Aspectual Structure and the Syntax of Chinese Durative Phrases

Liao, Wei-wen

1. Introduction

The article is aimed to inspect the aspectual structure through the study of the durative phrases in Chinese. I propose that aspect should be decomposed into three layers of representations, which I call the **Three-layered Aspectual Structure** (henceforth, TAS). This proposal corresponds to the big picture in Tenny (2000), who also argues for three aspectual positions in syntax but with little direct evidence (lower, middle, and higher aspects in her terms). There is, however, solid evidence for such a claim in Chinese. First of all, I argue that each of the three aspectual layers is morphologically realized (by the aspect markers –le’s, for example) in Chinese. Second, the interpretations of the durative phrases in Chinese can be classified into three categories (see Liao 2004). I propose that a syntax-semantics mapping is at work between the interpretations of the durative phrases and the representations of aspectual structure.

Section 2 reviews the ramification of aspectual theory, including Reichanbach (1947) and Hornstein (1990), Smith (1991), and Klein (1994). I inductively conclude that aspect should be decomposed into three levels of representations. The first (lowest) level is the compositional aspect level, the information of which is provided by the inherent lexical contents of the verbs. The second (middle) level is the viewpoint level, where the lexical content is restricted in a given topic time in the sense of Klein (1994). The third (highest) level being more functional and tense-related, I call this level the temporal aspect, where the temporal structure is punctualized into point, and a reference point is essential in this aspect level.

Supporting evidence comes from the durative phrases. I point out in Liao (2004) that the durative phrases can have three kinds of interpretations, which is compliant with three levels of the aspectual representations. The interpretations of the durative phrases are classified into the Process-related (P-related, measuring the process), the Target State-related (TS-related, measuring the target state), and the Reference Time-related (RT-related, measuring the time with respect to the reference time) interpretations, as in (1):

(1) a. Zhangsan du-le zhe-ben shu san-ge xiaoshi.
   ZS read-LE this-Cl. book three-Cl hour
   ‘(literally:) ZS read the book for three hours.’ [P-related]

   ZS hit-open window three-Cl. hour
   ‘ZS opened the window for three hours.’ [TS-related]

c. Zhangsan gai-wan zhe-dong fangzi san-tian le.
   ZS build-finish this-Cl. house three-day LE
   ‘It has been three days since ZS built the house.’ [RT-related]

The last category is an innovation by Liao (2004). The RT-related interpretation measures the time span of the event or the post-event state with respect to a reference point. Opposed to the former studies on Chinese durative phrases (Ernst 1987, Li 1987, and Lin
2. The Three-layered Aspectual Structure Hypothesis

In this section I first review three approaches of aspect. However, I focus more on the
notion ‘perfective’ than ‘imperfective,’ due to the better association between the durative phrases and aspect.

My main proposal is in section 2.2, where I propose the Three-layered Aspectual System (TAS), which is a syntax-semantics mapping from the theories I have reviewed. I propose three aspectual levels, which I call composition level, viewpoint level, and temporal level.

2.1 Theory of Aspect
The study of aspect has been the penumbra in the history of generative grammar. Much works have looked into the semantics of the aspect, though few have considered the syntax of aspect. In this subsection I review proposals of Reichanbach (1947) (defended by Hornstein 1990), Smith (1991), and Klein (1994).

2.1.1 Aspect as Tense: Reichanbach (1947) and Hornstein (1990)
Reichanbach (1947) deals with aspect as a part of the tense system, henceforth the Basic Tense System (BTS). BTS is composed by three primitives and two relations. The three primitives are speech time (S), event time (E), and reference time (R). The two relations are simultaneity (represented by ‘,’) and precedence (represented by ‘_’). The proposal of the R point is crucial. The relation between S and R is tense, and between E and R aspect:

\[ \text{Tense} \quad R \quad \text{Aspect} \]

Therefore, three possible combinations can be found, plus three aspect relations, shown as follows:

\[
\begin{array}{c|c|c}
\text{TENSE} & \text{ASPECT} \\
R,S / S,R & \text{present} & E,R / E,R & \text{simple} \\
R_S & \text{past} & R_E & \text{prospective} \\
S_R & \text{future} & E_R & \text{perfect} \\
\end{array}
\]

Reichanbachian tense system characterizes the elements of the tense system in natural languages in a simple and straightforward way, but it is weak in some respects. One of the problems is its weakness in describing the variety of the aspectual patterns. Note especially that the difference between perfect and perfective aspects is obvious, though they both denote closed events (Chung and Timberlake 1985, Binnick 1991, and Smith 1991). However, perfect aspect is present-related (or reference-time-related, to be exact), while the perfective aspect is not necessarily so. The perfective aspect concerns the wholeness of the situation.
regardless of the tense structure or a given reference time.\(^1\) In my opinion, this results in the overlook of the internal temporal structures. The overlook causes problems. Take Chinese examples; different aspectual values in (6) fall into the same temporal structure, as in (7):

(6) a.  Ta jintian-zaoshang du-wan na-ben shu.
   he today-morning read-finish that-Cl book
   ‘He finished reading that book this morning.’

   b.  Ta jintian-zaoshang du-le na-ben shu.
   he today-morning read-LE that-Cl book
   ‘He read that book this morning (regardless of the completion).’

