Furthermore, the repeated gestures accompanying the durative aspect utterances may also result from the implied meaning in the discourse as (5b) and Fig. 3.12 (5b). He was describing the situation where people think about whether it is better to deposit money or invest the money to buy a house. When speaking of ná zhe yí-gè fángzì (are having a house), he
moved his two palm-down open hands up and down alternatively and repeatedly three times, as if weighing things on a balance scale.

(5) a. [[//<0.5s> wǒmen sān-gè zài shàng-tou a nǎ zhe shéng]  
we three-CLF at above PRT hold DUR rope

kàn zhe dīxià nǎ-gē].
look DUR below that.

"We three stood at the top of hill, holding ropes and looking down at that below."

b. [ná zhe yì-gè fángzi].
hold DUR one-CLF house.

“(When they) are having a house.”

The first use of durative aspect is typical, while the last two are not so common. Thus, durative aspect speech is more likely to co-occur with single movement gestures.

The perfective aspect takes a holistic view towards the event, so co-speech gestures with it are not likely to be repeated. However, gestures with iteration in the actualized and experiential aspect speech contexts were motivated by the repetitive or multiple readings of utterances, as in example (6a) and (6b).

(6) a. [zījī hái shǒu yòng fánfúzi chǎo le]  
yourself also hand use traditional Chinese characters copy ACTL

"You also copied the whole book with hand in traditional Chinese characters."

b. [nǐ jīn nián yàn guo de yī xīlìè diànyīng dāngzhōng]
Among the series of films you played in recent years.

Fig. 3.13 (6a) Gesture with “copy-ACTL”

Fig. 3.13 (6b) Gesture with “play-EXP”

As shown in (6a) and Fig. 3.13 (6a), the speaker’s thumb touched the index finger on her right hand was molded like a pen head, and it moved laterally from the middle to her right side twice. The mimetic gesture of “writing action” expressed by chāo (copy) here preceded the speech and synchronized with the thematic elements zìjǐ hǎi shǒu yòng fāntǐ (you also (copied it) in the traditional Chinese characters by hand). As copying a script involves a lot of writing action, the repetitive construal implied here is demonstrated by the repeated gesture. In example (6b), the speaker moved her left open hand cyclically twice while uttering jīn nián yǎn guó de (you played in recent years), as Fig. 3.13 (6b) shows. The gestural repetitions were caused by the repetitive readings of the objects yī xīliè (a series of) mentioned later.

3.4.2 Grammatical aspect and gesture complexity

In addition to probing the hypothesis about the relationship between grammatical aspect and gesture forms, this chapter also examines whether macro grammatical aspects and micro aspectual categories differ in gesture complexity during event conceptualization. The findings provide a mixed answer to this question. Whereas imperfective aspect speech contexts were accompanied by more iconic gestures than perfective aspect contexts, actualized aspect-marked utterances also more frequently co-occurred with
iconic gestures. This result thus suggests that when the speakers chose the actualized aspect, a subtype of perfective aspect, to construe events in speech, they still tended to focus on the internal features of them in conceptualization, which were realized via the co-speech gestures. However, the perfective aspect was less likely to occur with representation of the internal structure of an event (perhaps because it prevented conceptual access to it), but rather with highlighting of the resulting static endpoint (e.g., Anderson et al. 2008; Anderson et al. 2013; Comrie 1976; Dowty 1977; Huette et al. 2014). With the other subtype of perfective aspect – the experiential aspect – speech contexts were mainly accompanied by pragmatic and deictic gestures, which encoded little information about the internal features of events. The experiential aspect emphasizes the completion of an event and the effect the past event had on the entities, so the gestures here were more pragmatically related to the immediate discourse and deictically related to the agent or location where an event took place. In this study, the actualized aspect was found to behave similarly to the imperfective aspect rather than to the experiential aspect in terms of gesture complexity about the events. It shows that the actualized aspect did not block access to the mental simulation of the internal features of a described event in language production. This may be interpreted as “aspect may not always indicate mental construal of event structure” to some extent (Parrill et al. 2013: 150). But there are many other possibilities for this result. One possibility is that the types of verb aspect, the inherent aspect of predicates (Vendler 1967), in the actualized and imperfective (the progressive and durative aspect) aspect speech contexts were more similar to each other than to those in the experiential aspect speech contexts. The interactions between the grammatical aspect and the verb aspect may affect event construal not only in speech but also in gesture (Becker et al. 2011). Another possibility is that although the gestures accompanying the actualized, progressive, and durative aspect speech contexts were all more likely to be iconic, they may be iconic in relation to different elements of an event, like the entities (agents, patients), action, path or settings. These points will be examined thoroughly in the following chapters.

In addition, the results also showed that the iconic gestures accompanying the
actualized, experiential and durative aspect speech contexts, which appeared more often in abstract events, were more likely to be metaphoric, while the iconic gestures with the durative aspect utterances, more frequently compatible with concrete events, were more concretely iconic. Why are there such preferences concerning grammatical aspect and event type? This may be related to our cognitive preferences. People tend to have more access to the concrete events which keep the action or maintain the resultant states of that action for a period of time. We may experience the durativity of concrete events like holding something or looking at how others hold something now and then in everyday life. Therefore, the concrete iconic gestures accompanying the durative aspect embody our recurrent durative experiences.

3.5 Conclusions

This study adds to our understanding of how grammatical aspect relates to gestural production in event construal. Gestures accompanying the progressive aspect were more likely to be iterative than those with the other grammatical aspects. These results suggest that the progressive aspect activates the dynamic intermediate phase of an event with more action. In addition, the gestures co-occurring with the imperfective aspect, including both the progressive and the durative aspect, and the actual aspect of the perfective aspect, were more complex with iconicity for the internal features of events than those with the experiential aspect of the perfective aspect. Such results have some implications for aspect studies. First, conflating the sub-aspect categories into the basic aspectual distinctions between perfective and imperfective aspects will probably cover some important differences in their gestural behaviors. Thus, it is of great significance to investigate the grammatical aspect at macro- as well as micro-levels. Second, the effects of grammatical aspectual distinctions on their co-speech gestures are gradient rather than categorical in language production. While constructing situation models, speakers may mentally simulate past events, which they acted in or perceived, for language production. The prototypical features of the events’ internal structure are
represented in speech via the imperfective aspect and co-occurring co-speech gestures, but they are also depicted in the gestures of perfective aspect, particularly the actualized aspect, to some extent. Third, the gestural construal of events was likely to be affected not only by the grammatical aspect, but also by other elements like verb aspect, arguments, and the event type itself. The following chapter will examine the relationship between grammatical aspect and gestures in different event types in terms of concreteness.