Ramesh Kumar Mishra

Interaction Between Attention and Language Systems in Humans

A Cognitive Science Perspective

Springer

Interaction Between Attention and Language Systems in Humans

Ramesh Kumar Mishra

Interaction Between Attention and Language Systems in Humans

A Cognitive Science Perspective

Springer

Ramesh Kumar Mishra
Centre for Neural and Cognitive Sciences
University of Hyderabad
Hyderabad, Andhra Pradesh
India

ISBN 978-81-322-2591-1 ISBN 978-81-322-2592-8 (eBook)
DOI 10.1007/978-81-322-2592-8

Library of Congress Control Number: 2015946096

Springer New Delhi Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer (India) Pvt. Ltd. is part of Springer Science+Business Media (www.springer.com)

To my father

Preface

Most first books originate from experience of teaching. The idea of this book probably originated when I started examining the interaction of language and vision few years ago at the Centre of Behavioural and Cognitive Sciences, University of Allahabad and later at the Centre for Neural and Cognitive Sciences, University of Hyderabad. I used eye tracking to link linguistic processing with visual and attentional processes as my main method. Soon it was clear to me that language and attention interaction was dynamic, way beyond what many Psycholinguistic text-books had taught me. This led to further research and teaching in and around these topics and also very fruitful collaborations that have been instrumental in the conceptualization of this book. As a psycholinguist it was important for me to comprehend fully the enormously important framework of cognitive psychology. Ideas of various chapters came later with my own diversification as a researcher and also with the rapid explosion of research in the last decade. They all had to be summarized in some sense, at least for the beginning student and also for the expert as a point of view.

I have reviewed others' work as far as I was aware of them and I could connect to their relevance in each chapter. I have used figures and illustrations from several sources to enrich comprehension which I assume the interdisciplinary student will find helpful. I did not intend to write a textbook on psycholinguistics in the disguise of a monograph since many good textbooks already exist. My own findings appear sporadically at many places. In no sense this selection of literature is complete or can be complete. This is an exciting and dynamic area of research and every other day new findings are coming in. Every other day one finds papers published in each one of the areas that this book has covered. It is a field that is moving at a very rapid pace at the moment. I have tried to provide an overall sense of achievement and concern to the reader that I have felt myself for themes covered here without often sounding like a text book writer.

Many collaborators and colleagues have read chapters and have commented on drafts. I only hope that I could take advantage of their advice. Most remaining flaws are mine. I particularly wish to thank Raymond Klein, Daniel Burnston, Cristian

Olivers, Andriy Myachykov and Rick Dale. Their comments have helped me to sharpen my thought in a major way. I also thank Seema Prasad, my graduate student for her extensive help in preparation of the chapters. I thank Subha who helped in editing some chapters. And of course a big thanks you to Shinjini, the editor at Springer who has been very encouraging and whose sustained comments helped me achieve clarity in writing.

Most of my own research described in this book has been made possible by several grants from the Department of Science and Technology under their Cognitive Science Initiative and I am very thankful.

I have also benefited enormously through my many visits to the Max Planck Institute for Psycholinguistics, Nijmegen as a visitor. I gratefully acknowledge the friendly support of Falk Huettig and the MPI directors for these visits. These visits and interactions certainly have influenced my thinking about the psychological basis of language. Dalhousie University for which I am grateful to Ray Klein. This led to collaborative research described in this book. Much of my own work described in this book happened at the Center for Behavioral and Cognitive Science, University of Allahabad. The supportive atmosphere at the center and conversations with Narayanan Srinivasan in particular on attention and related topics must have influenced my thinking.

I must also thank the anonymous reviewer of the manuscript whose comments and observations led to a much improved manuscript in many ways.

Thanks to Bidisha, my wife and Ritwika, my daughter, for standing by me, when they were there and not there. Many of the chapters were written and revisions done amid the lakes and greens of Indian Institute of Technology, Guwahati where Bidisha teaches.

Contents

About the Author

Ramesh Kumar Mishra, Ph.D., is an associate professor of cognitive science at the Centre of Neural and Cognitive Sciences, University of Hyderabad. He has previously taught at the Centre for Behavioural and Cognitive Sciences, University of Allahabad. Dr. Mishra has published internationally and has edited/co-edited books in the areas of psycholinguistics, language-vision interaction and attention. His areas of interest include language–attention interaction, visual processing and action control. He has been a scientific visitor to the Max Planck Institute for Psycholinguistics, Dalhousie University and the Institute for Cognitive Science Studies, Tehran, among others. Dr. Mishra is on the editorial board of *PLOS One*, *Frontiers in Cognition*, and *Journal of Experimental and Theoretical Artificial Intelligence*. He is also a co-editor of the cognitive science journal *Brain, Cognition and Culture*. Dr. Mishra is a fellow of the Psychonomic Society, USA.

Chapter 1
Linking Language to Attention

The study of language and attention is on the rise in several disciplines such as cognitive psychology, psycholinguistics, cognitive science, language disorders as well as philosophy of mind. However, there is no coherent framework that helps make sense of the diverse results observed in these fields. While it is no longer fashionable to say language does not require attention, it is also not clear how attention and language influence one another during cognitive processing. It is not an easy thing to write about a field that is in fast motion. Language and attention influence one another during cognitive processing in many different ways. My attempt in this book is to look closely at psycholinguistics and cognitive psychology to see how concepts of attention have been used in the explanation of everyday language processing and how different levels of language affect attentional deployment during cognitive processing.

This book deals with those important cognitive processes that make language processing seemingly effortless. Attention is one of them. In his celebrated *Aspects of Theory of Syntax*, Chomsky (1965) wrote, 'ideal speaker-listener in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of this language in actual performance' (p. 3). This book shows that these conjectures about the linguistic capacity of the ideal native user of language may not be true. Experimental evidence from myriad disciplines such as psycholinguistics, cognitive psychology and cognitive neurosciences in the last several decades has shown that human language performance is constrained by cognitive processes. Effortless actual linguistic performance is only possible if there is harmony among important cognitive processes such as memory, attention and motivation. It is now well accepted that any holistic theoretical understanding of human language capacity must include a thorough investigation of those cognitive processes that make linguistic performance a success. This book explores the locus of language and attention interaction in different psycholinguistic processes.

The issue of how language and attention interact is of enormous importance to anyone who wants to know how the mind works. This is because if language helps express thoughts, attention selects the most relevant ones to work on. Wundt had

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_1

investigated how executive control processes and attention influence sentence processing (see Blumenthal 1975). Similarly, one can find related concerns in the writings of Wittgenstein (see Stern 1995). Further, recent evidence from a range of linguistic and cognitive phenomena suggest that attention mechanisms affect how humans process language (see Kurland (2011) for a review and also Mishra (2014)). Attentional deficits often lead to delay in language development (Ebert and Kohnert 2011). Language structure itself seems to reflect aspects of human attentional mechanisms (Ibbotson et al. 2013). That is, what is attended most often becomes the grammatical subject of the sentence that is expressed (Wallace 1982).

Since attention is an important modulator of cognition, its interface with language is logical. While the exact nature of attentional involvement in different language processes remains specific to that process, there seems to be a broad consensus that attention influences much of linguistic processing in general (Gong and Shuai 2012; see Shtyrov et al. 2012 for a different view). Many assume that 'attention' is omnipresent in any and every cognitive process. Therefore, there is no need to propose a separate theory on the interaction of attention with other cognitive processes. The truth is that, there have been long-held views in psycholinguistics and cognitive sciences that see linguistic processing as unaffected by other processing and as encapsulated modules (Fodor 1983; see also Peretz 2012). This has led to a tacit assumption that some core syntactic processing of language can happen without attentional control and are like reflexes (Pulvermüller et al. 2008). However, in other domains of language processing such as picture naming (Roelofs 2007) or sentence generation (Myachykov et al. 2011) attentional involvement has been observed. It is the aim of this book to examine 'what' type of attentional involvement is more prominent in some linguistic processes and not in others. The study of the interaction between language and attention is also one way to examine critically the encapsulation argument for modules or even the concept of domain specificity. One important aim of this book is to examine these different observations and arrive at some consensus. Either attention and language interact more fully at some levels or they do not.

The book's chapters explicitly discuss different linguistic domains, including the influences of individual differences and culture where attentional mechanisms with regard to language processing have been investigated. This is not a psycholinguistics textbook in the conventional sense, but it is for researchers who work broadly in the fields of psycholinguistics and cognitive psychology as well as in the neuroscience of language, and who particularly look at language and attention in their research. The book is structured in a way so as to offer a comprehensive treatment of topics that have received serious research attention related to attention and language. I have only considered topics where substantial progress has been made both in theory and empirical research on the question of language and attention. Many a time researchers have not directly examined if attention affects the linguistic process but have used attention as an explanation of the data. The book examines how attention–language interface manifests in listening, speaking and reading as well as how individual differences such as bilingualism and literacy level influence such processing.

The book should also be helpful to students who wish to know the state of the art in research in the area of language and attention. Topics covered discuss research coming from multiple methodologies and a diverse range of participants to show how complex attention and language can be. Theoretically, this book motivates an anti-modular but interactionist position on cognitive operations relevant for language and attention. The book also shows that not a single theory or proposal can account for the multiple ways in which attention and language influence one another. Different human linguistic activities like reading, speaking and listening call for different involvement of attention and executive processes (see Mishra 2013 for a review). Different researchers have used different facets or conceptualization of attention, i.e. filtering process versus focusing device, to account for different linguistic processes, as indicated earlier. Others have examined attention and language together invoking concepts like 'automaticity' of cognitive operations. The book examines how these conceptualizations have been used to account for psycholinguistic processes in theory and model building.

Recent theorizations on attention and multi-modal processing look at how different cognitive systems interact dynamically during processing that includes language, vision and attention (Spivey 2007). Such interaction of cognitive processes represents a non-modular cognitive architecture (Dehaene et al. 1998). Both linguistic as well as non-linguistic information influence one another during processing (Mishra 2009). Therefore, the discussion of attention–language interaction has to include the discussion of modularity of cognitive systems as well as several other variables such as individual differences and task effects. Attention has been shown to be a core causal process whose deficits can explain a linguistic deficit in different cognitive disorders (Kurland 2011). Contemporary research shows language, attention, vision, memory and perception influence one another to produce cognition (Spivey 2007; Jackendoff 2012). For example, listening to spoken words can lead to changes in attentional states, driving eye movements towards objects that are only very remotely similar to what the spoken word refers to (Mishra 2013). Therefore, language processing in the visual world depends on attentional mechanism. This book discusses how attentional mechanisms are involved in language processing and how language in turn affects attentional strategies.

A main reason behind writing this book was to explore if there is any pattern and similarity among the different ways in which attention has been implicated in different linguistic processing. These could be attention as a process, as a focusing device, as something that leads to conscious perception and as a mechanism which filters undesirable stimuli. Researchers who have used eye tracking as a methodology explicitly link attention with oculomotor movement, which in turn presumably reflects some aspects of language processing (Henderson and Ferreira 2013; Mishra 2009). Reading researchers have linked attentional shifts during reading with lexical access and oculomotor control (Rayner 1998). Specifically, while some have viewed a serial allocation of attention during reading (Reichle et al. 2003), others have conceived of a parallel mode of deployment of attention on several words at a time (Engbert et al. 2005). Research in language production has linked attention with focusing on the object, which leads to phonological access (Griffin and Bock 2000). Researchers in the domain of sentence processing have looked at ‘automaticity’ with regard to syntactic versus semantic processing (Pulvermüller et al. 2008). In the domain of second language acquisition, attention has been linked with consciousness and motivation to learn (Tomlin and Villa 1994). Finally, infant language researchers have looked at joint attention and its links with early language development (Tomasello and Farrar 1986). Apart from this, attention has been also used as a device in the examination of discourse processes (Brennan 1995). Therefore, attention has been linked to particular tasks and domains of processing. One major rationale of this book is to seek some general explanation which can be used as a framework while discussing the language–attention interface for different linguistic processes.

The multifarious use of the concept ‘attention’ is partly due to the confusion that exists in cognitive psychology on the nature of attention. Currently, there are different theories on the nature of attention. Some have looked at attention as a process which leads to formation of perceivable objects from features (Treisman 1998), while others have looked at attention more as a manner of biasing focus towards one entity compared to another which leads to recognition of targets (Desimone and Duncan 1995). Yet others have further fractionated attention into internal and external formats that sub-serve different cognitive operations (Chun et al. 2011). Attention has also been looked at in terms of space and time as well as modes of deployment, i.e. exogenous and endogenous (Theeuwes 1991). There are also recently offered views that totally negate the very concept of attention (Anderson 2011). Philosophers of mind with interests in the link between consciousness and the role of attention in it have considered attention as a mechanism which transforms a phenomenal awareness into a more objective perception (Block 2007). Very recently, it has been proposed that attention should always be viewed as related to a task, i.e. what attention does for a task (Koralus 2014). These diverse views thus entail different theorizations on the nature of attention and language interaction. For our purpose, I will consider some views and aspects of attention that are more widely accepted and that have been used in the explanation of empirical phenomena in the domain of language processing. The attempt will not be to redefine what attention is, but to examine which view of attention is best used in the explanation of particular linguistic processing.

Substantial progress in disability research shows a clear link between attention and language/cognitive development. For example, dyslexic researchers now view reading deficits arising from a visual and attention incapacity to process information (Pavlidis 1981). Attentional and executive function deficits have been implicated in specific language impairment (Finneran et al. 2009). Auditory attention has been implicated in the comprehension of complex sentences (Montgomery et al. 2009). Aphasia is considered to be a higher language disorder that also seems to be linked with attention deficit (Tseng et al. 1993). This evidence suggests that attention is a critical factor in language development and its deficits could lead to language and cognitive disorders. Nevertheless, it is important to point out that there is no homogeneous conceptualization of attention in the explanation of these various disorders that have diverse symptoms. This book will consider evidence from such disorders in order to see if one can learn something about attention–language interface from them.

1.1 Structure of the Book

First, the book presents some historical background on the development of psycholinguistic research. It also discusses disciplinary differences that hindered the development of research in this area. After the introductory chapters, each successive chapter deals with one important psycholinguistic process and examines the role of attention in it. This organization was thought rational since I have already indicated that researchers studying different linguistic functions have used the notion of attention differently. Below, I provide some foundational points that are useful to look at before getting into the debate of attention–language interaction. These themes have guided me in examining issues discussed in each chapter. After that I have given details of the chapters to be presented in the book. Since there is no consensual theory as of now on the nature of language and attention interaction, I attempt to offer a view which integrates major findings across domains in the last chapter. I also mention below a few important themes that have helped in opening up research avenues for an understanding of attention–language interaction in the recent past. Some of them are discussed throughout the book.

1.1.1 Non-modular Cognition

A fundamental concern among cognitive scientists on the question of mind has been its modules and how they may interact. How are different types of cognitive functions accomplished? Seeing and speaking are different functions. But, is there any similarity between them from a cognitive perspective? Fodor's theory of a modular mind has been the most influential statement of classical symbolic cognitive science (Fodor 1983). Fodor suggested that modules have evolutionary advantages since each module can process information which it is privy to without being vulnerable to outside information. Furthermore, no top-down influences affect language processing including cognitive states and belief structures. However, much of recent evidence suggests that visual information affects linguistic processing and language input constrains what is visually perceived (Henderson and Ferreira 2013; see also Mishra 2009 for a review). Fodor stressed the informational encapsulation since it enhances the operation of a system. He said:

> Imagine a computational system with a proprietary… database. Imagine that this device operates to map its characteristic inputs onto its characteristic outputs… and that, in the course of doing so, its informational resources are restricted to what its proprietary database

> contains. That is, the system is "encapsulated" with respect to information that is not in its database…. That's what I mean by a module. In my view, it's informational encapsulation, however achieved, that's at the heart of modularity. (Fodor 1983, p. 63)

This book will show that much of this is not true as far as language and attention interaction is concerned.

Cognitive encapsulation and cognitive penetration can mean different things. As for Fodor, perceptual modules are cognitively impenetrable. Many have argued against it, citing evidence from neurophysiological and psychophysical tasks showing that perception is penetrable. Burnston and Cohen (2014) suggest that this distinction is not an all-or-none scenario. There are perceptual mechanisms which are cognitively penetrable and there are others whose behaviour conforms to Fodor's view. Burnston and Cohen (2014) thus attempt to suggest that it is important to look at each individual process and examine what class of outside information can affect it to understand how encapsulation and modularity are working. It is probably possible to keep both vital points of classic modularity and still talk of information integration for certain cases. Importantly, Burnston and Cohen (2014) do not cite data that deals with interaction of language and vision.

Fodor thought of language and other systems like vision as modular input systems. These systems did not need one another or influenced one another. Therefore, there was no issue of linking language behaviour to other cognitive structures. In contrast to this, there are frameworks inside linguistics that explicitly propose the interaction of language and vision (Jackendoff 1997). The modularity thesis remained a very restrictive ideology for many creative psycholinguists for a very long time. Today we know that nothing in the mind is modular, at least the main systems (Bellugi et al. 2014). It is difficult to appreciate the findings from attention–language interaction with a strictly modular organization of mental processes (Anderson et al. 2011). The non-modular approach to cognition has been a necessary and conducive framework for understanding this interaction (Robbins 2013). In several chapters I show that attention and language as well as vision interact in a non-modular fashion. Importantly, there seems to be not much of 'informational encapsulation' since linguistic and visual representations affect one another dynamically (Mishra and Marmolejo-Ramos 2010). While the issue of modularity in cognition is much broader and is linked to evolution of specialized cognitive modules, this book considers instances where linguistic and non-linguistic information interact.

1.1.2 Disciplinary Boundaries

Cognitive psychologists who developed theories of attention did explicit analysis of visual perception, particularly object recognition (Biederman 1995), while many others have conceived attention as a spotlight which illuminates certain aspects of the world for intensive processing (Eriksen and James 1986). Models of visual object recognition have dealt with identifying simple objects under different task constraints (Treisman 1969) and the idea has dominated thoughts about the nature of selective attention for a long time. This has led to the perception that such models cannot explain complex linguistic stimuli such as sentences. Not much of these experimental approaches dealt with linguistic stimuli. Psycholinguists did not develop models of human language processing keeping attentional and perceptual factors into account (see Chap. 3 for details). These disciplinary differences in methods and use of stimuli have led to constraints in the fruitful exploration of the interaction of language, attention and vision. This book argues for a position where true interdisciplinary interaction between cognitive psychology and psycholinguistics can lead to more holistic theories of human cognition. I draw evidence from both cognitive/experimental psychology and psycholinguistics and allied areas to illustrate the points. Throughout the book, cognitive psychological concepts related to attention will be used for explaining linguistics behaviour as they have been done for simple visual forms. The book aims to show that attentional mechanisms as studied and used by cognitive psychologists can be fruitfully used by language scientists in their own work.

1.1.3 Sensorimotor and Embodied Theories

There has been an impressive rise in embodied and sensorimotor theories of cognition in recent times (Wilson 2002). Some have even referred to these approaches as an alternative to classical cognitive science that focused on a modal symbol processing and representation (Shapiro 2011). This new wave in cognitive science has certainly affected the way the attention–language interaction has been viewed (Spivey 2007). The notion of embodiment in cognitive processing emphasizes the importance of the body and the environment as well as sensorimotor systems in cognition (Shapiro 2011). As a sub-theory of the larger enterprise of embodism in cognitive science, there has been rapid growth in linking mental and sensorimotor simulation to language processing (Zwaan 2014; also see Rommers et al. 2013). While classical cognitive science looks at cognition arising from algorithmic computations in an amodal manner, embodied cognition considers the mind, body and the environment together to account for cognition of the agent. Researchers who work within an embodied cognition framework believe that cognition is not just amodal symbolic processing or computation but is a manifestation of the interaction between the organism and its environment (Gibbs 2006). For example, several findings suggest that humans comprehend language using mental simulation described by language (Zwaan 1999; also see Caramazza et al. 2014 for an alternative account). Embodied cognitive science programme also proposes that critical cognitive operations like linguistic or visual cognition use re-enactment of sensory and motor representations (similarly, attention is being perceived as a real-world phenomenon, linked to our daily awareness and cognition: see Posner 2012). Sensor–motor activations related to language processing show interconnected neural networks. Similar to the embodied cognition approach, cognitive linguists have also started to look at the attentional and perceptual bases of language. These alternative approaches to the study of language and cognition have offered suitable avenues to discuss how attention and language can interact.

1.1.4 Methodological Developments

No science progresses without corresponding methodological developments. Early studies of psycholinguistics used the reaction time method as a measurement for mental processing. Currently, multiple methods are used to observe brain dynamics as well as behavioural action in the study of both linguistic and non-linguistic processing. Online methods such as eye tracking and recording of Event Related Potentials (ERPs) have led to a good deal of understanding of temporal scales at which information is processed and integrated over time (Jegerski and VanPatten 2013). Psycholinguists and cognitive psychologists use similar methods today to study processes like attention and language. For example, eye movement measures are used to examine attention mechanisms during visual search (Olivers et al. 2014) as well as linguistic processes like reading (Rayner 2009) and listening (Singh and Mishra 2013). Therefore, it has become possible to see similarities in mechanisms and apply paradigms across disciplines using similar dependent variables. This has led to much fruitful collaboration and mutual use of similar jargon in the description of experimental facts. For example, the term 'interference' during processing is equally accepted and understood similarly by language researchers and psychologists as well as concepts like 'top-down' or 'bottom-up' influences. In eye movement paradigm, dependent variables like 'fixation duration' and 'saccade latency' are used commonly to refer to processing constraints to both linguistic and non-linguistic stimuli (Mishra 2009). From neurobiological investigations, we know that processing language also leads to the activation of several motor, attentional and visual areas (Pulvermüller 2005). These developments have made it possible to study language–attention interaction in humans with precision and talk about the interaction between language and attention using similar terms.

1.1.5 Rise of Alternative Theories in Linguistics

As an alternative to the Chomskyan model of language, several other alternative frameworks like cognitive linguistics (Ungerer and Schmid 2013) have made it possible to directly examine the language–attention interaction. Cognitive linguistics studies language with reference to different cognitive systems such as perception, attention and vision. Cognitive linguists propose that using language includes activating the many sensory experiences associated with the language. Therefore, symbolic processing becomes meaningful and serves communication since language users situate language into their experience and sensations. While I do not cover extensively this view, it is used here to suggest that these frameworks have led to a rapid rise in empirical research that explicitly links attention and language including sensorimotor cognition (Jackendoff 1996). Therefore, by definition it has to include attention, vision and perception into the explanation of linguistic behaviour. Notably, Talmy has written exclusively on the attentional attributes of language (Talmy 1988; see also Robinson and Ellis 2008). According to him, language itself has several components like the deictic system and figure–ground mechanisms to channelize attention (Talmy 2000). This type of work shows how language arises as an interaction between basic cognition and symbolic structures. Particularly, experimental work on linguistic phenomena such as fictive motion has shown how comprehension of linguistic structures could influence visual attention (Mishra 2010).

Different chapters of this book explore how attention and language interact in different core linguistic processes such as speaking, listening and reading. This resembles how most psycholinguistic phenomena are described in terms of modality of use or as input–output systems. I do not discuss attentional involvement in writing, since not much on this has been done in psycholinguistics itself. The chapters aim to present the most up-to-date research to show that attention and language functions interact in many situations. They also present data that challenge this assumption. I have examined data that have directly looked at attention and language interaction and where most likely attention has been manipulated as a variable; several research findings have dealt exclusively with language–attention interaction. Since the focus is on attention and language, the chapters do not contain detailed discussions of individual models of processing as they can easily be found in any good psycholinguistic textbook. Therefore, this book assumes a basic knowledge of psycholinguistics or cognitive psychology so that the reader can appreciate and critically reflect on the arguments.

Chapter 2 offers an exclusive overview of attention as a cognitive psychological concept. It reviews recent and important theoretical developments that should be useful in the following specific discussions in later chapters. I have not covered all the major theories but only those that have received consensus and are widely used in basic and applied research in cognitive science and psycholinguistics. This chapter also discusses various important facets of attention and its links with eye movements and consciousness. The chapter should offer enough jargon to the reader to follow discussions in later chapters. The chapter is written keeping the non-psychologist in mind. At many places in this chapter there is more emphasis on the eye tracking methodology as well as how eye movements are related to attention because many experimental facts that are discussed later deal with these elements. The chapter is pedagogical since I think it is critical to first know how attention has been viewed by psychologists.

Chapter 3 is devoted to the issue of attention–language interaction. The chapter provides a historical background to the interaction between cognitive psychology and psycholinguistics. This tells the reader how the two fields have matured in recent decades and their interactions. There is also discussion of alternative approaches like cognitive linguistics and emergentism, where researchers have examined attentional involvement in language processing. The chapter aims at establishing the broad background against which the interaction of attention and language must be looked at. These two chapters could be considered introductory in the sense that they initiate the reader into the main points.