(7) E, R _ S \[this\ morning\] E, R _ S
   \[this\ morning\]

2.1.2  Two Parameters of Aspect: Smith (1991)

Smith (1991) proposes that the aspectual system is two-folded. In one respect, it concerns speakers’ viewpoint on the situation. In the other, it is related to the lexical contents, or the situation types.

The aspectual viewpoints in Smith’s (1991: 91) own words, ‘act like the lens of a camera, making objects (situations) visible to the receivers.’ This definition corresponds to the traditional definition that aspect is how speakers view the situations. Situations are observed metaphorically from its internal structure, i.e., the interval where it takes place and the two endpoints. The endpoints, including initial point (S\(_I\)) and final point (S\(_F\)) (S stands for the situation) play a crucial role in this theory. According to whether S\(_I\) and/or S\(_F\) stand, the aspectual viewpoints fit into three categories:

(8) The Viewpoint Categories in Smith (1991)

| +S\(_I\), +S\(_F\) | perfective |
| +S\(_I\), −S\(_F\) | neutral |
| −S\(_I\), −S\(_F\) | imperfective |

The other ‘parameter’ of aspect is the situation aspect. The situation aspect (also known as situation types or Aktionsart ‘kind of action’) refers to the distinction of the lexical contents of words in terms of the inherent temporal information in the lexicon.

According to Smith (1991), the aspect of a sentence combines a particular situational aspect and a given viewpoint aspect. The process is that the situation locates in a time interval, where there is a viewpoint as well. The speaker then focuses on the situation in that viewpoint.

Smith (1991) tries to combine the two independent but related concepts into one single theory. However, there is a theory-internal contradiction. If, on the one hand, the time intervals in two aspectual systems are split into two parts, a time interval for viewpoint aspect

---

\(^1\) On the other hand, Klein (1994) proposes that the difference between perfect and perfective aspects can be reduced to different temporal representations. The review of his proposal is in the latter section.
and another time interval for situation aspect, it seems that the time interval of situation aspect disappears when combined with the viewpoint aspect. For example, in the imperfective aspect, neither endpoint obtains at the time interval; therefore, in propositions like [John run], or [John sing a song], although the initial point obtains, it disappears (or is simply ignored) when it is combined with the imperfective viewpoint. However, it is not clear what the status is of the interval of the situation aspect. If the interval has a nature to be ignored, why do we have to make the proposal in the beginning? On the other hand, if the two intervals are in the same time axis, the contradiction between viewpoint aspect and situation aspect still exists. That is, the initial point (sometimes the final point as well) holds and does not hold at the same time (see Klein 1994). It is therefore not appropriate to define the two aspectual systems in the same axis of the temporal structure.

2.1.3 Time-related Analysis of Aspect: Klein (1994)

Klein (1994) proposes a time-related analysis for tense and aspect. Like Reichanbach, he proposes three primitives for the temporal structure, utterance time (TU), situation time (TSit) and Topic time (TT). Klein’s major departure from the Reichanbachian system is the proposal of TT. The reference time is a point. It either co-exists or splits from the other two points. However, in Klein’s approach, TT is understood as having a complicated internal structure. In his own words, ‘TT is the time span to which the speaker’s claim on this occasion is confined. This time span can be long … but it may also be relatively short’ (from Klein 1994: 4).

The notion of aspect, in Klein’s term, characterizes ‘the relation between TT and TSit—the way, or sometimes the ways, in which some situation is hooked up to some TT’ (Klein 1994: 6). For Klein, any situation is originally an infinite set; it contains only its particular inherent lexical content. Tense is the process which narrows down the assertion (the finite counterpart of a situation), and aspect characterizes the relationship between situation and its assertion, or the way how TSit is linked to TT.

Another innovation of the Klein is the characterization of the situation types, which are seen as different inherent lexical contents. Klein himself distinguishes three different kinds of lexical contents, called 0-phase, 1-phase, and 2-phase lexical contents, respectively (in Klein’s et al. 2000 terms).

Simply put, Klein’s approach proposes a revision for the lack of lexical content in Reichanbachian approach and the lack of consistency in Smith’s approach. However, Klein’s approach is not without problems. In terms of Klein (1994), the difference between perfective and perfect is a matter of the viewpoint that the speaker takes in a particular situation. In the perfective aspect, the speaker takes the view on the hinge of two phases (from the minus to the plus areas), while in the perfect aspect the speaker takes the view on the post-phase of the event. However, in the perfective aspect, TT itself may be measured. In the perfect aspect, on the other hand, TT does no involve duration, and is always regarded as a punctual point. This can be illustrated by the following contrastive examples:

(9)  
  a. John swam **for ten minutes.**  
  b. John has swum **for ten minutes.**
According to Klein, TT provides the only accessible information for aspectual marking. Therefore, the durative adverb in (10a) arguably modifies the topic time, which covers the TSit in this case. The durative adverb indicates the duration of the TT and the identical TSit. However, in (10b) the durative adverb modifies the duration between the TSit and the TT, not the TT itself, and hence shows a contrast between the two types of sentences. In Klein’s approach, it should be possible to modify the duration of the topic time since the TT is a time span as well even in the perfect aspect. However, this is contrary to the fact. Illustrations in (10) show the temporal representation of (9), respectively:

(10)  

a.  

