The semantics of $yuè V_{\text{non-gradable}} yuè A$ in Mandarin Chinese: coercion and the necessarily temporal reading

Xiao Li & Carlos A. Fasola

Abstract

While the existing literature on $yuè...yuè$ in Mandarin Chinese has mostly focused on examples where both $yuè$ appear in front of a gradable predicate (i.e., a gradable adjective or a gradable verb) (hence, $yuè A/V_{\text{gradable}} yuè A/V_{\text{gradable}}$), in the paper we focus on a group of $yuè...yuè$ sentences where the first $yuè$ appears in front of a non-gradable verb, and the second $yuè$ appears in front of a gradable adjective ($yuè V_{\text{non-gradable}} yuè A$). We show that this type of $yuè...yuè$ sentence expresses a necessarily temporal reading that $yuè A/V_{\text{gradable}} yuè A/V_{\text{gradable}}$ does not have. We argue that this necessarily temporal reading falls out as a result of coercion that forces the $VP_{\text{non-gradable}}$ to have an ordered domain parallel to that of a gradable adjective, where events like degree intervals are totally ordered under the proper subinterval relation and have a common starting point. Our analysis of $yuè$ suggests that the proper subinterval relation is not only compatible with comparisons of degree intervals but also compatible with comparisons of event intervals. Therefore, it encodes a more general notion of comparison than the greater than relation.

Key words: $yuè...yuè$, gradability, coercion, the necessarily temporal reading
1. Introduction

yuè...yuè sentences in Mandarin Chinese, exemplified in (1), have recently received much attention in the literature (Chao 1968, Li and Thomas 1981, Hsiao 2003, Lin 2007, Liu 2008a and references therein). It is generally assumed that they are parallel to the English the –er/more ... the –er/more construction in describing a positive correlation between two sets of degrees. For instance, (1a) describes a positive correlation between the degree to which an apple is big and the degree to which an apple is sweet (i.e., an increase in the size of an apple is accompanied with an increase in its sweetness).

(1) a. Píngguǒ yuè dà, yuè tián.
   apple big sweet
   ‘The bigger an apple is, the sweeter it is.’
   b. Zhāngsān yuè xǐhuān Lìsì, wǒ yuè tǎoyàn Lìsì.
      like I resent
   ‘The more Zhangsan likes Lisi, the more that I resent Lisi.’

Interestingly, yuè appears not only in front of a gradable predicate (e.g., a gradable adjective like da ‘big’ or a gradable verb like xǐhuān ‘to like’), but also in front of a non-gradable verb such as pāo ‘to run’ or tiào ‘to jump’, as shown in (2).

(2) a. Zhāngsān yuè pāo yuè kuài.
      run fast
      ‘Zhangsan ran faster and faster.’
   b. Zhāngsān yuè tiào yuè gāo.
      jump tall/high
      ‘Zhangsan jumped higher and higher.’
      ‘Zhangsan became taller and taller from jumping.’

The gradability of a predicate can be decided by whether it can be modified by a degree modifier such as hěn ‘very’, as shown by the contrast between (3) and (4).

(3) a. zhè gè píngguǒ hěn dà.
      this Cl apple very big
      ‘This apple is very big’
   b. Zhāngsān hěn xǐhuān chí píngguǒ.
      very like eat apple
      ‘Zhangsan likes eating apples.’

(4) a. *Zhāngsān hěn pāo.
      very run
   b. *Zhāngsān hěn kū.
      very cry

In the paper we focus on yuè ... yuè sentences like (2) where the first yuè appears in front of a non-gradable verb and the second yuè appears in front of a gradable adjective (hence, yuè V_non-
gradable yuè A).\(^1\) We show that this type of yuè ... yuè sentence is semantically distinct from yuè ... yuè sentences like (1) where both yuè appear in front of a gradable predicate (hence, yuè A/V\(_{\text{gradable}}\) yuè A/V\(_{\text{gradable}}\)). The former has a necessarily temporal reading that the latter do not have. Their semantic difference can be seen by comparing the truth-values of (2a) in the two scenarios in (5).

\[(5)\] Zhangsan runs on treadmill three times per week. Below was his training schedule for the past week.

a. Scenario A:
   Day 1: Zhangsan ran 5 miles; his running speed was 5.3 mph.
   Day 2: Zhangsan ran 4 miles; his running speed was 5.2 mph.
   Day 3: Zhangsan ran 3 miles; his running speed was 5.1 mph.

b. Scenario B:
   Day 1: Zhangsan ran 3 miles; his running speed was 5.1 mph.
   Day 2: Zhangsan ran 3 miles; his running speed was 5.2 mph.
   Day 3: Zhangsan ran 3 miles; his running speed was 5.3 mph.

The two scenarios in (5) differ in that in scenario B speed increases over time, whereas in scenario A, it does not. Intuitively, (2a) is true in scenario B but not in scenario A. In fact, (2a) can only be true in scenarios in which speed increases with time, and so may be described as having a necessarily temporal reading.\(^2,3\)

So the question is: why is it the case that when yuè precedes a non-gradable verb, the sentence receives a necessarily temporal reading? The goal of this paper is to provide an answer for this question. We argue that the necessarily temporal reading falls out as a result of coercion that forces a non-gradable VP (e.g., Zhangsan ran) to have a totally ordered domain parallel to that of a gradable predicate, where events like degree intervals on a degree scale, share a common starting point and stand in a proper subinterval relation.

---

\(^1\) In the appendix of the paper, we discuss, albeit briefly, the semantics of two other yuè ... yuè patterns: yuè A yuè V where the first yuè appears in front of a gradable predicate and the second yuè in front of a non-gradable verb, and yuè V yuè V where both yuè appear in front of a non-gradable verb.

\(^2\) The necessarily temporal reading of (2a) can be alternatively expressed through the idiomatic expression yuè lái yuè, as shown in (i). yuè lái yuè, with the first yuè preceding the non-gradable verb lái ‘to come’, only precedes a gradable predicate and adds a temporal reading to the sentence. Liu (2008b) provides a detailed discussion on this expression.

(i) Zhāngsān páo-de yuè lái yuè kuāi.
   ‘Zhangsan ran faster and faster.’

\(^3\) It is important to note that the necessarily temporal reading discussed in this paper is distinct from the ‘temporal reading’ that Lin (2007) claims the yuè...yuè example in (ii) to have. The sentence in (ii), where both yuè appear in front of a gradable adjective, describes a positive correlation between the degree of hotness associated with times and the degree of my uncomfortableness. It does not require the degree of hotness to increase over time.

(ii) Tiānqì yuè rè, wǒ jiù yuè būshūfu.
   ‘The hotter the weather is, the more uncomfortable I feel.’ (Lin 2007: 195)
We will compare this analysis to two alternative proposals. One assumes that non-gradable verbs can lexicalize a time argument but not a degree argument, while gradable predicates can lexicalize a degree argument but not a time argument. When yuè composes with a V\textsubscript{non-gradable}, the result is a set of pairs of events ordered based on times (Li and Fasola, 2010). The other assumes that a non-gradable verb can be associated with a particular degree that measures the cumulative amount of an event. When yuè composes with a V\textsubscript{non-gradable}, the result is a set of pairs of events ordered based on the cumulative amounts they are associated with. We show that although both analyses can successfully capture the necessarily temporal reading of yuè V\textsubscript{non-gradable} yuè A, they face problems which are not easy to solve.

The rest of the paper is structured as follows. Section 2 provides a brief review of Lin (2007)’s analysis of yuè … yuè, where the limitations of his analysis are discussed. Section 3 introduces the main assumptions and puts forward a coercion-based account. We argue that the necessarily temporal reading falls out as a result of coercion that forces the non-gradable VP to have a totally ordered domain parallel to that of a gradable predicate. Section 4 formalizes the analysis and discusses two welcome results that follow from it. Section 5 compares our account to two alternative analyses, and shows that the former fares better. Section 6 considers the implications of our analysis. In the appendix, we briefly look at the semantics of the two other yuè … yuè structures: yuè A yuè V\textsubscript{non-gradable} and yuè V\textsubscript{non-gradable} yuè V\textsubscript{non-gradable}.

2 Lin (2007)’s analysis of yuè…yuè

The most detailed semantic analysis of yuè … yuè is due to Lin (2007), who essentially models the meaning of yuè … yuè after that of the -er/more…the –er/more in Beck (1997). He argues that yuè … yuè sentences are bi-clausal, and they express a relation between two sets of situations (individuals or times) ordered based on degrees. Specifically, yuè has the meaning in (6).\footnote{In the paper, we make use of the following types: type e for individuals, type d for degrees, type i for times, type v for events, type s for situations, and type t for truth-values.}

\begin{equation}
[[\text{yuè}]] = \lambda P_{<d, <s, t>} \lambda d_{1}d_{2}\lambda s_{1}s_{2}[P(d_{1})(s_{1}) \land P(d_{2})(s_{2}) \land d_{1} < d_{2}]
\end{equation}

On this semantics, the first argument of yuè is a predicate, $P$, that denotes a relation between degrees and situations (of type $<d, <s, t>$). When yuè composes with $P$, the result is, essentially, an ordering on the set of situations corresponding to the order of the degrees that they are related to via the input property.

