Pictorial narratives and temporal refinement

Recent investigations into the semantics of pictorial narratives draw parallels with natural language (Ab14, AR17, RA17, among others). A recurrent theme is the significance of viewpoint in analyzing a picture as a projection. An argument for viewpoint centering (due to Ross) is repaired in "lineland" by RA17. Can we extend the notion there of viewpoint projection smoothly to capture narrative temporal progression? A step towards a more comprehensive conception of viewpoint, with dimensions for time as well as space (mixing pictures and language), is taken by the account of temporal projection proposed below. Complications with temporal succession studied in Ab14 and AR17 are analyzed by lifting the string perspective from linelands to the sequence $p_1 p_2 \cdots p_n$ of pictures p_i constituting a pictorial narrative.

Ab14 adopts Do86's rule of temporal succession to interpret $p_i p_{i+1}$ as p_i *and then* p_{i+1}, which we draw as (**D**) below to mark temporal boundaries.

(**D**) $\boxed{p_i}\boxed{p_{i+1}}$ (**F**) $\boxed{p_i, p_{i+1}}$ (**I**) $\boxed{p_i, p_{i+1}}\boxed{p_i, p_{i+1}}$

An alternative to (**D**) is (**F**), p_i *and simultaneously* p_{i+1}, with p_i and p_{i+1} *f*rozen in the same box. Let us assume a box carves out an interval of time, and intervals of adjacent boxes abut. Now, inasmuch as pictures are stative (Ab14), and statives are inertial (Do86), inertia suggests that in (**D**), p_i persists forwards and p_{i+1} persists backwards unless a force intervenes. That is, in the absence of an opposing force, inertia turns (**D**) into (**I**), which cumulativity reduces to (**F**). The conclusion is that if (**D**) and (**F**) are distinct, it is because a barrier between p_i and p_{i+1} is erected by some force. But what force? Can we make this account precise, and develop it further?

Let the semantic value $[\![p]\!]$ of a picture p be the set of world-viewpoint pairs that project to p (RA17). We say p *clashes with* p' if no world can have a viewpoint projecting to p and also one projecting to p'

$$p \text{ clashes with } p' \iff [\![p]\!]_\exists \cap [\![p']\!]_\exists = \emptyset$$

where $[\![p]\!]_\exists$ is the set of worlds w such that for some viewpoint v, $(w, v) \in [\![p]\!]$. Now, p_i clashing with p_{i+1} is a sufficient condition for a force to block inertial flow in (**D**) (in the account above). To say more about such a force, it is useful to bring in formulas φ, with semantic values $[\![\varphi]\!]$ that collect the world-viewpoint pairs satisfying φ. We say a formula φ *describes* a picture p if every world-viewpoint pair projecting to p satisfies φ

$$\varphi \text{ describes } p \iff [\![p]\!] \subseteq [\![\varphi]\!].$$

Free perception pairs $p_i p_{i+1}$ where p_i depicts an agent that sees p_{i+1} have "closely similar examples in natural language narratives" consisting of "an eventive clause that describes someone looking, followed by a stative clause describing what is seen" (AR17). How does p_i's eventive clause square with Ab14's conclusion that "all propositions contributed by pictures are formally stative"? At stake in the eventive/stative question is the dynamic/static contrast (KR93) that is discernible between (**D**) (the two boxes in which represent change) and (**F**) (where time stands still). Although (**F**) is arguably a natural reading for the veridical case of free perception (where p_i and p_{i+1} are projections of the same world), (**F**) fails to represent the essential point that veridical or not, p_{i+1} enters the agent's mental state as a result of the act of perception depicted by p_i. To represent this point in (**D**), p_{i+1} must be understood there as being in the agent's head, and not necessarily the external world. Returning to forces, the general conclusion is that actions described by eventive clauses serve as forces that keep boxes from collapsing into a single time. This is *not* to say that in free perception $p_i p_{i+1}$, the picture p_i is eventive as a whole. The claim rather is that given a set Φ of formulas, a picture p can be decomposed to the subset

$$\Phi[p] \;:=\; \{\varphi \in \Phi \mid [\![p]\!] \subseteq [\![\varphi]\!]\}$$

of Φ consisting of formulas describing p, some of which may be stative, and others eventive (*pace* Ab14).

Breaking a picture p down into a set $\Phi[p]$ of formulas allows us to refine our account of inertial flow so that a picture p need not flow inertially as an unbroken whole (all-or-nothing) to an adjacent box with picture p'. Instead, inertial flow from p to that box is restricted to the stative formulas $\varphi \in \Phi[p]$ that do not clash with other stative formulas already in that box, such as the stative formulas in $\Phi[p']$. Specifying when stative formulas clash is complicated by viewpoint centering, under which a formula φ and its negation $\neg\varphi$ may hold at the same world, with a change in viewpoint. This leads to boxes that consist of neither pictures p nor formulas φ on their own, but rather pairs (p, φ). The possibility that φ is eventive necessitates a modification of block compression in Fe16, used to equate, for instance, (**F**) and (**I**) above (assuming cumulativity and divisiveness). These and other problems are addressed in the full paper, where the set Φ of formulas is allowed to vary over finite sets, yielding a notion of time with bounded but refinable granularity, thereby sidestepping some difficulties with strong pictorial contents (RA17).

References

Ab14 Abusch, D. Temporal succession and aspectual type in visual narrative. In *The Art and Craft of Semantics: A Festschrift for Irene Heim*, 2014.

AR17 Abusch, D. & M. Rooth. The formal semantics of free perception in pictorial narratives. In *21st Amsterdam Colloquium*, 2017.

Do86 Dowty, D. The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics & Philosophy* 9(1):37–61, 1986.

Fe16 Fernando, T. Prior and temporal sequences for natural language. *Synthese* 193(11):3625–3637, 2016.

KR93 Kamp, H. & U. Reyle. *From Discourse to Logic*, Kluwer, 1993.

RA17 Rooth, M. & D. Abusch. Picture descriptions and centered content. In *Sinn und Bedeutung 21*, 2017.