Chapter 4 deals with attention in sentence processing. Here, attentional issues during spoken sentences are discussed (since there is another chapter on reading later). The chapter explores how attention has been explored in spoken word and spoken sentence processing, including the processing of speech. Several auditory sentence processing research has indicated that syntax processing may not need attention (Maidhof and Koelsch 2011), whereas semantics may. Similarly, research with ERPs on speech samples shows that the human brain automatically traces speech segments (Pakarinen et al. 2013), even when attention is deployed elsewhere. As per methodologies, I have liberally used evidence from both ERPs and fMRI where it was needed. The chapter also includes discussions on the effect of cognitive load on sentence processing. Data from neurological patients have been discussed that show motor disturbances cause problems in sentence processing.

Chapter 5 addresses the issue of attentional involvement during speaking. The chapter discusses studies with the picture-word interference paradigm and issues related to serial versus parallel access of information during naming. The chapter looks at the evidence which suggests that naming pictures require attention at least the ones with pictures. Attention in this scenario leads to activation of phonological and conceptual information. Recent work using eye tracking during speaking has revealed how intricately attention is linked to lexical access (Meyer and Lethaus 2004). The chapter also discusses how attentional cues can affect structural choice during sentence generation.

Chapter 6 examines how different patterns of language use can lead to changes in one's attentional capabilities. Particularly, research on bilinguals and illiterates has been examined that show the influence of individual cognitive difference on attention. Bilinguals with their extensive practice in the handling of two languages show enhancement in several executive functions (Singh and Mishra 2013). Bilinguals with higher fluency have been shown to have better conflict resolution ability as well as monitoring. Interestingly, this is seen with non-linguistic tasks in both manual and oculomotor domains. As for literacy, it is seen that subjects with low level of literacy show poor coordination of visual target chasing (Olivers et al. 2014). These individuals also show poor anticipatory eye movements during cross-modal language processing (Mishra et al. 2012). These studies show how a particular style of language use can have influence on the attentional and control mechanisms. The chapter also includes discussions on the role of attention during learning a second language. Much extensive theorization in this area has shown that focused attention to linguistic details enhances learning a second language. This chapter shows how individual differences affect language–attention interaction in many cases.

Chapter 7 shows how vision and language interact dynamically during cognitive processing. Several studies have shown that language input can dynamically influence allocation of visual attention to objects (see Huettig et al. 2011 for review). The role of attention in this interaction has been variously studied including populations with language disorder. The chapter discusses findings from the visual world paradigm that show attentional shifts on visual scenes with spoken language input (Huettig et al. 2011). Much recent studies of psycholinguistics and cognitive psychology have found that listeners immediately orient their visual attention towards objects that the spoken language mentions (Mishra 2009). Visual world studies thus show attentional capture and reorienting with speech input, which happens within 100 ms. These data using the visual world eye tracking paradigm have shown probably the most robust evidence so far of multi-modality of cognition.

Chapter 8 is on reading and attention. The chapter discusses different models of reading that show a link between oculomotor programming and lexical access. Many fluent readers think that reading is an automatic process and therefore does not need attention. The chapter discusses two influential models of reading to show how shifts of attention are linked to lexical processing. Readers shift attention to nearby words much before they actually shift their eyes (Rayner 1998). This covert shift of attention helps in acquiring semantic information from the word not yet looked at. The chapter also discusses attentional failure in dyslexia. Many recent studies have shown that it is primarily a visual and attentional deficit which is at the core of dyslexia (Farmer and Klein 1995). The chapter shows how different models of reading have used different aspects of attentional mechanism to account for oculomotor control.

Chapter 9 looks at attention and language interaction in the spatial domain. Particularly, the evidence from cross-linguistic research where language is shown to modulate attention has been examined. Much of this data, collected from speakers of different languages, show that language structure affects our spatial perception (Majid et al. 2004). It has been observed that speakers of typologically distinct languages use different spatial words to describe space. These words, i.e. preposition and deictic terms, help us locate things in space (Levinson 2003; Hickmann and Robert 2006). The chapter therefore looks at attention–language interaction from a cultural point of view including sensory–motor cognition.

Chapter 10 sums up what has been said in the previous chapters and offers a theoretical framework. An attempt has been made to organize different forms of attention–language interactions observed under broad frameworks like voluntary and involuntary forms of attention as well as endogenous and exogenous modes of attentional channelization. The chapter therefore is a summary of things described before with a look towards the future.

1.2 Thesis of This Book

In this section I make some observations that have both historical and contemporary importance and are crucial in making sense of much of the experimental data examined today in the domain of attention and language. The book argues that the symbolic structure of language functions on a support system of attention and other cognitive processes. It also argues that language reflects some of the evolutionary cognitive constraints of the human mind and therefore mirrors cognitive states. Language is a symbolic system which helps humans communicate thoughts. Philosophically speaking, language reflects the intentional nature of thoughts embedded within its symbolic structures. However, most fundamentally, language refers to objects and events in the world. For this to happen, it uses words and puts them into some kind of syntactic structures. Much has been written about the mathematical nature of human language syntax in the past five decades. Linguists have always been in awe of the abstract symbolic structure of language and the similarities that are seen in the world's languages. Linguistic theorization has led to greater understanding of how sounds are related to concepts. The psychological or physiological bases of this symbolic behaviour were not of much appeal to linguists of different persuasions, till we came to Lenneberg (1966). His concept of 'critical period' had the undercurrents of an interactive position, although Lenneberg did not directly remark anything on attention or memory and their roles in language. Before him, Edward Sapir, in his classic monograph titled *Language*, wrote:

> … accordingly it must be clearly understood that this introduction to the study of speech is not concerned with those aspects of physiology and physiological psychology that underlie speech. Our study of language is not to be one of the genesis and operation of a concrete mechanism; it is, rather, to be an inquiry into the function and form of the arbitrary systems of symbolism that we term languages. (Sapir 1921, p. 7)

In contrast to this, language today is conceived in terms of its biological, computational and cognitive attributes.

Long back, Wilhelm Wundt considered attention as that central mechanism which made us conscious and also helped in the execution of several cognitive goals (see Chap. 3). Wundt did not conduct many experiments on this topic. However, his introspectionist psychology did have a pervasive influence on later thinkers. Many contemporary psycholinguists are trying to bring back Wundtian insights related to the study of attention and language (see Piai et al. 2011). Wittgenstein's prophetic aphorisms in his *Tractatus Logico-Philosophicus* (1922) contain several direct and indirect references to attention. He was quite categorical in saying that referring to an object with a name involves blending of the sounds with the concept. Now we know how pointing and gaze patterns reveal the innermost functioning of language and such dynamic blending calls for attention in many cases (see Chap. 7). These writers were more psychologically inclined and thought of language functions with regard to the interactions it has with other mental attributes. The Russian psychologist Vygotsky in his approach to the study of cognition also was vocal about the relationship between language and thought, including executive functions (Vygotsky 1979). These insights have helped in generating workable hypotheses on the psychological basis of language functioning. In recent years, there have been renewed efforts to link attention and language with experimental demonstrations in different tasks (Tomlin 1997; Huettig and Altmann 2007).

Much of early psycholinguistics that was devoted to finding the psychological realities of language processing operated in insulation. Grammar was about the intuitions of the speaker and not performance (Chomsky 1972). For the generative grammarians, no additional cognitive influence was necessary for the operations of grammar and construction of a theory of language (see Chap. 3). Researchers who developed theories of speech perception, word naming or sentence processing did not consider attention or other cognitive mechanisms seriously. The theoretical climate was not conducive for undertaking interactionist research immediately after the cognitive revolution. Psychology was slowly coming out of the behavioural and materialistic grip. At the same time, linguists were under the hypnotic spell of the Chomskyan way of studying language. It is important to note here that when Leonard Bloomfield, one of the founding fathers of American linguistics, had returned from Wundt's lab at Leipzig and wrote his first book on language, he did discuss other cognitive systems that could influence language (Bloomfield 1983). However, later Bloomfield became much of a behaviourist referring little to the mentalist notions of language. This book is not about historical fallacies on this research but is a statement on its current form.

The term 'attention' is often referred to from a perspective of its 'limited resource' attribute. Attention is mainly understood as a limited capacity cognitive system. Attention modulates goal-relevant stimuli to get processed with higher efficiency. Attention also plays a critical role in executive control processes that in turn affect language processing (Ye and Zhou 2009; see also Mishra 2014). As far as language processing is concerned, many think that it is an unconscious activity. We are not often aware of the sentences we produce, how we produce them or even how effortlessly we understand speech. Therefore, the concept of selection or goal-directed action is not easily understood in the context of such language performances. Language comes naturally to us as we wake up, like our consciousness. Most who have read Chomsky carefully will know that nativists believe that language is not taught, since its acquisition is a spontaneous process triggered by a genetic mechanism. The child acquires a very sophisticated symbolic system without having undergone any training. The same does not happen with learning painting or singing. In this sense, language acquisition may not need 'attention'. Many think that activities that require voluntary control also require attention (Lau et al. 2004). Language behaviour is so spontaneous that we have trouble thinking of it as a controlled process. This view has often led to the feeling that one cannot learn much about language by looking at attention or memory. Many chapters in this book will demonstrate that attention is an indispensable cognitive mechanism involved in language.

How does one conceptualize the interaction of attention and language as seen in different cognitive processes? Both attention and language have their own complex histories and definitional issues. There is no single agreed upon definition of language and it is the same for attention. However, a systematic study of language and attention interaction can be undertaken with the following assumptions: One can start by postulating that language processing is autonomous and hence it does not call for attentional resources. These researchers can undertake experiments that show that without attention language still functions. Many such experiments are discussed in the chapter on auditory sentence processing (Chap. 4). Alternatively, one can assign a modulatory role to attention for language processing. That is to say, depending upon the attentional resources available language processing will be affected. Several studies of reading (Franceschini et al. 2012) and picture naming (Piai et al. 2011) have shown this. In this book, it is argued that the interaction of language and attention is bidirectional—attention and language affect one another depending on the task and the situation. Further, one view of attention is not applied indiscriminately for different language functions. Attentional resources are called upon depending on the modality and the task demands. Evidence suggests that individual differences affect such interaction (Olivers et al. 2014). It has also been argued that the structure of language itself indicates attentional involvement (Talmy 2000).

The position that language processing must be automatic and therefore any talk of attention is undesirable is non-sustainable in the face of current data. This book accepts the notion that language is first and foremost a product of human cognition in its broadest sense. Attention is one of the most important driving forces of cognition. Therefore, attention could modulate different aspects of language processing. The word 'modulate' is used with caution, since it can mean many things. A careful look at recent work on psycholinguistics and attention literature suggests that attention influences language processing to different degrees. However, there may be some internal modules of language that operate more automatically, i.e. early phrase structure processing (Hahne and Friederici 1999). As per the different domains of language use concerned, one sees differential effects of attention. For example, attentional requirements for reading and speaking are not similar. Similarly, other linguistic processes involve different aspects of attention. Attentional involvement with language processing is constrained by levels and task demands.

For example, during reading, attention shifts and eye movements are linked to lexical access (Rayner 1998), whereas during listening, spoken words cause attentional shifts in space rather involuntarily (Mishra et al. 2012). Furthermore, attentional requirement during processing of complex and simple sentences may differ (Prat and Just 2011). Importantly, differential use of language can lead to strengthening of attention and executive control skills as in bilinguals (Singh and Mishra 2012). Therefore, it is likely that attention and language interaction should be based on the nature of task as well as the particular cognitive profile of the individual. Individual differences certainly have a major role to play in attention and language interaction. Linguistic entities can cause changes in attentional states and attentional changes lead to differential processing of language. Further, a multi-modal and interactionist framework is assumed while referring to attention–language interaction. Since language is always studied in the form of different levels—phonology, semantics, phonology and pragmatics—it is important to examine if all these levels or only a few interact with attention. Research on auditory sentence processing has shown syntactic structure building to be automatic and not semantics (see Steinhauer and Drury 2012 for recent overview).

Does attention work differently on linguistic and non-linguistic objects? By linguistic one means words and by non-linguistic, let us consider pictures. When someone is looking at a picture and trying to generate its name, attention is focused on the picture, which helps activate the concept and it leads to its name (Levelt 1998). Similarly, when asked to identify a blue ball among balls of other colours, attention must be allocated to this object. Therefore, it is difficult to explain how attention works for linguistic and non-linguistic searches (Huettig et al. 2011). Many themes covered in this book will show that possibly there is a single attentional mechanism in the brain, which acts similarly for the difference between linguistic and non-linguistic stimuli (Mesulam 1990). Basic functions of attention like executive control, selection and goal-directed action, all work during linguistic and non-linguistic processing (It would therefore be odd to think that one can find a different attentional mechanism which is exclusive to certain language functions.). There is evidence which suggests that a domain-general working memory system sub-serves both sentence processing and non-linguistic processing (Fedorenko et al. 2006). Attention and working memory are now thought to involve the same mechanisms. Keeping something in working memory also means attending to the information (Engle 2002).

Research on language and attention is a fast-moving field and many old concepts and worldviews are likely to be soon modified with new data. This book should be considered in such spirit. While experts may find ideas and arguments mentioned here routine or controversial, the newcomer could see them as pointers towards the future work. Can one fruitfully apply theories and principles of cognitive psychology to linguistic stimuli? My answer is yes, only if one considers a cross-disciplinary approach and is not biased by discipline–internal definitional constraints. This should be the scenario in the long run if one wants to understand how vision and attention interact with language. Talking of disciplinary differences, Carl Hempel once said:

> The division of science into different areas rests exclusively on differences in research procedures and direction of interest; one must not regard it as a matter of principle. On the contrary, all the branches of science are in principle of one and the same nature; they are branches of the unitary science, physics. (Hempel 2000, p. 178)

What Hemple says is interesting since probably both cognitive psychologists and psycholinguists are merely dealing with stimuli and procedures of different types. At the end, they are dealing with the mind. Therefore, any theoretical development which is constrained by a stimulus or methodology will not be holistic. This book shows that what we know about attention from cognitive psychology must be immediately relevant for language, to most extents. If theories of attention are only limited to simple visual search, then they are no good to psycholinguists.

References

Anderson, B. (2011). There is no Such Thing as Attention. *Frontiers in Psychology, 2*, 246. doi:10.3389/fpsyg.2011.00246.

Anderson, S. E., Chiu, E., Huette, S., & Spivey, M. J. (2011). On the temporal dynamics of language-mediated vision and vision-mediated language. *Acta Psychologica, 137*(2), 181–189.

Bellugi, U., Blumstein, S., Fodor, J., Mann, V., & Samuel, A. (2014, January). Panel Discussion: The Modularity of Speech and Language. In *Modularity and the Motor Theory of Speech Perception: Proceedings of A Conference To Honor Alvin M. Liberman* (p. 359). Hillsdale, New Jersey. Psychology Press.

Biederman, I. (1995). *Visual object recognition* (Vol. 2). Cambridge: MIT press.

Block, N. (2007). Overflow, access, and attention. *Behavioral and Brain Sciences, 30*(5–6), 530–548.

Bloomfield, L. (1983). *An introduction to the study of language: New edition* (Vol. 3). New York: John Benjamins Publishing.

Blumenthal, A. L. (1975). A reappraisal of Wilhelm Wundt. *American Psychologist, 30*(11), 1081.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive processes, 10*(2), 137–167.

Burnston, D., & Cohen, J. (2014). Perceptual integration, modularity, and cognitive penetration. *Cognitive Influences on Perception: Implications for Philosophy of Mind, Epistemology, and Philosophy of Action. Oxford University Press, Oxford. forthcoming*.

Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror neurons: A critical assessment. *Annual Review of Neuroscience, 37*, 1–15. doi:10.1146/annurev-neuro-071013-013950.

Chomsky, N. (1972). *Language and mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. (1965). Aspects of the theory of syntax. In *Multilingual Matters*. Cambridge: MIT Press.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101.

Dehaene S., Kerszberg M., Changeux J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. In *Proceedings of the National Academy of Sciences*, U.S.A. 95, 14529–14534. doi:10.1073/pnas.95.24.14529.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Ebert, K. D., & Kohnert, K. (2011). Sustained attention in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 54*(5), 1372–1384.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Eriksen, C. W., & James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics, 40*(4), 225–240.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review, 2*(4), 460–493.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language, 54*(4), 541–553.

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 52*(4), 915–929.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge: The MIT Press.

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology, 22*(9), 814–819.

Gibbs, R. W. (2006). *Embodiment and cognitive science*. New York: Cambridge University Press.

Gong, T., & Shuai, L. (2012). Modelling the coevolution of joint attention and language. In *Proceedings of the Royal Society B: Biological Sciences*, rspb20121431.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11* (4), 274–279.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11* (2), 194–205.

Hempel, C. G. (2000). *Selected philosophical essays*. Cambridge: Cambridge University Press.

Henderson, J., & Ferreira, F. (2013). *Interface of language*. Vision and Action: Psychology Press.

Hickmann, M., & Robert, S. (Eds.). (2006). *Space in languages: Linguistic systems and cognitive categories* (Vol. 66). Amsterdam: John Benjamins Publishing.

Huettig, F., & Altmann, G. T. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*(8), 985–1018.

Huettig, F., Mishra, R. K., & Olivers, C. N. (2011). Mechanisms and representations of language-mediated visual attention. *Frontiers in psychology*, *2*.

Huettig, F., Rommers, J., & Meyer, A. S. (2011b). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151–171.

Ibbotson, P., Lieven, E. V., & Tomasello, M. (2013). The attention-grammar interface: eye-gaze cues structural choice in children and adults. *Cognitive Linguistics, 24*(3), 457–481.

Jackendoff, R. (1996). Conceptual semantics and cognitive linguistics. *Cognitive linguistics, 7*, 93–129.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge: MIT Press.

Jackendoff, R. (2012). Language as a source of evidence for theories of spatial representation. *Perception-London, 41*(9), 1128.

Jegerski, J., & VanPatten, B. (2013). *Research methods in second language psycholinguistics*. New York: Routledge.

Koralus, P. (2014). The erotetic theory of attention: Questions, focus, and distraction. *Mind and Language, 29*(1), 26–50.

Kurland, J. (2011). The role that attention plays in language processing. *SIG 2 Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders 21*(2), 47–54.

Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science, 303*(5661), 1208–1210.

Lenneberg, E. H. (1966) The natural history of language. In F. Smith & G. A. Miller (Eds.), *The genesis of language*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1998). The genetic perspective in psycholinguistics. Or where do spoken words come from? *Journal of Psycholinguistic Research, 27*, 167–180.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity* (Vol. 5). Cambridge: Cambridge University Press.

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience, 23*(9), 2252–2267.

Majid, A., Bowerman, M., Kita, S., Haun, D., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in cognitive sciences, 8*(3), 108–114.

Mesulam, M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of neurology, 28*(5), 597–613.

Meyer, A. S., & Lethaus, F. (2004). The use of eye tracking in studies of sentence generation. *The interface of language, vision, and action: Eye movements and the visual world* (pp. 191–211).

Mishra, R. K. (2009). Interaction of language and visual attention: evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K. (2010). Effect of intentional sentences on visual attention to human and animal pictures: evidence from eye movements. *Psychological Studies, 55*(3), 219–229.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing, 11*(4), 295–305.

Mishra, R. K. (2013). Attention in sentence processing: A review. *Psychological Studies, 58*(3), 308–317.

Mishra, R. K. (2014). Listening through the native tongue: A review essay on Cutler's Native listening: Language experience and the recognition of spoken words. *Philosophical Psychology*, (ahead-of-print), (pp. 1–15).

Montgomery, J. W., Evans, J. L., & Gillam, R. B. (2009). Relation of auditory attention and complex sentence comprehension in children with specific language impairment: A preliminary study. *Applied Psycholinguistics, 30*(01), 123–151.

Myachykov, A., Thompson, D., Scheepers, C., & Garrod, S. (2011). Visual attention and structural choice in sentence production across languages. *Language and Linguistics Compass, 5*(2), 95–107.

Olivers, C. N. L., Huettig, F., Singh, J. P., & Mishra, R. K. (2014). The influence of literacy on visual search. *Visual Cognition, 22*(1), 74–101.

Pakarinen, S., Teinonen, T., Shestakova, A., Kwon, M. S., Kujala, T., Hämäläinen, H., et al. (2013). Fast parametric evaluation of central speech-sound processing with mismatch negativity (MMN). *International Journal of Psychophysiology, 87*(1), 103–110.

Pavlidis, G. T. (1981). Do eye movements hold the key to dyslexia? *Neuropsychologia, 19*(1), 57–64.

Peretz, I. (2012). Music, language, and modularity in action. *Language and music as cognitive systems*, (pp. 254–268).

Piai, V., Roelofs, A., & Schriefers, H. (2011). Semantic interference in immediate and delayed naming and reading: Attention and task decisions. *Journal of Memory and Language, 64*(4), 404–423.

Posner, M. I. (2012). Attentional Networks and Consciousness. *Frontiers in Psychology, 3*, 64. doi:10.3389/fpsyg.2012.00064.

Prat, C. S., & Just, M. A. (2011). Exploring the neural dynamics underpinning individual differences in sentence comprehension. *Cerebral Cortex, 21*(8), 1747–1760.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*(7), 576–582.

Pulvermüller, F., Shtyrov, Y., Hasting, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language, 104*(3), 244–253.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(04), 445–476.

Robbins, P. (2013). Modularity and mental architecture. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(6), 641–649.

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological science*, 0956797613490746.

Robinson, P., & Ellis, N. C. (Eds.). (2008). *Handbook of cognitive linguistics and second language acquisition*. London: Routledge.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language, 57*(2), 232–251.

Sapir, E. (1921/1970). *Language*. London: Harcourt, Brace and World.

Shapiro, L. A. (2011). *Embodied cognition*. London: Routledge.

Shtyrov, Y., Smith, M. L., Horner, A. J., Henson, R., Nathan, P. J., Bullmore, E. T., & Pulvermüller, F. (2012). Attention to language: Novel MEG paradigm for registering involuntary language processing in the brain. *Neuropsychologia, 50*(11), 2605–2616.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(04), 771–781.

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: Evidence from Hindi-English bilinguals. *Frontiers in psychology*, *4*.

Spivey, M. (2007). *The continuity of mind*. Oxford University Press.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language, 120*(2), 135–162.

Stern, D. G. (1995). *Wittgenstein on mind and language*. Oxford: Oxford University Press.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive science, 12*(1), 49–100.

Talmy, L. (2000). *Towards a cognitive semantics, volume 1: Concept structuring systems*. Cambridge, MA: MIT Press

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development*, 1454–1463.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception and Psychophysics, 49*(1), 83–90.

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. *Language and conceptualization*, 162–189.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in second language acquisition, 16*(02), 183–203.

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 353*(1373), 1295–1306.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76*(3), 282.

Tseng, C. H., McNeil, M. R., & Milenkovic, P. (1993). An investigation of attention allocation deficits in aphasia. *Brain and Language, 45*(2), 276–296.

Ungerer, F., & Schmid, H. J. (2013). *An introduction to cognitive linguistics*. Routledge.

Vygotsky, L. S. (1979). The development of higher forms of attention in childhood. *Journal of Russian and East European Psychology, 18*(1), 67–115.

Wallace, S. (1982). Figure and ground: The interrelationships of linguistic categories. *Tense-aspect: Between semantics and pragmatics*, *1*, 201–23.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636.

Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience and Biobehavioral Reviews, 33*(8), 1168–1177.

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models.

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in cognitive sciences*.

Chapter 2
The Many Shades of Attention

2.1 Movement of Attention

It is probably not any more an interesting idea to ask the question, what is attention? William James emphasized the fact that all of us have some intuitive notion about this baffling phenomenon but none can define it. This is because attention does so many things and it is not a single process that can easily be tracked down. As indicated in Chap. 1, attention has been conceived as a binding glue which aids in object perception (Treisman and Gelade 1980), as a biasing mechanism which helps in selection of the task-relevant target (Desimone and Duncan 1995), as related to the action system (Norman and Shallice 1986), as a contributing agent in consciousness (De Brigard and Prinz 2010), etc. The attention system in the brain has been viewed as a network comprising several distinct functions such as alerting, orienting and executive control (Fan et al. 2005). Deployment of attention has also been viewed as goal directed or stimulus driven (Corbetta and Shulman 2002). This distinction refers to top-down guidance of attention specific to cognitive goals of the agent, or bottom-up capture of attention by external stimuli. This basic distinction of attentional control has been examined using the endogenous and exogenous cuing paradigm (Theeuwes 1991). This different conceptualization of attention relates to how attention has been implicated in different tasks and cognitive processes. Another crucial aspect of attention is its causal links with consciousness and its ability to bind diverse perceptions together (Cohen et al. 2012). Today, theories from object recognition to consciousness and social cognition, all use attention as a mechanism to explain a wide range of data. There is hardly any cognitive process which has not been shown to be influenced by some attentional manipulation.