To demonstrate how (6) works, let us consider the example in (7). Lin assumes that (7) has the LF in (8).

\begin{equation}
\text{7) } \text{nǐ yuè shēngqì, tā (jiù) yuè gāoxìng.} \\
\text{you angry then he then happy} \\
\text{‘The angrier you are, the happier he is.} \\
\text{(Lin 2007: 169)}
\end{equation}
The structure in (8) consists of two clauses CP₁ and CP₂, which are c-commanded by a covert universal quantifier ∀.

Following the degree-based approach to the semantics of the adjective, Lin assumes that the adjectives shēngqì ‘angry’ and gāoxìng ‘happy’ denote a relation among individuals, degrees and situations, as shown in (9). When yuè composes with an IP, CP₁ denotes a relation between degrees (i.e., d₁ and d₂) and situations (i.e., s₁ and s₂) such that you are angry to degree d₁ in s₁ and you are angry to degree d₂ in s₂; d₂ is greater than d₁, as shown in (10a). CP₂ denotes a relation between degrees, d₃ and d₄, and situations, s₃ and s₄, such that he is happy to degree d₃ in s₃ and he is angry to degree d₄ in s₄; d₄ is greater than d₃, as shown in (10b).

(9) a. \[[\text{shēngqì}] \equiv \lambda x_1 \lambda d_1 \lambda s_1. \text{angry}(x_1)(d_1)(s_1)\]
b. \[[\text{gāoxìng}] \equiv \lambda x_2 \lambda d_2 \lambda s_2. \text{happy}(x_2)(d_2)(s_2)\]

(10) a. \[[\text{yuè nǐ shēngqì}] \equiv [[\text{yuè}]][\lambda d_1 \lambda s_1. \text{angry}(\text{you})(d_1)(s_1)] \text{ and } \lambda d_2 \lambda s_2. [\text{angry}(\text{you})(d_1)(s_1) \wedge \text{angry}(\text{you})(d_2)(s_2) \wedge d_2 > d_1]\]
b. \[[\text{yuè tā gāoxìng}] \equiv [[\text{yuè}]][\lambda d_3 \lambda s_3. \text{happy}(\text{he})(d_3)(s_3)] \text{ and } \lambda d_4 \lambda s_4. [\text{happy}(\text{he})(d_3)(s_3) \wedge \text{happy}(\text{he})(d_4)(s_4) \wedge d_4 > d_3]\]

The universal quantifier has the semantics in (11a). It combines with CP₁ and CP₂, and yields the result in (11b):

(11) a. \[[\forall] \equiv \lambda G_{\text{ad.}}. \lambda d_1 \lambda s_1 \lambda d_2 \lambda s_2. \forall d_1 d_2 s_1 s_2 [G(d_1)(d_2)(s_1)(s_2)] \rightarrow \exists d_3 d_4 s_3 s_4 [Q(d_3)(d_4)(s_3)(s_4)]\]
b. \[[\forall \text{yuè nǐ shēngqì, jǐu yuè tā gāoxìng}] \equiv \forall d_1 d_2 s_1 s_2 [\text{angry}(\text{you})(d_1)(s_1) \wedge \text{angry}(\text{you})(d_2)(s_2) \wedge d_2 > d_1] \rightarrow \exists d_3 d_4 s_3 s_4 [s_1 \leq s_3 \wedge s_2 \leq s_4 \wedge \text{happy}(\text{he})(d_3)(d_4) \wedge \text{happy}(\text{he})(d_4)(s_4) \wedge d_4 > d_3 \wedge R(\langle d_1, s_1 \rangle, \langle d_3, s_3 \rangle) \wedge R(\langle d_2, s_2 \rangle, \langle d_4, s_4 \rangle)]\]

The formula in (11b) says: for any pair of degrees d₁ and d₂, and any pair of situations s₁ and s₂ such that you are angry to degree d₁ in s₁, and you are angry to degree d₂ in s₂, and d₂ is greater than d₁, there exists a pair of degrees, d₃ and d₄, and a pair of situations, s₃ and s₄, such that s₃ is an

---

5 In Lin’s analysis, R is a causal relation that relates degrees in CP₁ to degrees in CP₂. Liu (2008a) points out that R does not have to be causal. For instance, (1a) does not necessarily involve a causal relation between the size of an apple and the degree of its sweetness.
extended situation of \( s_1 \) and \( s_4 \) is an extended situation of \( s_2 \). He is happy to degree \( d_3 \) in \( s_3 \), and he is happy to degree \( d_4 \) in \( s_4 \). \( d_4 \) is greater than \( d_3 \). In short, (11b) says: there is a positive correlation between the degree to which you are angry and the degree to which he is happy, which correctly captures the meaning of (7).

Given the semantics in (6), \( yuè \) is a degree modifier that composes with predicates that contain a degree argument. For \( yuè \ldots yuè \) sentences like (2), Lin follows Doetjes (1997) in assuming that verbs are also gradable: they have a logical form parallel to gradable adjectives, with a degree argument that measures the situation, or event, of the verb. Verb \( pào \) ‘to run’, for example, denotes a relation among individual \( x \), degree \( d \) and situation \( s \) such that \( x \) has done \( d \)-quantity of running in situation \( s \) (Lin 2007: 187).

(12) \[
\text{[[pào]]} = \lambda x. \lambda d. \lambda s. \text{run}(x)(d)(s)
\]

With this assumption, the \( yuè \ldots yuè \) sentence in (2a)(repeated below) can be interpreted just like (7). Below we demonstrate how the semantics of (2a) is derived under Lin’s analysis. Given that Lin assumes that all \( yuè \ldots yuè \) sentences are bi-clausal, (2a) has the LF in (13), parallel to (8). CP₂ contains a null subject pro co-indexed with Zhangsan.

(2) a. Zhāngsān yuè pào yuè kuài.

‘Zhangsan ran faster and faster.’

\[
\text{[[pào]]} = \lambda x. \lambda d. \lambda s. \text{run}(x)(d)(s)
\]

(13) \[
\begin{array}{c}
\forall \text{CP} \\
\text{CP}_1 \\
\text{yuè} \\
\text{IP} \\
\text{NP} \\
\text{Zhangsan}_i
\end{array} \quad \begin{array}{c}
\text{CP} \\
\text{yuè} \\
\text{IP} \\
\text{NP} \\
\text{NP}_i
\end{array} \quad \begin{array}{c}
\text{VP} \\
\text{run} \\
\text{AP} \\
\text{fast}
\end{array} \quad \begin{array}{c}
\text{NP}_i \\
\text{pro}_i
\end{array}
\]

(14) a. \[
\text{[[pào]]} = \lambda x. \lambda d. \lambda s. \text{run}(x)(d)(s)
\]

b. \[
\text{[[yuè]]} = \lambda P_{d, s_1, s_2} \lambda d_1 \lambda d_2 \lambda s_1 \lambda s_2 [P(d_1)(s_1) \land P(d_2)(s_2) \land d_2 > d_1]
\]

c. \[
\text{[[yuè Zhāngsān pào]]} = \lambda d_1 \lambda d_2 \lambda s_1 \lambda s_2 [\text{run}(Zh)(d_1)(s_1) \land \text{run}(Zh)(d_2)(s_2) \land d_2 > d_1]
\]

d. \[
\text{[[yuè pro}_i \text{kuài]]}^{g_1 \rightarrow \text{Zhāngsān}} = \lambda d_3 \lambda d_4 \lambda s_3 \lambda s_4 [\text{fast}(Zh)(d_3)(s_3) \land \text{fast}(Zh)(d_4)(s_4) \land d_4 > d_3]
\]

e. \[
\text{[[\forall \text{Zhāngsān}_i \text{yuè pào pro}_i \text{yuè kuài]}]}^{g_1 \rightarrow \text{Zhāngsān}} = \\
\forall \exists d_3 d_4 s_3 s_4 [s_1 \leq s_3 \land s_2 \leq s_4 \land R(<d_1, s_1>, <d_2, s_2>, <d_3, s_3>, <d_4, s_4>) \land \text{fast}(Zh)(d_3)(s_3) \land \text{fast}(Zh)(d_4)(s_4) \land d_4 > d_3]
\]

(14e) says: for any pair of degrees \( d_1 \) and \( d_2 \), and any pair of situations \( s_1 \) and \( s_2 \) such that Zhangsan has done \( d_1 \)-quantity of running in \( s_1 \), and Zhangsan has done \( d_2 \)-quantity of running in
s_2, and d_2 is greater than d_1, there exists a pair of degrees d_3 and d_4 and a pair of situations s_3 and s_4 such that s_3 is an extended situation of s_1 and s_4 is an extended situation of s_2. Zhangsan is fast to degree d_3 in s_3, and Zhangsan is fast to degree d_4 in s_4. d_4 is greater than d_3. In other words, (14e) expresses a positive correlation between the amount of running that Zhangsan did and his running speed.