\[ + + + + + + + + + + \]  

pretime \{ \text{TSit} \} \quad \text{posttime}  

\[ \text{10 min.} \]  

b.  

\[ + + + + + + + + + + [ ] \]  

pretime \{ \text{TSit} \} \quad \text{posttime}  

\[ \text{10 min.} \]  

b’ *  

\[ + + + + + + + + + + [ ] \]  

pretime \{ \text{TSit} \} \quad \text{posttime}  

\[ \text{10 min.} \]  

2.2 Three-Layered Aspectual Structure

Inductively speaking, the theories of Aspect, in general, deal with how speakers view the situations. Speakers observe the situation to assert them. The extraction of the basic element for such observations reduces the complexity to an observer, an observee, and a basis for observation. The observers are the speakers, and the observees the situations. The bases for observation vary, and at least three levels of representations are involved, discussed as follows.

The first level is the composition level. This level corresponds to what traditionally is regarded as the lexical aspect or \textit{aktionsart}. The basis for the observation in this level is the lexical features, which involves the notion of ‘change.’ Theoretically speaking, this level provides the archetype of the lexical information, corresponding to the \( \alpha \)-phase lexical content in Klein’s approach.

The second level is the viewpoint level, corresponding to what is traditionally called grammatical aspect, or the aspectual viewpoint. The basis for observation in this level is the TT, and the internal structure of the situation, including the initial (SI) and final (SF) points of the situation.

The third level is the temporal aspect. The basis for the observation is the Reference point. The reference point functions as an anchor from which the speaker views the situation.

\[ ^2 \text{Klein proposes that TSit is what durative phrases measure (Klein 1994: ch.10). However, I suggest that an adjustment is needed to avoid the above mentioned contradiction. As I show in this work, a fine-grained analysis of the interpretations of the durative phrases argues against Klein’s proposal that the durative only measures the TSit.} \]
Note that the internal structure of the situation becomes invisible. Metaphorically, the internal structure of the situation is ‘condensed’ to punctual point. This level of aspect is best understood in the Reichanbachian approach. Therefore, it is the relative positions between E and R that is at work in this level. The general frame of the hybrid three-layered aspectual structure (henceforth, **TAS**) looks as follows:

(11) Three-layered aspectual system (TAS)

<table>
<thead>
<tr>
<th>aspect</th>
<th>classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>compositional aspect</td>
<td>lexical feature</td>
</tr>
<tr>
<td>viewpoint aspect</td>
<td>[± change]/[ADD TO]</td>
</tr>
<tr>
<td>temporal aspect</td>
<td>basis for observation</td>
</tr>
<tr>
<td></td>
<td>[± telic])</td>
</tr>
<tr>
<td></td>
<td>perfective</td>
</tr>
<tr>
<td></td>
<td>neutral</td>
</tr>
<tr>
<td></td>
<td>imperfective (incl.</td>
</tr>
<tr>
<td></td>
<td>progressive &amp; habitual)</td>
</tr>
<tr>
<td></td>
<td>simple</td>
</tr>
<tr>
<td></td>
<td>perfect</td>
</tr>
<tr>
<td></td>
<td>prospective</td>
</tr>
</tbody>
</table>

The interaction between the three layers is not rigid and subject to specific language strategies. Consider the perfective aspect in English. English is a language without explicit perfective markers. However, the concept of perfectivity is represented by means of the other two levels (past tense plus accomplishment). On the other hand, Chinese employs the resultative verb compounds (RVC) or the secondary predicates to mark the perfective events, while the past simple tense/aspect does not necessarily represent a perfective viewpoint (Smith 1991, Kang 1999, Klein et al. 2000, among others):

(12) a. John ate a hamburger.
     b. *John ate a hamburger, and he is still eating it.

(13) a. Zhangsan chi-wan yi-ge hanbao.
     ‘ZS ate a hamburger.’
     b. Zhangsan jin-tian zao-shang chi(-le) yi-ge hanbao,
     dao xianzai hai zai chi.
     to now still Prog. eat
     ‘(literally) ZS ate a hamburger this morning and is still eating (that hamburger).’
     c. Zhangsan ba yi-ge hanbao chi-de ganganjingjing.
     ‘ZS ate a hamburger completely.’

We have not yet dealt with the question how TAS is mapped to a syntactic model. Recent studies of various languages have shed light on the nature of the functional and the aspectual projections (see Cinque 1999). First consider the proposal by Tenny (2000). From the behaviors between adverbial modifications and core events, she argues that the functional projections are not a one-to-one mapping mechanism. The ‘big picture’ hence looks like the
The interesting point with respect to the proposed theory is that the lexical cycle contains two aspectual levels and the functional cycle contains the other level. This is straightforwardly predicted in TAS. The former two levels (compositional, and viewpoint) corresponds to the inner and middle aspect. The root verb or other resultative projections incorporate/conflate with V₂ provides the archetype of the lexical content (Hale and Keyser 1993, 1997, 2002). The middle aspect provides the boundary of the topic time. These two aspectual levels are subject to the lexical cycle. The next step is to build the functional cycle, where the higher aspectual value is set with respect to a reference time.