One of the most striking discoveries about attention is that it can orient in space, without the movement of the eyes. Posner (1980) first demonstrated that attention can be summoned to a location using a peripheral cue. Similarly, one can direct

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_2

attention towards a location or an object voluntarily, generally following a symbolic arrow cue. These peripheral cues were called exogenous cues and the central arrows, endogenous. The cuing paradigms have revealed that humans can deploy attention wilfully towards certain aspects of the stimuli to achieve immediate goals and this is called endogenous attention. Mostly, things in the world capture our attention and we orient towards them involuntarily, called exogenous capture. Posner made these distinctions using simple cuing tasks and these have remained influential in thinking about attention (Posner and Peterson 1990; Posner 1980). Posner thought in terms of the covert and overt movements of attention in space.

Posner used a simple detection paradigm where one has to detect a target appearing in one of the two place holders that are equidistant from a central fixation cross (Fig. 2.1). On some trials, one of the boxes is flashed briefly and this summons attention towards it reflexively. In some other trials, a central arrowhead directs attention "voluntarily" towards one of the boxes. One critical issue in this paradigm is the percentage of trials where either the brief brightening of one of the boxes or the arrow head predicts the appearance of the target. A typical finding is that, when a target appears in one of the boxes where attention had already "moved", target detection is faster. This classic paradigm is now regularly used to study covert and overt shifts of attention as well as executive control.

There are three distinct stages in such movements of attention. First, there is an orienting phase. In this phase, attention can move towards a location without eye movements. Then, depending on the perceiver's goal, attention can be engaged at that location or disengaged from the object concerned. This operates like a cycle in

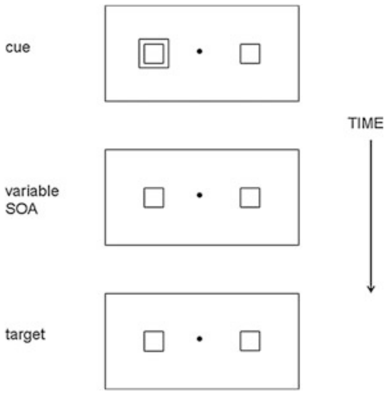

Fig. 2.1 A simple cuing paradigm showing exogenous movement of attention. One of the peripheral boxes is briefly cued. After some interval a target appears and the participant has to identify it. Target processing is facilitated when the target appears immediately at the location of the cue. However, if there is some delay, i.e. 200 m, its processing is inhibited. This is called the Inhibition of Return Effect (IOR). This figure does not show central endogenous cues. Generally, the target is a solid object which is not shown in this figure (Reproduced from Colzato et al. 2010, p. 3)

everyday situations and these distinct phases of movement of attention is linked to our overall cognitive behaviour. As agents we are constantly in need of surveying our visual environment. Klein (2000) observed that we need to constantly engage and disengage attention since we need to be constantly updating information in our visual world. Therefore, it makes sense to imagine that attention is not to be fixed at a place if nothing more is going to happen at that location.

If attention orients towards any location or object in space, because of some external sensory stimulation or even internal desire, then it is likely that objects that appear in that location later will receive priority processing. However, attentional mechanism seems to be reluctant to return to a location where it has been recently. This might be because of our novelty seeking tendency. This observation has led to a remarkable finding in the field known as the "inhibition of return" (IOR) phenomena (Posner and Cohen 1984; Klein 2000). It is a robust finding that if attention has moved towards a location and it has been disengaged from that location, then there is inhibition of further movement of attention towards the same location. However, a recent survey of attention experts of their opinions and conceptualization of "IOR" suggest that there are many different interpretations of this mechanism (Dukewich and Klein 2015). In most cases, it is seen that detecting a target object at a location where attention has moved is enhanced if this object appears immediately but detection is delayed if the target appears at a gap of, say 250–300 ms. Attention is in a constant state of motion and one can see its effect in a simple cuing paradigm with costs and benefits as function of the scope of attention (SOA) between the appearance of the cue and the target. There is, though, some controversy if both exogenous and endogenous cues give rise to IOR. It is widely believed that one may not see IOR with endogenous cues such as centrally presented arrows. Conceptually, it makes sense to imagine that with central arrow cues attention moves voluntarily and, therefore, disengagement is costly. Therefore, there might not be immediate inhibition at that location. However, IOR is commonly observed when the cues are exogenous and if attention had moved reflexively (Klein 2000). IOR indicates the constraints in the movement of attention as well as the relationship between attention and control mechanisms.

Most studies use an arrow for triggering endogenous attention shift. It is natural that arrow heads will immediately move attention in their direction since they are learnt social stimuli. Similarly, eye gazes also move attention in the direction of their gaze. By definition, endogenous attention shift should be under the control of the perceiver. However, if these social cues shift attention reflexively because of their over learnt nature; it is difficult to separate out the exogenous component from the endogenous. Hommel et al. (2001) showed that words like "left" and "right" when presented as non-predictive central cues led to attention capture. These social stimuli therefore move attention like arrows and these movements could be reflexive. Arrows and eye gaze are socially learnt symbolic cues. Therefore, the way these cues orient attention could be entirely different from pure endogenous and exogenous forms of attention (Ristic and Kingstone 2012). This way of putting it also puts attentional control into the social domains. Humans orient towards things that they know and have experienced. Automatically orienting towards the direction of an arrow or gaze indicates immediate processing of the semantics of such cues. Imagine the role arrows have on street sign boards play. This type of orienting and attentional engagements may require a very distinctive type of control different from the traditional dichotomy (Berger et al. 2005).

Interestingly, language can also move attention. Words and sentences refer to things in the world or even to internal mental states. Direction words like 'up' and 'down' shift attention reflexively and affect eye movements (Singh and Mishra 2012). Spoken words can cause visual attention to move in space towards things that they refer to and this seems to be rather automatic (Mishra et al. 2012; Salverda and Altmann 2011). The cross-modal culling of information moves attention and helps in search behaviour.

Recently, Klein and Lawrence (2011) have made an effort in grouping and classifying the various modes of attention allocation in different situations. In this framework, modes of allocation of attention is crossed with domains of attention, i.e. space, time, sense and task. This taxonomy is both intuitive and it captures a lot of experimental findings in different domains. Considering language, a performance that humans engage intentionally, it appears to be the case that both these modes of attention are utilized differently by different language activities. For example, naming a picture would require selective attention and visual object identification while listening to speech in noise may call for more filtering and control. This taxonomy reflects both the selectional and control properties of attention in different situations. The top-down and bottom-up views of attentional operations reveal something about the ways in which humans function intelligently in the world. Another way to think about the several attributes of attention is to think of it in terms of external and internal attention (Chun et al. 2011). This view acknowledges attention as comprising multiple cognitive and control mechanisms (Pashler and Sutherland 1998). Chun, Golomb, and Turk-Browne write,

> … 'external attention refers to the selection and modulation of sensory information, as it initially comes into the mind, generally in a modality-specific representation and often with episodic tags for spatial locations and points in time. This sensory information can be organized by features or into objects, which can themselves be targets of external attention. Another way to think of external attention is as perceptual attention. Internal attention refers to the selection and modulation of internally generated information, such as the contents of working memory, long-term memory, task sets, or response selection (p. 77)'.

The external aspect of attention is all about attending to things in the external world. Interestingly, one can also specifically deploy attention towards some features or aspects of objects for higher cognition and action (Wolfe and Horowitz 2004). This exogenous and endogenous distinction in the deployment of attention has proved to be crucial in how attention functions and interacts with other cognitive systems. On the other hand, internal attention can help control action. It helps in maintaining a task-specific goal for a long time and in monitoring eventualities.

2.2 Temporal Selection and Capacity Limitation

Selective attention is the key to acquisition of robust information. We move our eyes and fixate on things that we want to inspect more carefully. Selective attention on one object can block information flow from other surrounding objects. Even small insects select the targets to follow for food while ignoring others (Frye 2013). Selective attention could mean several processes like staying focused on something, ability to ignore irrelevant distractors as well as to monitor the ongoing activities for better performance.

Selective attention is necessary for processing some aspect of the stimuli while ignoring other aspects. When it comes to selection of the task-relevant stimuli from the myriad varieties that are constantly bombarding themselves on our senses, it makes sense to ask if such selection is early or late. An influential view in the psychology of attention has been that, attention basically functions as a filtering mechanism (Broadbent 1977). Attention works like a gate that controls the inflow of information for further processing. Broadbent proposed two stages of attentional filtering of incoming information. At the early stage, all stimuli undergo some filtering based on some physical measures such as pitch or location of sounds. Participants are not conscious of such early process since they happen at a very early time scale. The second stage operated in more serial manner and filtered stimuli only on the basis of some content such as meaning. This two-stage model of selective attention included the concepts of 'limited resource' and 'filtering'. The irrelevant stimuli are excluded from further processing much early in the processing cycle or do they all receive same processing for some time and then final selection follows. This dichotomy has been the most talked about theoretical issue in the modern history of attention. One way to show the functionality of the limited capacity system is to consider a filtre that allows only the most useful information for further meaningful processing (Broadbent 1958).

In this 'early' selection view, the non-attended or filtered out information will have no subsequent cognitive importance. In this conceptualization which was an 'early' selection model, the ignored stimuli were completely blocked out and no memory traces were formed (see Driver 2001 for a review). It could also be the case that all the features of ignored stimuli receive full consideration, but they are 'attenuated' (Treisman 1964). Perceivers can always separate out different messages based on their physical features, i.e. a male voice from a female voice. Often, meaning plays a crucial role in this perceptual distinction. It is not necessarily the case that the system filtres them out totally, one can still see some influence of these stimuli in later processing. Accentuation of features makes them more available to awareness and they can thus slip into memory and affect performance. Another extreme position could be that, all the signals that present themselves to the sensory processing system initially receive full processing (Deutsch and Deutsch 1963). Most of these early studies that examined filtering and selective attention in dual task situations used auditory-verbal material. The evidence that distractors can influence processing and attract attention has been considered as a support for the later selection view (Eriksen and Eriksen 1974).

Rock and Gutmann (1981) asked people to focus attention on only one colour of shapes while the visual stimuli were superimposed by other colours. A memory retrieval test showed that participants could not report the unattended features. This result suggested that the unattended features were not coded during early perceptual processing although the visual system had processed them. These results are similar to the in attentional blindness studies in the sense that unattended information escapes consciousness and subjects fail to report. However, from these results it is not possible to completely rule out that the ignored information had absolutely no effect on later processing. Later studies showed that even ignored stimuli could enter into awareness and affect response. This led to reformulation of the early models and arrival of the 'late' section models (Deutsch and Deutsch 1963). Tipper (1985) observed that participants were slower in naming a picture in a particular colour when a previously ignored distractor becomes the target in the current trial. This could only happen if the distractor had received full perceptual processing whose features are still active in influencing decision.

When is selective attention essential? One suggestion (Lavie 1995, 2000) has been that heavy perceptual or cognitive load ensures the engagement of selective attention. Load can be understood in terms of increase in the number of distractors or additional tasks. Low visual quality as well as multiple tasks demanding limited attention can also increase perceptual load. Lavie (2005) attempted to explain the distractor effect by using the 'load' metaphor. She argued that distractors may not be that influential if the main task demands sufficient amount of selective attention and taxes the cognitive system. If the task is of low load and is easy, then distractors will exert their influence. According to this theory if perceptual load is higher, there are many targets and the stimuli is degraded, then distractors may not have much influence. That is to say, if the task is attentionally demanding, people can ignore the distractors. Without a demanding task, distractors can occupy the available space and thus affect response. Lavie et al. (2004) further observed that where as a high perceptual load reduced distractor interference, high cognitive load (under a dual task) increased distractor interference, thus suggesting two modes of selective attention mechanism. Contradictory to the claims of the perceptual load theory, Roper and Vecera (2013) found flanker effects even under high perceptual load. Likewise, Yeshurun and Marciano (2013) found flanker effects even when the stimuli were perceptually degraded. These results suggest that the central claims of the load theory are still to be thoroughly evaluated and may not be as readily observable in all situations. Task demands, level of stimuli degradation, type of perceptual or cognitive load can influence target processing.

Attentional capacity at any given point in time is limited. The limited capacity of the attentional system is evident in a phenomenon called attentional blink. This arises when two different visual targets appear very close to one another temporally.

In this situation, detection of the second target is delayed (Raymond et al. 1992). Attentional blink is higher when the time lag between the first and the second target is very small, i.e. 100 m. However, with the increase in this duration, responses towards second target improve. The attentional dwell time hypothesis claims that resources are unavailable for both the targets when the first target is getting processed, particularly when the second target is presented very shortly (Ward et al. 1996). In laboratories, it is generally tested by presenting letters or numbers constantly at the same location. The participants normally detect the identity of one letter whose colour makes it distinguishable. Interestingly, when another letter quickly follows this target, often participants have difficulty in recalling the second letter. Among the many explanations of attentional blink, it seems likely that it could be a result of capacity limitation of attention. It is interesting to note that when the second target is highly salient, such as a capital letter among small letters there is no attentional blink (Shapiro et al. 1997). When attention is deployed at one location or on an object, it is likely that other events are not registered. Others have viewed this phenomenon arising from a lack of cognitive control and change in attention set. Attention set comprises the objects or entities that one wishes to temporally hold in working memory for some goal-driven action. In psychological experiments, task instructions influence attention set. For example, if the stimuli are all numbers or all words of a certain type then this constitutes an attention set.

Di Lollo et al. (2005) did not notice attentional blink when the items belonged to the same semantic category, i.e. all numbers used in the trial. However, attentional blink appeared when targets were from different semantic field. This account of attentional blink emphasizes on the maintenance of a certain attentional set for task. Di Lollo et al. proposed that there is an input filtre working at the commands of a central controlled that passes the first target based on a criteria, i.e. word or letter. If the second target is similar to the first with regard to these features, then it is easily processed. Since the input filtre does not have to change its settings for it one does not observe attentional blink for this situation when both the targets are from the same category. If the second target is from another category, then the input filtre's criteria have to be modified and this can cause processing delay. Therefore, attention blink will be observed when the two stimuli are from different categories. This evidence suggests that similarities and differences among stimuli can affect efficiency of processing as well as memory retrieval. The relevance of the task and its complexity could guide how evenly one divides the available attention resources (Kahneman 1973). Similarities and differences between the tasks could also affect performance (Allport et al. 1972). If tasks belong to the same modality, division of attention could be problematic. Limited capacity arises because of certain very core cognitive constraints. For example, one can store only four items in the visual short-term memory for some time (Luck and Vogel 1997). Things kept in working memory can only remain for a short amount of time before they vanish (Baddeley 2003). Even if we accept a limited capacity model of cognition, it is important to know how the correct stimuli are selected.

How any stimuli become our goal and how our attentional mechanisms select them for processing remain a mystery. A bright red Ferrari captures your attention while you are window-shopping. This kind of attention capture by an exogenous cue could be the result of salience of the object—its distinctiveness from its surrounding (Olivers and Humphreys 2003). However, attention captured by an exogenous cue can also be disengaged with the use of cognitive control. Mishra et al. (2012) demonstrated that highly fluent bilinguals could disengage attention capture faster in a cueing paradigm. However, distractors can catch attention and it might be difficult to pursue targets (Yantis 1993; Theeuwes et al. 1999). Maintenance of an attentional set helps target selection quickly and disengages attention from the irrelevant stuff. In any case, selective attention plays a crucial role in many such events of control. Attention during visual search can be directed to objects based on their distinctive visual features. This mode of feature-based attentional allocation is different from spatial attention. For example, searching for a particular friend in the crowd is possible as specific features of this person's face can lead to attentional bias in the visual field (Zhou and Desimone 2011). Importantly, this feature-based bias is seen when eye movements are planned to some other location (Bichot et al. 2005). It has been found that feature-based attention leads to inhibition of distractors within 100 m of target onset (Moher et al. 2014). The biased competition account of attention is an influential account which treats attention as emergent phenomena of neural mechanism and not as an agent (Duncan 1996). The theory predicts competition among visual stimuli for neural representation. In this model, multiple stimuli compete for selection at any given time and finally attention is deployed on one. This can happen through both top-down and bottom-up influences (Beck and Kastner 2009). Any pre-activated object feature can boost representation of that object and bias attention in a top-down manner. Similarly, various low-level visual features and visual organization can lead to bottom-up bias. This account does not assume a filtering model of attention; neither does it consider attention as an agent. Competition is central to its mechanisms. A lot of data from single cell recording show competition and bias leading to selection. Directing attention towards any one object leads to suppression of activity in nearby elements (Kastner et al. 1998). This bias and simultaneous suppression is affected by both top-down and bottom-up mechanisms. Desimone and Duncan (1995) influential theory of 'biased competition account'suggests that attention moves towards the targets as a function of competition between targets and distractors. When distractors are more salient, then they capture and hold attention. Even though distractors compete with targets, attention gradually gets biased towards targets.

Thus, selection through competition as a mechanism through the interplay of top-down and bottom-up factors have been considered as a central mechanism for object selection during visual search. The central question is, how, given a set of targets and distractors, selective attention is directed towards targets. Building on the basic premises of the biased competition account of target selection, an integrationist account of visual selection and action has been offered by Bundesen (1990).

According to this theory, attention allocation during visual processing is a competitive allocation of resources between targets and distractors. Given a template that consists of targets and distractors, computation of attentional weights for different category of objects take place. Attentional weights are assigned based on prior consideration of features, i.e. green objects. Importantly, in this model, only a certain number of objects can be maintained in the visual short-term memory since it has very limited space. This model also does not make a distinction between attention allocated and task-relevant or task-irrelevant stimuli initially, except the weights that are attached to them. Unlike the load theory, which emphasizes the perceptual load, the theory of visual attention (TVA) does not account for the distribution of attention to targets and distractors in terms of load or task difficulty. Thus, the load theory predicts that when there is high perceptual load, number of distractors will have no effect on processing. On the other hand, TVA proposes that the number of distractors will have an effect on target processing irrespective of perceptual load, since it considers visual working memory limitations as a major factor in selection (see Giesbrecht et al. 2014).

It is not enough to suggest that top-down goals and bottom-up factors decide what gets selected: the question is, how this mechanism works for different situations. The situation becomes complex when there are multiple targets competing with multiple distractors. Figural similarities can further complicate the matter. Bundesen (1990) proposes that attentional selection happens through two different mechanisms of 'filtering' and 'pigeon holing'. Filtering helps recognize the object and the latter processes its features. That is how initially attention is biased towards a certain object with a certain feature. When such a filtering has happened, objects that fulfil these criteria are immediately encoded into the visual short-term memory. Since the capacity of the visual short-term memory is limited (Luck and Vogel 1997), competition for selection could be stiff. While it is still not clear if distractors receive full processing before they are discarded as undesirable, it is clear that even identification of the simplest of features requires focused attention of some sort (Luck and Ford 1998). Selective attention modulates processing of items that are important.

2.3 One or Many Focuses?

The classical view on selective attention is that it has a single focus. However, if it is shown that people can attend to many objects at the same time, then one has to accept the view that attention can be deployed to several objects at the same time. Most studies with the visual search paradigm ask participants to find out only one target among some competitors. In this situation, the test participant has to keep features of this lone target in mind and search. However, on many occasions we search for many things or keep their identities in mind, although the number of

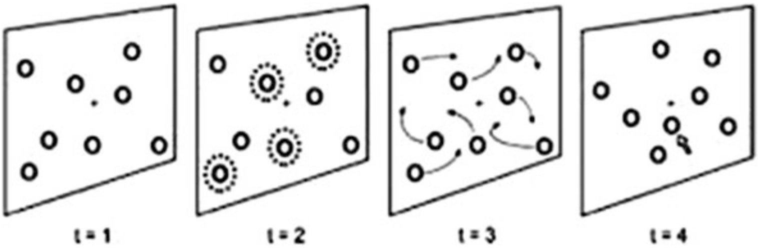

Fig. 2.2 The figure shows a multiple object tracking experiment. First a display of 8 identical objects is shown (t = 1). Then a subset of 4 'targets' are briefly flashed to make them distinctive (t = 2). Following this the objects stop flashing so the 'target' set becomes indistinguishable from the other objects. All objects then move in a random fashion for about 10 s (t = 3). Then the motion stops (t = 4) and the observer's task is to indicate all the tracked objects by clicking on each one using a computer mouse (Reproduced from Scholarpedia, 2(10): 3326)

items we can keep in our visual working memory seems limited to four (Luck and Vogel 1997). It looks like most can comfortably track a few objects and the changes they go through for some time (Pylyshyn 2003). This is like attending to any two or three children and tracking them as they are playing in a park where there are other children around. There could be a large amount of visual similarities among the target and distractor children and the number of distractor children could affect the tracking efficiency.

In the multiple objects tracking experimental set-up (Fig. 2.2), some objects are first indicated as targets and then the targets and distractors move about in the visual field for some time. The task of the participant is to say after some time if a particular item was the target. Therefore, the important measure is if one is able to simultaneously follow several targets over time. Participants actually can follow about four or five objects and do very well on the test. This can only happen if one assumes multiple focusing attributes for attention. Or, one assumes that several targets that share colour or move in the same direction with same velocity are treated as a single object (Kahneman et al. 1992). Once attention glues these various targets and treats them as one object, then tracking them can become easy. It is like tracking a flock of birds as they fly in some arrangement and cross the path of another flock. However, the question is when attention is on the flock as an object, could it still be deployed on each bird that makes the flock? The issue has consequences for some major theorization about the capacity limitation of selective idea of attention. Pylyshyn has proposed (Pylyshyn and Storm 1988) this mechanism in terms of FINST (Fingers of INSTantiations). This means, some indexes are attached to each target which then dynamically move with it. Indexes allow attention to keep track of targets.

In any case, it is likely that the available resource will be shared by all the targets. When the targets are pre-attentively indexed and tracked, later processing shows facilitation (Sears and Pylyshyn 2000). Participants are generally better in judging any change if the object undergoing change was tagged as a target. Interestingly, number of non-targets seems not to affect this process. In most multiple object tracking experiments, objects move around or change their form. In traditional visual search experiments, the display is static and changes very little. The deployment of visual short-term memory and attention is different for both cases. Similarly, the space between objects seems to be critical in tracking experiments as this has an influence on what are treated as an object (Franconeri et al. 2010a). This evidence indicates that humans can focus attention on multiple objects and carry out tasks. In the multiple object tracking studies, the important issue is related to the identity of the targets, which apparently remains constant over the tracking time. That is subjects track the same targets based on their initial identities. Franconeri et al. (2010b) observed that when targets come too close to one another, i.e. spatial distance between them is less, and then subjects have difficulty in reporting them individually. In multiple object tracking studies, subjects individuate each target and track them. However, does this mean that they keep in memory all the features of the targets that they track? Bahrami (2003) showed that even if subjects have been tracking some targets successfully they fail to notice featural changes in targets during the tracking. This means, targets are tracked with their original identity and this remains so. Bahrami (2003) also observed that subjects are better at probe detection on targets than non-targets.

The multiple object tracking studies have challenged the belief that attention has a single focus as well as the capacity limitation. The studies described above demonstrate that subjects can track around four objects over time accurately. However, it is still not clear if attention operates serially or indeed has multiple focuses (Cavanagh and Alvarej 2005 for a review).

2.4 Attention and Consciousness

Attention is needed not only for working out on current plans and goals and execute them, it also enriches our conscious awareness of such objects (Cohen et al. 2012; see also Koch and Tsuchiya 2012). One can be phenomenally conscious of certain things but needs attention to have specific knowledge about them (Block 2011). The rich environment around can influence us in many ways and we seem to have some idea about things in general but we need to focus attention on them to develop explicit and objective knowledge about them (Block 2005). Attention and consciousness are related although they could be different processes (Lamme 2004). Subjects can be conscious of certain objects without paying any attention (Koch and Tsuchiya 2007). Several researchers have tried to examine if focused attention is necessary for conscious perception. Or conscious awareness of objects can arise in the near absence of attention. One important paradigm to study these mechanisms has been priming. Bussche et al. (2010) presented primes that were either subliminal or were clearly visible. They also manipulated attention by asking participants to focus on the primes or not. The results showed that prime visibility and top-down attention had a significant influence on target recognition. This suggests that attention elevates conscious perception. These views suggest that selective attention is required only for in-depth processing of some stimuli but not always. Objects that are attended to receive rich representation in memory and are also retrieved faster (Buckner et al. 1999). Attending to something also enhances the salience of objects and makes them more prominent. When attention illuminates objects, we can act on them in a goal-oriented manner. Attention is necessary for conscious recollection of events from memory (De Brigard 2012).

