

Workshop on the Interaction of Vision and Language in Cross-Modal Comprehension

Presentation 2

Multimodal Comprehension of Information Graphics and Language by Cengiz Acarturk

Documents containing information in different representational modalities (i.e. text, tables, pictures, figures, diagrams, graphs) as well as different sensory modalities (i.e. visual, auditory and haptic representations) are used frequently in newspapers, educational settings and scientific literature. Nevertheless, the illustration-type specific differences in multimodal processing have been seldom considered in detail. More specifically, whereas many design guidelines for illustrations and information graphics have been published regularly, well-founded research on cognitive and perceptual processing of these representational modalities is comparatively rare.

Pictorial illustrations can informally be characterized by their visual resemblance to the objects they stand for. Nevertheless, pictorial illustrations do not possess an internal syntax in the sense of representational formats as discussed by Kosslyn (1980; p.31). On the other hand, information graphics are representational artefacts, which possess internal syntactic structures. Thus the syntactic analysis of a graph is fundamental for succeeding processes of semantic and pragmatic analyses in graph comprehension (Kosslyn, 1989; Pinker, 1990; Tversky, 2004). These differences between pictorial illustrations and information graphics can also be summarized from another perspective: the former has a lower level of abstraction than the latter. Pictorial illustrations and their referents have spatially similar layout (i.e. iconic similarity), which for example in the case of photographs guarantees an optically veridical mapping from the visual-world object to the external representation. Although maps are abstract as well, the level of abstraction is different from graphs in that maps have iconic similarity to what they represent (Tversky, 1995). In other words, similar to pictorial and diagrammatic illustrations, maps use the spatial layout in a similar way to what they represent. Graphs are abstract in the sense that they depict abstract relations between variables that refer to real-world entities. Graphs lie between words of the language and pictorial illustrations (Winn, 1987; Schmidt-Weigand, 2005).

In particular from the perspective of communication, conceptual representations are the pivot of human cognition: The level of conceptual representations, which encodes meaning independent from any particular language, is the content-specifying level in language comprehension as well as in language production. Furthermore, conceptual structures are an essential part of the interfaces between language and perception as well as action (cf. Jackendoff's interface architecture 1997, 2002). Concerning the use of information graphics in multimodal documents, if information graphics and text constitute a multimodal document, both modalities have to contribute to a common conceptual representation. As described in (Habel & Acarturk, 2007), the interaction between information graphics and language is mediated by common conceptual representations, in particular a set of basic spatial concepts, which is fundamental for the terminology of graphs as well as for the terminology of the specific domain in which an individual graph is applied.

In this workshop slot, the work in progress in multimodal comprehension of text-graph documents will be discussed from the interdisciplinary perspective of cognitive science, psychology, computer science and linguistics. There are at least three primary dimensions to systematize the scientific investigation of comprehending multimodal text-graphics documents. The first is the type of graphs (e.g. line graphs, bar graphs etc.). The second is about the means for providing interaction between text and graphics, in particular elements on the graph (e.g. axis labels, annotations etc.). On the one hand, textual annotations in statistical graphics play the role of bridging the graph region and the main text (e.g. paragraphs) of a multimodal document (Acarturk et al., submitted 2008a); on the other hand the graphical means for annotations, such as a thin line connecting the annotation text to the relevant part of the graph, play important role for the construction of the co-reference relations between the annotation and the graph elements within the graph region of the document (Acarturk et al., submitted 2008c). The third dimension

considers the conceptual structures and the contents that are communicated by text and graphics, e.g. events, their temporal properties and their causal relations (Acarturk et al., submitted 2008b). Depending on the time slot of the workshop, the interaction between spoken language and visual graphs will be discussed based on the findings of the studies in progress (Acarturk et al., work in progress 2008).

References

- Acarturk, C., Habel, C., Cagiltay, K., & Alacam, O. (submitted 2008a). Multimodal Comprehension of Language and Graphics: Graphs with and without annotations.
- Acarturk, C., & Habel, C. (submitted 2008b). *Multimodal Text-Graphics Comprehension: The Role of Annotation Position in Causal Attributions*.
- Acarturk, C., Habel, C., & Cagiltay, K. (submitted 2008c). Multimodal Comprehension of Graphics with Textual Annotations: The Role of Graphical Means Relating Annotations and Graph Lines.
- Acarturk, C., Habel, C., Cagiltay, K., & Alacam, O. (work in progress 2008). The role annotations and sensory modality on learning from graph and text.
- Habel, C., & Acarturk, C. (2007). *On reciprocal improvement in multimodal generation: Co-reference by text and information graphics*. Paper presented at the Workshop on Multimodal Output Generation (MOG 2007), Aberdeen, United Kingdom.
- Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT-Press.
- Jackendoff, R. (2002). *Foundations of language. brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Kosslyn, S. M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.
- Kosslyn, S. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185–225.
- Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial Intelligence and the Future of Testing* (pp. 73-126). Hillsdale, NJ: Erlbaum.
- Schmidt-Weigand, F. (2005). *Dynamic visualizations in multimedia learning: The influence of verbal explanations on visual attention, cognitive load and learning outcome*. Justus-Liebig-Universität Gießen.
- Tversky, B. (1995). Cognitive Origins of Graphic Conventions. In F. T. Marchese (Ed.), *Understanding Images* (pp. 29-53). New York: Springer-Verlag.
- Tversky, B. (2004). Semantics, syntax, and pragmatics of graphics. In K. Holmqvist & Y. Ericsson (Eds.), *Language and visualisation* (pp. 141-158). Lund: Lund University Press.
- Winn, B. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration* (Vol. 1, Basic research pp. 152-198). New York: Springer-Verlag.