

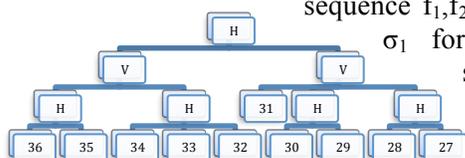
Temporal Succession and Aspectual Type in Visual Narrative

Dorit Abusch

Following Greenberg (2011), the propositional content of a picture P is the set of situations σ that project to the picture via a geometric projection G , $\llbracket P \rrbracket = \{\sigma \mid G(\sigma) = P\}$. Silent comics, which are comics without speech bubbles, thought bubbles, or captions, are used here as a pure case of visual narrative constructed as a sequence of pictures. As in linguistic applications of temporal discourse representation theory, we aim to account for the semantic content of such picture sequences by adding temporal relations to the conditions on described situations that are provided by the semantics. In order to construct a linear sequence from panels laid out on a page, a comic



is parsed into a tree structure labeled with horizontal and vertical stacking operations, as in the example at the left (Tanaka et al. 2007, Cohn 2008). Thus a comic is formally comparable to a natural language narrative, with a syntactic sequencing of objects (sentences or panels), each of which has a propositional semantic content.



Comics are governed by a default isomorphic interpretation of panel sequencing – given panel sequence f_1, f_2 at least by default described situation σ_1 for f_1 temporally precedes described situation σ_2 for f_2 . However, McCloud



(1994) pointed out examples of panels with apparent stative or activity aspect, whose described situations



seem to overlap the described situations for adjacent panels. For instance in the sequence at the right from Tanaka's *Gon*, the described situation with a bobcat on his back seems to overlap the described situation with some objects (understood to be eagles) flying overhead. This suggests a discourse representation construction rule similar to Kamp and Rohrer (1983) or Kamp and Reyle (1993) that is sensitive to aspectual type of the propositional content.

We argue formally that all panels have stative literal content, using properties including cumulativity: if σ_1 satisfies picture P , σ_2 satisfies P , σ_1 and σ_2 are temporally sequenced and temporally contiguous, then $\sigma_1 + \sigma_2$ satisfies P (Krifka 1989). This makes it impossible to condition DRS construction on aspect. A second problem is that even panels with apparent overlapping interpretation have temporally specific information that is not understood as constant through the adjacent panels in realistic models, for instance the eagles being in a specific location. This supports an account similar to Dowty (1982), with a rigid interpretation of temporal sequencing of described situations in the model, accompanied by inferred or implicated overlap, but only for an entailment of the literal content. The overlap is an enrichment of literal content.

Much of the difference between language and picture sequences comes from the literal content of pictures being weak. On the basis of geometric projection, the panel on the right is consistent with scenes with three real boys, scenes with acrylic statues of three boys, and scenes with billboard pictures of three boys. However the author intends and the reader construes an enriched content, where described situations contain three real boys. Similarly, the author intends and reader construes an enriched content that entails that in a super-situation σ_2 of the described situation σ , the boys move and one of them sinks a basket. This enriched content is eventive. Only at the level of enriched content can distinctions of temporal constitution be made.