One may have the illusion that one has sufficient knowledge about things in our visual world, but it is not true. At best this could be very vague, a snapshot only. To generate objective ideas about things, selective attention must be deployed. Often, it is difficult to detect changes in the environment, if selective attention is not paid (Rensink et al. 1997). Attention foregrounds objects that we are interested in and pushes others into the background. Selection of action on particular objects of choice happens through these constant shifts of attention from one to the other. This gives us a sense of control on our action and our goal directedness (e.g. Frith 2002). Many a times, we just look at some objects since they capture our attention for reasons beyond our control. It has also been argued that without the deployment of attention, objects escape our consciousness awareness (Simons 2000; Wu 2011). Termed as in attentional blindness, this phenomenon means that without paying attention one cannot perceive an object (Mack 2003). People fail to detect changes or another event occurring in the visual field, if they are attending selectively to some specific object or event (Simons 2000). Drew et al. (2013) asked expert radiologists, who have spent years examining images and apparently have the ability to detect small changes in them, to detect a gorilla in some of the pictures. The gorilla was too large to be missed. Interestingly, 83 % of the radiologists could not see the gorilla; however, their eye movement record showed that they had fixated on the gorilla.

This evidence suggests that without selective attention even a very large and salient visual event can escape our consciousness. Furthermore, participants can sometimes report events and objects if they are primed and if a suitable context is provided. Slavich and Zimbardo (2013) observed that a majority of participants could not locate the figure of a suicidal woman in a picture that was suspended mid-air even though they saw the picture many times. However, when primed with a story related to suicide, some of them could locate the woman. This suggests that past experience and context can direct attention towards relevant locations in the visual field and thus enable us to locate objects. Listeners cannot detect changes even in their native language when their attention is not cued (Neuohff et al. 2014). All this evidence suggests a close link between selective attention, conscious perception and our ability to observe.

2.5 Visual Search

How do we search for things in the real world? Sometimes the search for targets is easy since the target could be sufficiently salient from its environment. It is very easy to find out the X in Fig. 2.3a since it pops up. This does not call for deployment of selective attention and search may progress in parallel (Treisman and Gelade 1980). This kind of search is purely bottom-up. In contrast to this, searching the inverted T in Fig. 2.3b is difficult and may involve serial search. This involves blending of two features, i.e. identity and its orientation. When the set size, i.e. number of other items that are not targets increase, search times also increase differently for these two types of searches. What is also crucial to note here is the feature similarities between targets and distractors. The more similar the distractors are to the targets, the more features they share and the search requires higher focused attention. Attention also works as a glue for different types of perception and leads to some holistic gestalt (Treisman and Gelade 1980). If one is looking for a red cube whose boundaries are also solid, then, it is necessary to keep these two features glued for efficient search. This is what probably Wittgenstein had in mind

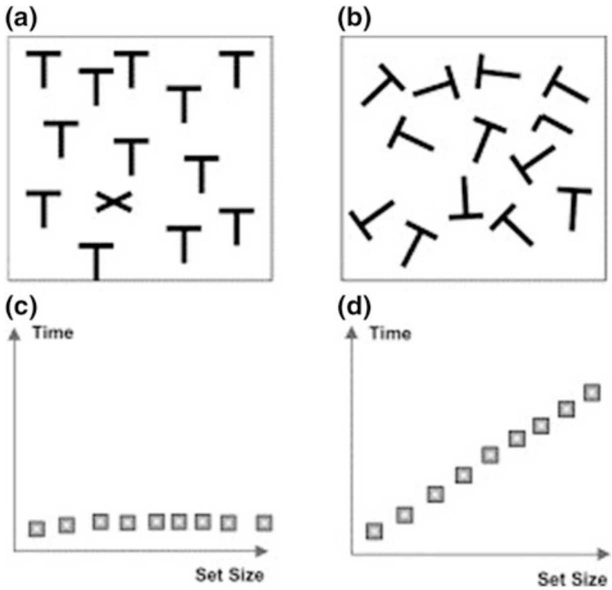

Fig. 2.3 The *top left panel* shows a simple visual search where the target 'pops out'. The *top right* shows a sample of what is known as a 'conjoined' search. Here finding the target, L involves keeping the letter identity as well its orientation in mind and this search requires serial application of selective attention. The bottom results show increase in target detection speed for these two types of searches. Set size refers to the total number of items on the search display. While the increase in set size affects the efficiency of conjoined search, it does not affect much the simple search (Reproduced from Yang et al. 2002, p. 293)

when he said that a child needs to pay attention to both the spoken form and the visual object in order to grasp the essence of the object (Wittgenstein 2010). Two different types of signs thus could be unified and represent an object. Different features of visual objects are tied together with the involvement of attention (Wolfe 2012; see also Di Lollo 2012).

Top-down and bottom-up attention control mechanisms have been found to engage different brain networks. In a study with young children and older adults, Li et al. (2013) observed that older adults recruited more frontal and parietal regions for both serial and parallel search tasks. Children recruited more parietal areas for the pop-out search whereas adults recruited the parietal and frontal areas more equally. This evidence suggests that cognitive control mechanisms interact with these two types of searches. Buschman and Miller (2007) observed frontal neurons signalling top-down attention control early on in search task and partial neurons for bottom-up task in primates. This indicates neural specialization for these two types of attention allocation. Top-down attention takes time to act while bottom-up attention is immediately activated.

An important question in this debate between top-down and bottom-up attention control has been if salient stimuli capture attention unintentionally. If that is so, this could demonstrate the fallibility of top-down attention control. This debate has also been around the issue of the time scale at which top-down and bottom-up attention exert their influence. Human agents deploy attention towards task-relevant stimuli wilfully. This has been hailed as top-down control of attention. On the other hand, attention is also captured unintentionally by several objects in the environment. This has been variously called as bottom-up capture of attention. The earlier example of parallel search where the salient target captured attention can be considered an example of bottom-up capture. However, this can also be considered top-down since the participants knew 'what' object they were searching for Kim and Cave (1999) asked observers to search for a unique target while the display had a salient coloured distractor. Data showed that this distractor captured attention early on but attention returned to the target later. This suggested that top-down factors did not prevent attention capture by distractor. However, Lamy et al. (2003) found that when a conjoined search task is given, i.e. search cannot be accomplished with the pop-out mode, then the salient distractors did not capture attention. This suggested that top-down attention can act early on and prevent attention capture. However, it is not clear from these studies if the type of search, i.e. serial vs. Parallel is the key to contingent capture. There is evidence which suggests any salient stimuli will capture attention irrespective of the relationship it has with the target (Theeuwes 1992). In contrast to this, 'contingent capture' theorists suggest that a singleton will capture attention if it matches in features with the target (Folk and Remington 1998). Theeuwes (1992) asked observers to respond to the orientation of line inside a target diamond surrounded by other distractor circles. On some trials one of the distractor circles was in green colour. This led to attention capture and delay in judgement. This happened when the colour of the circle had no similarity with the diamond's colour. Thus suggesting that such a pure capture of attention is involuntary and operates in a bottom-up manner very early during visual processing. This depends on the salience of the distractor and is pre-attentive. Folk and Remington (1998), however, observed that if distractors and targets match in colour or in some feature, then there is attention capture. When distractors were in green and targets were in red, then there is no capture of attention by the distractor. This suggests that early on attention exerts top-down influence on visual processing.

Another group of studies have shown that contents of working memory can guide attention automatically and any item matching to the working memory contents can lead to attention capture. However, there are conflicting facts around these phenomena. Olivers et al. (2006) found that singletons capture attention more strongly when they matched items held in working memory. Further, these singletons attracted eye movements but fixation durations were not affected. This suggested that participants could disengage attention from contingent capture soon. In another study, Soto et al. (2005) asked participants to identify a tilted line among vertical distractors. This line was within a shape with a unique colour. Importantly, on some trials this shape matched an item held in working memory. Search was efficient when there was a match between the shape and working memory shape. However, on a similar visual search design, Woodman and Luck (2007) did not find any attention capture by distractors that matched working memory contents. Rather, they observed that participants could avoid looking at the distractors, suggesting a flexible manipulation of working memory contents for action. Han and Kim (2009) made the search difficult by manipulating the perceptual salience and also delayed the search onset. They found that in such cases there is no automatic guidance of attention from working memory. The authors suggested that a perceptually difficult search can eliminate the automatic capture and delay caused cognitive control to operate top-down effects on search. Therefore, while the issue of automatic guidance remains controversial it is important to understand that attention capture can happen through many means and working memory plays a key role. An item matching in features with the target can capture attention and this is called contingent capture distractors (Wyble et al. 2013). Visuo-spatial attention works as a filtering agent exerting top-down influence on attention capture (Theeuwes et al. 2010). An object's visual distinctiveness makes it a good candidate for attention selection (Itti et al. 1998; Serences and Yantis 2006). Cognitive agents bias their attention towards task-relevant objects or locations and avoid distractors using top-down experiential knowledge (Gazzaley and Nobre 2012).

2.6 Scope of Attention

So far we have seen that attention has been viewed as a limited capacity filtering agent as well as a component in cognitive control. We also saw the manner in which attention can be deployed during visual search, i.e. top-down and bottom-up. Another important property of visuo-spatial attention is its aperture. Attention has always been imagined like a light beam which illuminates aspects of the world for the cognitive system to process. It appears that one can wilfully ‘broaden’ or ‘narrow’ down the spatial scope of attention during visual perception. For example, while looking at the scene of a forest, one can look at the entire forest, the clusters of trees and other things without specifically paying attention to any particular object. This broad view helps in perceiving the whole object; provides a kind of holistic vision. Similarly, one can be asked to count the number of leaves on a plant and so on. To do so, the subject has to specifically focus on each leaf. One can deploy attention on a ‘global’ scale as well as to ‘local’ elements.

Global attention is diffused and is generally spread across multiple objects. These attentional states are induced by asking participants to process these hierarchical figures in one or another mode. And also in other situations, participants can themselves deploy these modes in a top-down manner. An analogous event in the domain of language will be to ask someone to just focus on the overall meaning of a sentence or its specific parts. It can be then assumed that in the former case, a kind of global attentional mode operates while in the latter case a local mode is relevant.

A long running idea in this domain has been that visual perception operates at a global level by default. That is we see the forest first and then the trees. Navon proposed the ‘global precedence’ hypothesis to suggest that the visual perception system processes global features more immediately and with some primacy. Navon did early experiments on this idea (1977) with hierarchical figures. Hierarchical figures are made out of simple figures (Fig. 2.4). For example a large five made out of small fives [congruent] or made out of small sixes [incongruent]. One is asked to count the number of letters that compose the large letter and it is assumed that in doing so one is using the ‘narrow’ focus of attention. If asked about just the identity of the large number then global attention can be sufficient. Navon (1977) found that

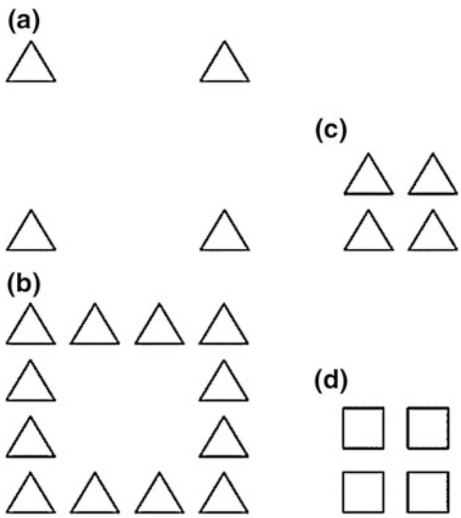

Fig. 2.4 Different representations of global and local levels. A square made out of small triangles could be problematic to perceive as compared to a square made out of squares. This compatibility between global and local levels has been important for object recognition. *Source* Navon (2003, p. 276)

participants had interference in naming an auditorily presented letter while focusing on the global level of a hierarchical figure. Subjects also suffered interference when the global level of a hierarchical figure was incongruent with the local level. This suggested that attention operates at a global level in the beginning and then at a local level. Navon (1981) observed that with exposure duration of 150 m, participants were good at identifying the global letters and were also good at discriminating a tone but were slow with the local letters. Navon, argued that the global processing is a default processing mode and precedes the local processing mode. The global processing mode thus employed the broad scope of attention while the local level called for more focused and effortful attention. De Cesarei and Loftus (2011) presented high and low passed filtered images of objects, faces and hybrid images to participants for identification. A global presence was found thus confirming that the theory generalizes to all types of stimuli and even to perceptually degraded stimuli.

Brain imaging studies have shown that selection of visual scenes for either ‘global’ or ‘local’ type of processing happens in distinct cortical sites. The right lingual gyrus is activated when one is deploying ‘global’ attention to a scene, i.e. looking at the forest, while the left inferior occipital cortex gets activated when one is in a state of ‘local’ attention, i.e. counting the trees (Fink et al. 1996). If one is asked to first process a scene under a ‘global’ mode and then some other stimuli is presented which also demands a ‘global’ mode of processing then there is facilitation (Ward 1982). There is conflict when the two modes of attention are different. Switching between global and local attention states influences psychological processing. Mishra and Singh (2012) presented hierarchical figures and then written sentences for judgement. The hypothesis was that attending to either global or local level should affect overall reading times and judgement. There were sentences with morpho-syntactic problems as well as sentences with semantic problems and normal sentences. The result showed that participants were slower in reading as well as judging sentences of all types when they focused their attention at the global level but were faster when they focused attention at the local level. This suggested that global and local attention states can affect processing of linguistic stimuli differently. However, this study did not find any interaction with the mode of attention of type of sentences.

The zoom lens view assumes that attention can be shifted from narrow scope to broad scope voluntarily depending on the task at hand. When under diffused state of attention, one can discriminate only very broad features of objects, e.g. colour, but a focused beam is required when more details about the object is required. Selective attention is deployed during the narrow or focused state and objects or features of objects are ‘foveated’. Scope of attention can be thought of as a gradient, spread out over the entire visual field (LaBerge and Brown 1989). Visual acuity falls off gradually as one moves from the centre to the periphery and this affects processing and discrimination. During focusing on a word while reading a sentence, it is possible to be vaguely aware of the words to the left or right but it is difficult to be sure of their details. This means, in order to know their details, eyes have to be moved towards them to bring them under foveal processing (Rayner 1998). Movement of the eyes is essential to bring newer objects under the fovea and deploy selective attention. Therefore, the gradient view of attentional shifts views movement of attention as a dynamic thing. Thus, the studies with hierarchical figures show that attention has different levels of scope and these levels affect perceptual processing.

2.7 Attention, Control and Action

Attention is first and foremost linked to our limited capacity to process information. Michel Posner notes that attention should not just be viewed as a selection mechanism but a way of the mind to act in the world (Posner 2012). The brain has to select the most relevant information for processing among competing stimuli for optimal cognitive processing. Amid the constant flux of myriad sensory events, one needs to find goal-specific information, select objects and 21hannelizi action (Neumann et al. 1986). If this were not the case, then we would not be in control of our thoughts and action. Nevertheless in spite of our top-down control, attention is often captured by peripheral cues that may not be task relevant (Johnston et al. 1990). Selection and goal-directed actions have been conceived as crucial attributes of attention (Posner 2011; Cohen 2014). Selection of the goal-relevant stimuli is affected by many other competing stimuli in the environment. Therefore, the need to select the most relevant stimuli is crucial for successful cognition and survival. Our cognitive limitations do not allow for simultaneously coordinating many actions. Thus attention helps filtre and focus on the task-relevant phenomena.

One can select, focus and act wilfully because of attention. It is the view that attention is useful for action, and that goal-directed action is important for the understanding of language–attention interaction. That is because language is also an action-based system (Pulvermuller 1995), though symbolic. Broca's area has been found to be active not only during language processing but also in action observation, suggesting a high degree of overlap between the linguistic and action systems (Willems et al. 2007). Selecting the right language itself might call for attention selection as it has been observed in bilingual language processing (Green 1998). Therefore, attention and language interaction has to be understood in terms of selection, control as well as focusing features of the attention. The control of 21hanneliz or its regulation is important for cognition. Attention affects the way our thoughts and actions synchronize (Posner 2012). Actions crucially depend on selection of the right goal, maintenance of information and sorting out conflicts timely.

Humans can control their action by resisting interference from 21hannelizin and selecting the correct response. For example, to pick an apple one has to control interference from oranges, since they resemble an apple in shape and also belong to the same semantic field. Studies have shown that unintentionally the action system gets briefly biased towards 21hannelizin in such situations (Duran et al. 2010).

Several important brain networks have been implicated that suggest the regulatory function of attention in such situations (Posner 1994). Both top-down and bottom-up factors can affect such control processes. Top-down effects may include current goals, motives as well as task instructions; whereas, bottom-up influences arise from the salient properties of the stimuli themselves. Attention plays a role in controlling distractor interference (Serences et al. 2005; Lleras et al. 2013 for a recent review on this issue). Therefore, apart from the selectional role, attention aids in cognitive control.

Current cognitive goals are required to be translated into action and 22hanneliz. Attention plays a crucial role in translating these goals into action (Monsell and Driver 2000; Posner and Fan 2007). This aspect of attention is important for linguistic processes like speaking and discourse (Brennan 1995). Often goals have to be maintained for a longer period of time so that there is no interruption in the execution of action. This maintenance of information and action schema in the working memory is taken care of by the executive control system (Miller and Cohen 2001). Attention's role as a selection mechanism is linked to this conceptualization. Second, attention's role in goal-directed 22hanneliz has often been used by language researchers for a long time (see Allport 1993).

Attention first and foremost influences control on our thoughts and action. Actions that are habitually performed appear automatic. Such tasks do not call for voluntary attention. Reading, speaking and listening look automatic to all from this perspective. The dichotomy between 'controlled' vs. 'automatic 'processing indicate differential involvement of attention (Kahneman 1973; Shiffrin and Schneider 1977). Controlled processes require attentional engagement and they are often voluntary. On the other hand, automatic processes require no attentional resources and get instantiated and complete their cycle, sometimes even unconsciously (Norman and Shallice 1980; Posner 1978) However, if there is change in the action plan or a conflict arises where the habitual response might lead to more problem or when one has to stop a wrong action plan, then cognitive control becomes necessary. Tasks that are effortful require attention.

In a comprehensive conceptualization of automaticity, Bargh (1994) separated automaticity into four components. They were awareness, intention, efficiency and control (see also Moors and De Houwer 2006). More recently, Bargh has written, 'Automaticity means the direct environmental control over internal cognitive processes involved in perception, judgment, behaviour, and goal pursuits' (Bargh 2011, p. 629). There are some processes that seem automatic in the sense that they are activated and processed without the requirement of focal attention. For example, in a priming task, the activation and later influence of the prime on target processing is automatic. This happens when the prime is perceived subliminally. In some situations, focal attention is required for identification and perception. Automatic processes could be just reflexes (Anthony and Grahme 1985). Reflexes are things that our cognitive systems have acquired by experience and spreading activation leads to their automaticity. For example, the temptation to look at the direction of an arrow is a reflex since an arrow is a socially learnt stimulus. Cognitive control is necessary to stop an action that has come about because of some reflex.

Attention, apart from being a selective and a filtering mechanism, becomes crucial in inhibition of task-irrelevant responses. Many a times one has to focus attention on one aspect of a task while ignoring other aspects. For example, in a Stroop task, attention has to be given to the colour of the word while ignoring the meaning of the task. Therefore, conflict resolution in such tasks calls for selective attention. The executive control system exercises inhibitory control to deal with such a situation (Botvinick et al. 2001). The prefrontal cortex and the anterior cingulate cortex show activity during sustained attention tasks (Miller 2000). An eventual conflict then can be easily handled with higher alertness and change in the response. Many important theories of conflict resolution have accorded a central role to attention (Botvinick et al. 2001). Task-related goals must be kept in working memory. Attending to this information held in working memory requires executive attention (Engle 2002).

2.8 Attention in Scene Perception

The studies described so far have employed well-defined laboratory-based tasks where participants gave responses to stimuli under experimental conditions. The studies with attention blink as well as with attention scope showed different facets of attention. However, one wonders how attention operates when one is viewing a real-life complex scene, i.e. a photograph. What kind of top-down and bottom-up factors influence scene viewing? The use of eye tracking has led to important discoveries about the mechanism of scene perception, attention as well as memory (Henderson 2003). The real world is full of objects and often they are embedded against one another in a rich and scattered canopy. Objects in the real world do not come organized as they are often in psychological tasks. Even features that we have to blend together for search might be very diverse. Then, what guides visual search in the real world? This issue is crucial to understand how attention interacts with language since language refers to things in the real world and influences attention. Visual semantics and scene context could guide attention during a scene search. Interestingly, objects consistent with the scene schema are identified faster and reported better (Davenport and Potter 2004). Similarly, when objects do not go well with the scene semantics, they pop out and capture attention (Hollingworth and Henderson 1998). However, this popping out is because of very early activation of conceptual knowledge that make the scene-inconsistent object salient and not so much due to its visual features. Scene semantics is rooted in our knowledge of objects and their everyday locations. The cognitive goals of the perceiver and task goals influence where people look in a scene and in what order they look (Buswell 1935). There is a close relationship between what people know in general about things and what experience they have had with such scenes (see Intraub 2012).

An important debate in this domain of research has been on the issue of attentional involvement in scene perception. Does one need focal attention for scene perception? And a second issue has been around the top-down and bottom-up factors in scene perception. Many studies that have investigated real-world search have used complex scenes, mostly photographs. It is known that within a very brief amount of time, viewers can register the gist of scenes (Fei-Fei et al. 2007). Often, the claim is that this sort of gist is acquired without attentional engagement (Fei-Fei et al. 2002). It is natural that with a brief glance there is not much time to look around and get details. The knowledge received from a gist could be very vague, e.g., the scene of a beach. You would not be able to tell the colour of the cap of the boy who was standing near the palm tree. For this to happen, you should deploy selective attention to the object concerned. Report of information is always better when some amount of attention is paid to the scene (Mack and Clark 2012). However, Cohen et al. (2011) argued that when the primary task is demanding, when participants are asked to recognize scenes under a dual task situation, then attention is necessary for perception. Therefore, whether scene perception requires attention depends on the task demands and cognitive goals of the perceiver. Even if it is possible to acquire some gist in a brief amount of time, one needs to deploy focal attention for more objective processing of scenes.

As for the issue of top-down versus bottom-up factors in scene perception, there have been two dominant views. One view has been that the cognitive goals of the perceiver guide scene perception (Henderson 2003), while others have offered a more stimuli-based account of scene perception, suggesting that eyes move around salient locations in a scene. Itti et al. (1998) showed that scene saliency attracts eye movements. It is evident that human observers do not look at things predicted by computational models that only consider the visual salience of things (Foulsham and Underwood 2008). Underwood et al. (2006) directly examined of scene saliency attracts eye movements during scene perception. They presented scenes that contained objects that were either highly salient or had low salience value. In the first task, participants had to look at the scene for a later memory task while in the second, task participants were asked to identify a low salient target in the scene. In the first task, participants looked more often at the highly salient objects while in the second task, the highly salient object failed to attract attention. The study thus demonstrated that when perceivers view scenes with a cognitive task, low-level visual factors do not dominate visual perception. Bottom-up theorists believe that low-level salience of things capture and guide attention (Wolfe 1994). However, many think it is the cognitive goal of the perceiver and context that guides attention in the real world (Henderson et al. 2007). Depending on what information one is looking for, one will look around the scene in a particular manner (Henderson 2003).

Visual context and top-down knowledge seem paramount for 25hannelizing attention during real-world scene perception. Language interacts with scene semantics and affects attention dynamically (Knoeferle and Crocker 2010). Selective visual attention to different objects in a scene is an outcome of what we know about the scene, what we are looking for and what we listen about it. It is apparent that these events are vastly complicated compared to the simple search tasks described before.

2.9 Attention, Language and Eye Movements

It is not enough to talk about attention with data from a single modality but it is also important to look at the contexts where information from multiple modalities interact and influence one another. A rich history of psychological research tells us that attentional shifts in one modality are sensitive to events in another (Driver and Spence 1998). Psycholinguists have shown that language input influences attention shifts. Eye movements are crucial indicators of such shifts in internal and external attention. For example, the spoken word ‘apple’ can drive eye movements towards a fruit basket. This sort of saccadic eye movement is both automatic and show sensitivity to both visual and linguistic context (see Mishra et al. 2013). Most researchers study basic eye movements like saccades, fixations and smooth pursuit as well as micro-saccades in certain specific situations. These days many different types of eye trackers are easily available to study human cognitive functions (Rayner 1998).