Syntactically, the building order of aspectual levels reminds me of the strict cycle condition (SCC) in the structure building process. In the minimalist spirit, especially in Chomsky (2001), the strict cycles (or phases) are taken to be CP and vP. The status of CP as a phase is to introduce functional projections (e.g. sentential level operator). The vP is the projection in which lexical information is built up. In TAS, the compositional level and the viewpoint level have close interactions, and the levels show productive lexical-level verbal morphology (that is, the RVC and the verb-\textit{-le}). On the other hand, little lexical interaction is found in the temporal level, and the system is subject to sentence-final particle (e.g. sentence-\textit{-le} in the sentential level). The strict cyclicity of the TAS corresponds to the syntactic nature of phases. That is, they are cyclically applied, and the latter operations cannot refer back to the past ones. This may suffice as a theoretical support for the syntax of TAS, though many technical details still need working out. In the next section, however, I provide empirical evidence from Chinese durative phrases.
3 The Interpretations of the Durative Phrases in Chinese

The first subsection makes a distinction between three types of durative phrases. I argue that the distinction supports the TAS since the three types of durative phrases measures the time of the three kinds of aspects. Further confirmation comes from the interaction between the overt aspectual markers and the durative phrases, which is discussed in 3.2.

3.1 Three Kinds of Durative Phrases

Contrary to the former analyses (Li 1987 and Lin 2003a), Liao (2004) proposes that the durative phrases should be classified into three types in Chinese. The three types are P-related, TS-related, and RT-related (for expository convenience, I assume readers’ familiarity with theories of Klein and Reichanbach):

(15) P-related (process)
   a. ZS zhe-ben shu du-le liang-ge xiaoshi.
      ZS this-Cl. book read-LE two-Cl. hour
      ‘ZS read this book for two hours, (whether he finished or not.)’
   b. [———]+++ …
      pre-reading { reading } post-reading
      < > or > → two hours

(16) TS-related (target state)
   a. ZS ba chuang da-kai san-ge xiaoshi.
      ZS BA window DA-open three-Cl. hour
      ‘ZS opened the window for three hours.’
   b. ——— ++ …
      pre-opening { opening } post-opening
      < > → three hours

(17) RT-related (reference time)
   a. ZS du-wan shu san-tian le.
      ZS read-finish book three-day LE
      ‘It has been three days since ZS read this book.’
   b. E ____ R, S
      < > → three days

The classification draws a distinction between the two post-event durative phrases, namely, the TS-related and the RT-related. The TS-related durative phrase measures the time span of the target state, in the sense of Parsons (1990), and the RT-related durative phrase measures the time with respect to a given reference time.

As far as TAS is concerned, the distinction brings positive outcome. In the bottom-up order of the TAS, first, the TS-related durative phrase is admissible only to the lexical content which generates a “target state” (or result state in terms of Piñón 1999.) The concept of target state is proposed by Parsons (1990). Take the famous example John threw a ball onto the roof. There are two states involved in this event. One is the resultant state [John has thrown the ball], and the other is the target state [the ball is on the roof]. The resultant state does not
change over time, while the target state may or may not change over some period of time.\(^3\) Therefore, in the case of opening the window in (25), the target state is the window being (resultatively) open, (the window can be closed again.) On the other hand, in cases such as [x swim] and [x love y], no target states are built due to the lexical contents of ‘swim’ and ‘love.’ Therefore, the reasonable assumption is that TS-related durative phrase is linked to the lexical content, or the compositional aspect, in the model of TAS.

The P-related durative phrase corresponds to the viewpoint aspect. The P-related concerns the perfectivity of the situation. In Chinese, the P-related durative phrases concern the time from the initial point to the final point (either arbitrary or natural), which is brought about by the aspectual marker verb-\textit{le}. I discuss the details in the next subsection.

The RT-related durative phrase is mapped to the temporal aspect. We measure the time with respect to a given reference point. There are three possible ‘time axes’ that we can measure, ongoing event, target state, and resultant state, respectively:

1. \((18)\) ZS pao-le san-tian le. \([\text{ongoing event}]\)
   ZS run-\textit{LE} three-day \textit{LE}.
   ‘ZS has been running for three days (until now).’

2. \((19)\) ZS da-kai chuanghu san-tian le. \([\text{target state}]\)
   ZS DO-open window three-day \textit{LE}
   ‘ZS has opened the window for three days (until now).’

3. \((20)\) ZS xie-wan zuoye san-tian le. \([\text{resultant state}]\)
   ZS write-finish homework three-day \textit{LE}
   ‘It has been three days since ZS finished his homework.’

### 3.2 Evidence from Chinese Aspectual Markers

One piece of evidence in favor of the TAS hypothesis comes from the interaction between aspectual markers and the durative phrases. In Chinese overt aspectual markers are found. In the hierarchical (bottom-up) order of TAS, the compositional aspect level often shown by the resultative verb compound or the secondary predicate, as in (21). The viewpoint aspect level can be represented by the verb-\textit{le}, as in (22), and the temporal aspect level by the sentence-\textit{le}, as in (23):

1. \((21)\)
   a. ZS ba chuanghu da-kai. \([\text{RVC}]\)
      ZS BA window DO-open
      ‘ZS opened the window.’
   b. ZS tiao jin-shui-li. \([\text{secondary predicate}]\)
      ZS jump into-water-(in)
      ‘ZS jumped into the water.’