However, this is incorrect. As we have seen in the introduction (i.e., 5), (2a) expresses that Zhangsan’s running speed increases over time. This temporal reading is not captured by the formula in (14e). What (14e) expresses is the meaning of (15), which, unlike (2a), is true in scenario A in (5a).

\[(15) \quad \text{The more Zhangsan ran, the faster he went.}\]

\[(5) \quad \text{Zhangsan runs on treadmill three times per week. Below was his training schedule for the past week.}\]

\[\text{a. Scenario A:}\]
\[
\begin{align*}
\text{Day 1: Zhangsan ran 5 miles; his running speed was 5.3 mph.} \\
\text{Day 2: Zhangsan ran 4 miles; his running speed was 5.2 mph.} \\
\text{Day 3: Zhangsan ran 3 miles; his running speed was 5.1 mph.}
\end{align*}
\]

\[\text{b. Scenario B:}\]
\[
\begin{align*}
\text{Day 1: Zhangsan ran 3 miles; his running speed was 5.1 mph.} \\
\text{Day 2: Zhangsan ran 3 miles; his running speed was 5.2 mph.} \\
\text{Day 3: Zhangsan ran 3 miles; his running speed was 5.3 mph.}
\end{align*}
\]

We conclude from the above that Lin’s analysis does not capture the semantics of the necessarily temporal reading of (2a). So how can we explain the semantic difference between (2a) and (15) while maintaining the distinction between gradable and non-gradable predicates? In the following section, we propose that the necessarily temporal reading arises when non-gradable VPs (e.g., \([\text{VP Zhangsan ran}]\)) are coerced to have an ‘ordered’ domain parallel to that of a gradable predicate, where events like degree intervals on a degree scale are (i) totally ordered under a subinterval relation, and (ii) aligned on their starting point.

### 3 The necessarily temporal reading: a coercion-based account

This section lays out a coercion-based account that explains the necessarily temporal reading of (2a). Before delving into the analysis, it is necessary to state our assumptions. There are two main assumptions that we draw on in our account: one is the notion of coercion discussed in de Swart (1998) and Sawada and Grano (2011), and the other is the idea that degrees should be modeled as intervals rather than points (Kennedy 2001, Schwarzschild and Wilkinson 2002, a.o.).

#### 3.1 Coercion
Coercion is a general term for contextual re-interpretation. It is triggered by violations of constraints (de Swart 1998, Sawada and Grano 2011). De Swart (1998: 360) observes that in English “coercion is triggered if there is a conflict between the aspectual character of the eventuality description and aspectual constraints of some other element in the context”. Consider the two examples in (16).

(16)  a. John played the sonata.
     b. John played the sonata in 3 hours. (event)

In (16), the verb phrase *play the sonata* describes an event that has an inherent end point. It is compatible with an *in*-adverbial. The verb phrase *live in Paris* in (17), on the other hand, describes a state that does not have an inherent end point. It is compatible with a *for*-adverbial.

     b. John lived in Paris for 3 years. (state)

It is interesting to note that the verb phrase *play the sonata* in (16a) can also appear with a *for*-adverbial. But in this case the VP assumes an iterative reading – John played the sonata again and again.

(18) John played the sonata for 3 hours.

De Swart argues that this iterative reading is due to the presence of an implicit coercion operator C, which maps an event to a (homogeneous) state, as shown in (19).

(19)  [PAST[FOR 3 hours[C[John play the sonata]]]].

She further suggests that the coercion operator C is different from a grammatical operator in that the former is syntactically and morphologically invisible. C is governed by implicit contextual reinterpretation mechanism triggered by the need to resolve aspectual conflicts.

More recently, Sawada and Grano (2011) show that coercion is also responsible for the differential interpretation of measure phrases in Japanese. Measure phrases in Japanese can receive two distinct interpretations depending on what adjectives they combine with. They can receive an absolute interpretation when in front of adjectives whose scale contains a minimal element (i.e., a lower-closed scale)(e.g., 20), or a differential interpretation when in front of an adjective with an open scale (with no minimal elements). In (21), the measure phrases specify the difference between two degrees--the degree to which the property holds of the subject and a contextually supplied standard.

(20)  a. Kono sao-wa 5-do magat-teiru. (Absolute measure)
     this rod-to 5-degree bend-PERF
     ‘This rod is 5 degrees bent.’
     b. Kono fusuma-wa 3-senti ai-teiru. (Absolute measure)
     this sliding door-top 3-centimeter open-PERF

---

6 The typology of scale structures is based on Kennedy and McNally (2005) and Kennedy (2007).
'This door is 3 centimeters open.'

(21) a. Kono tana-wa 2-meetoru takai (Differential measure)  
     this shelf-top 2-meter tall  
     ‘This shelf is 2 meters taller.’

b. Kono roopu-wa 5-inchi nagai.  (Differential measure)  
     this rope-top 5-inch long  
     ‘This rope is 5 inches long.’

Sawada and Grano (2011) argue that the differential interpretation of the measure phrases in (21) is due to the presence of a coercion operation $C_s$, whose function is to force gradable adjectives with an open end scale to become predicates with a contextually determined standard.

(22) a. $[[takai]] = \lambda x. \text{height}(x)$  

b. $[[C_s]]([ [takai]])) = \lambda x. \text{height}_{\text{heigh}}^s (x)$  
   (where $s$ stands for a contextually determined object.)

In (22a), the adjective ‘tall’ denotes a measure function that maps from individual $x$ to $x$’s height. In (22b), the adjective combines with $C_s$, and the result is a measure function that measures the difference between $x$’s height and a contextually determined standard (i.e., $s$’s height).

We take de Swart and Sawada and Grano’s study as suggestive that coercion is pervasive cross-linguistically. We argue that it also accounts for the necessarily temporal reading of yuè $V_{\text{non-gradable}}$ yuè $A$ in Mandarin Chinese.

### 3.2 Degrees as intervals

Another important assumption in our analysis concerns the ontological representation of degrees. Traditional analyses take degrees as points. Kennedy (2001) and Schwarzschild and Wilkinson (2002) showed that such an analysis cannot explain some important patterns in comparative constructions, one of which is so-called cross-polar anomaly (CPA), exemplified in (23).

(23) *Alice is shorter than Carmen is tall.

(23) is a comparative formed out of a pair of a positive and a negative adjective tall and short. Kennedy (2001) argues that treating degrees as points does not explain the semantic anomaly of (23). If we model degrees as intervals, then it is possible to attribute the anomaly of (23) to the incommensurability of degrees of opposite polarities. That is, positive adjectives are associated with positive degrees. They are intervals starting with 0 and ending with $n$. Negative adjectives are associated with negatives degrees. They are intervals starting with $n$ and ending with infinity. These two types of degrees are demonstrated in (24).
If degrees are intervals, a comparative relation between two degrees $d_1$ and $d_2$ can be captured by a proper subinterval relation ($\sqsubseteq$). If $d_1$ and $d_2$ are positive degrees, it is a proper subinterval relation between two intervals that share the same starting point 0; if $d_1$ and $d_2$ are negative degrees, it is a proper subinterval relation between two intervals that share the same ending point $\infty$. This is illustrated in (25).

Let us define two specific subinterval relations, $\sqsubseteq_{\text{start}}$ and $\sqsubseteq_{\text{end}}$, to capture the comparisons in (25a) and (25b).

In the following section, we show that $\sqsubseteq_{\text{start}}$ not only can be used to capture the comparative relation between two degree intervals on a degree scale, but also can be used to capture the comparison between two event intervals.

### 3.3 A coercion-based analysis

Given the two assumptions above, we are now ready to provide an explanation for the necessarily temporal reading of $\text{yuè } V_{\text{non-gradable}}$ $\text{yuè } A$. We propose that this reading arises when the selectional restriction of $\text{yuè }$ in (27) is violated.

In Mandarin Chinese, $\text{yuè}$ can only felicitously combine with a predicate whose domain is totally ordered under the proper subinterval relation $\sqsubseteq_{\text{start}}$ or $\sqsubseteq_{\text{end}}$.

$\text{yuè}$ can felicitously combine with a gradable adjective (of type $<d, et>$), whose domain consists of a set of positive/negative degrees. It is however infelicitous to combine with a non-gradable VP whose domain is a set of events. Events, unlike degrees, are not totally ordered. Assuming
that verbs take an event argument, a non-gradable VP like [Zhangsan ran] denotes a set of events of Zhangsan’s running, which can be modeled as a lattice in (28).7

(28) 

In (28), events are partially ordered (e.g., there is no ordering among events on the same level). To satisfy the selection restriction of yuè in (27), we propose that there is a coercion operator Ce that mediates between yuè and the non-gradable VP [Zhangsan ran],8 whose function is to turn the lattice in (28) to a scale structure like (30). In (30), events are totally ordered under the subinterval relation ⊏start just like positive degrees on a degree scale.9

(29) The LF of (2a): [yuè1 [Ce[VP Zhangsan ran]]] [yuè2 fast]

(30) Coerced ‘event scale’

\[ e_1 \downarrow \]
\[ e_1 \downarrow e_2 \downarrow e_3 \]

On this analysis, (2a) means: for any pair of events e₁ and e₂, if e₁ and e₂ have a common starting point and e₁ is a subinterval of e₂, e₁ is slower than e₂. This amounts to saying that the longer Zhangsan ran, the faster he went.

This analysis correctly predicts (2a) to be true in scenario B but not in scenario A in (5)(repeated below).