Eye movements over a scene and their patterns can also indicate the cognitive goals of the perceiver. It is known since the early studies of Yarbus (1967) that human participants look around a space depending on their goals. Yarbus asked participants to search for specific information in a painting and each time he found very different scan paths (sequential movement of eyes). This was interpreted suggesting that scan paths indicate the ‘route’ goal-directed plans take. Yarbus’s simple experiment was a powerful way to examine how our top-down goals and experiences affect the way we survey the world around us—suggesting that eye movements are controlled by intentional factors and can indicate cognitive Eyestates. Eye movements, therefore, largely reflect goal-driven cognitive plans of perceivers. Measurement of eye movements in cognitive tasks can tell us about certain ‘cognitive plans’ of the participants even when they themselves may not be aware of these movements. Eye movements are unconscious most of the time and participants do not always remember where they were looking. Once it is established that eye movements reflect the locus of cognitive processing, it makes sense to use them in understanding how attention interacts with language processing in different contexts.

The use of eye movements as dependent measures in the exploration of many cognitive tasks like reading, speaking, listening and scene perception can reveal facts about attentional mechanisms (Rayner 2009; Mishra 2009). With sophisticated and less expensive eye trackers available, it is now easy to measure these fast movements in a variety of real-life and experimental situations and even with infants and neuropsychological patients. Eye movements to a location follow covert attention shifts to that location (Sheliga et al. 1994). Some have argued that when one shifts attention covertly and does not make an eye movement immediately, it is as good as having made a saccade. Covert attention shifts are nothing but unexecuted saccades (Rizzolatti et al. 1994). This means one can select a target for possible saccades but may not look there. Therefore, covert orienting towards a

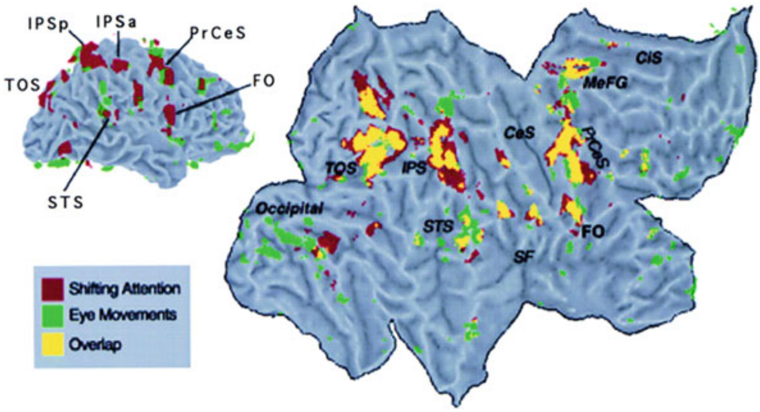

Fig. 2.5 Figure showing neural activation in the *left* and the *right* hemispheres in eye movement and attention tasks. The *red areas* indicate activation of cortical areas during attention tasks and *green* indicate areas active during eye movement tasks. The *yellow areas* indicate common areas active during both the tasks. In this study, participants shifted attention to a location with or without eye movements. Several *common areas* indicate a single common mechanism for these two cognitive systems (Reproduced from Corbetta et al. 1998, p. 834)

location or object in space is a by-product of the mechanisms of the oculomotor system. Not many though agree today with the strong version of this proposal but the idea remains quite attractive to those researchers who use eye movements in their researches to measure 'attentional' shifts. They tactically assume that looking at something means the subject is paying attention and therefore eye movements as a response system are cognitively interesting.

Eye movements and attention shifts could be independent systems, at least functionally (Klein 1980). This proposal suggests that attention shifts and eye movements arise for different reasons and thus should not be treated as the same mechanisms. However, other theorists believe that eye movements are nothing but a reflection of overt shifts of attention and one must treat them as the same system. Neuroimaging studies have observed significant overlap between the neural structures responsible for eye movements and attention shifts (Corbetta 1998; see Fig. 2.5).

An important issue here is whether prior attention shift is necessary for saccades towards an object or special location. By far, the strongest evidence that suggests a causal link between visuo-spatial attention and saccade generation comes from an experiment by Hoffman and Subramanium (1995). The researchers asked participants to programme a saccade towards a location while they were required to detect a visual target presented just before the actual eye movement. Participants were faster and more accurate in target detection when the location of the programmed saccade was the same as the location of the target. This result suggested that visuo-spatial attention and its movement are functionally related to the execution of saccades. Further experiments showed that it is practically impossible to move the eyes in one direction while shifting attention in some other direction. Similarly, Shepherd et al. (1986) attempted to separately manipulate attention shifts and eye movements. Participants were asked to programme a saccade towards a peripheral target while a centrally presented arrow indicated the location of the possible target. Centrally presented arrows have been known to cause voluntary shifts of attention in the arrow direction as well as indicating the likely appearance of the target. Participants were faster in making a correct eye movement towards a peripheral target if it appeared in the direction of the arrow. This suggested that preparing to make a saccade towards a location substantially brings down the saccadic reaction time to targets presented at that location and this indicates the effect of spatial attention shift. Saccadic reaction times could be higher if the targets appear somewhere else, indicating a cost. Therefore, this evidence might suggest that making an eye movement indicates allocation of attention and this shift of attention precedes eye movements. Deubel and Schneider (1995) examined if attention focus and saccade programming are interlinked. Deubel and Schneider (1995) observed that when participants were asked to programme a saccade towards a specific location, later target discrimination at that location improved. This could suggest that programming a saccade leads to covert attention shift to a location, even when no eye movements are made.

Attention and eye movements seem to be linked intimately and one can study them both using neurophysiological and behavioural methodologies. In an important study done with primates, Moore and Fallah (2001) stimulated the primates' frontal eye field neurons while the participants were asked to make eye movements to visually presented targets. It was observed that stimulation of these oculomotor neurons also caused shifts in spatial attention. Participants were faster in detecting the object when the object occupied a space within the receptive field of the stimulated oculomotor neuron. Further evidence for a direct role of oculomotor neurons in visual target selection has come from studies with primates where micro-stimulation of superior colliculus neurons have been shown to directly modulate saccade target selection. It has been long known that superior colliculus neurons play a critical role in saccade generation. Carello and Krauzlis (2004) trained monkeys to make saccades towards visual targets that they selected based on colour discrimination. This design was novel in the sense that this included a direct manipulation of intention given the fact that the primates selected their next saccadic target based on the colour feature. The visual targets appeared in the opposite hemi-fields. Micro-stimulation of the superior colliculus neurons resulted in faster selection of the appropriate target and a decrease in reaction time.

These data suggest that saccades necessarily indicate attentional shifts. Eye movements cannot be made without already committing attention to a location. However, attention can move without eye movements. This, of course, flows from the influential pre-motor theory that considers covert spatial attention as a by-product of an oculomotor programme. Covert orienting of spatial attention can be independently studied without the simultaneous initiation of an oculomotor response (Klein 1980; Klein and Pontefract 1994).

Recently, Belopolsky and Theewwes (2012) explored the issue further with a visual search task. It was investigated if covert shifting of attention is always followed by a saccade to that location. Participants were asked to search for numbers among letters that came at different locations on the computer screen. Moreover, participants were asked to 'covertly' search for the number and then make an eye movement towards that location. Each number, i.e. 1, 2, 3 and 4 were linked up with four locations. It was observed that covert shifts of attention were coupled with the activation of a saccadic programme but an eye movement did not always result. However, a very interesting pattern of results emerged when participants were asked to 'maintain' attention covertly at the location. There was suppression of the oculomotor programme. This evidence suggests that shifts of visuo-spatial attention, and target selection is functionally linked with those very neurons that move the eyes in space. Sometimes a random peripheral object attracts attention in a bottom-up manner and we move our eyes towards this object. In this case, the eye movement is not under the volitional control of the subject and indicates a shift of attention, leading to a saccade.

A considerable amount of research in psycholinguistics and related disciplines has been on the issue of language and its interaction with attention and vision (see Ferreira and Henderson 2004; Spivey et al. 2001). This research attempts to integrate findings and theoretical frameworks from vision, perception, attention and language. A commonly used paradigm has been to use spoken language and visual simultaneously and measure scanning behaviour. For example, with a display that contains four objects—the picture of a candle, a candy, a towel and a fish, when the word 'candle' is presented, eyes also move towards the picture of a 'candy' (Allopena et al. 1998; see also Huettig and Altmann 2005). This covert shift of attention and saccades towards an object that is not relevant for the listener indicates attention shift because of incremental processing of phonological information. These eye movements also may reflect consideration of multiple items before the target best matching to the input is selected. Objects consistent with the input attract attention. Others have found that even objects that share the same shape or even colour can attract attention and listeners look at them (Huettig and Altmann 2011). This type of eye movements during search with real-world objects indicates some type of 'unintentional' activation of the oculomotor system. This can also happen even when perceivers have been explicitly asked to search only one object and move their eyes towards it and not to other items (Salverda and Altmann 2011). Language thus mobilizes attentional shifts in the visual world. This is accomplished with the referential functions of language and a close match between linguistic input and visual concepts.

These eye movements reflect emerging attentional bias because of incremental processing of language and visual information (Mishra and Marmalejo-Ramos 2010). Spoken words have been shown to cause a great deal of interference with visual tasks (Salverda and Altmann 2011). Similarly people look at unrelated and task-irrelevant objects when their primary goal is to look at something else (see Mishra 2009 for a review). No one wants to look around apples when you want to buy oranges. It appears that objects can therefore capture attention unintentionally and once attention is captured then there is an obligatorily eye movement. Many so called ‘visual world’ eye tracking studies that present visual displays and spoken words simultaneously exploit such a situation. Interestingly, this covert shift of attention towards a related object is completely unconscious and happens for a brief amount of time (Mishra et al. 2012). This situation is also similar to where the onset of a distractor captures attention and saccades are programmed towards them. It is another matter that once the perceiver realises that the object is not the target he/she quickly corrects the saccade.

At other times, eye movements and therefore shifts in attention towards some location or objects can be unintentional and anticipatory. Eye movements reflect cognitive operations that are ongoing and not yet complete (Spivey 2008). For example, given few objects to look at, people often inspect unrelated objects briefly before making a final saccade towards the target. This transitory aspect of eye movements and looking behaviour reflects how dynamically cognitive states evolve in real time. This means, eye movements can be linked to those stages of evolution of a decision making process. Moreover, eye movements can indicate anticipatory cognitive processing in many situations. We have seen so far that in experimental psychological tasks, participants are often given very fixed instructions to attend or move their eyes to certain locations and they often do a task. Be it the cuing paradigm or even tasks that measure covert shifts of eye movements, participants operate under fixed task instructions and attention shift is then contingent on such task goals. However, this may not be always the case in some situations.

Human agents use their real-world knowledge and also structural knowledge of language to shift attention towards objects that are task relevant. In the first demonstration of ‘anticipatory’ eye movements during spoken sentence processing, Altmann and Kamide (1999) observed that listeners quickly move their attention towards objects that are compatible with the sentence context and even with the grammatical tense of the verbs used. Altmann, Kamide and Heyood (2003) used clipart scenes that depicted a boy, a man, a mug of beer and a cookie. When the sentence ‘the boy will drink/eat ‘was presented participants quickly oriented their attention and looked at the picture of the ‘cookie’ and when the sentence was ‘the man will drink’, immediately with the onset of the verb, participants looked at the beer mug. This suggests that attention shifts during real-world scene viewing is often contingent with real-world knowledge and event structures. People know already what things are and how they are related to real-life experience. Many more studies have ever since investigated these effects and there is now a great deal of consensus that attention shift can be anticipatory in many real-world situations and they can then cause shifts in eye movements. Therefore, a complete understanding of the nature of attention and how it is related to cognitive processing must go beyond viewing it just as some sort of limited resource or a filtering mechanism. It is still not clear why we look at things even when we know that they are not objects representing our immediate goals.

2.10 Conclusion

Considerable theoretical and experimental progress in attention research has shown that attention is an all-pervasive system that aids successful cognition. The chapter presented some of the most important attributes of attention. It is important to note that considerable controversies exist in each of these domains as far as results are concerned. However, some of these attributes will be crucial in understanding its interaction with language. I will end the chapter with another strong position recently echoed by Brit Anderson. Anderson says that "attention never causes anything. Because there is no such thing as attention. There are many empirical findings that can be accurately labelled attentional. In a phrase, attention is more adjectival than nominal" (Anderson 2011, p. 2). This sums up the current confusion in the field of attention as it is. Attention is a process that basically allows us to channelize our limited cognitive resources towards the most immediate issues in the environment (Carrasco 2009). Attentional processes affect the way we survey our natural world and perceive it. Attention modulates our consciousness experience. Even though there are disagreements on what attention is or what processes need it, there is also quite a bit of agreement among researchers about its essential nature. Such as, it is a limited resource; one needs it to process something more accurately and attention shifts to a location without eye movements can enhance perceptual discrimination. There is also currently much work on the relationship between attention and working memory. For example, many believe that to keep something in working memory for a later action is equivalent to 'attending' to it (Engle 2002). The demonstration that manipulating the scope of attention can lead to differential processing of other stimuli is an elegant notion that can be used fruitfully to study language–attention interaction. Psycholinguists have been currently using several of these basic ideas mentioned in this chapter to understand aspects of language processing and also how language interacts with vision and memory to give us our cognition. We will see in the later chapters that attentional issues and their involvement with language processing together with visual processing are becoming quite a popular enterprise lately. Lastly, whatever may be the theoretical characterization of attention, this book will argue that no psycholinguistic modelling can be complete without a serious consideration of this baffling phenomena. This we have already started to see in several models of reading, speaking as well as sentence processing. However, it will be seen later, there is a strong urge among language researchers to view many aspects of language processing occurring 'automatically' and therefore attention may not be necessary.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419–439.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. *Attention and Performance, 14*, 183–218.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *The Quarterly journal of experimental psychology, 24*(2), 225–235.

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology, 2*, 1–8.

Anthony, B. J., & Graham, F. K. (1985). Blink reflex modification by selective attention: Evidence for the modulation of 'automatic' processing. *Biological Psychology, 21*(1), 43–59.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual cognition, 10*(8), 949–963.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A. (2011). Unconscious thought theory and its discontents: A critique of the critiques. *Social Cognition, 29*(6), 629–647.

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research, 49*(10), 1154–1165.

Belopolsky, A. V., & Theeuwes, J. (2012). Updating the premotor theory: The allocation of attention is not always accompanied by saccade preparation. *Journal of Experimental Psychology, 38*, 902–914.

Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology: General, 134*, 207–221. doi:10.1037/0096.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science, 308*(5721), 529–534.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, *9*(2), 46–52. British Journal of Psychology, *92*, 53–78.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Science, 15*, 567–575.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes, 10*(2), 137–167.

Broadbent, D. (1958). *Perception and communication*. London: Pergamon Press.

Broadbent, D. E. (1977). Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology, 29*, 181–201.

Buckner, R. L., Kelley, W. M., & Petersen, S. E. (1999). Frontal cortex contributes to human memory formation. *Nature Neuroscience, 2*(4), 311–314.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4), 523.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science, 315*(5820), 1860–1862.

Bussche, E. V. D., Hughes, G., Humbeeck, N. V., & Reynvoet, B. (2010). The relation between consciousness and attention: An empirical study using the priming paradigm. *Consciousness and Cognition, 19*(1), 86–97.

Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Carello, C. D., & Krauzlis, R. J. (2004). Manipulating intent: Evidence for a causal role of the superior colliculus in target selection. *Neuron, 43*, 575–583.

Carrasco, M. (2009). Cross-modal attention enhances perceived contrast. *Proceedings of the National Academy of Sciences, 106*(52), 22039–22040.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*(7), 349–354.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101.

Cohen, R. A. (2014). Cognitive science of attention: Current concepts and approaches. In The neuropsychology of attention (pp. 55–68). US: Springer.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science, 22*(9), 1165–1172.

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). Response to Tsuchiya et al.: Considering endogenous and exogenous attention. *Trends in Cognitive Sciences, 16*(11), 528.

Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: identical, independent, or overlap-ping neural systems? *Proceedings of the National Academy of Sciences of the United States of America, 95*, 831–838.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Colzato, L. S., Pratt, J., & Hommel, B. (2010). Dopaminergic control of attentional flexibility: Inhibition of return is associated with the dopamine transporter gene (DAT1). *Frontiers in Human Neuroscience, 4*, 53. doi:10.3389/fnhum.2010.00053.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15*(8), 559–564.

De Brigard, F. (2012). The role of attention in conscious recollection. *Frontiers in Psychology, 3*, 29.

De Brigard, F., & Prinz, J. (2010). Attention and consciousness. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(1), 51–59.

De Cesarei, A., & Loftus, G. R. (2011). Global and local vision in natural scene identification. *Psychonomic Bulletin & Review, 18*(5), 840–847.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Deutsch, J., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80–90.

Di Lollo, V. (2012). The feature-binding problem is an ill-posed problem. *Trends Cognitive Science, 16*, 317–321.

Di Lollo, V., Kawahara, J. I., Ghorashi, S. S., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research, 69*(3), 191–200.

Drew, T., Võ, M. L. H., & Wolfe, J. M. (2013). The invisible gorilla strikes again sustained in attentional blindness in expert observers. *Psychological Science, 24*(9), 1848–1853.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*(1), 53–78.

Dukewich, K. R., & Klein, R. M. (2015). Inhibition of return: a phenomenon in search of a definition and a theoretical framework. *Attention, Perception, & Psychophysics*, 1–12.

Duncan, J. (1996). Cooperating brain systems in selective perception and action. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI* (pp. 549–578). Cambridge, MA: MIT Press.

Duran, N. D., Dale, R., & McNamara, D. S. (2010). The action dynamics of overcoming the truth. *Psychonomic Bulletin & Review, 17*(4), 486–491.

Driver, J., & Spence, C. (1998). Cross–modal links in spatial attention. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 353*(1373), 1319–1331.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*(1), 143–149.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage, 26*(2), 471–479.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*(2).

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene?. *Journal of Vision*, *7*(1).

Fei-Fei, L., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 14.

Ferreira, F., & Henderson, J. M. (2004). Introduction to the interface of vision, language, and action. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. (1996). Where in the brain does visual attention select the forest and the trees? *Nature, 382*, 626–628.

Folk, C. L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 847.

Franconeri, S. L., Jonathan, S. V., & Scimeca, J. M. (2010a). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science, 21*(7), 920–925.

Frith, C. (2002). Attention to action and awareness of other minds. *Consciousness and Cognition, 11*(4), 481–487.

Frye, M. A. (2013). Visual Attention: A cell that focuses on one object at a time. *Current Biology, 23*(2), R61–R63.

Franconeri, S. L., Jonathan, S. V., & Scimeca, J. M. (2010b). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science, 21*(7), 920–925.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*(2), 129–135.

Giesbrecht, B., Sy, J., Bundesen, C., & Kyllingsbæk, S. (2014). A new perspective on the perceptual selectivity of attention under load. *Annals of the New York Academy of Sciences, 1316*(1), 71–86.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(02), 67–81.

Han, S. W., & Kim, M. S. (2009). Do the contents of working memory capture attention? Yes, but cognitive control matters. *Journal of Experimental Psychology, Human Perception and Performance*, *35*(5), 1292.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., Mack, M., Fischer, M., Murray, W., & Hill, R. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. *Eye movements: A window on mind and brain*, 537–562.

Hoffman, J. E., & Subramanium, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57*, 787–795.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127*, 398–415.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science, 12*(5), 360–365.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), 23–32.

Huettig, F., & Altmann, G. (2011). Looking at anything that is green when hearing 'frog': How object surface colour and stored object colour knowledge influence language-mediated overt attention. *Quarterly Journal of Experimental Psychology, 64*(1), 122–145. doi:10.1080/17470218.2010.481474.

Intraub, H. (2012). Rethinking visual scene perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(1), 117–127.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on, Pattern Analysis and Machine Intelligence, 20*(11), 1254–1259.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119*(4), 397.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science, 282*(5386), 108–111.

Klein, R. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. M. Klein (Ed.) (2000), Inhibition of return. *Trends in cognitive sciences*, 4(4), 138–147.

Klein, R. M. (2000). Inhibition of return. *Trends in cognitive sciences*, *4*(4), 138–147.

Klein, R. M., & Lawrence, M. A. (2011). The modes and domains of attention. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (2nd ed.). New York: Guilford Press.

Klein, R. M. & Pontefract, A. (1994). Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In C. Umilta & M. Moscovitch (Eds.), *Attention & performance XV: Conscious and unconscious processing* (p. 333–350). Cambridge: MIT Press.

Kim, M. S., & Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. *Perception and Psychophysics, 61*(6), 1009–1023.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, *11*(1), 16–22.

Koch, C., & Tsuchiya, N. (2012). Attention and consciousness: Related yet different. *Trends in Cognitive Sciences, 16*(2), 103–105.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96*, 101–124.

Lamme, V. A. (2004). Separate neural definitions of visual consciousness and visual attention: a case for phenomenal awareness. *Neural Networks, 17*(5), 861–872.

Lamy, D., Tsal, Y., & Egeth, H. E. (2003). Does a salient distractor capture attention early in processing? *Psychonomic Bulletin & Review, 10*(3), 621–629.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451.

Lavie, N. (2000). Selective attention and cognitive control: Dissociating attentional functions through different types of load. In *Attention and performance XVIII* (175–194). Cambridge: MIT Press.

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences, 9*(2), 75–82.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339.

Li, L., Gratton, C., Fabiani, M., & Knight, R. T. (2013). Age-related frontoparietal changes during the control of bottom-up and top-down attention: An ERP study. *Neurobiology of Aging, 34*(2), 477–488.

Lleras, A., Buetti, S., & Mordkoff, J. T. (2013). When do the effects of distractors provide a measure of distractibility. *The Psychology of Learning and Motivation, 59*, 261–316.

Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. In *Proceedings of the National Academy of Sciences* (Vol. 95, pp. 825–830). USA.

Luck, S. J., & Vogel, E. J. (1997). The capacity of visual working memory for features and conjuctions. *Letters to Nature, 390*, 279–281.

Mack, A. (2003). Inattentional blindness looking without seeing. *Current Directions in Psychological Science, 12*(5), 180–184.

Mack, A., & Clarke, J. (2012). Gist perception requires attention. *Visual Cognition, 20*(3), 300–327.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Miller, E. K. (2000). The prefontral cortex and cognitive control. *Nature Reviews Neuroscience, 1* (1), 59–65.

Mishra, R. K. (2009). Interface of language and visual attention: Evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K., Huettig, F., & Olivers, C. N. (2013). Automaticity and conscious decisions during language-mediated eye gaze in the visual world. In N. Srinivasan & V. S. C. Pammi (Eds.), *Progress in brain research: Decision making: Neural and behavioral approaches*. Amsterdam: Elsevier.

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing*,. doi:10.1007/s10339-010-0363-y.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability—evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*, 1–10.

Mishra, R. K., & Singh, J. P. (2012). *Does scope of attention influence sentence processing?*. Riva del Garda: Poster at AMLAP conference.

Moher, J., Lakshmanan, B. M., Egeth, H. E., & Ewen, J. B. (2014). Inhibition drives early feature-based attention. *Psychological Science, 25*(2), 315–324.

Monsell, S., & Driver, J. (2000). *Control of cognitive processes: Attention and performance XVIII*. Cambridge, MA: MIT Press.

Moore, T., & Fallah, M. (2001). Control of eye movements and spatial attention. *Proceedings of the National Academy of Sciences, 98*(3), 1273–1276.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353–383.

Navon, D. (1981). The forest revisited: More on global precedence. *Psychological Research, 43* (1), 1–32.

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica, 114*(3), 273–309.

Neuhoff, J. G., Schott, S. A., Kropf, A. J., & Neuhoff, E. M. (2014). Familiarity, expertise, and change detection: Change deafness is worse in your native language. *Perception, 43*, 219–222.

Neumann, O., Heijden, A. H. C., & Allport, D. A. (1986). Visual selective attention: Introductory remarks. *Psychological Research, 48*(4), 185–188.

Norman, D. A., & Shallice, T. (1986). *Attention to action* (pp. 1–18). US: Springer.

Olivers, C. N., & Humphreys, G. W. (2003). Visual marking inhibits singleton capture. *Cognitive Psychology, 47*(1), 1–42.

Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1243.

Pashler, H. E., & Sutherland, S. (1998). *The psychology of attention* (Vol. 15). Cambridge, MA: MIT press.

Posner, M. I. (1978). *Chronometric explorations of mind* (p. 271). Hillsdale (New Jersey): Lawrence Erlbaum Associates.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences, 91*(16), 7398–7403.

Posner, M. I. (Ed.). (2011). *Cognitive neuroscience of attention*. New York: Guilford Press.

Posner, M. I. (2012). *Attention in a social world*. New York: Oxford University Press.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance* X: *Control of language processes*, 32, 531–556.