2. \((22)\) ZS xie-le zuoye. \([\text{verb-}\text{le}]\)
ZS write-LE homework
‘ZS wrote his homework.’

(23) ZS xie zuoye le. [sentence-le]
ZS write homework LE
‘ZS has written his homework.’

In the compositional aspect level, the RVC and the secondary predicate bring about the target state in Chinese. The finding corresponds to Piñón (1999), who calls the two target states ‘open’ type and ‘in the water’ type.

Turn to the other two levels. Although the semantic content of the two le’s are controversial, I assume that verb-le denotes a ‘realization’ aspect (Liu 1988), while the sentence-le denotes perfect aspect (see Liao 2004 for full argumentation).

The realization analysis of verb-le argues that verb-le marks only the initial point of the situation, and the completion is not explicitly marked, but is borne only by implicature:

(24) a. ZS zuotian du-le yi-ben shu, keshi mei du-wan.
ZS yesterday read-LE one-Cl. book, but not read-finish
‘ZS read a book yesterday, but he did not finish it.’
b. ZS gai-le yi-dong fangzi, keshi meigai-wan.
ZS buile-LE one-Cl. house, but meibuild-finish
‘ZS built a house, but he didn’t finish it.’

As a perfect marker, the sentence-le introduces the reference point to the temporal-aspectual structure. The contrast between (25a) and (25b) can well illustrate this point. A given reference point ‘1492’ is introduced in (25b), while in (25a) ‘1492’ is the event time (as well as the reference point):

(25) a. 1492 nian, Gelunbu faxian-le xin-dalu.
1492 year Columbus discover-LE new-continent
‘In 1492, Columbus discovered the new continent.’
b. #1492 nian, Gelunbu faxian xin-dalu le.
1492 year Columbus discover new-continent LE
‘In 1492, Columbus had discovered the new continent.’ (contrary to the fact)

Take the durative phrases into the structure. We find a close relation between the aspectual markers and the durative phrases. Given the appearance of le, the P-related usually occurs with the verb-le, the TS-related can occur without either le’s, and the RT-related must occurs with sentence-le, as in (15) to (17). The appearance of le’s can be explained by the fact that the P-related durative phrases corresponds to the viewpoint level, and verb-le is in the charge of this level, marking out the initiation of the situation. The TS-related modify the compositional level; therefore, either secondary predicate or RVC appears so that the modification is ensured. The RT-related, on the other hand, accompany the sentence-le because the time is measured with respect to a given reference point (provided by sentence-le).

4 The Syntax of Aspect and the Durative Phrases
This chapter discusses how Chinese aspect and durative phrases are represented in syntax. However, due to the close relation between aspect markers and durative phrases, we need an understanding of the syntax of the Chinese aspectual markers, especially the structures of two le’s. Other issues concern the structural configuration of the Chinese durative phrases. First, concerning the directionality of adjuncts in Chinese, which are conventionally leftward, we find that the durative phrases are idiosyncratically rightward. The other issue is about the ordering between the objects and the durative phrases. The common ordering between verbs (V), durative phrases (D), and objects (O) is V-O-D, while in certain cases, Chinese allows the ordering V-D-O, with D-O forming a DP constituent (linked by prenominal modifier marker de).

The syntax of aspectual markers is discussed in 4.1. I suggest that the verb-le heads a light verb projection within the framework of the light verb syntax proposed by Lin (2001). On the other hand, the sentence-le hosts the higher aspectual projection, AspP, and the AspP is a raising category, which triggers the comp-to-spec movement. The motivation of this movement is for the dynamicity checking proposed in Shen (2004).

Once the structure of aspect is clear, the placement of the durative phrases into the structure follows. I illustrate the structures in 4.2, with evidence from scope tests.

### 4.1 The Structure of Chinese Aspect

Under the proposed analysis, the durative phrases have close relationship with the projections of Chinese le’s, the structures of which also raise many controversies (see Cheng 1990, Sybesma 1997, 1999, and Jo-wan Lin 2000, 2003b). For conciseness, I assume the light verb syntax framework in Tzong-hong Lin (2001). The verb-le contributes the REALIZE eventuality to the main predicate through ‘conflation’ (in the sense of Hale and Keyser 1993, 1997, 2002). The level of the conflation is language-specific. Lin (2001) assumes that the conflation applies in the sentence level (s-syntax) in Chinese (lexicon level in English, l-syntax). The general picture of the structure is as follows:

\[(26) \quad \text{vP} \quad \rightarrow \quad \text{argument-selecting light verb}\]

\[
\text{v} \quad \text{vP} \quad \rightarrow \quad \text{aspectual light verb = viewpoint aspect}\]

\[
\text{v} \quad \text{VP} \quad \rightarrow \quad \text{matrix verb}\]

\[
\text{V} \quad \text{VP} \quad \rightarrow \quad \text{secondary predicate}\]

Verb-le as well as other aspectual markers, –zhe and –guo, is treated as the head of the (viewpoint) aspectual light verb. The verb is base-generated in the matrix V, a series of head movements leads to the surface form. The matrix V\(^0\) incorporates (or conflates) with object v, aspectual v, and subject v, cyclically, subject to the Head Movement Constraint (HMC). See (27):

\[(27) \quad \text{a. Ta xie-wan-le zhe-feng xin.}\]
he write-finish-LE this-Cl letter
‘He wrote the letter.’

b. vP
   /   \
  v'  vP
     /   \ 
    v   le
     |   V
     |   VP
     |   |
     |   xie
     |   wan

This analysis has the advantage that it is mapped to the TAS in a homomorphism. Notice that in Klein (1994) and Klein et al. (2000), verb itself owns an X-phase lexical content (X=0, 1, 2) and then the lexical content is linked to a given TT (provided by the aspectual light verb). The proposed framework works in just same way.