(5) Zhangsan runs on treadmill three times per week. Below was his training schedule for the past week.

a. Scenario A:
   Day 1: Zhangsan ran 5 miles; his running speed was 5.3 mph.
   Day 2: Zhangsan ran 4 miles; his running speed was 5.2 mph.
   Day 3: Zhangsan ran 3 miles; his running speed was 5.1 mph.

b. Scenario B:

---

7 ⊕ is a two-place operation called ‘join’. We assume that for any two elements x and y in a set S, x⊕y is defined, and x⊕y ∈ S.
8 There is more discussion on the LF in section 4.1.
9 There is a question of why Ce cannot coerce the set of events denoted by the non-gradable VP to be ordered under ⊏end. We think that this might have to do with the fact that Negative degrees are infinite (ending with ∞), but the events under comparison are closed intervals, having both a starting and an ending point.
Starting with scenario B, the three running events on day 1, day 2 and day 3 are temporally ordered. Based on them, we can define three cumulative events, e₁, e₂, and e₃, such that e₁ is the running event that John did on day 1; e₂ is the running event that John did on day 1 and day 2; and e₃ is the running event that John did on day 1, day 2 and day 3. These three events are totally ordered under $\sqsubseteq_{\text{start}}$, as illustrated in (31). Assuming that each cumulative event is associated with its final speed, e₃ is 5.3 mph; e₂ is 5.2 mph, and e₁ is 5.1 mph. Our analysis predicts that Zhangsan ran faster and faster is true in scenario B.

\[
\begin{array}{ccc}
\text{Day 1} & \text{Day 2} & \text{Day 3} \\
| e₁ | & & \\
& e₂ & \ \\
& & e₃ \\
\end{array}
\]

With respect to scenario A, the three running events on day 1, day 2, and day 3 are also temporally ordered and we can similarly define three cumulative events. However, associating each of these with their final speeds, e₃ is 5.1 mph, e₂ is 5.2 mph, and e₁ is 5.3 mph. Therefore, our analysis predicts that (2a) is not true in this scenario.

To summarize, in this section we argued that the necessarily temporal reading of $\text{yuè V}_{\text{non-gradable}} \text{yuè A}$ arises when the constraint in (27) is violated. That is, $\text{yuè}$ is infelicitous to combine with a non-gradable VP whose domain is not totally under the subinterval relation $\sqsubseteq_{\text{start}}$ or $\sqsubseteq_{\text{end}}$. The non-gradable VP is coerced to have a domain parallel to that of a gradable adjective where events are totally ordered under $\sqsubseteq_{\text{start}}$. In the following section, we will formalize this analysis and show how the necessarily temporal reading of (2a) is compositionally achieved.

**4. Formal Interpretations**

This section provides a formal account of (2a) based on the analysis in section 3.3. In section 4.1, we examine the syntax of (2a), and argue that it is mono-clausal; in section 4.2, we compute the meaning of (2a); in section 4.3 we discuss two welcome results that follow from the proposed analysis.

**4.1 Syntax**

We argue that (2a) has a mono-clausal structure. Our main evidence for this claim comes for examples like (2b)(repeated below), which is ambiguous between two readings.

\[
\text{(2) b. Zhāngsān yuè tiào yuè gāo.} \\
\text{jump} \quad \text{tall/high} \\
\text{‘Zhangsan jumped higher and higher.’}
\]

12
‘Zhangsan became taller and taller from jumping.’

We propose that the two readings are related to the two syntactic structures in (32): (32a) is a mono-clausal structure, which is associated with the reading ‘Zhangsan jumped higher and higher’; (32b) is a bi-clausal structure, which is associated with the reading ‘Zhangsan became taller and taller from jumping.’

(32) a. [CP Zhāngsān yuè tiào yuè gāo]. mono-clausal
b. [CP Zhāngsān yuè tiào], [CP pro1 yuè gāo]. bi-clausal

(32) is supported by the distribution of the morpheme jiù ‘then’ and the future aspect marker huì ‘will’. In a bi-clausal yuè ... yuè sentence, jiù ‘then’ and huì ‘will’ can only appear in front of yuè2, but not yuè1, as shown by the contrast between (33) and (34) below.

(33) a. nǐ yuè shēngqì, tā jiù yuè gāoxìng.
   you angry he then happy
   ‘The angrier you are, then the happier he is.

b. nǐ yuè shēngqì, tā huì yuè gāoxìng.
   you angry he will happy
   ‘The angrier you are, the happier he will be.

(34) a. *nǐ jiù yuè shēngqì, tā yuè gāoxìng.
    you angry he then happy
    ‘The more Zhangsan jumped, the taller he then became.’

b. *nǐ huì yuè shēngqì, tā yuè gāoxìng.
    you angry he will happy
    ‘The more Zhangsan jumps, the taller he will become.’

For (2b), jiù ‘then’ and huì ‘will’ can appear in front of either yuè1 or yuè2. In each case, however, a different reading results, as shown in (35) and (36).

(35) a. Zhāngsān jiù yuè tiào yuè gāo.
   then jump tall/high
   ‘Zhangsan then jumped higher and higher.’

b. Zhāngsān huì yuè tiào yuè gāo.
   will jump tall/high
   ‘Zhangsan will jump higher and higher.’

(36) a. Zhāngsān yuè tiào, jiù yuè gāo.
   jump then tall/high
   ‘The more Zhangsan jumped, the taller he then became.’

b. Zhāngsān yuè tiào, huì yuè gāo.
   jump will tall/high
   ‘The more Zhangsan jumps, the taller he will become.’

In (35), jiù ‘then’ and huì ‘will’ appear in front of yuè1. (35a) means: ‘Zhangsan then jumped higher and higher’; (35b) means: ‘Zhangsan will jump higher and higher’. In both readings, gāo is predicated of the jumping events. In (36), jiù ‘then’ and huì ‘will’ appear in front of yuè2,
(36a) means: ‘Zhangsan then became taller and taller from jumping’; (36b) means: ‘Zhangsan will become taller and taller from jumping’. In these two readings, gāo is predicated of the subject Zhangsan. For (36a-b), we can add an overt subject tā ‘he’ in front of jiù ‘then’ and huì ‘will’, as shown in (37):

(37) a. Zhāngsān yuè tiào, tā jiù yuè gāo.
   jump he then tall/high
   ‘The more Zhangsan jumped, the taller he then became.’

b. Zhāngsān yuè tiào, tā huì yuè gāo.
   jump he will tall/high
   ‘The more Zhangsan jumps, the taller he will become.’

Assuming that tā ‘he’ is co-indexed with Zhangsan, (37a-b) have the same interpretations as (36a-b). We can conclude from the above that the sentences in (36) are bi-clausal.

The yuè … yuè sentences in (35), on the other hand, have a mono-clausal structure. They do not allow specification of a direct object following the verb, as shown in (38).

(38) a. Zhāngsān jiù yuè tiào shèng yuè kuài.
   will jump rope fast
   Intended: Zhangsan will jump rope faster and faster.

b. Zhāngsān huì yuè tiào shèng yuè kuài.
   will jump rope fast
   Intended: Zhangsan will jump rope faster and faster.

(38) can be improved by either adding a reduplicated verb tiào before yuè, to take the object shèng ‘rope’, as shown in (39a), or fronting the object before yuè, as shown in (39b).

   jump rope will jump fast

   rope will jump fast

The contrast between (38) and (39) is resonant of a well-known restriction on the ‘post-verbal complement construction’ in Mandarin Chinese -- a verb cannot take an object and an adverbial complement at the same time. (40) shows that an adverbial complement is attached to a verb through the morpheme de. (41) shows that a verb cannot take an object and an adverbial complement at the same time.

(40) Zhāngsān tiào-de hěn kuài.
   jump-De very fast
   ‘Zhangsan jumped very fast.’

   jump rope De very fast

10 We thank Lisa Cheng and an anonymous reviewer for drawing our attention to examples like (38).
Intended: ‘Zhangsan jumped rope very fast.’

Like (38), (41) can also be repaired by either adding a reduplicated verb to take the object (e.g., 42a), or fronting the object before the verb (e.g., 42b).

(42)  

jump rope jump-De very fast


rope jump-De very fast

The parallel between yuè V yuè A and the post-verbal complement construction points to a common underlying structure. Following Chao (1968), Li and Thompson (1981), Wei (2006), and Tsai (2012), we assume that in the complement construction, the adverbial complement is the main predicate, and the remaining material (represented by XP) constitutes the sentential subject, as illustrated in (43).

(43)

\[
\begin{array}{c}
\text{S} \\
\text{XP} \\
\quad \text{Zhangsan pào-de} \\
\text{AP} \\
\quad \text{hěn kuài}
\end{array}
\]

4.2 Semantics

Based on (43), we propose that the yuè V yuè A sentence in (2a) has the LF in (44).

(44)

\[
\begin{array}{c}
\forall \\
\text{yuè}_1 \\
\quad \text{C}_e \\
\quad \text{yuè}_2 \\
\text{S} \\
\quad \text{Zhangsan run} \\
\quad \text{fast}
\end{array}
\]

In (44), there is a covert universal quantifier that c-commands both yuè phrases. C\_e is a coercion operator that mediates between yuè\_1 and the non-gradable VP. The combination of C\_e and the VP

---

11 It is not clear to us why in yuè V yuè A de is not required (in fact, not allowed) to connect the adverbial complement to the verb.

12 An alternative view analyzes the complex verb V-de as the main predicate, and the adverbial complement as the secondary predicate (Huang 1988, Zhang 2001, Lim 2005, a.o.). The differences between these two analyses do not directly concern us here.
yields a set of subevents totally ordered under the proper subinterval relation \( \preceq_{\text{start}} \), as shown in (45).