Posner, M. I., & Fan, J. (2007). Attention as an organ system. In J. Pomerantz (Ed.), *Neurobiology of perception and communication: From synapse to society. De lange conference iv*. Cambridge, UK: Cambridge University Press.

Posner, M. I., & Petersen S. E. (1990). The attention system of the human brain. *Annual Review of Neurosciences*. *13*, 2542.

Pulvermüller, F. (1995). Agrammatism: behavioral description and neurobiological explanation. *Journal of Cognitive Neuroscience, 7*(2), 165–181.

Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain. *Trends in cognitive Sciences, 7*, 3.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849–860.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372.

Rayner, K. (2009). Eye movements in reading: models and data. *Journal of Eye Movement Research, 2*(5), 1–10.

Rensink, R. A., O'Regan, J. K., & Clark, J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368–373.

Ristic, J., & Kingstone, A. (2012). A new form of human spatial attention: automated symbolic orienting. *Visual Cognition, 20*(3), 244–264.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 231–265). Cambridge, MA: MIT Press.

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*(2), 275.

Roper, Z. J. J., & Vecera, S. P. (2013). Response terminated displays unload selective attention. *Frontiers in Psychology, 4*, 967. doi:10.3389/fpsyg.2013.00967.

Salverda, A. P., & Altmann, G. T. M. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1122–1133.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. detection, search, and attention. *Psychological Review, 84*, 1–66.

Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 54*(1), 1–14.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science, 16*(2), 114–122.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences, 10*(1), 38.

Shapiro, K. L., Caldwell, J., & Sorensen, R. E. (1997). Personal names and the attentional blink: A visual *cocktail party* effect. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 504–514.

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research, 98*(3), 507–522.

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relation-ship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology, 38*, 475–491.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualis: Language and Cognition, 15*(4), 771–781.

Simons, D. J. (2000). Attentional capture and in attentional blindness. *Trends in Cognitive Sciences, 4*(4), 147–155.

Slavich, G. M., & Zimbardo, P. G. (2013). Out of mind, out of sight: Unexpected scene elements frequently go unnoticed until primed. *Current Psychology, 32*(4), 301–317.

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 248.

Spivey, M. (2008). *The continuity of mind* (Vol. 40). USA: Oxford University Press.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science, 12*(4), 282–286.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception and Psychophysics, 49*(1), 83–90.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception and Psychophysics*, *51* (6), 599–606.

Theeuwes, J., Olivers, C. N., & Belopolsky, A. (2010). Stimulus-driven capture and contingent capture. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 872–881.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1595–1608.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology, 37*(4), 571–590.

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior, 3*(6), 449–459.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology, 18*(03), 321–342.

Ward, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 562–581.

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology, 30*(1), 79–109.

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex, 17*(10), 2322–2333.

Wittgenstein, L. (2010). *Philosophical investigations*. Oxford: Wiley-Blackwell.

Wolfe, J. M. (1994). Visual search in continuous, naturalistic stimuli. *Vision Research, 34*(9), 1187–1195.

Wolfe, J. M. (2012). The binding problem lives on: comment on Di Lollo. *Trends in Cognitive Sciences, 16*(6), 307.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*, 495–501.

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 363.

Wu, W. (2011). Confronting many-many problems: Attention and agentive control. *Nous, 45*, 50–76. doi:10.1111/j.1468-0068.2010.00804.x.

Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 861.

Yang, G., Dempere-Macro, L., Hu, X., & Rowe, A. (2002). Visual search: Psychophysical models and practical applications. *Image and Vision Computiong, 20*, 291–305.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 676–681.

Yarbus, L. (1967). *Eye movements and vision*. New York: Plenum Press.

Yeshurun, Y., & Marciano, H. (2013). Degraded stimulus visibility and the effects of perceptual load on distractor interference. *Frontiers in Psychology, 4*, 289. doi:10.3389/fpsyg.2013.00289.

Zhou, H., & Desimone, R. (2011). Feature-based attention in the frontal eye field and area V4 during visual search. *Neuron, 70*(6), 1205–1217.

Chapter 3
Attention in Language: Historical Perspectives

3.1 Wundt and the Beginnings of Experimental Psychology

It is important to draw a sketch of the history of psychology, particularly the development of experimental psychology, which had a formidable influence on later research on language, attention and vision. However, I will deal with this briefly since many others have dealt with this theme extensively (e.g., Mandler 2007; Leahey 1987). In this book, we will explore experimental investigation of language and attention. Therefore, it is important to take a look at the historical context in which empiricism evolved in psychology. The earliest developments in this direction were made with the publication of Wilhelm Wundt's (1832–1920) *Principles of Physiological Psychology* and Franz Brentano's (1838–1917) *Psychology from an Empirical Standpoint*, both in 1874. These two books indicate a radical turn in the basic vision of psychology as the science of mental events in different ways (Crane 2014). Brentano was a philosopher while Wundt was a psychologist. Wundt established empirical methods for doing psychology while Brentano wrote philosophical treatises on the essential object of psychological studies. Wundt's student Titchener (1921) notes the similarities and differences between the thoughts of Brentano and Wundt, the two most important precursors of modern experimental psychology.

Among other things, Brentano is known to have introduced the concept of 'intentionality' into modern psychology (Brentano 1874). For Brentano, mental states were intentional states. That is, every mental state has 'aboutness' in it. Every thought is about something. Apart from his conceptual contribution on the topic of intentionality, Brentano was also deeply interested in the issues related to the disciplinary boundary within which psychology should function. Brentano held the view that psychology was distinct from physiology. At that time it was not clear if psychological objects could be studied empirically without resorting to introspection. In his *Psychology from an Empirical Standpoint* (Brentano 1874), Brentano

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_3

called for a naturalistic investigative path for psychological phenomena (see also Fancher 1977). Thus, mental and psychic phenomena were to be investigated as natural phenomena just like in the natural sciences. However, Brentano explicitly believed that the aim of psychology is to examine the experiential nature of mental phenomena. He also indicated the unitary role of consciousness in all mental phenomena (Benjamin Jr. 2007). He suggested that psychology should study mental processes that are conscious and are experienced by agents. While Brentano's ideas were philosophical in nature, they had a profound influence on later psychologists.

Wundt has been hailed as the father of modern experimental psychology. The first ideas relating language to psychological processing and human cognition go back to him. Wundt established the methods for studying human psychological processes such as attention, perception, memory and human performance. He considered both physiological and psychological methods of examination to proceed in parallel. Importantly, Wundt emphasized the pivotal role of experiments in the conformation and replication of human psychological processes, knowing fully well that these processes have subjective bases and are often physically untraceable (Wundt 1907). However, Wundt did not think that much experimental work could happen in the studies of human linguistic behaviour (Blumenthal 1970; Kess 1981). Many consider Wundt's influence on psycholinguistics as vital (Levelt 2012). Wundt had suggested a role of executive function on language processing (see Reber 1987). Wundt viewed linguistic expressions as more related to rules and intentions of speakers and less with the physical manifestation of speech. Wundt recognized the role of 'selective voluntary attention' in human cognition which is now referred to as top-down attention and cognitive control (Blumenthal 1975). Wundt had used the word 'appreciation' to refer to volitional attention which he thought is controlled by the frontal lobe of the human brain (Blumenthal 1970). He believed that while human psychological processes can be fruitfully studied using the methods of physical sciences, at the core, human psychological processes remain governed by the goals and intentions of an active agent (Danziger 2001a, b). Today, this remains a very strong assumption of action sciences. Wundt theorized on many aspects of linguistic behaviour and influenced such scholars like William James (1842–1910) and Ferdinand de Saussure (1857–1913).

Many contemporary psycholinguists view language processing to be a goal-driven process. Humans select and act on things in the world armed with knowledge and certain values. That is, top-down influences affect language that is used to describe thoughts (see Roelofs 2007). As for language, Wundt thought that generativity in language use comes out of the activities of attention (Blumenthal 1975). Generativity refers to our ability to produce an infinite number of sentences, sometimes even completely new ones. Wundt did not consider a sentence as merely a collection of grammatically appropriate strings of words but as a structure arising from our volitional systems. He elaborated on the idea that human language activity has in its roots an executive control system that governs several of its dimensions. There is currently a resurrection of Wundtian ideas of attention in language processing at least in some domains of psycholinguistics (Roelofs 2007). Wundt never did experiments on his ideas related to attention and language. But his ideas and references to volitional control and psychological processes, and in particular human linguistic behaviour, has had its desired effect on generations of linguists and psycholinguists. For example, the American structural linguist Leonard Bloomfield (1887–1949) showed appreciation for Wundtian psychology in his first book on language (Bloomfield 1933). In his *Die Sprache* (Wundt 1900), Wundt was particularly explicit about the mental structures that were key to learning and the of language. Blumenthal writes of Wundt:

> The fundamental unit of Wundtian linguistics is the sentence, which he defined in mentalistic terms and which is prior to words and word orderings. The particular form that spoken sentences take is a result, in Wundt's system, of universal characteristics of human mental processes, namely the temporal and information handling constraints on selective attention and immediate memory. (Blumenthal 1987, p. 317)

Thus, language functions for Wundt arose from basic cognitive capabilities (see Blumenthal (1987) for a fuller description). This explicit characterization of sentence processing in terms of attentional and memory structures is something that has influenced generations of cognitively minded psycholinguists ever since.

It has been noted that many important psycholinguistic frameworks that originated during the second half of the twentieth century were influenced by Wundt's conceptualizations (see McCauley 1987). Wundt's theories of sentences and linguistic behaviour were much more explicit and were given a psychological explanation. He talked about general cognitive structures that are at the heart of linguistic perception. In summary, one sees an early interest in looking at human linguistic ability in harmony with other cognitive systems in the work of Wundt.

3.2 Behaviourism and Cognitive Psychology

Wundt's blend of introspections and empiricism did not culminate in a dominant framework for psychology soon after. In America, it was soon replaced by a radical mechanistic approach known as behaviourism. Mandler (2007) discusses the chronology of the development of psychological thought from the days of Wundt, Brentano and James through behaviourism and finally the appearance of cognitive psychology. Behaviourism was the dominant psychological paradigm that influenced all experimental work in psychology and remained so till the 1950s. John B. Watson (1913) was instrumental in creating the new framework of behaviourism in psychology on the lines of natural sciences. For Watson, psychology should examine human behaviour in all its complexity using scientific methods similar to the natural sciences (Watson 1913). It is important to note that behaviourism explained all human actions as stimulus response. Gilbert Ryle's famous book *The Concept of Mind* (1949) finely captures the sentiments of this time. To the behaviourist, there was no such thing as a mental state above and beyond observable behaviour, which was easily explained as responses to stimuli. Thus, behaviourism in psychology was rigorous in its methods but did not focus on the study of mental processes which Wundt and Brentano had prescribed. The behaviourist approach to the study of behaviour in animals was more strongly enforced by Skinner (1953), who analyzed behaviour in organisms as systematic responses to stimuli and scheduled reinforcements. This behaviour towards stimulus did not arise as a result of any conscious states or motivations (Skinner 1956). Skinner also produced an account of language behaviour in humans following the frameworks of stimulus and response (Skinner 1957). Noam Chomsky's scathing critique of Skinner's book, *Verbal Behavior* (Skinner 1957), is now well regarded as a turning point in the history of cognitive science and the modern discipline of linguistics (Chomsky 1959). Chomsky highlighted the enormous creativity that is seen in child language acquisition without much external teaching and instruction; whereas, for Skinner, all forms of linguistic utterances were mere responses towards learnt stimuli. Chomsky's critique is largely known to have put behaviourism to an end as far as the analysis of mental events was considered (Stemmer 1990).

However, it would be absurd to conclude that all psychologists around the world during the first half of the twentieth century—the heyday of behaviourism—were blind towards the study of the mind and its intractable mental states. More recent historical analysis of this period suggests that psychologists such as Jean Piaget (1896–1980) in Switzerland and Lev Vygotsky (1896–1934) in Russia were more cognitively inclined and pursued interdisciplinary studies into mental phenomena (Vauclair and Perret 2003). Piaget studied concept development in young children and created a theory which systematically recorded its different stages (Piaget and Inhelder 1969). Vygotsky (1978) examined the development of mental skills as an interaction between the individual and different cultural and societal constraints (Frawley 1997). Therefore, it is possible to see roots of cognitive approaches to the study of the mind in several European centres of learning during the heyday of behaviourism in America. It is also important to mention that although behaviourists did not pay attention to the mind, they certainly turned introspectionist psychology into a hard, natural science with objective methods. This has had important influence even on later work when psychologists studied the mind but using empirical methods (Mandler 2007).

By the end of 1950s, psychology had thrown the yoke of behaviourism and was trying to look at the mind as an information processing device. A collective revolution in intellectual thought took place at this time, popularly known as the *cognitive revolution*.[1] A multi-disciplinary and concerted effort was launched to study the mind. Neisser (1967) published his important book on cognitive psychology, which suggests that cognitive psychology aimed to study important mental processes such as attention, memory, visual perception and the psychology of language. Cognitive psychologists also focused on higher processes such as consciousness and human decision making. They applied the rigour of experimental methods, studying seemingly intractable mental processes. They also aimed to explain the laws that governed human mental behaviour and studied motivation and

[1]Significant historical details of this time can be found in Gardner (2008) and in Boden (2006).

the affective bases of human cognition. Foundational discoveries were made on the nature of selective attention (Broadbent 1958) and visual short-term memory (Sperling 1960). These early experimental demonstrations of components of human cognition had lasting influence on later researchers. More fundamental discoveries followed on the nature of the orientation of attention by Posner (Posner and Boies 1971; Posner 1980). Atkinson and Shiffrin (1968) proposed their model of the human memory system with three distinct components. This model proposed that initially all stimuli is initially available as sensory memory. Only information that is attended to enters into short-term memory and eventually into long-term memory. This radical shift in the outlook of psychology led to a new era in experimentation into the nature of human cognition.

3.3 Structuralism in Linguistics

As psychology at the turn of the twentieth century moved from philosophical speculation to experimentation, important changes took place in linguistics. If Brentano had emphasized that the object of psychology was to study mental phenomena, linguists focused on the structure of language and its description. Much work to be described later in this book deals with psychological processing of linguistic structure. The singular effect of Ferdinand de Saussure on the structuralist tradition of modern linguistics cannot be overemphasized. Saussure developed his theory of language around the concept of signs that moved away from psychological influences (Koerner 1973). Saussure's lectures, posthumously collected and published by his students as *Course in General Linguistics* (1917), is seen as a seminal text in linguistic theory. For Saussure, language is an intricate web of signs. A linguistic sign is composed of a signifier and a signified. The Saussurian concept of 'signified' was conceptual and psychological (de Saussure and Baskin 2011). Linguistic signs are also arbitrary. For example, the animal English speakers refer to as 'horse' has no resemblance to the animal, so it could have been any other word. That is, the concept described by the word 'horse' has no connection with it. So, language users refer to concepts through signs. Saussure also introduced the important dichotomy of '*langue*' and '*parole*' (Chapman and Routledge 2005). He described *langue* as the mental and social construct of linguistic signs that all the speakers have internalized as rules and systems, whereas *parole* is the spoken manifestation of such internal rules in speakers. This distinction has been seminal in the psychological studies of language, as psycholinguists study verbal behaviour in order to understand the underlying mental representation of language.

Saussure's influential theory of language led to the growth of structural linguistics as a framework which dominated the scientific study of language till the 1950s. In this period, Bloomfield published his important epic *Language* as a manifesto of structural linguistics (1933). Bloomfield considered the description of language following structural rules as pivotal to the study of language. It is important to note here that he had attended the lectures of Wundt at the University of Leipzig (Blumenthal 1987), but his theory of language was not mentalist. Rather, Bloomfield subscribed to the then dominant behaviourist framework of psychology (see Levelt 1972; de Lourdes and Passos 2007). However, Bloomfield's framework of structural linguistics dominated American linguistics till the 1960s and even later (Robins 2013). Thus, while during this period the emphasis was on the description of language, linguists did not show much concern about its psychological basis. Levelt (1972) notes that during the period through which structural linguistics flourished in America, the interaction of linguistics and psychology was non-existent and this continued till the 1950s. Thus, although many linguists were exposed to developments in Psychology not all included psychological methods in their analysis of language behaviour.

3.4 The Chomskyan Turn

Many think that modern linguistics began with the publication of *Syntactic Structures* by Noam Chomsky (Chomsky 1957). Chomsky's approach was very radical given that he was trained in the structuralist tradition (see Barsky 2011). To Chomsky, the scientific study of language involved not simply describing languages but to exploring *how* humans acquire such complex knowledge. His questions concerned the nature of language, its relation to the human mind and how language is acquired (Chomsky 1972). He also emphasized that all human languages may differ on the surface but essentially follow the same laws of structural computation (Newmeyer 1997). His approach to language, known as 'generative grammar', discussed the mental computations that human beings intuitively use to acquire the knowledge of language o and its innateness. That is, such ability is part of the biological make up of humans and no extra effort is needed to learn a language. More fundamentally, Chomsky brought mentalism into the study of language which was absent during the preceding decades (Chomsky 1966a, b). Needless to say that the Chomskyan approach to linguistics provoked a wide range of research activities in psycholinguistics and other fields. Here, by Chomskyan linguistics, we mean the dominant approach of transformational generative grammar (Chomsky 1957). Critical assumptions of Chomsky's thoughts, like the biological basis of language, nativism and rationalist thinking on the question of the mind influenced psycholinguistic work. However, at this time linguists did not want to consider the psychological implications of their theories, and cognitive psychologists were wary of linguistic theories.

Chomsky recommended linguistics to be a branch of cognitive psychology (Chomsky 1972). However, this did not mean that linguists who worked with a Chomskyan framework paid explicit attention to psychological theories. Chomsky projected language as a universal and innate symbolic structure that is an integral part of the human mind. However, he was not sure if all linguistic theories were to be developed keeping psychological facts in mind. Chomsky sought clear autonomy for syntax and made it supreme in the explanation of the human mind. Many workers considered psychology and linguistics as separate disciplines and their union impossible, or even undesirable. Psychologists who took up studies on language examined how linguistic theories could be tracked down to empirical demonstrations (see Blumenthal 1970). While Chomsky consistently asked for linguistics to be understood as a larger science of psychology and biology, empirical researchers could not find enough theoretical links (see Seuren et al. 1998; and also Levelt 2012). Different phases of the modern history of psycholinguistics have shown the uneasy relationship between psychological and linguistic theories (Kess 1992).

It is important to note that Chomsky had heavily emphasized 'competence' and thought that investigations of performance—i.e. actual language behaviour—were valueless (Newmeyer 1997). Cognitive psychologists were investigating performance and explored information processing under carefully controlled empirical experiments (Neisser 1976). The link that was still tying linguists and psychologists was the information processing approach to the study of mental functions. Chomsky's mathematical characterization and strong reference to biology and human cognition attracted workers outside of linguistics to research on language. Many followers of the generative enterprise believed that studying the structure of language could therefore say something about our core mental faculties (Chomsky 1986). That is, if language acquisition is an innate mechanism and its use is a creative process, understanding its psychological basis was thought to be a very fruitful research exercise.

3.4.1 Early Psycholinguistics

A significant body of psycholinguistic work triggered by Chomsky's approaches to language in the 1960s and 1970s was driven to seek empirical support for various proposals of the generative grammar framework. Psychologists wanted to explore if syntactic facts as proposed by linguists were psychologically real. This could only be done by experimentally investigating human performance on different tasks. Psycholinguists conducted experiments to explore psychological evidence for the idea that syntax processes operated at deep and surface structure levels across languages, which had processing consequences. For example, Ford and Holmes (1978) found that during sentence production, speakers took longer to respond to a tone when the tone appeared at the end of deep structure clauses rather than surface structure clauses. This evidence suggested that, as predicted by transformational generative grammar, speakers were conscious of the two types of syntactic structures during language production. In click displacement experiments, it was found that listeners tend to report clicks at the junction of sentence clausal boundaries even though physically the clicks were somewhere else (Fodor and Bever 1965). In these studies, participants listen to a sentence and also some click sounds that appear randomly. The participants are asked to report the location of clicks in the sentence. The assumption was that syntactic processing would psychologically affect the location of perception of clicks. However, Abrams and Bever (1969) did not find any difference in reaction times to clicks around the clause endings as this study did not ask participants to report where they had perceived the clicks. This also demonstrates an early effort to use a range of psychological experimental techniques to explore very fine-grained aspects of language processing as a function of psychological performance, in the form of responses in judgement or reporting tasks.

These studies clearly show a very serious psycholinguistic enterprise around syntactic theory and also tremendous developments in methodologies and techniques used in psychological investigations of language. Taking stock of significant episodes in the history of psycholinguistics, Dell and Chang recently (2014) noted that during the so-called classical period of psycholinguistics (1960–1970), study of language production was done in consonance with language comprehension. Studies carried out during this period used psychological methods and performance was measured. However, this changed later and performance did not count as a valid indicator of linguistic reality, and it was not what linguistic theory emphasized (Dell and Chang 2014). This shows a gradual change within psycholinguistics in the last few decades of the twentieth century when issues related to modularity of different linguistic processes as well as the validity of performance data became important.

The fall of structuralism in linguistics and the rise of Chomskyan linguistics were also parallel to the rise of cognitive psychological theories and the information processing approach to cognition in psychology (Mandler 2007). The idea that language is computationally and biologically grounded was well received by both linguists and psychologists who were participating in the cognitive revolution (Chomsky 2009), during which many cognitively inclined psychologists, linguists, computer scientists and philosophers began studying the nature of mind (Miller 2003). However, unlike linguists, cognitive psychologists were empirical workers and did experiments to understand human cognitive behaviour (Newmeyer 1997). Linguists did not feel accountable to psychologists and even if they collaborated, this did not lead to significant cross-disciplinary research (Miller 1990). The important task for cognitive psychologists was to know how intuition was psychologically grounded, or how competence and performance interacted and constrained one another (Harman 1967). Some psychologists took the study of language very seriously and developed cognitive models (Gentner and Stevens 1983). Developments in the studies of attention, memory, vision and artificial intelligence had little or no effects on linguists who were studying abstract structures of sentences and intuitions of native speakers.

In spite of the general thinking that the mind was to be conceptualized as an information processing device that worked on abstract algorithms (Pylyshyn 1984), linguists and psychologists were wary of one another for a long time. Inner confusions in linguistics and psychology also played a role in their unsuccessful collaboration. Much of this sentiment was captured by Roeper when he wrote, 'when psychological evidence has failed to conform to linguistic theory, psychologists have concluded that linguistic theory was wrong, while linguists have concluded that psychological theory was irrelevant' (Roeper 1982, cited in Newmeyer 1983, p. 45). Language remained outside the discourse of all psychological work and discoveries on phenomena like attention, visual perception as research on cognitive architecture went on. George Miller wrote:

> It is certainly true that grammarians are more interested in what could be said than in what people actually say, which irritates psychologists, and that psychologists insist on supplementing intuition with objective evidence, which irritates linguists. But these are precisely the differences that patient collaboration should over-come. (Miller 1990, p. 321)

Miller traces the mismatch between psychologists and linguists to their very different ways of theorizing about the same entity. While linguists love to talk of abstract and innate laws governing language behaviour, psychologists seek an objective and observer-independent grounding of these laws. Linguists remained indifferent to psychology or one reason or another (Gerald et al. 1985; cited in Miller 1990).

It was not clear to psychologists interested in language, how linguists could account for their theories in terms of human cognitive and perceptual data. Some psycholinguists argued that linguistic structure reflected our general cognitive tendencies (Bever 1970). Bever asked a pertinent question: 'How does the instinct to communicate integrate the distinct components of perception, cognition and motor behaviour into human language?' (p. 4). Language behaviour in its formative stages in infants can be seen as an outcome of several sensory, perceptual and cognitive attainments. Language is, therefore, an outcome of these very processes that make us intelligent and thinking beings. Infants bring in several cognitive strategies to learn language and these strategies are independent of language (MacNamara 1972).

3.4.2 A Psychological Explanation of Language?

Psychological theorizing on human linguistic behaviour has come a long way. It has been accepted that if linguistic theories are statements about human mental intuitions, then they must conform to psychological attributes (see Soames 1984 for an alternative view). P.N. Johnson-Laird aptly asked, 'Why is it that arguments about the psychology of language based on linguistic considerations have tended to outweigh arguments based on psychological findings?' (Johnson-Laird 1984, p. 75). Psychological theories are used to explain language learning (Gleitman 1990), thinking for speaking (Slobin 1996), attentional and working memory constraints in sentence processing (Just and Carpenter 1992), and sentence production (Bock 1982), as well as word naming (Levelt 1989). The emergence of new brain imaging tools like *event-related potentials* have grounded psycholinguistic theories in brain dynamics (Kutas and Van Petten 1994). Currently, syntactic theories are being compared to human action systems (Moro 2014). Several empirical works show how vision and attention affect language processing (e.g., Henderson and Ferriera 2004).