As for the sentence-`le (le2), Shen (2004) notes that there is a dynamicity [±Dyn] agreement between Chinese sentence final particle (SFP) and the eventuality predicates (in the sense of Lin’s light verbs). Therefore, the sentence-`le agrees with a dynamic light verb, while the sentence-`ne agrees with a static one. Compliant with the temporal aspect, the sentence-`le heads an AspP, which is higher than the whole light verb construction in (32). I propose that the AspP is a raising category in the sense of Kayne (1994) due to the strong dynamic feature of the light verb, which is checked by the Asp0 through head-spec relation, according to Chomsky (1995):

As Tzong-hong Lin (p.c.) points out, there is empirical evidence for this claim. Note that CED (Huang 1982) directly prohibits movements from the domain of the moved vP. Assume Tsai’s (1994) theory of A’-dependencies.4 Adjectival wh-phrases, but not nominal wh-phrases,

4 There is another consequence in Tasi’s (1994) proposal. That is, subjacency and CED also applies at LF, contrary to the proposal in Huang (1982).
must move at LF, instead of being unselective bound by an operator at CP-level. The proposed analysis observes the proposal in Tasi (1994). The VP-level (manner adverbs) adjectival wh-phrase zenme ‘how’ cannot move out of the moved vP:

(29) a. Laowang zenme zhu niurou (*le)?  
Laowang how(manner) cook beef LE  
‘How does Laowang cook the beef?’

d. ZS zenme chi-wan-le hanbao (*le)?  
ZS how eat-finish-LE hamburger LE  
‘How did ZS eat the hamburger?’

The structure of the TAS is now clear. The compositional aspect shows different lexical content, which is represented in Chinese by the main verb and the composition of secondary predicate or RVC construction. The viewpoint aspect is represented by the verb-le (and guo, zhe) projection, the function of which is to link the lexical content to a given viewpoint (by observing the initial and final point of the lexical content), and syntactically, the linking is achieved through head incorporation of V⁰ to v⁰. The temporal aspect is shown by the sentence-le. Semantically, it condenses the internal structure to a point (and does not look back to the internal structure again), and linked it to a given reference time. The complement of AspP, the vP, then raises to Spec, AspP for dynamicity checking. Since the vP is also an impenetrable domain (Chomsky 2001), the semantic property is presumably deduced.

4.2 The Structure of the Chinese Durative Phrases

First, the P-related durative phrases are licensed by the verb-le, the structure looks like (30):

(30) a. ZS [v xie-le zhe-feng xin san-tian.]  
ZS write-LE this-Cl letter three-day  
‘ZS wrote the letter for three days.’

The durative phrase is licensed by the v head which requires a homogenous event (Moltmann 1991). This is achieved by the conflation from V to v. Notice that the adverb is illicit when the verb chunk is not homogenous in nature. Therefore, xie-wan-le ‘write-finish-LE’ would be ruled out in presence of a durative adverb. The verb chunk further raises to subject-selecting light verb, hence the correct S-V-(O)-D ordering in the surface form (D for durative adverb).
Consider next the TS-related durative phrases. I have argued that the TS-related durative phrase requires a target state. Syntactically, two sources of the target state are available, corresponding to the [open] and [in the water] types of explicit result states in Piñón (1999). The first source is the lexically denoted target state. This can be illustrated by verbs like da-kai ‘open,’ da-pao ‘dispell,’ ju-gao ‘raise up,’ etc.\(^5\) Another syntactic source of the target state comes from resultative secondary predicate. For example, (tiao) jin-shuili ‘(jump) into the water,’ (duo) dao-shanshang ‘hide in the mountains,’ (pa) dao-shushang ‘(climb) onto the tree’ (see also Ramchand 2003 for a lexical decomposition approach of the RP). Two kinds of structure involved in the TS-related durative phrases are shown in (31) and (32):

(31) a. ta da-kai chuanghu wu-fenzhong.
   ta do-open window five-minute
   ‘He opened the window for five minutes.’

   b. 
   \[
   \text{VP} \\
   \text{V} \quad \text{VP} \\
   \text{V} \quad \text{AP} \\
   \text{da} \quad \text{ADVP} \\
   \text{[DO]} \quad \text{wu-fenzhong} \quad \text{kai}
   \]

   he jump-into water-in five-minute
   ‘He jumped into the water for five minutes.’

   b. 
   \[
   \text{VP} \\
   \text{V} \quad \text{VP} \\
   \text{V} \quad \text{PP} \\
   \text{tiao} \quad \text{ADVP} \quad \text{reanalysis} \\
   \text{wu-fenzhong} \quad \text{P} \quad \text{NP} \\
   \text{jin} \quad \text{shuili}
   \]

