

LOG OF CHANGES – 'Space and the Vision-Language Interface: A Model-Theoretic Approach'

In this new draft I have attempted to make my position as explicit and formal as possible, as both reviewers point that the main flaw of the previous draft was a clear stance on the central topics of the paper. Consequently, this version is substantially different from the one I previously submitted. Now the paper has the following structure, presented below per section:

1. Introduction
2. Brief outline of previous literature
3. Review of “static” and “dynamic” Vision; formal treatment of Vision; Review of spatial Language with DRT (formal) treatment;
4. Formal proposal on the Vision-Language interface; Discussion of its empirical consistency; Discussion of broader topics;
5. Conclusions

I have made the following general and particular changes, in order to address points raised by both reviewers. Hopefully, these changes will answer the questions raised by reviewers, in thorough detail.

First, I follow one reviewer's point and treat *Language* and *Vision* as two distinct (but “connected”) modules of *Cognition*, and define in some detail their properties and their relation. I thank this reviewer in the main text body for pointing out this mistake I made in the first draft. Second, the current draft is slightly longer than the previous draft, but I have 'compressed' all reviewing sections, in order to include the new sections and discussions. For instance, the sections in which I review the literature on 'static' and 'dynamic' vision (current section 3.1; 3.3 corresponding to previous sections 3.1;3.2;3.3) have been approximately halved. I believe that this is a good compromise between offering a precise proposal and related discussion, and compressing those sections in need of a synthesis.

Third, both reviewers ask a series of questions regarding the nature of the sets V and DRSs, and how the notion of transduction is related to the relation between these two levels of comprehension. In other words, both reviewers ask me to be explicit about my assumptions and positions. In order to answer these questions, I spell out a formal treatment of the elements in the set V and their relation with DRSs, by adding three novel sections in the paper (section 3.2; 3.4; 4.1). I basically use a simplified First Order Logic (what I call “Logic of Vision”) to explicitly represent visual computations (section 3.2; 3.4). I then compare this logic to the one represented by DRT, and propose a two-way relation between this novel logic and DRT, suggesting that this relation formalises my earlier, sloppy approach to transduction.

Fourth, both reviewers suggest that I should have addressed the relation between DRSs and conceptual (visual) content, especially since the analysis of this relation can tell us a lot about the more general relation between modules of cognition. I believe that I address these topics in this proposal (especially in the new section 4.2; 4.3), and more in general I spell out a specific stance on how these two modules are related, and what are the consequences for the broader FLN/FLB debate. There are other relatively 'minor' points that I address in the new draft.

Fifth, both reviewers criticise my discussion of 'merge'. I have removed this discussion, and only discuss the parallels between DRT's 'merge' operation and the notion of set union. There is no other use of this concept in the paper, as it would not be appropriate, as the reviewers correctly point out.

Sixth, one reviewer points out that my discussion in the paper would be best thought as dealing with *spatio-temporal* relations and how they are represented in Vision and Language. I point out this fact early in the paper and thank the reviewer. One reviewer also points out that my proposal is really focused on English, so claims about 'Language' are a bit premature. I also point out this limitation early on, and thank this reviewer accordingly.

Seventh, two native speakers of a dialect of English have proof-read the paper. I hope that such proof-reading process will suffice to make the current draft virtually flawless, from a stylistic point of view. Overall, I sincerely think that the current draft is a much improved work, because it aims to answer to several, if not all, of the questions rightfully (and insightfully) raised by the reviewers, hopefully in a clear and unambiguous way. If there are further issues that the reviewers think I should address, I will certainly do so in the next round of the reviewing process.