Attention and Vision in Language Processing
Attention and Vision in Language Processing
Preface

It is perhaps obvious that language interacts with vision and attention. In many everyday situations people give or receive directions or instructions for action or talk about things in their surroundings, say, a photo or a painting, the state of the kitchen sink, or an unfolding sports event. Moreover, our shifts in eye gaze are often (consciously or subconsciously) directed by spoken language input. A parent may tell a child to look at the beautiful leaf or a visitor may ask about a new gadget he has spotted in our living room. In short, the ability to integrate visual and auditory input with stored linguistic and nonlinguistic mental representations is a hallmark of human cognition. It follows from these observations that it is very likely to be a mistake to study language, vision, and attention independently of one another. It is surprising therefore that many of the cognitive processes occurring under such circumstances traditionally have been investigated in isolation although they are all involved when language is used.

The traditional approach to investigate language separately from other cognitive systems is a consequence of the theoretical development in the language sciences. Hockett, for instance, developed his list of features of human language (Hockett and Altmann 1968) and displacement (the fact that concepts need not refer to an object that is physically present) was considered a (or the) key feature of human language. Recent eye-tracking research however suggests that when the object of spoken language is actually physically present individuals express a strong tendency to refer to it. The cognitive system does this by orienting its visual sensory apparatus toward the object thereby linking a linguistically activated type representation to a specific perceptual token in the outside world. Cooper (1974), for instance, simultaneously presented participants with spoken fictional stories such as a safari in Africa and a visual display containing nine line drawings of concrete objects (e.g., a lion, a zebra, a tree, a camera, etc.). Participants in the study were asked to just listen to the stories. Cooper found that listeners’ fixations of the objects were very closely time-locked to the unfolding speech input. Whenever a spoken word referred to an object participants rapidly shifted their overt attention to the object or similar objects even though this was not required
for any task. Studies within the visual world paradigm (Tanenhaus et al. 1995) later confirmed that participants tend to very rapidly fixate objects which are visually similar (Dahan and Tanenhaus 2005; Huettig and Altmann 2004, 2007), similar in meaning (Huettig and Altmann 2005; Yee and Sedivy 2006) or whose names are consistent with the unfolding speech signal (Allopenna et al. 1998). In other words, although language need not refer to an object that is physically present, it is often used in such a way. Moreover, spoken language often guides visual orienting without volitional control. Language-mediated eye movements appear to be fast, unconscious, and largely overlearned (i.e., automatized through extensive practice, cf. Logan 1988). It seems that some prior conditions need to be met for language to be able to drive eye movements (for instance, to actively listen to the speech and a willingness to make eye movements, i.e., to look around rather than to force oneself to fixate one location). But once these conditions are met the available experimental evidence suggests that the integration of language with oculomotor behaviour may be unstoppable (Mishra et al. 2013, for further discussion).

Many researchers in psycholinguistics in the 1980s and 1990s have ignored vision and attention because of the widely held view that vision and language are separate informationally encapsulated modules (Fodor 1983). If the language system is encapsulated then there may be few reasons for language scientists to read work on vision and attention and study how people produce and understand utterances about objects and events they see. Fodor’s notion of the modularity of mind however has been strongly challenged in the past 20 years. Psycholinguists have shown how linguistic and nonlinguistic processes jointly determine how the language user understands language (see for instance the special issues in the Journal of Memory and Language, Ferreira and Tanenhaus 2007; and Acta Psychologica, Hartsuiker et al. 2011).

This volume is timely in the best sense in that it explicitly puts the focus on the interactions between language processing and other cognitive processes such as attention and vision. Key issues concern the functional architecture of the mind, how linguistic and nonlinguistic processes jointly determine language comprehension and production, and how the linguistic system interfaces with perceptual systems and attention.

Part I focuses on attention and vision in spoken language comprehension and production. Rigoli and Spivey (Chap. 1) set the stage by arguing that language is part of a continuous perception-action loop that a person develops with his/her environment. They make the case for a fundamental extension of cognition onto the surrounding environment. Rigoli and Spivey argue that by defining language and cognition as separate from other individuals and from our actions in the environment much of the field has been overlooking the very essence of language itself.

Altvater-Mackensen and Mani (Chap. 2) present the results from a novel experiment using a visual priming paradigm to assess the extent to which toddlers retrieve sub-phonemic detail during lexical access. They observe that both the retrieval of an object’s label and toddlers’ recognition of a word involve activation of not only phonemic but also sub-segmental information associated with the lexical representation of this word.
Hintz and Huettig (Chap. 3) present three eye-tracking experiments investigating the impact of the complexity of the visual environment during language-mediated visual search. Their results demonstrate that phonological word-object mapping is dependent on the nature of the visual environment. These findings add to a growing body of evidence that the nature of our visual surroundings induces particular modes of processing during language-mediated visual search.