It is also important to accept that psychologists have not found consensus within the field about mental phenomena like attention, memory and perceptual processes and how perception is linked to cognition (Carlson and Tanenhaus 1988). For example, a serious and widely used concept like ‘attention’ is now under revision and even sometimes questioned (Anderson 2011). This has created problems in grounding research questions linking language to attentional mechanism. Even very recent philosophical and psychological work on ‘attention’ shows doubts on its mechanism (Mole 2013). Unless one knows what constitutes attention and what activities manifest it, it is difficult to examine such processes in linguistic behaviour, which is even more abstract and elusive. Christopher Mole (2013) in his examination of attention suggests that attention is just a special mode of processing or acting, given many more alternatives. That is, attention is not to be thought of as a special brain mechanism but is linked to the agent’s understanding of the task at hand and how s/he wants to go about it. There is no dominant version of a psychological theory which the linguist will consult and bring his/her data to account for. Similarly, within linguistics, even in recent times there have been doubts about the use of psychological methods in the explanation of competence of the native speaker, since linguists still assume that intuitions about the language do not depend on performance data. For example, psycholinguists who study sentence processing collect grammatical judgement data, which shows this mind-set. Recently, there have been many discussions on the issue of methods of data collections in linguistics and their suitability for theorizing about the mind and language connection. Gibson and Fedorenko (2010) have claimed that traditional methods of data collection by linguists could be faulty and therefore one should be sceptical of linguistic theories that have been derived out of such data. The allegation was that linguists do not use verifiable and objective methods as used by experimental psychologists. This was a direct attack on a strongly held view that native speakers’ judgements, which lead to glimpse of the linguistic algorithms in the mind, are sacrosanct. However, this was challenged by Sprouse and Almeida (2012), who took data from a linguistic textbook and re-analyzed data from another linguistic article to show that the methods used by linguists, such as eliciting native speakers’ judgements, were good enough and comparable to other methods used in psychology. What this shows is that even now there is some struggle in brining language completely into the empirical/experimental fold.

3.5 The Issue of Modularity

Although the Fodorian notion of ‘modularity’ is not a hot favourite among psycholinguists these days (Henderson and Ferreira 2004; see also Burnston and Cohen, in press),it is still important to consider debates around the notion of modularity and encapsulation of mental operations, to appreciate how language and attention might interact. The modularity thesis projected cognitive systems such as language and vision as separate modules that had their own mandatory functioning and specific inputs. Modules were informationally encapsulated and thus were not vulnerable to outside influence. Further, modules had fixed neural architecture in the brain. Fodor's views of modularity and its various interpretations (Fodor 1983) fostered the 'autonomous' linguistics idea and processing. Modules were encapsulated from outside influence: for example, the visual system (Fodor and Pylyshyn 1988) demonstrates autonomy of early vision. A modular architecture of the mind was supportive of the information processing view of the mind. That is because if different specialized modules handle their own cognitive operations, then it makes the computational demands easier. An offshoot of the modularity notion was also to deny any type of outside influences or 'cognitive penetration' of the workings of core modules. For example, Fodor considered that the influence of top-down information on phoneme restoration was an example of top-down influence but was not an example of cognitive penetration of the language faculty (Coltheart 1999). This top-down influence was 'within' the language module and therefore these examples were not evidence against modularity, which meant that language can be seen as comprising several modules and each module has top-down influences affecting it. Fodor also proposed that information processing inside modules could be only bottom-up and run in mandatory fashion (Fodor 1983). These modules did not have any access to information computed outside them (see Appelbaum 1998). Under these views, it was logical to assume that attention and the executive control system or even visual perception could not affect what happened in the domain of language.

One main aim of this book is to break the myth that language is 'cognitively impenetrable' and operates autonomously. It is now known widely that information content of modules influence one another during processing. Different linguistic levels influence one another apart from the fact that language as a whole is influenced by other cognitive processes like attention and memory. For example, Altmann and Steadmann (1988) showed that semantic information affects syntactic processing. Data from child language acquisition shows that modules do not operate in isolation. It was also shown that action related to visual information acquired through vision could influence how verbs are learnt (Gleitman 1990). Psychological investigations of generative grammar showed interactionism (Fodor et al. 1974). We know now that every aspect of language processing is an outcome of both top-down and bottom-up processes that include information acquired from vision (Spivey 2007). Human language users use knowledge external to language to construct goals and action, and this plays a crucial role in language processing. For a long time, the received view was that sentence processing is 'automatic' which is to say that it does not depend on the volitional acts of the language user (Anderson 2009). Today, this climate has changed and we know the enormous importance context and semantics play on parsing sentences (see Altmann 2013 for a review). Action-related events in visual scenes influence comprehension of language (Knoeferle and Crocker 2006). Anticipatory eye movements towards probable visual entities indicate use of prediction during language–vision interaction (Mishra 2009; Altmann and Kamide 2007). In the domain of sentence processing, semantic information and pragmatic knowledge seem to affect early parsing (Altmann and

Steedmann 1988). These examples show that language is cognitively penetrable and the Fodor and Pylyshin (1988) type formulations for vision may not work for language.

In sum, this brief and somewhat scattered historical analysis suggest that psychological analysis of language phenomena in terms of human performance as an academic discipline has taken many decades of collaboration to evolve. Development in cognitive psychology and other areas of cognitive science have led to increased attention to language as a cognitive process. In the last few decades, several alternative and novel approaches have been developed that ground language to its cognitive basis. The following section is an attempt to see how these developments have offered a conducive approach to studying language and attention interaction. Both inside psycholinguistics and outside it, language processes are being viewed with regard to concepts like attention, memory and executive control with questions ranging from child language acquisition to adult language processing and language disorder.

3.6 The Fall of the Classical Cognitive Science and Psycholinguistics Research

3.6.1 Embodiment

In this section, an attempt will be made to show how several other alternative theoretical frameworks have looked into the question of language and its relationship with important cognitive processes and particularly the motor system. Recently, several authors have argued for the influence of higher order control on language processing (Federnko and Thompson-Schill 2014). Frameworks like embodied cognitive science (Shapiro 2011) and sensorimotor theories of language and cognitive processes (Pulvermuller and Fadiga 2010) have made the study of language–attention more informative. Embodied approaches refer to perception and action without including any intermediate symbol 'transduction' or representation. Embodiment also refers to the idea that cognition is an outcome of the interaction of the whole organism (body) with its environment and not just the mind. Many view the rise of embodiment a radical shift in cognitive science from the classical framework that considered mental operations akin to a digital computer doing symbol manipulations, (Gibbs 2006) where symbols represented sensory information. These views ground language and cognition in the environment of the cognitive agent (Noe 2004). Embodiment is a general theory which assumes a central role of the body as opposed to the mind in cognition (Clark 1999). Sensorimotor events and the overall environment affect the way cognition works. The sensorimotor view considers processing of language in consonance with simulation of motor system. It has been observed that when one listens to action verbs it leads to the activation of the fronto-parietal motor cortex (Tettamanti and Weniger 2006).

One thread of this view has led to several demonstrations of mental simulation in language processing (Kaschak et al. 2005). Understanding language involves mental simulation of events and experiences that the language describes, thus involving memory, motoric activation and action systems in the brain (Zwaan and Taylor 2006). These views go directly against the Fodor and Pylyshyn (1988) thesis of encapsulation. If language faculty is a module and its workings are encapsulated, then the memory and action systems would not affect comprehension. Experimental demonstrations suggest that even eye movements during visual perception reflect mental retrieval of memory traces from the environment (Spivey 2007). A theory of language constructed on these norms allows one to examine the effects of higher control systems like attention and memory as well as multi-modal interactions.

It is evident that the above discussed views do not consider cognitive processes as encapsulated symbol manipulations away from the experience of the user.

In the domain of visual perception, Gibson's psychology of direct perception (Gibson 1986) led to the understanding that human perception need not have symbolic and a modal computation as an intermediate stage between perception and cognition. Gibson proposed a powerful theory of vision which emphasized the role of visual environment and affordances. That is, the way objects are in the environment affect the way we perceive them. This was in contrast to David Marr's theory of vision which was a computational theory of vision. While Marr had emphasized on various computational stages in visual perception (Marr 1982), Gibson's account was based on ideas of 'affordances' and optic arrays in the visual system. Properties of objects in the world are picked up directly by the perceiver's visual system without the intervening symbolic computation. This stance against symbolic computation in vision has influenced the emergence of the sensorimotor approach to vision (O'Regan and Noe 2001). Affordances relate to the way objects constrain the manner in which we interact and perceive them. For example, ERP studies have shown that within 100–200 m of the presentation of an object, brain waves indicate action preparation much before an actual response is made (Leuthold et al. 2004). That is, the motor system prepares itself for action considering what the object to be manipulated is like. It is important to note that we cannot do things with objects if our physiology does not support such actions. Therefore, the way we see and cognize objects is dependent on how they lend themselves to us. For example, the handle of my cup affords my holding it in some manner, which is independent of me. Because of this even a Chimpanzee can hold and lift this cup, although both he and I will get different sensorimotor feedback. Affordances, thus, can be conceived of as what the organism can do with the available features of the environment (Chemero 2003).The idea that human cognition is not inside the head but outside has opened up new ways to look at language and cognition interaction (see Gentner 2010; see also Wilson 1993). Sensorimotor learning and the constant interaction of the whole organism with various aspects of the environment leads to knowledge (Clark 1999; Wilson 2002; also see Mahon and Caramazza 2008). Current views indicate that social and contextual information influence attention in human cognition (Posner 2012).

Human agents act cognitively perceiving the context and changes in the world around them. Thinking beings are situationally aware of their environment and their action is contingent on these events (Anderson 2003). Although the idea of embodied cognition looks appealing, embodied accounts of sensorimotor activations need to be accepted with some caution (Markman and Brendl 2005).

The rise of embodiment as an alternative theory of the mind and human cognitive processing is increasingly becoming popular in both psycholinguistics and cognitive psychology (see Varela et al. 1991). Embodied theories predict a direct link between language and perceptual processes via sensorimotor activations (Fischer and Zwaan 2008; Mishra and Marmalejo-Ramos 2010). This includes roles for motor system, the attentional system and vision. Mental simulation during language comprehension has been shown to affect visual scanning and attentional processes (Matlock 2004; Mishra and Singh 2010). Simulation refers to the re-enactment of perceptual experience (Chemero 2009). It may also include recreating the visual depiction of the actual event. This simulation is rich in sensorimotor activation of events and therefore, any later motor movement compatible with the action described will be faster. In contrast to this, when the motor movement to be executed is different from the action and orientation perceived in the sentence, people are slower (see Borreggine and Kaschak 2006). However, recent replication studies have raised questions on a strict version of the mental simulation account of language perception (Rommers et al. 2013).

It is not the aim here to indicate that sensorimotor theories of cognition stand a better chance of explaining the language–attention interaction compared to the classical approach or they have completely replaced the classical approaches. But to suggest the range of frameworks and theories that one can look for when one wants to look at language and attention. Both extreme forms of symbolic computational cognitive sciences and embodied simulation-based accounts can be problematic. If the allegation against the 'symbolists' has been that these models cannot explain how symbols refer to things, then purely embodied [anti-representational] accounts need some type of symbol manipulation for meaning creation. Therefore, both views of cognition have their strengths and limitations. Stephen Levinson has complained that classical cognitive science went wrong because it constructed theories around an 'ideal' human being and emphasized uniformity (Levinson 2012). Human cognition cannot be uniform. The diversity in cognition arises from linguistic and cultural, as well as biological factors. Both psychologists and psycholinguists have to account for how things outside the brain affect things inside. In this view, the study of attention, language and vision must also be pursued, keeping their diversity in mind.

3.6.2 Emergentism

It is natural to wonder how the interaction of attention with language began. In order to understand how complex systems come about and interact it is important to know something about how they evolve in the first place. As opposed to nativism that proposes language arises as an innate structure in the human species, emergentism is a theoretical framework which considers that a complex system like language evolves from the interaction of several other systems like perception, memory, biology and also social systems (O'Grady 2008). Researchers who work within this approach look at the evolution of cognitive and communication mechanisms in both complex biological and artificial systems. A serious advocate of this programme, MacWhinney wrote, 'the goal of emergentism is the construction of models that avoid stipulation regarding specific hard-wired neural activity. In the place of stipulation, emergentism provides accounts in which structures emerge from the interaction of known processes' (2013, p. x). The core assumption is that complex behaviour arises out of statistical learning and neurons know this through their mutual activation states. Computer simulations with simple neurons have shown the emergence of complex grammatical knowledge (Christiansen and Chater 1994). Thus, language behaviour is an outcome of learning and not innate properties of the brain.

While this could be the case, it is important to know what emergentists have to say about the cognitive processes themselves which interact with language. Is it possible that attention and visual perception were crucial evolutionary systems that had an important causal role on the emergence of language? Available evidence suggests that attention glued together diverse perceptions and made possible the emergence of complex symbolic structures like sentences. For example, the development of declarative pointing gestures is thought to have a key role in development of communication systems (Tomasello and Camaioni 1997). Pointing in space is aimed at expression of intention to another participant and is central to joint attention. Pointing certainly arises out of attentional selection of the object. Linguistic elements like demonstratives are often used to point at things (Levinson 2004). It is conceivable that the development of such intentional gestures and language used attentional resources. It seems likely that attention must have played a role here in reinforcing these diverse perceptions of events, objects and actions.

Another area that shows the influence of central resources on language development is working memory. Working memory plays an important role in both attention and executive control processes (Engles 2002). Working memory consists of the central executive, the phonological loop and the visuo-spatial sketch pad as its important components (Baddeley and Hitch 1974; see also Baddeley 2003 for recent discussions). The central executive monitors information flow into systems. The phonological loop maintains verbal information through rehearsal, while the visuo-spatial sketchpad deals with images. Ellis makes the point explicit with reference to cross-modal association when he writes 'in addition to implicit learning within input modalities, attentional focus in the working memory can result in the formation of cross-modal associations' (Ellis 2001, p. 42). Thus, during development, association of different perceptions and their blending require working memory and attention. Even in adults, cross-modal information processing, for example from linguistic and visual domains requires working memory (Huettig et al. 2011).

According to the emergent tradition, learning and using a language is actually a problem akin to sequence learning. Sequence learning has been an important area of research in cognitive psychology that explores how humans remember and manipulate sequences of information (Cohen et al. 1990). Importantly, sequence learning involves attention (Mayr 1996). Sequence learning may involve predictive knowledge. For example, in the domain of language when we speak a sentence, we know in what sequence certain words will appear and what would be their unique identities. Predictive processes as they are in sequence learning tasks are seen in language processing (Dell and Chang 2014). One can predict the upcoming units based on an analysis of what one has been exposed to. Sequence learning is perfect when one retrieves from long-term memory the chunks of information that are necessary while manipulating information from working memory (Jimenez and Mendez 1999).

Similarly, attention plays a central role in allowing the learner to select relevant information and focus on it for long-term use. The child must make a decision about what would be the most appropriate association to make between the verbal and the visual material, select the relevant information, and code it for speaking. In this sense, attention could be influential in regulating goal-directed verbal behaviour during development (Rueda et al. 2005). While the emergent theory might view attention as a supervisory system that helps focus information in the working memory, it also seems that attention forges a bond between spoken and visual cues in a dynamic manner instigating foraging behaviour (Klein 2000). Shifts in attention for searching through eye movements arise from our need to constantly look out, and indicates our foraging behaviour. While attention could perform an implicit function in creating and maintaining associations, it is also crucial for goal-directed action during search and selection (Corbetta and Shulman 2002). Therefore, language and its interaction with attention and memory systems have an evolutionary basis.

Evidence of early coupling of attention and language processing comes from studies that show anticipatory eye movements during simultaneous processing of linguistic and visual objects. There is evidence which suggests that very young infants show anticipatory eye movements to simple linguistic and visual stimuli (McMurray and Aslin 2004). In their study, McMurray and Aslin (2004) trained infants to anticipate locations of targets depending on cues. Results showed that infants could quickly programme eye movements towards correct locations using information from cues as well as novel stimuli. Within milliseconds, young children orient their attention towards objects whose names partially match the spoken input. In the domain of language processing, Mani and Huettig (2012) have shown that 2-year-olds can predict upcoming words in a sentence using contextual information and programme anticipatory saccades. In these tasks, participants often orient attention towards visual referents much before they have fully heard the spoken words. Many now argue that anticipation and prediction are key to language processing and much of cognition (Pickering and Garrod 2013). Children can predict linguistic information related to speakers and look in the appropriate direction (Borovsky and Creel 2014). Anticipation to linguistic stimuli has been shown to rapidly evolve and stabilize by the time children are 2 years old (Fernald et al.

1998). This evidence suggests that attentional mechanisms and language processing quickly collaborate during childhood.

How does a child know that an object among many others is a 'tree' and orients towards it when listening to the word 'tree'? The tree as an object is not as such visually salient, since many other objects surrounding it that may resemble a tree in many manners. It is possible that the spoken word helps increase the salience of this particular object over others (Lupyan and Ward 2013). One aspect of attention is to bind multiple events and maintain a single representation as well as combine information originating from two different modalities (see Mishra and Marmalejo-Ramos 2010 for discussions). When action sequences are repeated often, attentional selection becomes automatic and we have the illusion that we just find out objects in the world as soon as we listen to their names. Indirectly, we can say that linguistic input drives attention towards useful referents.

3.7 A View from Cognitive Linguistics

Chomskyan linguistics and classical psycholinguistics examined the human innate capacity by looking at linguistic competence. However, there was limited exploration of the executive resources that fuel language processing or other cognitive systems. Language was not viewed as a cognitive event with close ties with other systems such as vision, attention or memory. Cognitive linguistics has evolved into a discipline that examined language from a cognitive science perspective (Langacker 1993). It began as a rebellion against the modular Chomskyan style of language processing models and has attracted many followers among cognitive psychologists and others ever since. It studies language functions as a result of basic human cognitive processes such as perception, attention and emotion (Johnson and Lakoff 2002). Cognitive linguistics thus offers a very favourable landscape for the study of attention and language interaction in a broad sense. Cognitive linguists do not consider language to be a modular faculty with insular operations (Croft and Cruse 2004). There is heavy emphasis on meaning as it is learned through use in social circumstances.

Talmy, among others, has made explicit characterization of attentional mechanisms in language interactions within a cognitive linguistics framework (Talmy 1988). In Talmy's framework, language has its own structure that it uses to organize perception while it draws heavily from other systems like attention and visual perception. Talmy has made use of the figure–ground dichotomy, popular in the gestalt psychological tradition (Kohler 1969) to explain attentional mechanisms in language. Figures are embedded against backgrounds. Many think that figures are attended first and memorized better (Qiu et al. 2007). Talmy wrote,

> The figure is a moving or conceptually movable entity whose site, path, or orientation is conceived as variable, the particular value of which is the relevant issue. The ground is a reference entity, one that has a stationary setting relative to a reference frame, with respect to which the figure's site, path, or orientation is characterized. (2000, p. 184)

Imagine that someone sees a bicycle parked against a wall. For Talmy, the bicycle is the figure and the wall serves as the ground. This is a matter of how the visual system organizes things for our perception (Metzger and Spillmann 2006). How are figure–ground relationships reflected in language? Talmy suggests that the structuring of language, the way we make sentences, shows an effect of attentional events during perception. The good thing about this view is that it considers visual perception, attention and retrieval from long-term memory as key events during formation of concepts. Figure–ground distinctions influence perception and event construction and language. Figures and grounds are always in close functional relation to one another. Consider these two sentences taken from Talmy (2011):

1. The pen lay on the table.
2. The pen fell off the table.

In both these cases the object, pen, functions as the figure and the table as the ground. Often figures are mentioned first in sentence constructions and they could also be grammatical subjects (Wallace 1982). Figures are also active agents but need not be always so. For example, a figure–ground asymmetry in the following sentence could make the whole sentence odd to speak.

3. The house is near the bike.

The house is not readily conceptualized as a figure here since it is the reference point with respect to which the location of the bike needs elaboration and not the other way around. It has been widely assumed that figure–ground segregation occurs before attentional deployment (Peterson and Gibson 1994). This is to say, observing a visual scene leads the system automatically to parse the figure with respect to the ground based on some low-level visual cues (Kienker et al. 1986).

Figures stand out from the background and often are under attentional focus. Figure–ground distinctions have been suggested as innate perceptual strategies by the Gestalt psychological school (Wagemans et al. 2012). However, when we apply this distinction to linguistic categorizations, not always do things fall in place since language manifests such perceptual distinctions differently (see Wallace 1982).

Figure–ground perception is obviously not a linguistic issue. Cognitive psychological studies have shown that focal attention may not be necessary for figure–ground extraction during perception (Kimchi and Peterson 2008). It is still not entirely clear if figure–ground assignments are affected by manipulation of visuo-spatial attention (Vecera et al. 2004). For instance, in the bike and wall example discussed earlier, if attention is first drawn towards the wall instead of the bike, there is a possibility that one might consider the wall as figure since it receives focal attention. Talmy proposed that the attentional mechanism helps in the segregation of figures from the ground during visual scene perception and this, in turn, reflects how people speak about the scenes. This can only be true if figures draw attention automatically. It appears that cuing attention towards one object could have decisive effect on the formulation of a sentence with that object (Tomlin 1997). Attention drawn towards any object makes this object more salient and this object could become the subject of the sentence uttered. Figures make themselves salient and summon attention while this leads to the ground being blurred (Qiu et al. 2007).

It also appears that language has certain devices that help with 'windowing' attention (Talmy 2003). That is, certain constructs of language modulate attention towards objects. Not every portion of a sentence demands equal attentional focus, and listeners always know where to pay more attention. When speakers want to emphasize certain events or certain actors, they put them in some attentionally prominent position in the sentence, generally in the beginning as subjects. Talmy proposes that language exploits only some aspects of the attentional system and not all. It is often not clear if Talmy posits a general attentional system that serves all types of cognitive functions as a selectional and supervisory system, or an attentional system specific to language; or whether the way language is structured indicates an organization similar to attentional selection. Talmy says: 'the language-related faculty of the brain evolved to its present character in the presence of other already existing cognitive domains …and no doubt developed in interaction with their mechanisms of functioning, perhaps incorporating some of these' (Talmy 2003, p. 96).

While cognitive linguists propose that the structure of language and some of the devices it uses—like the deictic and prepositional systems (system of words that language uses to refer to space)—reflect how attention orients in space. It is one thing to find out figure–ground types of segregations in a sentence with respect to some visual scene, and another to state that attention might have played a role in this. It is, however, reasonable to assume that attention helps cut and organize visual perception in many ways and its selectivity reflects in the structure of language. What to say and what to give prominence to is a matter of attentional selection and not language. Of course, language can optimize these selected perceptions for most economical use in communication.

3.8 Attention and Language in Disorders

Many critical insights into the understanding of the attention and language relationship have come from studies of language and cognitive disorders. If attention is key to aspects of language development, then its malfunctioning can cause delay in linguistic development. One view considered language perception an informationally encapsulated input module—it was believed that linguistic and other cognitive disorders could be dissociated (Bellugi et al. 1993). It was thought that a condition such as specific language impairment (SLI) is a pure case of morpho-syntactic disorder, which does not involve deficits in non-linguistic domains (Van der Lely 1997). Specific language impairments were associated with particular deficits in morpho-syntax and grammar. Similarly, in theories about Williams syndrome (a condition where cognition is impaired with excellent verbal ability intact), there was a tendency to see a separation between non-linguistic and linguistic deficits (Bellugi et al. 1990). However, closer examination suggests that children with SLI (Specific language impairment) and William's disorder suffer from one or another aspect of verbal deficits that are related to both accessing information from long-term memory and morpho-syntactic computation (Clahsen and Almazen 1998). Many children who show some form of grammatical deficits also show problems on executive control tasks that demand selective attention (Finneran et al. 2009). Children with SLI seem to be slower in visuo-spatial processing, speed of orientation and related behaviour (Schul et al. 2004; Stevens et al. 2006). What these data may suggest is that attention and executive control could explain a range of childhood learning and language-based disorders. Such evidence shows that attentional involvement in linguistic disorders is prevalent. Talking of disorders, a very common disorder is dyslexia, which is characterized by an inability to learn to read. It is a disorder that primarily manifests in a deficit in grapheme to phoneme conversion, and children with dyslexia do not show very good phonological awareness. Graphemes are letters used in denoting sounds in a language whereas phonemes are the basic units of speech perception. In such tasks, typically the participant is asked to repeat a word without a sound, for example, the word 'grasp' without 'r' which the dyslexics find difficult. Graphemes are orthographic units of language represented through symbols. That is, a dyslexic will have a difficult time in separating some sounds and analyzing them as independent units of speech. There is no disagreement that it is primarily a language-based disorder where the awareness of units of spoken language suffers. However, there is ample evidence that visuo-spatial attention and a range of visual factors could be at the roots of dyslexia (Boose et al. 2007; Facoetti and Molteni 2001).