Following Hale and Keyser (1993, 1997, 2002), these phrases are base-generated as VP-complement, denoting the result state. If Lin (2001) is on the right track, and conflation

applies in Chinese directly in the level of s-syntax, I propose that the TS-related durative adverb is base-generated in the Spec position of the root phrase or the secondary predicate.6

Turning to the RT-related durative phrases, we find that the RT-related durative phrases must associate the measured states with a reference point. Possible candidates of the measured states are the resultant state, process, or the target state. Since the reference point is provided by the sentence-le, it is reasonable to declare that the RT-related durative phrases are licensed under the AspP (see 33):

(33)

Empirical evidence in favor of the proposed analysis is plentiful. Consider the negative scope first. Empirically, mei-you ‘not-have’ negates the main predicate, and syntactically the NegP is base-generated between AspP and subject-selecting vP. Accordingly, it also raises to the Spec, AspP. Therefore, both the P-related and the TS-related are predicate-internal, and are supposed to be negated along with the predicate. On the other hand, the RT-related, being predicate-external, is immune from the negative scope. See the following examples:

(34)

a. ZS mei-you chi hen-jiu.
   ZS not-have eat very-long
   ‘ZS did not eat for a long time.’

b. ZS mei-you kai chuanghu hen-jiu.
   ZS not-have open window very-long
   ‘ZS did not open the window for a long time.’

c. ZS mei-you zhe-yang zuo hen-jiu le
   ZS not-have this-way do very-long LE
   ‘It has been long since ZS did so last time.’

(35) The RT-related durative phrases: immune from negation

AspP

6 To make things more straightforward, the subject position of the secondary predicate is omitted in this discussion. Actually, this is an issue that deserves a serious argumentation. For simplicity, I assume the generalized control approach in Huang (1984). In (34b), this leaves a Pro (bound by the object) in the subject position of the secondary predicate, to which the adjectival head attribute its property. Another possibility to consider this structure is that X’ (X= P or A) itself undergoes reanalysis, as in Larson (1998), and move as a head, as illustrated in (35b).
On the other hand, consider the modal scope. Suppose that modalP is base-generated superior to AspP. It is quite interesting that the modal scope is over the P-related, TS-related, and RT-related durative phrases. The fact that RT-related durative adverb also falls into the modal scope indicates that the RT-related must be generated not higher than modal but higher than NegP, leaving only the AspP, as desired:

(36) a. ZS dagai du daxue wu-nian-duo le.
   ZS probably read college five-year-more LE
   ‘ZS probably have studied in college for more than five years.’

b. ZS keyi yong zuo-shou xie-zi henjiu le.
   ZS can use left-hand write-word very-long LE
   ‘It has been long that ZS can use his left hand to write.’

(37)

5 Consequences and Remarks

Turn to the semantic issue of telicity in Chinese. An asymmetry between Chinese and English telic expressions is found. Chinese allows, while English rejects, the existence of the durative phrases with a telic expression:

(38) a. ??John ate the hamburger for three hours.

b. Zhansan chi-le zhe-ge hanbao san-ge xiaoshi.
   ZS eat-LE this-Cl. hamburger three-Cl. hour
   ‘ZS ate the hamburger for three hours.’

The problem becomes less complicated once the durative phrases are in theory related to aspect. Under the proposed theory, this puzzle can be handled with in a straightforward manner. The verb-\textit{le} denotes a realization aspect (Liu 1988), and therefore licenses the P-related durative phrases (specify the duration of the realization). To be more general, the verb-\textit{le}, being a realization marker which specifies the initial but not the natural final point,
introduces an arbitrary final point of the situation by default (as in the case of English activities). This arbitrary final point allows the events to be activity-like (Smith 1991). The same logic applies to English accomplishments as well, where the arbitrary final point of the situation is introduced by contexts or by other syntactic means (durative preposing, for example). The arbitrariness then ‘atelicizes’ the telic events, causing accomplishments to become activities. Therefore, the durative phrases can access to the atelicized events in the following sentence:

(39) For three hours, John ate the hamburger.7
    a. ‘John brought about the event that the eating of the hamburger lasts for three hours.’
    b. *‘John naturally entered the state that the hamburger was eaten for three hours.’

In (41a) an arbitrary final point is introduced, and the sentence put emphasis on John’s arbitrariness to eat the hamburger intentionally for three hours (Verkuyl calls it ‘forced reading’). (41b) sounds bizarre since the arbitrary final point has replaced the natural point, and this is why the sentence loses its ‘accomplishment’ flavor.

As for syntax, the most obvious question is why Chinese legitimates the durative phrases as the only rightward adjuncts. Especially, the empirical data seems to violate the conclusion met in Lin and Liao (2003), which argues that Chinese adjuncts observe the uniform leftward directionality. The fact that Chinese does not allow rightward adjunction is obvious in many cases (see Tang 1990 for detailed analysis on Chinese adjuncts). For example, locative, temporal (both deictic and anaphoric), and manner adverbs are instances of left adjuncts. Assume they are vP-level (written as PrP in Bowers 1993 and Tang 1990) or IP-level, following Tang (1990). The structures hence look like (40a) and (40b). CP-level adjuncts are also labeled as left adjuncts, such as (40c):

(40) a. ZS [vP zai-jia-li [\v kan dianshi]].
    ZS in-home-in watch TV
    ‘ZS is watching TV at home.’
    b. ZS [IP mingtian [\p zai-jia-li xie baogao.]]
    ZS tomorrow in-home-in write paper
    ‘ZS is going to write his paper tomorrow.’
    c. [cong jintian-qi, [\c ZS shi boshi le.]
    from today-on ZS BE PhD LE
    ‘From today on, ZS is a Ph.D.’