Knoeferle (Chap. 4) reviews the literature on real-time visual context effects and visually situated language comprehension in children and in healthy young adults. She argues that visual context should play an active role in child language comprehension and that children benefit from a similarly rapid interplay of visual attention and language comprehension as young adults.

Norcliffe and Konopka (Chap. 5) consider the extent to which the planning processes involved in producing sentences are fine-tuned to grammatical properties of specific languages. They argue that incrementality is a general principle of production that applies cross-linguistically and that incremental encoding can be controlled by those aspects of the language that are responsible for linearization, namely grammar. Norcliffe and Konopka conclude that differences in language-specific grammatical constraints on word order result in differences in the order of encoding operations performed to produce grammatically correct utterances.

The second part of the book focuses on attention and vision processes in reading. Saint-Aubin and Klein (Chap. 6) discuss whether printed words are identified the same way when presented in isolation and in connected texts. They review the literature on the cognitive processes involved in reading connected texts with a special focus on the combined use of eye movements and the letter search task.

While there have been a larger number of studies on visual and attentional processes involved in reading of European languages using predominantly the Latin script, there is very little work on writing systems from other regions of the world. Winskel (Chap. 7) addresses this gap by focusing on visual and attentional processes involved in reading Thai script. Using Thai she illustrates how the distinctive features of an orthography (reading without interword spaces) can influence the visual and attentional processes involved in visual-word recognition and reading.

Part III contains chapters which focus on attention and vision in bilingual language processing. Hartsuiker (Chap. 8) asks how bilinguals select the right language for the particular context they are in and how do they do this. He first discusses the evidence on the degree to which bilinguals activate multiple languages and how this is constrained by linguistic cues. He then presents new lines of research that investigate visual language cueing (such as the language associated with familiar people’s faces, with the race of unknown faces, and with cultural icons).

Chabal and Marian (Chap. 9) provide an informative review of eye-tracking work on bilingual spoken word processing and parallel language activation. They make the case that multimodal investigations of language processing (e.g., using eye-tracking) are not only ecologically valid as they closely resemble real-world multimodal situations but also demonstrate how language interacts with other cognitive and perceptual systems in a non-modular mind.
Mishra and Singh (Chap. 10) present evidence for language nonselective activation in Hindi–English bilinguals using an oculomotor task. They show that bilinguals suffer interference during a simple visual task suggesting that they activate translations of spoken words unintentionally. Mishra and Singh discuss the detrimental influence such nonselective activation may have on cognitive processing.

Finally, Part IV focuses on language processing in a social context. Lev-Ari (Chap. 11) discusses how the identity of interlocutors influences which cues we attend to during language processing. She points out that adjustments in the allocation of attention can have cascading linguistic and social consequences. Lev-Ari argues that it is indispensable for our understanding of language processing to study cognitive processes in different social contexts.

Vinson, Dale, Tabatabaeian, and Duran (Chap. 12) further take up the issue of language processing and social context. They review the literature suggesting that social cues (including low-level perceptual variables, perception of another’s gaze, knowledge of another’s belief states) influence language processes. They make the case for a systematic research agenda to uncover how various processes work together to bring about multimodal coordination between two or more interacting persons.

This volume emanates from the first “Attentive Listener in the Visual World” workshop held in 2012 at the Centre of Behavioural and Cognitive Sciences at the University of Allahabad, India. Studies of language, vision, and attention are intrinsically related. We hope that this volume will encourage further workshops and research on this crucial topic for the language sciences.

Hyderabad, India
Allahabad, India
Nijmegen, The Netherlands

Ramesh Kumar Mishra
Narayanan Srinivasan
Falk Huettig

References


## Contents

### Part I  Attention and Vision in Spoken Language Comprehension and Production

1  **Real-Time Language Processing as Embodied and Embedded in Joint Action** ................................ 3  
   Lillian Rigoli and Michael J. Spivey

2  **Phonological Features Mediate Object-Label Retrieval and Word Recognition in the Visual World Paradigm** ............  23  
   Nicole Altvater-Mackensen and Nivedita Mani

3  **The Complexity of the Visual Environment Modulates Language-Mediated Eye Gaze** .............................. 39  
   Florian Hintz and Falk Huettig

4  **Visually Situated Language Comprehension in Children and in Adults** ................................................ 57  
   Pia Knoeferle

5  **Vision and Language in Cross-Linguistic Research on Sentence Production** ............................................. 77  
   Elisabeth Norcliffe and Agnieszka E. Konopka

### Part II  Attention and Vision in Reading

6  **Capturing Reading Processes in Connected Texts with Eye Movements and a Letter Detection Task** ................. 99  
   Jean Saint-Aubin and Raymond M. Klein
7 Reading in Thai: Visual and Attentional Processes ............... 111
Heather Winskel