(45)  
\begin{align*}
&\text{a. } \left[ [C_e] | [VP_{\text{non-gradable}}] \right] = \lambda e. P(e) \land P(E) \land e \preceq_{\text{start}} E \\
&\text{b. } e \preceq_{\text{start}} E \equiv e \preceq E \land \text{start}(e) = \text{start}(E)
\end{align*}

The step-by-step computation of (44) is provided in (46):

(46)  
\begin{align*}
&\text{a. } \left[ [\text{Zhangsan ran}] \right] = \lambda e. \text{run}(Zh, e) \\
&\text{b. } \left[ [C_e] \right] = \lambda P_{cv.} \land \lambda e_c [P(e) \land P(E) \land e \preceq_{\text{start}} E] \\
&\text{c. } \left[ [C_e \text{ Zhangsan ran}] \right] = \lambda e, [\text{run}(Zh, e) \land \text{run}(Zh, E) \land e \preceq_{\text{start}} E] \\
&\text{d. } \left[ [\text{yuè}] \right] = \lambda P_{cv.} \land \lambda e_1 \lambda e_2 [P(e_1) \land P(e_2) \land e_1 \preceq_{\text{start}} e_2] \\
&\text{e. } \left[ [\text{yuè } C_e \text{ Zhangsan ran}] \right] = \lambda e_1 \lambda e_2 [\text{run}(Zh, e_1) \land \text{run}(Zh, e_2) \land \text{run}(Zh, E) \land e_1 \preceq_{\text{start}} e_2 \land e_1 \preceq_{\text{start}} E \land e_2 \preceq_{\text{start}} E] \\
&\text{f. } \left[ [\text{fast}] \right] = \lambda d \lambda e, \lambda v, \text{fast}(d)(e) \\
&\text{g. } \left[ [\text{yuè}_2] \right] = \lambda P_{ed.} \land \lambda e_1 \lambda e_2 \exists d_1 d_2 [P(d_1)(e_1) \land P(d_2)(e_2) \land d_1 \preceq_{\text{start}} d_2] \\
&\text{h. } \left[ [\text{yuè}_2 \text{ fast}] \right] = \lambda e_1 \lambda e_2 \exists d_1 d_2 [\text{fast}(d_1)(e_1) \land \text{fast}(d_2)(e_2) \land d_1 \preceq_{\text{start}} d_2] \\
&\text{i. } \left[ [\forall \text{yuè}_1 C_e \text{ Zhangsan ran } \text{yuè}_2 \text{ fast}] \right] = \forall e_1 e_2 [\text{run}(Zh, e_1) \land \text{run}(Zh, e_2) \land \text{run}(Zh, E) \land e_1 \preceq_{\text{start}} e_2 \land e_1 \preceq_{\text{start}} E \land e_2 \preceq_{\text{start}} E] \rightarrow \exists d_1 d_2 [\text{fast}(d_1)(e_1) \land \text{fast}(d_2)(e_2) \land d_1 \preceq_{\text{start}} d_2] \\
&\text{a. } \left[ [\exists \forall \text{yuè}_1 C_e \text{ Zhangsan ran } \text{yuè}_2 \text{ fast}] \right] = \exists E \forall e_1 e_2 [\text{run}(Zh, e_1) \land \text{run}(Zh, e_2) \land \text{run}(Zh, E) \land e_1 \preceq_{\text{start}} e_2 \land e_1 \preceq_{\text{start}} E \land e_2 \preceq_{\text{start}} E] \rightarrow \exists d_1 d_2 [\text{fast}(d_1)(e_1) \land \text{fast}(d_2)(e_2) \land d_1 \preceq_{\text{start}} d_2]
\end{align*}

(46k) says: there is an event \( E \) of Zhangsan’s running; for any pair of subevents \( e_1 \) and \( e_2 \) which share the same initial point as \( E \), if \( e_1 \) is a proper subinterval of \( e_2 \), \( e_1 \) is slower than \( e_2 \).

In our analysis, \( \text{yuè} \) has two different semantics depending on what predicate it combines with. In (46e), \( \text{yuè}_1 \) combines with \([C_e + VP_{\text{non-gradable}}]\); the result is a set of events ordered under the proper subinterval relation \( \preceq_{\text{start}} \). In (46h), \( \text{yuè}_2 \) combines with a gradable predicate; the result is a set of events ordered based on degrees they are associated with. Degrees are inherently ordered under either \( \preceq_{\text{start}} \) or \( \preceq_{\text{end}} \). Despite the semantic ambiguity, it is important to note that the two semantics of \( \text{yuè} \) are consistent in the sense that the result of combining \( \text{yuè} \) with a gradable or a non-gradable predicate is always an ordering on a set of entities ordered based on \( \preceq_{\text{start}} \) or \( \preceq_{\text{end}} \).

Not only can \( \text{yuè} \) \( V_{\text{non-gradable}} \) \( \text{yuè} \) A be mono-clausal, it can also be bi-clausal. This is shown by (37a)(repeated below):

(37)  
\begin{align*}
&\text{a. } \text{Zhāngsān } \text{yuè tiào, tā (jiù) } \text{yuè gāo.} \\
&\text{jump \ he \ then \ tall/high} \\
&'\text{The more Zhangsan jumped, the taller he then became.}'
\end{align*}

We assume that (37a) has the LF in (47).

---

13 In (46f) kuài ‘fast’ here is interpreted as an adverb, denoting a relation between events and degrees.
(47) is bi-clausal; it allows jiù ‘then’ to appear in the second clause. Below we provide the computation of (47).

(48) a. \[\text{[[Zhangsan jump]]} = \lambda e. \text{jump}(Zh, e)\]
b. \[\text{[[C_d]]} = \lambda P_{c, d} \lambda e_i [P(e) \land P(E) \land e \sqsubseteq e_1]\]
c. \[\text{[[C_d Zhangsan jump]]} = \lambda e_i [\text{jump}(Zh, e) \land \text{jump}(Zh, E) \land e \sqsubseteq e_1]\]
d. \[\text{[[yuè_1]]} = \lambda P_{c, v} \lambda e_i \lambda e_2 [P(e_1) \land P(e_2) \land e_1 \sqsubseteq e_2]\]
e. \[\text{[[yuè_1 C_d Zhangsan jump]]} = \lambda e_1 \lambda e_2 [\text{jump}(Zh, e_1) \land \text{jump}(Zh, e_2) \land \text{jump}(Zh, E) \land e_1 \sqsubseteq e_2 \land e_1 \sqsubseteq e_2 \land e_2 \sqsubseteq e_2]\]
f. \[\text{[[he_1 tall]]}^{[[i \rightarrow Zh]]} = \lambda d_2 \lambda s_1 \text{tall}(Zh)(d)(s)\]
g. \[\text{[[yuè_2]]} = \lambda P_{c, d} \lambda s_1 \lambda s_2 \exists d_2 [P(d_1)(s_1) \land P(d_2)(s_2) \land d_1 \sqsubseteq d_2]\]
h. \[\text{[[yuè_2 he_1 tall]]}^{[[i \rightarrow Zh]]} = \lambda s_1 \lambda s_2 \exists d_2 [\text{tall}(Zh)(d_1)(s_1) \land \text{tall}(Zh)(d_2)(s_2) \land d_1 \sqsubseteq d_2]\]
i. \[\text{[[\forall ... jiù]]} = \lambda P_{c, v} \lambda s_1 \forall e_1 e_2 [P(e_1)(e_2) \rightarrow \exists s_1 s_2 [Q(s_1)(s_2) \land R(e_1, s_1) \land R(e_2, s_2)]]\]
j. \[\text{[[\forall yuè_1 C_d Zhangsan jump jiù yuè_2 he_1 tall]]}^{[[i \rightarrow Zh]]} = \forall e_1 e_2 [[\text{jump}(Zh, e_1) \land \text{jump}(Zh, e_2) \land \text{jump}(Zh, E) \land e_1 \sqsubseteq e_2 \land e_1 \sqsubseteq e_2 \land e_2 \sqsubseteq e_2 \rightarrow \exists s_1 s_2 [\text{tall}(Zh)(d_1)(s_1) \land \text{tall}(Zh)(d_2)(s_2) \land d_1 \sqsubseteq d_2 \land d_1 \sqsubseteq d_2 \land R(e_1, s_1) \land R(e_2, s_2)]]\]
k. \[\text{[[\exists yuè_1 C_d Zhangsan jump jiù yuè_2 he_1 tall]]}^{[[i \rightarrow Zh]]} = \exists E \forall e_1 e_2 [[\text{jump}(Zh, e_1) \land \text{jump}(Zh, e_2) \land \text{jump}(Zh, E) \land e_1 \sqsubseteq e_2 \land e_1 \sqsubseteq e_2 \land e_2 \sqsubseteq e_2 \rightarrow \exists s_1 s_2 [\text{tall}(Zh)(d_1)(s_1) \land \text{tall}(Zh)(d_2)(s_2) \land d_1 \sqsubseteq d_2 \land d_1 \sqsubseteq d_2 \land R(e_1, s_1) \land R(e_2, s_2)]]\]

(48k) says: there is an event E of Zhangsan’s jumping; for any pair of subevents e_1 and e_2 that share the same initial point as E, if e_1 is a proper subinterval of e_2, e_1 is associated with situation s_1 in which Zhangsan is tall to d_1, and e_2 is associated with situation s_2 in which Zhangsan is tall to d_2; d_1 is a subinterval of d_2.

