The absence/presence of aspectual markers in narrative discourse in Mandarin

Chou Mo & Bert Le Bruyn (Utrecht Institute of Linguistics)

Intro | Sun (2014) was the first to systematically study the constraints governing the absence/presence of aspectual markers in Mandarin. The generalization she arrives at is that unmarked forms can only receive a stative or generic interpretation. The upshot of this is that eventive predicates always take aspectual markers.

The focus of Sun was on out-of-the-blue sentences. We tackle sentences in context and focus on narrative discourse. We show that (i) Sun's generalization cannot be extended to narrative discourse in its current form and (ii) establish dynamic semantics as one of the crucial factors governing the absence/presence of aspectual markers.

Sun's generalization in narrative discourse | The role of aspectual markers in narrative discourse has been studied before (e.g. Smith & Erbaugh 2005). Sun's generalization has however not been put to the test. To ground our analysis in solid data we carry out a corpus study on an extended fragment of narrative discourse: the Mandarin version of the first Chapter of Camus' *L'Étranger* (421 verbs).

To check Sun's generalization we need to annotate aspect. Given the size of our corpus as well as the intrinsic difficulty in determining aspectual class we rely on annotation projection. This is a parallel corpus technique that annotates data from one language on the basis of data from other languages. The data we rely on for the annotation of aspect in Mandarin are based on tense use in the French source text and in its translations to Spanish, English and German. Even though there is no one-to-one correspondence between tense use and aspectual class, these data do give a good approximation. *Table 1* presents the main combinations of tenses in our corpus, adds abstract tense labels, the corresponding aspectual labels and corpus frequencies:

Table 1: Tense use for annotation projection			
Tense use in French, Spanish, German and English	**Abstract tense**	**Aspect**	**Frequency**
< passé composé, pretérito indefinido, perfekt, simple past>	Past Perfective	Event	265
< imparfait, pretérito imperfecto, präteritum, simple past >	Past Imperfective	State	87
< présent, presente, präsens, simple present>	Present	State	28
<imparfait, pretérito imperfecto, präteritum, past continuous>	Past Progressive	State	24
< passé composé, pretérito perfecto compuesto, perfekt, present perfect>	Present Perfect	Event	8

If the generalization Sun arrived at for out-of-the-blue sentences were valid for narrative discourse, we would expect to find unmarked forms primarily with Past Imperfective and Present and marked forms primarily with Past Perfective. *Table 2* shows that this prediction is not borne out:

Table 2: Use of unmarked forms		
Abstract tense	Unmarked form	%
Past Perfective	140	56
Past Imperfective	65	26
Present	24	10
Past Progressive	14	6
Present Perfect	3	1
Total	**246**	**100**

246 out of 421 forms are unmarked. This is close to 60%. An even more striking result is that the unmarked forms primarily pattern with Past Perfectives, a tense category we associate with events. These data invalidate Sun's hypothesis that aspectual markers obligatorily cooccur with events, and states remain unmarked.

Table 2 does not factor in contexts in which the different tenses pattern with aspectual markers. This does not impact our conclusion on Sun's generalization in its current form. It does however mask the fact that Sun's generalization carries over to narrative discourse in a weaker form, *viz.* as identifying aspect as one of the factors at play. Normalized Pointwise Mutual Information (NPMI, Bouma 2009) allows us to take into account bi-directional associations between tenses and aspectual markers. We provide the NPMI values for a range of tenses/aspectual markers in *Table 3*. NPMI ranges from -1 to 1. Positive values indicate a positive association, negative values a negative one. In the columns we find the most frequent aspectual markers. *Le1* is post-verbal *le*, *Le12* can – due to its position – be post-verbal or sentence-final *le*. On its strong form, Sun's

generalization predicts values close to 1 for <Past Imperfective, Unmarked form> and <Present, Unmarked form> and values close to -1 for <Past Perfective, Unmarked form>. On its weak form, it predicts positive values for the former and a negative value for the latter.

Table 3: Association strength between tenses and aspectual markers							
	Unmarked form	RVC	Le1	Le12	Zhe	Zai	Guo
Past Perfective	-0.1	0.1	0.2	0	0	-0.2	0
Past Imperfective	0.1	-0.2	-0.2	0	0	0	0
Present	0.1	-0.2	0	0	0	0	0
Past Progressive	0	-0.2	0	-0.1	0.3	0.5	0
Present Perfect	-0.1	0	0	0.2	0	0	0.6

Table 3 shows that there is a slight positive association for <Past Imperfective, Unmarked form> and <Present, Unmarked form>. We also find a slight negative association for <Past Perfective, Unmarked Form>. These tendencies are in line with a weak version of Sun's generalization. We conclude that aspectual class is likely to be a factor that is regulating the absence/presence of aspectual markers in narrative discourse but that it cannot be the only one.

From aspect to dynamic semantics | With aspect no longer being the only factor at play, we need to identify the other factors governing the presence/absence of aspectual markers. We argue that dynamic semantics is one of them.

The intuition we pursue builds on the insight that tense/aspect markers can sometimes be used to refer to a past event but are felt to be unnatural in a narrative context. The Dutch Perfect is a case in point: we can use it to say that we have done something (*Ik ben gisteren naar de winkel gegaan* litt. 'I have yesterday to the shop gone') but we cannot use it in a narrative context (de Swart 2007). We propose that aspectual markers in Mandarin are obligatory at the edges of narrative chains and more specifically to introduce new stages in the narrative that are not directly linked to the preceding discourse.

To test our hypothesis, we annotated the French original of our corpus for narrative stages, restricting ourselves to the first one hundred non-future forms. Simple examples of stages in the opening pages of *L'Étranger* include (i) an encounter with the protagonist's boss, (ii) lunch in a restaurant, (iii) a bus drive, etc. We marked verbs that introduced these stages as 'New-stage verbs' and the others as 'Within-stage verbs'. We then calculated the association between these two types of verbs and the absence/presence of aspectual markers based on their NPMI:

Table 4: Association strength between new/within-stage verbs and unmarked/aspectually marked forms		
	Unmarked form	Aspectually marked form
New-stage verb	-0.5	0.6
Within-stage verb	0.7	-0.4

Table 4 shows that – unlike the association between aspectual class and the presence/absence of aspectual markers – there is a strong association for <New-stage verb, Aspectually marked form> and <Within-stage verb, Unmarked form>. The negative values of the other combinations confirm these tendencies. We conclude that dynamic semantics in the form of narrative stages is one of the factors regulating the absence/presence of aspectual markers in Mandarin.

Conclusion | Our corpus study established that Sun's generalization only carries over to narrative discourse in a weakened form: eventive verbs have a tendency to take an aspectual marker but do not do so obligatorily. Aspect is thus only one of the factors regulating the presence/absence of aspectual markers in narrative discourse. A full understanding of aspect requires us to identify the other factors. One candidate with a high predictive value we have identified in our corpus is the notion of narrative stage. In the presentation we model this new notion in an event based semantics.

Sun (2014). Temporal construals of bare predicates in Mandarin Chinese. LOT: Utrecht. | Smith & Erbaugh (2005). *Linguistics, 43*(4), 713-756. | Bouma (2009). NPMI in collocation extraction. *Proceedings of GSCL*, 31-40. | De Swart (2007). *Journal of pragmatics, 39*(12), 2273-2307.