Part III Attention and Vision in Bilingual Language Processing

8 Visual Cues for Language Selection in Bilinguals ................ 129
Robert J. Hartsuiker

9 In the Mind’s Eye: Eye-Tracking and Multi-modal
Integration During Bilingual Spoken-Language Processing .......... 147
Sarah Chabal and Viorica Marian

10 Spoken Word Mediated Interference in a Visual Task:
Eye Tracking Evidence from Bilinguals ......................... 165
Ramesh Kumar Mishra and Niharika Singh

Part IV Language Processing in a Social Context

11 Adjusting the Manner of Language Processing
to the Social Context: Attention Allocation During
Interactions with Non-native Speakers ......................... 185
Shiri Lev-Ari

12 Seeing and Believing: Social Influences on Language Processing .... 197
David W. Vinson, Rick Dale, Maryam Tabatabaie and Nicholas D. Duran
Editors and Contributors

About the Editors

**Ramesh Kumar Mishra** has a Ph.D. from the University of Delhi and is currently associate professor and Head at the Center for Neural and Cognitive Sciences, University of Hyderabad. He earlier taught at the Centre of Behavioural and Cognitive Sciences, University of Allahabad, India. Dr. Mishra has published widely in the area of psycholinguistics and cognitive science of language and has edited books in the area of language and cognition. His current research focuses on how language and other important cognitive processes like attention and vision interact during cognitive processing. Dr. Ramesh Mishra is on the editorial board of journals like Journal of Theoretical and Artificial Intelligence, Frontiers in Cognition. He also co-edits the International Journal of Brain, Culture and Cognition.

**Narayanan Srinivasan** is currently Professor and Head at the Centre of Behavioural and Cognitive Sciences, University of Allahabad, India. Professor Srinivasan was a visiting scientist at the Riken Brain Science Institute (2006–2012). He has a master’s degree in Electrical Engineering from the Indian Institute of Science and subsequently earned his Ph.D. in Psychology from the University of Georgia, USA in 1996. He has been working at the Centre for more than a decade. His primary research interests are attention, perception, emotions, and consciousness. He employs multiple methodologies to study cognitive processes. Professor Srinivasan has edited six books and a special issue and has published more than a hundred journal articles, book chapters, and conference proceedings papers. He was the editor-in-chief of the International Journal of Mind, Brain, and Cognition and the associate editor of the journals, Frontiers in Cognitive Science, Cognitive Processing, Neuroscience of Consciousness, Royal Society Open Science, and Psychological Studies.

**Falk Huettig** is a Senior Investigator at the Max Planck Institute for Psycholinguistics in Nijmegen, Netherlands and a Visiting Professor at the University of Hyderabad, India. He received a B.Sc. and an M.Sc. from the University of Edinburgh and a Ph.D. in Psychology from the University of York, UK. His main
Editors and Contributors

Nicole Altvater-Mackensen  Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Sarah Chabal  Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA

Rick Dale  Cognitive and Information Sciences, University of California, Merced, CA, USA

Nicholas D. Duran  Arizona State University, Tempe, USA

Robert J. Hartsuiker  Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, Ghent, Belgium

Florian Hintz  Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Falk Huettig  Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; Donders Institute for Brain, Cognition, and Behavior, Radboud University, Radboud University, The Netherlands

Raymond M. Klein  Dalhousie University, Halifax, Canada

Pia Knoeferle  Research Group “Language and Cognition”, CITEC, Cognitive Interaction Technology Excellence Center, Bielefeld University, Bielefeld, Germany

Agnieszka E. Konopka  Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Shiri Lev-Ari  Nijmegen, Netherlands

Nivedita Mani  University of Göttingen, Göttingen, Germany

Viorica Marian  Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA

Ramesh Kumar Mishra  Center for Behavioural and Cognitive Sciences, University of Allahabad, Allahabad, India; Center for Neural and Cognitive Sciences, University of Hyderabad, Hyderabad, India

Elisabeth Norcliffe  Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

research interest is in multimodal cognition. Other main interests include the effect of cultural inventions such as reading on general cognition in children, illiterate adults, and individuals with reading impairments; and predictive language processing. Falk Huettig is an editorial board member of the Journal of Memory and Language and editor-in-chief of Brain, Cognition, and Culture.
Editors and Contributors

Lillian Rigoli  Cognitive and Information Sciences, University of California, Merced, CA, USA

Jean Saint-Aubin  Université de Moncton, Moncton, Canada

Niharika Singh  Center for Behavioural and Cognitive Sciences, University of Allahabad, Allahabad, India

Michael J. Spivey  Cognitive and Information Sciences, University of California, Merced, CA, USA

Maryam Tabatabaeian  Cognitive and Information Sciences, University of California, Merced, CA, USA

David W. Vinson  Cognitive and Information Sciences, University of California, Merced, CA, USA

Heather Winskel  Psychology, School of Health and Human Sciences, Southern Cross University, Coffs Harbour, NSW, Australia