An explanation is in order regarding the semantics of the covert universal quantifier \(\forall\) and the morpheme jiù in (48i). Here we follow Lin in assuming that jiù is a syncategorematic item that is interpreted together with \(\forall\). It contributes relation R that connects the set of pairs of events denoted by [yuè_1 + VP_{non-gradable}] in (48e) with the set of pairs of situations denoted by [yuè_2 + AP] in (48h). In Lin’s analysis, R indicates causality. However, as Liu (2008a) points out, R does not have to be causal. In (37a), for example, it can be coincidental that Zhangsan’s height happens to

---

14 We assume an ontological distinction between states and events. Adjectives and stative verbs (e.g., xihuān ‘to like’) have a state argument, while action verbs have an event argument.
increase while he jumps. Nonetheless, the value of $R$ is not unconstrained. It is subject to the following constraint:\footnote{An anonymous reviewer provides a good example to illustrate the temporal constraint on $R$.}

(49) For any pair of subevents $e_1$ and $e_2$, if $e_1$ is a proper subinterval of $e_2$, $R$ must associate $e_1$ with a state that temporally precedes the state associated with $e_2$.

(50) provides a concrete illustration of this constraint. In (50) there are three jumping events, $e_1$, $e_2$, and $e_3$, and three relevant height states $s_1$, $s_2$, and $s_3$. $R$ must associate $e_1$, $e_2$, and $e_3$ with $s_1$, $s_2$, and $s_3$ respectively. $R$ cannot associates $e_1$ with $s_2$, or $e_2$ with $s_1$. In (37a), we assume that each jumping event is associated with a temporally overlapping state. If an event overlaps with more than one state, it is the final state that the event is associated with (e.g., $e_2$ is associated with $s_2$ rather than $s_1$).

\begin{center}
\begin{tikzpicture}
    \node (e1) at (1,0) {$e_1$};
    \node (e2) at (3,0) {$e_2$};
    \node (e3) at (5,0) {$e_3$};
    \node (s1) at (0,0) {$s_1$};
    \node (s2) at (2,0) {$s_2$};
    \node (s3) at (4,0) {$s_3$};

    \draw (e1) -- (e2) -- (e3); % jumping events
    \draw (s1) -- (s2) -- (s3); % relevant states

    \node at (6,0) {time};

    \node at (1,1) {time};
\end{tikzpicture}
\end{center}

\subsection*{4.3 Two welcome results}

There are two welcome results that immediately follow from our analysis. First, it provides a straightforward explanation for the ‘unbounded condition’ observed in Liu (2008a). Liu observes that in the $yuē$ $V$ $yuē$ $A$ construction, $V$ can only be an atelic predicate (states or activities), but not a telic predicate (achievements and accomplishments), as shown by the contrast between (51) and (52).

\begin{enumerate}
\item[(51)] a. *tā $yuē$ dàodá shān-dǐng, $yuē$ gāoxìng. (achievement)
   he arrive mountain-top, happy
   
   b. *tā $yuē$ chí yí-gè-píngguó, $yuē$ gāoxìng. (accomplishment)
   he eat one-cl-apple, happy
\end{enumerate}

\footnote{An anonymous reviewer provides a good example to illustrate the temporal constraint on $R$.}

(i) Tā $yuē$ qǐpiàn tā, tā $yuē$ shāngxīn.
   he cheat she she sad
   ‘The more he cheated on her, the sadder she became.’

Suppose that there are two instances of cheating events ($e_1$ and $e_2$), and the wife finds out about the second cheating event ($e_2$) only after she finds out about the first cheating event ($e_1$). Therefore, the state of sadness associated with first cheating event ($s_1$) follows the state of sadness associated with the second cheating event ($s_2$). Even though the wife is sadder in $s_1$ than in $s_2$, intuitively (i) cannot be true in this scenario.
(52) a. Zhāngsān yuè xīhuān Lisi, wǒ yuè tǎoyàn tā. (state) angry I resent
   ‘The more Zhangsan likes Lisi, the more that I resent Lisi.’

b. Zhāngsān yuè pào yuè kuài. (activity) run fast
   ‘Zhangsan ran faster and faster.’

We can consider the ‘unbounded condition’ as a requirement on the VP
non-gradable to have the subinterval property (Bennett and Partee, 1972): if atelic predicates are true at some interval i, then they are true at every subinterval of i. States, and activities are atelic and have the subinterval property (e.g., if for a certain interval i, it is true that Zhangsan is running, then it is also true that for every subinterval of i, Zhangsan is running). Achievements, and accomplishments, on the other hand, describe events that contain an inherent culmination point; they do not have the subinterval property (e.g., if for a certain interval i, it is true that John ate an apple. It is not true that for every subinterval of i John ate an apple, as he only ate parts of that apple).

The requirement of the subinterval property on the VP non-gradable in yuè…yuè sentences follows directly from the semantics of the coercion operator C in (45a)(repeated below). That is, C requires that the VP non-gradable denote a property true of both the super-event E and its subparts.

(45) a. \[[C_e][[VP_{\text{non-gradable}}]] = \lambda e. P(e) \land P(E) \land e \subseteq \text{start } E\]

In addition, our analysis also extends a natural explanation for the different quantificational effects of the adverbial quantifier tōngcháng ‘usually’ in the yuè A/Vgradable yuè A/Vgradable structure (e.g., (1)) and the yuè V non-gradable yuè A structure (e.g., (2)).

Lin (2007) and Liu (2008a) observe that an overt adverbial quantifier such as tōngcháng ‘usually’ can overwrite the default universal quantifier ∀ in a yuè A/Vgradable yuè A/Vgradable sentence. For instance, (1a)(repeated as 53a) means that for all pairs of apples x₁ and x₂, if x₁ is bigger than x₂, x₁ is sweeter than x₂. With the presence of tōngcháng ‘usually’, the sentence then has the meaning in (54b), which says: for most pairs of apples, x₁ and x₂, if x₁ is bigger than x₂, x₁ is sweeter than x₂.

(53) a. píngguǒ yuè dà, yuè tián. apple big sweet
   ‘The bigger an apple is, the sweeter it usually is.’

b. ∀ x₁ x₂ (apple(x₁) \land apple(x₂) \land x₁ is bigger than x₂) \rightarrow [x₁ is sweeter than x₂]

(54) a. tōngcháng, píngguǒ yuè dà, yuè tián. usually apple big sweet
   ‘The bigger an apple is, the sweeter it usually is.’

16 Beck (2012) makes a similar observation about the adverbial quantifier moistens ‘usually’ in the temporally continuous comparative conditional like (i) in German.

(i) Otto rannte moistens schneller, je mehr er trainierte. Otto ran usually faster the more he practiced
   ‘Otto usually ran faster and faster, the more he practice.’

She observes that moisten ‘usually’ in (i) does not talk about the relevant subevents, but quantifies over big events (Beck 2012: 97). In her analysis, the temporally continuous reading is attributed to the presence of a plural sequence operator, PL seq. The interested reader should refer to Beck’s work for details.
b. MOST $x_1 x_2$  \[\text{apple}(x_1) \land \text{apple}(x_2) \land x_1 \text{ is bigger than } x_2\]
\[x_1 \text{ is sweeter than } x_2\]

In contrast, tōngcháng ‘usually’ in the $\text{yuè}\ V_{\text{non-gradable}}\ \text{yuè}$ A structure like (2a) does not quantify over subevents but super-events. Consider (55) in the scenario in (56):

(55) tōngcháng, Zhāngsān yuè pǎo yuè kuài.
usually run fast
‘Zhangsan usually ran faster and faster.’

(56) Scenario: Zhangsan participated in a running test, where his speed was recorded every 5 minutes. Below were the records.

<table>
<thead>
<tr>
<th>5th minute</th>
<th>10th minute</th>
<th>15th minute</th>
<th>20th minute</th>
<th>25th minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 mph</td>
<td>5.2 mph</td>
<td>5.3 mph</td>
<td>5.4 mph</td>
<td>5.5 mph</td>
</tr>
</tbody>
</table>

In (56), there is only one super-event of Zhangsan’s running, which consists of multiple subevents. Intuitively, (55) is not true in (56). The sentence cannot mean: for most pairs of subevents $e_1$ and $e_2$, if $e_1$ temporally precedes $e_2$, $e_1$ is slower than $e_2$.

In our analysis, this quantificational effect of tōngcháng ‘usually’ follows from the order of the existential quantifier that binds the super-event and the universal quantifier that binds the pair of subevents. The example in (2a), without tōngcháng ‘usually’, has the truth-conditions in (46k)(repeated below).

(46) k. [[\exists\forall\text{nord}\text{u} C_d \text{Zhangsan ran yuè}_d \text{fast}]] = \exists E \forall e_1 e_2[[\text{run}(\text{Zh}, e_1) \land \text{run}(\text{Zh}, e_2) \land \text{run}(\text{Zh}, E) \land e_1 \sqsubseteq_{\text{start}} e_2 \land e_1 \sqsubseteq_{\text{start}} E \land e_2 \sqsubseteq_{\text{start}} E] \text{d R } E \sqsubseteq \exists d_1 d_2[\text{fast}(d_1)(e_1) \land \text{fast}(d_2)(e_2) \land d_1 \sqsubseteq_{\text{start}} d_2]

In (46k) the universal quantifier is embedded under the existential quantifier. When tōngcháng ‘usually’ is present, it overwrites the existential rather than the universal quantifier.

To summarize, in this section we provided a formal analysis of the necessarily temporal reading of (2a) based on the discussion in section 3.3. As the reader might be aware, there are other possible analyses available to capture the meaning of (2a). In the following section, we address this concern by considering two main alternative approaches, and show that they are less desirable than our proposal.

