

Preface

 <https://doi.org/10.1075/hcp.23.01pre>

Pages xi–xiii of

Discourse, Vision, and Cognition

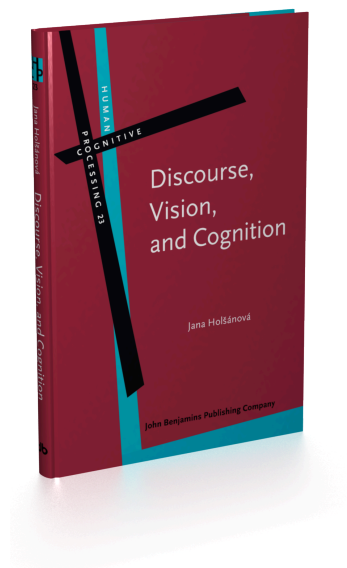
Jana Holšánová

[*Human Cognitive Processing*, 23] 2008. xiii, 202 pp.

© John Benjamins Publishing Company

This electronic file may not be altered in any way. For any reuse of this material written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

For further information, please contact rights@benjamins.nl or consult our website at benjamins.com/rights



Preface

While there is a growing body of psycholinguistic experimental research on mappings between language and vision on a word and sentence level, there are almost no studies on how speakers perceive, conceptualise and spontaneously describe a complex visual scene on higher levels of discourse.

This book aims to fill this gap by exploring the relationship between language, vision and cognition in spoken discourse. In particular, I investigate the dynamic process of picture discovery, the process of picture description and cognitive processes underlying both. What do we attend to visually when we describe something verbally? How do we construct meaningful units of a scene during the scene discovery? How do we describe a scene on different occasions for different purposes? Are there individual differences in the way we describe and visualise a scene?

The point of departure in this book are complex units in spoken language description and visual units during the process of picture viewing. Observers perceive the picture on different levels of detail, mentally group the elements in a particular way and interpret both WHAT they see and HOW the picture appears to them. All this is reflected in the stepwise process of picture viewing and picture description. The transcripts of spoken discourse – including prosody, pausing, interactional and emotional aspects – contain not only ideas about concrete objects, their shapes, qualities and spatial relations but also the describers' impressions, associations and attitudes towards them. This enables us to study the complex dynamics of thought processes going on during the observation and description of a scene.

In this book, I combine discourse analysis with a cognitively oriented research and eye movement protocols in order to gain insights about the dynamics of the underlying cognitive processes. On the one hand, eye movements reflect human thought processes. It is easy to determine which elements attract the observer's gaze, in what order and how often. On the other hand, verbal foci formulated during picture description are the linguistic expressions of a conscious focus of attention. With the help of a segmented transcript of the discourse, we can learn what is in the verbal focus at a given time. Thus, verbal

and visual data (the contents of the attentional spotlight) are used as two windows to the mind. Both kinds of data are an indirect source to shed light on the underlying cognitive processes.

It is, of course, impossible to directly uncover our cognitive processes. If we want to learn about how the mind works, we have to do it indirectly, via overt manifestations. The central question is: Can spoken language descriptions and eye movement protocols, in concert, elucidate covert mental processes? To answer this question, we will proceed in the following steps:

Chapter 1 presents a segmentation of spoken discourse and defines units of speech – verbal focus and verbal superfocus – expressing the contents of active consciousness and providing a complex and subtle window on the mind.

Chapter 2 focuses on the structure and content of spoken picture descriptions. It describes and illustrates various types of foci and superfoci extracted from picture descriptions in various settings. Chapter 3 takes a closer look at how speakers create coherence when connecting the subsequent steps in their picture descriptions. Chapter 4 discusses individual description styles.

While the first four chapters of this book deal with characteristics of picture descriptions in different settings, in the remaining four chapters, the perspective has been broadened to that of picture viewing. These chapters explore the connection between spoken descriptive discourse, picture viewing and mental imagery. The discourse segmentation methodology is therefore extended into a multimodal scoring technique for picture description and picture viewing, leading up to an analysis of correspondence between verbal and visual data.

Chapter 5 deals with methodological questions, and the focus is on sequential and processual aspects of picture viewing and picture description. The reader gets acquainted with the multimodal method and the analytical tools that are used when studying the correspondence between verbal and visual data.

In Chapters 6 and 7, the multimodal method is used to compare the content of the visual focus of attention (specifically clusters of visual fixations) and the content of the verbal focus of attention (specifically verbal foci and superfoci) in order to find out whether there is correspondence in units of picture viewing and simultaneous picture description. Both temporal and semantic relations between the verbal and visual data are investigated. Finally, clusters on different levels of the discourse hierarchy are connected to certain functional sequences in the visual data.

Chapter 8 focuses on the issue of visualisations in discourse production and discourse comprehension and presents studies on mental imagery associated with picture viewing and picture description.

The concluding Chapter 9 looks back on the most important issues and findings in the book and mentions implications of the multimodal approach for other fields of research, including evaluation of design, users' interaction with multiple representations, multimodal systems, etc.

This book addresses researchers with a background in linguistics, psycholinguistics, psychology, cognitive science and computer science. The book is also of interest to scholars working in the applied area of usability and in the interdisciplinary field concerned with cognitive systems involved in language use and vision.

Acknowledgements

I wish to thank Professor Wallace Chafe, University of Santa Barbara, for inspiration, encouragement and useful suggestions during my work. I also benefited from discussions and criticism raised by doctoral students, researchers and guest researchers during the eye tracking seminars at the Humanist Laboratory and Cognitive Science Department at Lund University: Thanks to Richard Andersson, Lenisa Brandão, Philip Diderichsen, Marcus Nyström and Jaana Simola. Several colleagues have contributed by reading and commenting on the manuscript. In particular, I want to thank Roger Johansson and Kenneth Holmqvist for their careful reading and criticism. I also wish to thank the anonymous reviewer for her/his constructive suggestions. Finally, thanks are due to my family: to our parents, my husband and my children Fredrik and Annika.

The work was supported by the post-doctoral fellowship grant VR 2002–6308 from the Swedish Research Council.