Durative phrases escape this convention and appear in the rightward position. How can this be? Of course we can stipulate an ad hoc rule that only durative phrases are rightward, but this is undesirable for theoretical elegance. Under the proposed theory, nevertheless, the surface ordering is actually derived from series of movements, leaving the durative phrases
‘stranded’ in the sentence-final position. Thus, the general structures shape like the following, in a unified form:

\[ [YP \ldots X^0 \ldots XP \ldots \text{RT-durative} \ldots Y \ [ZP \ldots P-/TS \ldots \text{durative} \ldots t \ i \ldots]] \]

In the case of P-related, the YP=VP, and X^0=V^0, which head-adoins to the v^0. In the TS-related, YP=VP, and ZP represents the resultative phrase, which moves up to VP and stranded the durative phrase. The only exception is the RT-related, in which the durative phrase appears in the YP, instead of ZP, and the XP=vP, which raises to spec of YP for dynamicity checking. Therefore, the uniform directionality of adjunct still holds in the proposed theory.

Another interesting question is to examine the co-existence of the durative phrases. Skeptical readers must have noticed that the durative phrases should be allowed to occur more than once in a sentence since the proposed theory makes a distinction of the hierarchical positions of the different durative phrases. However, the fact is more complicated than expected. Consider the following sentences:

\[(42)\]

   ZS  open-LE window five-minute five-day one-Cl-month LE 
   (P-related, TS-related, and RT-related)
b. *ZS kai-le chuanghu wu-tian yi-ge-yue  le. 
   ZS  open-LE window five-day one-Cl-month LE 
   (TS-related and RT-related)
   ZS  open-LE window five-minute five-day 
   (P-related and TS-related)
d. *ZS kai-le chuanghu wu-fenzhong yi-ge-yue  le. 
   ZS  open-LE window five-minute one-Cl-month LE 
   (P-related and RT-related)

The combinations of the P-related, the TS-related, and the RT-related above show the coexistence restriction, although we still can find the grammatical sentences like the following:

\[(43)\]

a. ZS meitian paopu ban-xiaoshi san-nian le. 
   ZS everyday run half-hour three-year LE 
   ‘For three years, ZS has been running for half hour every day.’ 
   (P-related and RT-related)
b. ZS meitian kai chuanguhu wu-fenzhong san-nian le. 
   ZS everyday open window five-minute three-year LE 
   ‘For three years, ZS has opened the window for five minutes every day.’ 
   (TS-related and RT-related)

After a close examination of the possible combinations of three kinds of durative phrases, the coexistence restriction actually turns out to be positive evidence for the proposed analysis.

Notice that the in the licit cases, the RT-related durative phrase selects a habitual event, which is taken out by the adverbial meitian ‘every day.’ How do we explain this fact? I think the
answer lies in the structure. The license of the RT-durative phrase is achieved through the head-spec agreement in the AspP. Habitual events arguably allow the RT-related durative phrases to modify the homogenous habitual state. The habit itself is homogenous, and can take either the P-related or the TS-related durative phrases. On the other hand, the other three states do not take a durative phrase itself because otherwise they become bounded per se:

(44) AspP
    \[\text{vP} \quad \text{Asp'}\]
    \{i. habitual
       ii. resultant state
       iii. target state
       iv. process\}
    RT-Dur. Asp'
    Asp t

Another question is about the coexistence restriction between the P-related and the TS-related. I attribute this fact to the general linguistic phenomenon. Higginbotham (1994) may have noted similar facts. Actually, he radically claims that the secondary predicate can, in some cases, function as the main predicate:

(45) &The boat floated under the bridge.
    a. ‘(atelic reading): Under the bridge, the boat floated.’
    b. ‘(telic reading): The boat went under the bridge in the manner of floating.’

Higginbotham (1994) claims that when under is taken as being telic, meaning ‘go under.’ The PP under the bridge functions as the main predicate:

(46) float (the boat, e1) & under (the bridge, e1, e2) & in an hour (e1, e2)

He claims that it is a language specific property of English that the locative PP can function as the main predicate. However, I suggest that Chinese is the same as English with respect to this property. Therefore, not only in the P-related, but in the TS-related interpretations as well, it is actually the main predicates that receive the aspectual marking, and are able to license the durative phrases. This point again supports the proposed analysis.

6 Conclusion
I have presented the syntax-semantics interface of the aspect and the durative phrases in Chinese. The interaction between the aspect and the durative phrases becomes more transparent. To summarize, the durative adverbs in Chinese should be classified into three kinds of interpretations, P-related, TS-related, and RT-related. This classification is isomorphic to the decomposition of the aspectual structure. The TS-related modify the root projection or the resultative predicate. The P-related modify the viewpoint aspect projection (verb-le). The RT-related modify the temporal aspect projection (sentence-le). The distinction not only solves the semantic issues of telicity, but also characterizes an elaborated architecture of aspect and the durative adverbs.

REFERENCES


Lin, Jo-wang. 2003a. Event Decomposition and the Syntax and Semantics of Durative