5 Two alternative analyses: association with degrees and association with times.

In this section we compare the coercion-based analysis to two alternative proposals. One assumes that non-gradable verbs can be associated with a cumulative degree through a null operator (section 5.1); the other assumes that non-gradable verbs lexicalize a temporal argument but gradable predicates cannot (section 5.2). We show that both analyses face problems that are not easy to solve.
5.1 Association with Degrees

The analysis to be discussed in this section is essentially a modification of Lin’s analysis.\(^{17}\) Recall that Lin’s analysis of (2a) fails because he assumes that non-gradable verbs lexicalize a degree argument that measures the quantity (i.e. either the cumulative or the non-cumulative amount) of an event. If, on the other hand, we assume that non-gradable verbs can only be associated with a degree that measures the cumulative amount of an event, then we can successfully capture the necessarily temporal reading of (2a). Suppose this cumulative amount is contributed by an implicit measure function \(\mu_{\text{cum}}\), as shown in (57).

\[(57)\quad \text{Zhāngsān yùè } [\mu_{\text{cum}} + \text{pāo}] \text{ yùè kuài.}
\]

\(\text{run fast} \quad \text{‘Zhangsan ran faster and faster.’}\)

On this analysis, (2a) means: for any pair of subevents of Zhangsan’s running, \(e\) and \(e'\), if the cumulative amount of \(e'\) is greater than that of \(e\), \(e'\) is faster than \(e\). This analysis predicts (57) to be true in scenario B in (5)(repeated below). In B there are three running events, \(e_1, e_2\) and \(e_3\), which are temporally ordered. They are part of a super-event \(E\), as demonstrated in (58). The cumulative amount of \(e_1\) in \(E\) is an amount measured from the beginning of \(E\) through \(e_1\), written as \(\mu_{\text{cum}}(e_1) = 3\) miles; the cumulative amount of \(e_2\) in \(E\) is an amount measured from the beginning of \(E\) through \(e_2\), written as \(\mu_{\text{cum}}(e_2) = 6\) miles, etc. As one can easily see, there is a correlation between a cumulative amount of an event \(e\) and its temporal precedence: the later \(e\) occurs in \(E\), the greater its cumulative amount is.

(5) Zhangsan runs on treadmill three times per week. Below was his training schedule for the past week.

a. Scenario A:
   Day 1: Zhangsan ran 5 miles; his running speed was 5.3 mph.
   Day 2: Zhangsan ran 4 miles; his running speed was 5.2 mph.
   Day 3: Zhangsan ran 3 miles; his running speed was 5.1 mph.

b. Scenario B:
   Day 1: Zhangsan ran 3 miles; his running speed was 5.1 mph.
   Day 2: Zhangsan ran 3 miles; his running speed was 5.2 mph.
   Day 3: Zhangsan ran 3 miles; his running speed was 5.3 mph.

\[(58)\quad \begin{array}{c}
\hline
\text{e}_1 \quad \text{E} \quad \text{e}_2 \quad \text{e}_3 \\
\hline
\end{array}
\]

\(^{17}\) We thank an anonymous reviewer for drawing attention to this possibility.
The main problem with this analysis is that the measure function \( \mu_{\text{cum}} \) is not sufficiently justified. There is no independent reason for why a non-gradable VP must be associated with a cumulative amount as opposed to a non-cumulative amount of some activity. Moreover, in such an analysis, \( yuè \) is treated on a par with other degree modifiers (e.g., \( hěn \) ’very’), as they all compose with predicates that contain a degree argument. This analysis abstracts away from an important distinction between \( yuè \) and other degree modifiers: the former can directly combine with a non-gradable verb whereas the latter cannot (cf. (3)).

5.2 Association with times

The other approach makes crucial reference to the distinction between gradable predicates and non-gradable verbs.\(^{18}\) It assumes that gradable predicates do not contain a time argument in their semantics while non-gradable verbs do, and that gradable predicates do contain a degree argument, while non-gradable predicates do not. This is shown by the semantics of the verb \( pāo \) ‘to run’ and the adjective \( gāoxìng \) ‘happy’ in (59).

\[
\begin{align*}
\text{59)} & \quad \text{a. } \llbracket [pāo] \rrbracket = \lambda x \lambda t \lambda s. \text{run}(x)(t)(s) \quad \langle e, <i, <s, t>> \rangle \\
& \quad \quad \text{b. } \llbracket [gāoxìng] \rrbracket = \lambda x \lambda d \lambda s. \text{happy}(x)(d)(s) \quad \langle e, <d, <s, t>> \rangle
\end{align*}
\]

In this analysis \( yuè \) has two different interpretations depending on whether the predicate it composes with is a gradable predicate or a non-gradable verb, as shown in (60):

\[
\begin{align*}
\text{(60)} & \quad \text{a. } \llbracket [yuè] \rrbracket = \lambda P_{<i,> \text{, } t> \text{, } d> \text{, } s>} \lambda e_1 \lambda e_2 \exists t_1 \exists t_2 \left[ P(t_1)(e_1) \land P(t_2)(e_2) \land t_2 > t_1 \right] \quad \text{Non-gradable} \\
& \quad \quad \text{b. } \llbracket [yuè] \rrbracket = \lambda P_{<d,> \text{, } s> \text{, } i> \text{, } t> \text{, } d>} \lambda s_1 \lambda s_2 \exists d_1 \exists d_2 \left[ P(d_1)(s_1) \land P(d_2)(s_2) \land d_2 > d_1 \right] \quad \text{Gradable}
\end{align*}
\]

On the semantics of (60a), \( yuè \) combines with a non-gradable verb; the result is a set of pairs of events ordered based on their temporal precedence.

Although the above analysis provides a straightforward explanation for the necessarily temporal reading of (2a), it is built upon a questionable assumption--gradable predicates (gradable adjectives and gradable verbs) cannot take a temporal argument. Lin (2009), based on comparatives in Mandarin Chinese, explicitly argues against this claim. He argues that in the comparative in (60) the adjective \( kāixīn \) takes both a temporal argument and a location argument, as shown in (62).

\[
\begin{align*}
\text{(61)} & \quad \text{tā zuótiān zài-xuéxiào bǐ wǒ jiàntiān zài-jīālǐ kāixīn.} \\
& \quad \quad \text{he yesterday at-school} \quad \text{I today at-home} \quad \text{happy} \\
& \quad \quad \text{‘He was happier at school yesterday than I am at home today.’}
\end{align*}
\]

\[
\begin{align*}
\text{(62)} & \quad \llbracket [kāixīn] \rrbracket = \lambda d \lambda i \lambda x. x's \text{ happiness at location } l \text{ at time } i \geq d. \\
& \quad \quad \text{(Lin 2009, 22)}
\end{align*}
\]

Moreover, it is well-know that in Mandarin Chinese there is no clear morphological distinction between adjectives and verbs. Both categories allow direct affixation of an aspect marker, as shown in (63).

\[18\] See Li & Fasola (2010) for a more detailed discussion on this approach.
If adjectives differ from non-gradable verbs in not being able to take a temporal argument, it is curious why they can bear an aspect marker just like non-gradable verbs. The question of what predicates can lexicalize a temporal argument is a complicated issue which deserves much more extensive discussion than we can offer here. Therefore, we will leave it for future work.

6 Conclusion and implication

In this paper, we have shown that in Mandarin Chinese the $yuè$ $V_{\text{non-gradable}}$ $yuè$ $A$ structure (e.g., (2)) is semantically distinct from the $yuè$ $A/V_{\text{gradable}}$ $yuè$ $A/V_{\text{gradable}}$ structure (e.g., (1)): the former has a necessarily temporal reading that the latter does not have. We argued that this necessarily temporal reading arises as the non-gradable verbal predicate (e.g., pào ‘to run’) is coerced to have a domain parallel to that of a gradable predicate, where event intervals, like degree intervals, are totally ordered under a proper subinterval relation and have a common starting point.

Our analysis, if on the right track, has several theoretical implications. First, it differentiates between the proper subinterval relation ($\sqsubseteq$) and the greater than relation ($<$), which are often used interchangeably in the description of comparisons. On the standard degree-based analyses, the truth-conditions of a simple comparative like (64) can be represented in two possible ways, as shown in (65). In (65a) the comparative relation is captured by a greater than relation between two degree points on a scale; in (65b) it is captured by as a proper subinterval relation between two degree intervals.

(64) John is taller than Mary is.

(65) a. $\exists d_1 \exists d_2 [\text{tall}(d_1)(J) \land \text{tall}(d_2)(M) \land d_2 < d_1]$ 
    b. $\exists d_1 \exists d_2 [\text{tall}(d_1)(J) \land \text{tall}(d_2)(M) \land d_2 \sqsubseteq d_1]$ 

Although (65a) and (65b) are extensionally equivalent, there is a non-trivial difference between them. That is, as our analysis of $yuè$ shows, the subinterval relation ($\sqsubseteq$) is compatible with both comparisons of degree intervals and comparisons of event intervals, whereas the greater than relation ($<$) is only limited to comparisons of degrees. In this regard, the former encodes a more general notion of comparison than the latter.

Moreover, our analysis also connects to the recent study on the semantics of cross-categorial modifiers (Anderson 2013, Bochnak and Csipak 2014). Anderson (2013) shows that in English hedge words sorta and kinda exhibit a cross-categorial distribution, like $yuè$: they modify gradable adjective (e.g., 66a) as well as non-gradable verbs (e.g., 66b).\(^{19}\)

\[^{19}\] It should be noted that sorta and kinda have a wider distribution than $yuè$: the former can also modify non-gradable adjectives such as extinct and non-gradable NPs such as fairytale.

(i) a. It’s sad [Chinese river dolphins are] (sorta) extinct.
    b. a sorta fairytale.
(66)  a. Bill is sorta tall.
    b. Bill sorta swam over to the boat.

Both (66a) and (66b) have an approximative reading. (66a) means: Bill’s height is close to the standard for being tall. (66b) means: Bill did something that was like swimming over to the boat. Based on these interpretations, Anderson argues that sorta is a degree modifier. When combined with a non-gradable predicate such as swim in (66b), the latter is coerced to have a degree argument that measures precision (of type <d, <e, t>>). The role of sorta is to saturate that degree argument of the predicate with a value that is close to, but less than, the contextually provided standard.

In Mandarin Chinese yuè has a similar distribution to sorta: it appears in front of a gradable as well as a non-gradable predicate. However, in our analysis, yuè is not analyzed as a degree modifier. When yuè is combined with a non-gradable predicate, the latter is coerced to have an ordered domain parallel to that of a gradable predicate rather than to be associated with a degree argument. If such an analysis is correct, it suggests that coercion can manifest itself in more than one way in the context of ‘cross-categorial’ modification: type-shifting (like sorta in English) or domain ordering (like yuè in Mandarin Chinese).

Appendix

This appendix offers a brief discussion of the semantics of two other types of yuè…yuè sentences, namely, yuè A yuè Vnon-gradable and yuè Vnon-gradable yuè Vnon-gradable. We show that the former does not have a necessarily temporal reading, while the latter does.

**yuè A yuè VPnon-gradable**

Comparing the yuè Vnon-gradable yuè A sentence in (67) and the yuè A yuè Vnon-gradable sentence in (68), they minimally differ from each other in the order of the predicates following yuè. Unlike (67) which only has a necessarily temporal reading, (68) can take either a temporal or a degree reading. For instance, When (68) is preceded by (67) like in (69), the sentence has a temporal reading: ‘the sadder he became, the more he cried’. But when uttered out of blue, (68) can be understood as ‘the sadder Zhangsan is, the harder he cries’.

(67) Zhāngsān yuè kū yuè shāngxīn yuè Vnon-gradable yuè A
    cry    sad
    ‘The more Zhangsan cried, the sadder he became.’

(68) ?Zhāngsān yuè shāngxīn yuè kū. yuè A yuè Vnon-gradable
    sad    cry
    (i)  ‘The sadder he became, the more he cried.’
    (ii) ‘The sadder he became, the harder he cried.’

(69) Zhāngsān yuè kū yuè shāngxīn, yuè shāngxīn yuè kū.
    cry    sad    cry
    ‘The more Zhangsan cried, the sadder he became; the sadder he became, the more he cried.’
The goal of this section is (i) to show that our analysis does not predict (68) to have a necessarily temporal reading like (67), and (ii) to provide an explanation for its degree reading.

Let us begin with considering the truth-conditions of (68). Under our analysis, (68) has the LF in (70). CP₁, 张三 yue shangxin, denotes a set of pairs of situations ordered based on Zhangsan’s sadness, as shown in (71a); CP₂, yue pro, kù, denotes a set of pairs of events ordered based on the subinterval relation, as shown in (71b).

(70) The LF of (66): \( \forall [\text{CP₁ yue₁ [Zhangsan sad]}, \text{[CP₂ yue₂[C_d[pro₁ cry]]]}] \).

(71) a. \([\text{yuè Zhangsan sad}] = \lambda s_1 \lambda s_2 \exists d_1 \exists d_2 [\text{sad(Zh)}(d_1)(s_1) \land \text{sad(Zh)}(d_2)(s_2) \land d_1 \sqsubseteq s_2 d_2]\)
b. \([\text{yuè C_d pro₁ cry}] = \lambda e_1 \lambda e_2 \exists E [\text{run(Zh, e₁)} \land \text{run(Zh, e₂)} \land \text{run(Zh, E)} \land e_1 \sqsubseteq e_2 \land e_1 \sqsubseteq E \land e_2 \sqsubseteq e_2]\)

When CP₁ and CP₂ compose with the implicit universal quantifier in (72a), (68) has the truth-conditions in (72b).

(72) a. \([\lambda P_{<s}, \lambda s_1 \lambda s_2 \forall Q_{<v}, \lambda v, \lambda d_1 \lambda d_2} \forall s_1 s_2 [P(s_1)(s_2) \rightarrow \exists e_1 \exists e_2 [Q(e_1)(e_2) \land R(s_1, e_1) \land R(s_2, e_2)]]\)
b. \([\lambda \exists s_1 \exists s_2 \exists d_1 \exists d_2 \exists \text{sad(Zh)}(d_1)(s_1) \land \text{sad(Zh)}(d_2)(s_2) \land d_1 \sqsubseteq s_2 d_2] \rightarrow \exists e_1 \exists e_2 \exists E [\text{run(Zh, e₁)} \land \text{run(Zh, e₂)} \land \text{run(Zh, E)} \land e_1 \sqsubseteq e_2 \land e_1 \sqsubseteq E \land e_2 \sqsubseteq E \land R(s_1, e_1) \land R(s_2, e_2)]]\)

(72b) says: for any pair of situations \(s_1\) and \(s_2\) such that Zhangsan is sad to \(d_1\) in \(s_1\), Zhangsan is sad to \(d_2\) in \(s_2\), and \(d_1\) is a proper subinterval of \(d_2\), there is a pair of subevents of Zhangsan’s running \(e_1\) and \(e_2\) such that \(e_1\) is associated with \(s_1\) and \(e_2\) is associated with \(s_2\), \(e_1\) is a proper subinterval of \(e_2\).

Because (71b) says nothing about the temporal ordering between \(s_1\) and \(s_2\), there are two possibilities: (i) \(s_1\) temporally precedes \(s_2\), which is an ordering that corresponds to the ordering of \(s_1\) and \(s_2\) on a scale of sadness, and (ii) \(s_1\) temporally follows \(s_2\), which is a reverse ordering of \(s_1\) and \(s_2\) on a scale of sadness. Below we consider these two possibilities respectively. The example in (69) illustrates the first possibility. The first clause of (69)(i.e., (67)) has the truth-conditions in (73). It says: for any pair of subevents of Zhangsan’s crying \(e_1\) and \(e_2\), if \(e_1\) is a proper subinterval of \(e_2\), \(e_1\) is associated with \(s_1\); \(e_2\) is associated with \(s_2\) Zhangsan is sadder in \(s_2\) than in \(s_1\). Because of the temporal constraint of \(R\) in (49), \(s_1\) temporally follows \(s_2\). This is illustrated in (74).

(73) \( [\exists \forall e_1 e_2 \exists E [\text{run(Zh, e₁)} \land \text{run(Zh, e₂)} \land \text{run(Zh, E)} \land e_1 \sqsubseteq e_2 \land e_1 \sqsubseteq E \land e_2 \sqsubseteq E \rightarrow \exists d_1 \exists d_2 \exists \text{sad(Zh)}(d_1)(s_1) \land \text{sad(Zh)}(d_2)(s_2) \land d_1 \sqsubseteq d_2 \land R(e_1, s_1) \land R(e_2, s_2)] \)

Given the temporal precedence of \(s_1\) and \(s_2\), the second clause of (69)(i.e., (68)) are truth-conditionally equivalent of (67). It says: if Zhangsan is sadder in \(s_2\) than in \(s_1\), \(e_2\) is a super-event of \(e_1\), as shown in (74).
On the second possibility, $s_1$ temporally follows $s_2$, as shown in (75). In such a case, the truth-conditions of (72b) predict that $s_2$ is associated with $e_2$ and $s_1$ is associated with $e_1$, which essentially contradicts the constraint of $R$ in (49).

We suspect that the degree reading of (68) arises in case like (75) where the temporal ordering between $s_1$ and $s_2$ is distinct from their ordering on a degree scale. To avoid this conflict, the audience may re-interpret (68) as either (76a) or (76b).

(76) a. Zhāngsān yuè shāngxīn, kū-de yuè lǐhai.
    sad cry-De hard
    ‘The sadder Zhangsan is, the harder he cries.’

b. Zhāngsān yuè shāngxīn kū-de yuè jiǔ.
    sad cry-De long
    ‘The sadder Zhangsan is, the longer he cries.’

$yuè$ VP\text{non-gradable} $yuè$ VP\text{non-gradable}$

Finally, our analysis predicts that the $yuè$ VP\text{non-gradable}$yuè$ VP\text{non-gradable}$ structure, exemplified in (77), has a necessarily temporal reading. Intuitively, (77) describes two concurrent events, the event of Zhangsan’s running and the event of Lisi’s chasing.

(77) Zhāngsān yuè pāo, Līsì yuè zhūī.
    run chase
    ‘The more Zhangsan ran, the more Lisi chased him.’

On our analysis, (77) has the truth-conditions in (78), which indicates that two sets of subevents are temporally overlapped, as shown in (79).
There is an event of Zhangsan’s running E. For any pair of subevents of E that share the same initial point, e₁ and e₂, if e₁ is a proper subinterval of e₂, e₁ is associated with e₃, and e₂ is associated with e₄. e₃ and e₄ are subevents of Lisi’s chasing E. They share the same initial point and e₃ is a proper subinterval of e₄.

References