It is gradually being realized that many of these linguistic or intellectual deficits show executive control deficits apart from their characteristic features (Barkley 1997). That is, attention, memory and other cognitive mechanisms could underlie language disorders. This could, in essence, suggest that an efficient executive control system, of which attention is a part, is necessary for smooth functioning of several other cognitive functions, including language (Kurland 2011). Co-morbidity of executive control deficits and linguistic and verbal deficits are a norm. This has led to new policies in diagnosis and therapy. This approach is interactionist and very conducive for a proper discussion of several such conditions. It is possible to give examples of many other linguistic deficits, for example, aphasia where an attention deficit has been implicated (Murray 1999). In sum, linguistic deficits are primarily cognitive deficits and many aspects of attention and other executive functions play a crucial role in their manifestation. Thus, this line of research has also been very useful in providing a complementary angle to the attention–language interaction debate.

3.9 Integrationist Psycholinguistics

Contemporary psycholinguistics investigates language processing in consonance with other cognitive processes. However, the critical question is, which linguistic processes are more interactive with attention than others? Psycholinguistics has become non-modular, interactionist and also truly cognitive in the last two decades. A close look at any contemporary handbook of psycholinguistics will show many chapters that link attention and other executive processes, including vision, with different language processes (Goldrick et al. 2014; Traxler and Gernsbacher 2011). This is also true of books that deal with different psycholinguistic methods (Podesva and Sharma 2014). The chapters covered in these books discuss how researchers have started to use multiple techniques to explore human language processing mechanisms from both psychological and neurobiological perspectives, importantly considering cognitive mechanisms like attention, memory and vision. Psycholinguistics has come out of the Chomsky–Fodor type of insulation. Use of terms like computational psycholinguistics suggests diversification of the discipline itself. This kind of psycholinguistics is much closer to developments in cognitive psychology and vice versa in many areas. Psycholinguistics has become multi-modal and integrationist. Currently, for example a wide range of research examines the interaction of language, vision, attention and memory together to understand mental processing (Huettig et al. 2011; Henderson and Ferreira 2004).

It is sensible to think of human cognitive action as a result of the combined effort of several types of mechanisms, both central and input systems. We have seen that under the modularity wave, no one was willing to examine the processing of language as a function of, or in combination with other cognitive processes. The alternative approach of multi-modal system in Jackendoff (2009) in recent times shows not only how different levels of language can interact but also how language interacts with systems such as vision. Jackendoff linked Marr's computational theory of vision and object recognition to linguistic expression (Jackendoff 1987). According to this thesis, visual perceptual systems and linguistic systems show large amount of representational overlap. When one sees an object in space, its spatial and other identities are mapped onto linguistic conceptual structures, and that is how we are able to talk about things we see through language so effortlessly. Humans can talk about things in the world in very precise terms since the spatial representations directly feed information into the language and the motor systems (Landau and Jackendoff 1993). This interface between the visual and haptic object recognition system and the linguistic system is crucial for naming and talking in general (see Herskovits 1997).

The first important point to be noted in Jackendoff's enterprise is that he explicitly brings vision and language together. This was never the case within linguistics or cognitive psychology previously. Secondly, this way of looking at event perception and description offers novel insights into how language interacts at the representational level with other systems. Further, in this framework, language does not merely mirror visually perceived events, but guides and constrains visual perception as well in gluing together disparate events in space and time (Dessalegn and Landau 2008). Language does not merely refer to things seen, but also brings out spatial relations between them through its own syntax. This framework later paved the way for more empirical work leading to the demonstrations of online interaction between language-vision and attention. This brief digression into Jackendoff shows that revolutions inside linguistics, mainly against the transformational generative framework, were motivated by developments in cognitive and computational sciences. Jackendoff brought the object recognition theories of Biederman (1987) and Marr very close to understand how language describes the world (see Jackendoff 2012 for recent discussions).

Attending to an object has many benefits in linguistic processing. Wittgenstein referred to 'attention' rather explicitly in his *Philosophical Investigations* (Wittgenstein 1953/2001; see also Campbell 2000). He thought that when we name an object or some feature of it what we do is to concentrate our attention on it. He writes,

> —How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign.—Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connection between the sign and the sensation. (p. 258)

However, Wittgenstein was often ambiguous about the essential role attention could play in language.

In contemporary research, the use of eye tracking has provided direct evidence of the coupling of language and attention as well as vision. A simple observation will suggest that language drives attention in the real world. When one listens to a word, attention immediately orients towards the object whose name matches with the linguistic input (Cooper 1974). Language users anticipate visual referents using contextual information (Altmann and Kamide 1999). Information from visual events is used to understand language in a dynamic manner (Knoeferle and Crocker 2006). This orienting can happen very automatically and sometimes language users are not aware of these. Interestingly, when someone hears the word 'apple', attention is also biased towards oranges, since they come from the same semantic field (Huettig and Altmann 2005). Alternately, listening to the word 'candy' can drive attention towards 'candle' (Allopenna et al. 1998; see also Mishra et al. 2012). This sort of orienting behaviour of attention is linked with processing language. Much work in cognitive psychology on multi-modal integration over a period of time now has revealed that representations from both audition and visual modalities dynamically interact. Information from multiple modalities interacts online and affects central systems. This is what Fodor was reluctant to accept.

The work described above has brought together vision researchers, psycholinguists, and neuroscientists (see Ferreira and Tanenhaus 2007). Information acquired from visual search and scene perception can be useful for understanding how language guides attention towards objects. We have noted previously that psycholinguists were once not very keen to apply theories of cognitive and experimental psychology in their daily work. Today, one can easily find references to such work in language research in many areas. Many everyday phenomena like searching, attending and holding things in memory are being understood as outcome of multi-modal processes. The use of eye tracking methodology has revealed how language and attention could interact in real time, thus leading to the emergence of cognition (Mishra 2009; Spivey and Dale 2006). Research into reading (Rayner 2009), spoken language comprehension (Tanenhaus et al. 1995), and word production (Meyer et al. 1998) shows intimate connection between attentional and linguistic processes.

This is what Jackendoff has in mind when he links language with vision. What this probably shows is that language has never operated in the real world without attention and vision. What is integrationist here? The integrationist development involves employment of theoretical predictions from different disciplines, with an unbiased understanding of the workings of the brain and basic cognitive functions. While some time back language researchers thought that psychological facts might constrain their theorizing, today, this link seems facilitatory.

3.10 Conclusion

Psycholinguistics has come a long way since Wundt's introspectionism, and Miller's cognitivism and has now become fully multi-modal modal. This journey has seen the fall of several paradigms and the rise of many others. I digressed into the history of interactions between linguistics and psychology/cognitive science was necessary to show the deep mistrust between the disciplines not so long back. The rise of alternative frameworks inside cognitive science and linguistics has opened up new ways to study the interaction of language with other cognitive systems. With the emergence of new methods in brain imaging and computational models this work has become very expansive. There is a good deal of hope that we can now understand how attention and language interact and guide action.

This book includes chapters on bilingualism and processing in illiterate subjects to show how cultural and socio-linguistic factors can exert influence on this interaction. It seems likely that the attention–language interaction must take the individual and his environment into account, since this interaction is only useful as far as it is used for action in specific contexts. A wide range of variability in colour perception, for example, has been shown with people from different places and who speak different languages (Thierry et al. 2009). However, we can only understand diversity once we know how the core interactions work.

We have seen how psycholinguistics has come from being 'insular' to 'integrationist'. In the following chapters, this book shows how current work has progressed on these lines in different sub-domains of language and cognition research and brings up vision and memory in discussions of this interaction.

References

Abrams, K., & Bever, T. G. (1969). Syntactic structure modifies attention during speech perception and recognition. *The Quarterly Journal of experimental Psychology, 21*(3), 280–290.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of language processing. *Cognitive Psychology, 6*, 84–107.

Altmann, G. T. (2013). Anticipating the garden path: The horse raced past the barn ate the cake. *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structures, 4*, 111.

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502–518.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191–238.

Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Anderson, B. (2011). There is no such thing as attention. *Frontiers in psychology*, *2*.

Anderson, J. R. (2009). *Cognitive psychology and its implications*. New York: Worth Publishers.

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence, 149*(1), 91–130.

Appelbaum, I. (1998). Fodor, modularity, and speech perception. *Philosophical Psychology, 11* (3), 317–330.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The psychology of learning and motivation, 2*, 89–195.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Baddeley, A., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation, 8*, 47–89.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(04), 577–660.

Barsky, R. F. (2011). *Zellig Harris: from American linguistics to socialist Zionism*(pp. xvii+–353). Cambridge, MA: MIT Press.

Bellugi, U., Marks, A. B., Bihrle, A., & Sabo, H. (1993). Dissociation between language and cognitive functions in Williams syndrome. *Language development in exceptional circumstances*, 177.

Bellugi, U., Poizner, H., & Klima, E. S. (1990). Mapping brain functions for language: Evidence from sign language. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps* (pp. 521–543). New York: Wiley-Liss.

Benjamin Jr, L. T. (2007). *A brief history of modern psychology*. Oxford: Blackwell Publishing.

Bever, T. G. (1970). The cognitive basis for linguistic structures. *Cognition and the development of language, 279*(362), 1–61.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115–147.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart, and Winston.

Blumenthal, A. L. (1970). *Language and psychology: Historical aspects of psycholinguistics*. Oxford: John Wiley.

Blumenthal, A. L. (1975). A reappraisal of Wilhelm Wundt. *American Psychologist, 30*(11), 1081–1088.

Blumenthal, A. L. (1987). The emergence of psycholinguistics. *Synthese, 72*(3), 313–323.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*(1), 1.

Boden, M. A. (2006). *Mind as machine: A history of cognitive science*. Oxford: Oxford University Press.

Borovsky, A., & Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. *Developmental Psychology, 50*(5), 1600.

Borreggine, K. L., & Kaschak, M. P. (2006). The action–sentence compatibility effect: It's all in the timing. *Cognitive Science, 30*(6), 1097–1112.

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104*(2), 198–230.

Brentano, F., (1874). *Psychology from an empirical standpoint*. (Rancurello, A. C., Terrell, D. B., and McAlister, L. L., trans.). New York: Humanities Press.

Broadbent, D. E., (1958). Perception and communication., (pp. 81–107). Elmsford, NY, US: Pergamon Press, v, 340 pp.

Burnston, D. C., & Cohen, J. (Forthcoming). Perceptual integration, modularity, and cognitive penetration. *Cognitive Influences on Perception: Implications for Philosophy of Mind, Epistemology, and Philosophy of Action* (Raftopoulos, A., and Zeimbekis, J., Eds.). Oxford University Press.

Campbell, S. B. (2000). Attention-deficit/hyperactivity disorder. In *Handbook of developmental psychopathology* (pp. 383–401). Springer US.

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. *Syntax and semantics, 21*, 263–288.

Chapman, S., & Routledge, C. (2005). *Key thinkers in linguistics and the philosophy of language*. Oxford University Press: Oxford

Chemero, A. (2003). Information for perception and information processing. *Minds and Machines, 13*(4), 577–588.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge: The MIT Press.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language, 35*(1), 26–58.

Chomsky, N. (1966a). *Cartesian Linguistics: Chapter in the Historz of Rationalist Thought*. New York: Harper and Row

Chomsky, N. (1966b). *Topics in the theory of generative grammar* (Vol. 56). New York: Walter de Gruyter, The Hague: Mouton.

Chomsky, N. (1972). *Language and mind* (p. 100). New York: Harcourt Brace Jovanovich.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger Publishers.

Chomsky, N. (2005). Three factors in language design. *Linguistic inquiry, 36*(1), 1–22.

Chomsky, N. (2009). *Cartesian linguistics: A chapter in the history of rationalist thought*. Cambridge: Cambridge University Press.

Christiansen, M. H., & Chater, N. (1994). Generalization and connectionist language learning. *Mind and Language, 9*(3), 273–287.

Clahsen, H., & Almazan, M. (1998). Syntax and morphology in Williams syndrome. *Cognition, 68*(3), 167–198.

Clark, A. (1999). An embodied cognitive science? *Trends in cognitive sciences, 3*(9), 345–351.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*(1), 17.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences, 3*, 115–120.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*(1), 84–107.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Crane, M. T. (2014). *Framing Authority: Sayings, Self, and Society in Sixteenth-Century England*. NJ: Princeton University Press.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.

Danziger, K. (2001a). *The Unknown Wundt* (pp. 95–120). US: Springer.

Danziger, K. (2001b).The Unknown Wundt. In *Wilhelm Wundt in History* (pp. 95–120).Springer US.

De Saussure, F., & Baskin, W. (2011). *Course in general linguistics*. New York: Columbia University Press.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634), 20120394.

Dessalegn, B., & Landau, B. (2008). More than meets the eye the role of language in binding and maintaining feature conjunctions. *Psychological Science, 19*(2), 189–195.

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). Cambridge: Cambridge University Press.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Facoetti, A., & Molteni, M. (2001). The gradient of visual attention in developmental dyslexia. *Neuropsychologia, 39*(4), 352–357.

Fancher, R. E. (1977). Brentano's psychology from an empirical standpoint and Freud's early metapsychology. *Journal of the History of the Behavioral Sciences, 13*(3), 207–227.

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in cognitive sciences, 18*(3), 120–126.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*(3), 228–231.

Ferreira, F., & Tanenhaus, M. K. (2007). Introduction to the special issue on language–vision interactions. *Journal of Memory and Language, 57*(4), 455–459.

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language and Hearing Research, 52* (4), 915.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*(6), 825–850.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge: MIT press.

Fodor, J. A., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*(5), 414–420.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1), 3–71.

Fodor, J. A., Bever, T. G., & Garrett, M. F. 1974. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw Hill.

Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition, 6* (1), 35–53.

Frawley, W. (1997) *Vygotsky and Cognitive Science. Language and the Unification of the Social and Computational Mind*. Cambridge: Harvard University Press.

Gardner, H. (2008). *The mind's new science: A history of the cognitive revolution*. New York: Basic books.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775.

Gentner, D., & Stevens, A. L. (1983). *Mental models*. NJ: Psychology Press.

Gerald, G., Klein, E., Pullum, G., & Sag, I. (1985). Generalized phrase structure grammar. Oxford: Basil Blackwell; Cambridge, Massachusetts: Harvard University Press; xii+276 pp.

Gibbs, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind and Language, 21*(3), 434–458.

Gibson, E., & Fedorenko, E. (2010). Weak quantitative standards in linguistics research. *Trends in cognitive sciences, 14*(6), 233.

Gibson, J. J. (1986). *The ecological approach to visual perception*. San Diego: Psychology Press.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55.

Gleitman, L. R., Gleitman, H., Landau, B., & Wanner, E. (1989). Where learning begins: Initial representations for language learning. Language: Psychological and biological aspects, 150–193.

Goldrick, M. A., Ferreira, V., & Miozzo, M. (2014). *The Oxford handbook of language production*. Oxford: Oxford University Press.

Gross, R. (2012). Psychology: The Science of Mind and Behaviour , 6th Edn. Hachette UK.

Harman, G. H. (1967). Psychological aspects of the theory of syntax. *The Journal of Philosophy, 64*(2), 75–87.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Herskovits, A. (1997). Language, spatial cognition, and vision. In *Spatial and Temporal Reasoning* (pp. 155–202). Springer Netherlands.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23–B32.

Huettig, F., Mishra, R. K., & Olivers, C. N. (2011).Mechanisms and representations of language-mediated visual attention. *Frontiers in psychology*, *2*.

Huettig, F., Olivers, C. N., & Hartsuiker, R. J. (2011b). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica, 137*(2), 138–150.

Jackendoff, R. (1987). *Consciousness and the computational mind* (pp. 253–256). Cambridge, MA: Mit Press.

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception, 26*(3), 195–204.

Jackendoff, R. (2012). Language as a source of evidence for theories of spatial representation. *Perception-London, 41*(9), 1128.

Jespersen, O. (1922). *Language, its nature, development, and origin*. New York: H. Holt.

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25*(1), 236.

Johnson, M., & Lakoff, G. (2002). Why cognitive linguistics requires embodied realism. *Cognitive linguistics, 13*(3), 245–264.

Johnson-Laird, P. M. (1984). Psychology and Linguistics. *Psychology and Its Allied Disciplines: Psychology and the humanities, 1*, 75.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition, 94*(3), B79–B89.

Kess, J. F. (1981). Wundt and Bloomfield: The Leipzig Connection. *Working Papers of the Linguistics Circle*, *1*(1), 123–151.

Kess, J. F. (1992). *Psycholinguistics: Psychology, linguistics, and the study of natural language* (Vol. 86). Amsterdam: John Benjamins.

Kienker, P. K., Sejnowski, T. J., Hinton, G. E., & Schumacher, L. E. (1986). Separating figure from ground with a parallel network. *Perception, 15*(2), 197–216.

Kimchi, R., & Peterson, M. A. (2008). Figure-ground segmentation can occur without attention. *Psychological Science, 19*(7), 660–668.

Klein, R. M. (2000). Inhibition of return. *Trends in cognitive sciences, 4*(4), 138–147.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science, 30*(3), 481–529.

Koerner, E. F. K. (1973). *Contribution au débat post-saussurien sur le signe linguistique: introduction générale et bibliographie annotée*. The Hague: Mouton.

Kohler, W. (1969). *The task of gestalt psychology*. NJ: Princeton University Press.

Kurland, J. (2011). The role that attention plays in language processing. *SIG 2 Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, *21*(2), 47–54.

Kutas, M. & Van Petten, C. (1994). Psycholinguistics Electrified: Event-related potential investigations. In: M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 83–143), Academic Press, San Diego.
Landau, B., & Jackendoff, R. (1993). 'What' and 'where' in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16*, 217.
Langacker, R. W. (1993). Reference-point constructions. *Cognitive, linguistics, 4*(1), 1–38.
Leahey, T. H. (1987). *A history of psychology: Main currents in psychological thought*. Englewood Cliffs, NJ: Prentice-Hall.
Leuthold, H., Sommer, W., & Ulrich, R. (2004). Preparing for action: inferences from CNV and LRP. *Journal of Psychophysiology, 18*(2/3), 77–88.
Levelt, W. (2012). *A history of psycholinguistics: The pre-Chomskyan era*. OUP Oxford.
Levelt, W. J. (1972). Some psychological aspects of linguistic data.*Linguistische Berichte*, *17*, 18–30.
Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
Levinson, S. C. (2004). Deixis. In *The handbook of pragmatics* (pp. 97–121). Blackwell. Oxford.
Levinson, S. C. (2012). The original sin of cognitive science. *Topics in cognitive science, 4*(3), 396–403.
Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences, 110*(35), 14196–14201.
Macnamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review, 79* (1), 1–13.
MacWhinney, B. (Ed.). (2013). *The emergence of language*. New York: Psychology Press.
Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*(1), 59–70.
Mandler, G. (2007). *A history of modern experimental psychology: From James and Wundt to cognitive science*. Cambridge, MA: MIT Press.
Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843.
Maria de Lourdes, R. D. F., & Passos, M. A. M. (2007). The influence of Bloomfield's linguistics on Skinner. *The Behavior Analyst, 30*(2), 133.
Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science, 16*(1), 6–10.
Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt and Co., Inc.
Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition, 32*(8), 1389–1400.
Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22*(2), 350.
McCauley, R. N. (1987). The not so happy story of the marriage of linguistics and psychology or why linguistics has discouraged psychology's recent advances. *Synthese, 72*(3), 341–353.
McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy, 6*(2), 203–229.
Metzger, W., & Spillmann, L. (2006). *Laws of seeing*. Cambridge, MA: MIT Press.
Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*(2), B25–B33.
Miller, G. A. (1990). Linguists, psychologists, and the cognitive sciences. *Language, 66*(2), 317–322.
Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences, 7*(3), 141–144.
Mishra, R. K. (2009). Interaction of language and visual attention: evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K., & Singh, N. (2010). Online fictive motion understanding: An eye-movement study with Hindi. *Metaphor and Symbol, 25*(3), 144–161.

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing, 11*(4), 295–305.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Mole, C. (2013). *Attention is cognitive unison: An essay in philosophical psychology*. Oxford: OUP.

Moro, A. (2014). On the similarity between syntax and actions. *Trends in cognitive sciences, 18*(3), 109–110.

Murray, L. L. (1999). Review attention and aphasia: Theory, research and clinical implications. *Aphasiology, 13*(2), 91–111.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. New York: WH Freeman/Times Books/Henry Holt & Co.

Newmeyer, F. J. (1983). *Grammatical theory: its limits and its possibilities*. University of Chicago Press.

Newmeyer, F. J. (1997). *Generative linguistics: a historical perspective*. New York: Psychology Press.

Noë, A. (2004). *Action in perception*. Cambridge: MIT press.

O'Grady, W. (2008). The emergentist program. *Lingua, 118*(4), 447–464.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(05), 939–973.

Peterson, M. A., & Gibson, B. S. (1994). Must figure-ground organization precede object recognition? *An assumption in peril. Psychological Science, 5*(5), 253–259.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(04), 329–347.

Podesva, R. J., & Sharma, D. (Eds.). (2014). *Research methods in linguistics*. Cambridge: Cambridge University Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Posner, M. I. (2012). *Attention in a social world*. OUP USA.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*(5), 391.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*(5), 351–360.

Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge, MA: MIT press.

Qiu, F. T., Sugihara, T., & von der Heydt, R. (2007). Figure-ground mechanisms provide structure for selective attention. *Nature Neuroscience, 10*(11), 1492–1499.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Reber, A. S. (1987). The rise and (surprisingly rapid) fall of psycholinguistics. *Synthese, 72*(3), 325–339.

Robins, R. H. (2013). *A short history of linguistics*. New York: Routledge.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language, 57*(2), 232–251.

Roeper, A. (1982). How the gifted cope with their emotions. *Roeper Review, 5*(2), 21–24.

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological science*, 0956797613490746.

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573–594.

Ryle, G. (1949). *The concept of mind*. New York: Barnes and Noble.
Sapir, E. (1929). The status of linguistics as a science. *Language*, 207–214.
Schul, R., Stiles, J., Wulfeck, B., & Townsend, J. (2004). How 'generalized'is the 'slowed processing'in SLI? The case of visuospatial attentional orienting. *Neuropsychologia, 42*(5), 661–671.
Seuren, P. A., Sevren, P. A., & Seuren, P. A. (1998). *Western linguistics: An historical introduction* (p. 42). Oxford & Malden, MA: Blackwell.
Shapiro, L. (2011). *Embodied cognition*. NY: Routledge Press.
Skinner, B. F. (1953). *Science and human behavior*. New York: Simon and Schuster.
Skinner, B. F. (1956). A case history in scientific method. *American Psychologist, 11*(5), 221.
Skinner, B. F. (1957). *Verbal behavior*. New York: AppletonCentury-Crofts.
Skinner, B. F. (1985). Cognitive science and behaviourism. *British Journal of Psychology, 76*(3), 291–301.
Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'.*Rethinking linguistic relativity, 17,* 70–96.
Soames, S. (1984). Linguistics and psychology. *Linguistics and Philosophy, 7*(2), 155–179.
Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied, 74*(11), 1.
Spivey, M. J. (2007). *The continuity of mind*. New York: Oxford University Press.
Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science, 15*(5), 207–211.
Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*(4), 235–241.
Sprouse, J., & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics, 48*(03), 609–652.
Stemmer, N. (1990). Skinner's verbal behavior, Chomsky's review, and mentalism. *Journal of the Experimental Analysis of Behavior, 54*(3), 307.
Stevens, C., Sanders, L., & Neville, H. (2006). Neurophysiological evidence for selective auditory attention deficits in children with specific language impairment. *Brain Research, 1111*(1), 143–152.
Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science, 12*(1), 49–100.
Talmy, L. (2003). Concept structuring systems in language. *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, 2*, 15–46.
Talmy, L. (2011, August). Figure and ground in complex sentences. In*Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (Vol. 1, No. 1).
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.
Tettamanti, M., & Weniger, D. (2006). Broca's area: A supramodal hierarchical processor? *Cortex, 42*(4), 491–494.
Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentivecolor perception. *Proceedings of the National Academy of Sciences, 106*(11), 4567–4570.
Titchener, E. B. (1921). Brentano and Wundt: Empirical and experimental psychology. The American Journal of Psychology, 108–120.
Tomasello, M., & Camaioni, L. (1997). A comparison of the gestural communication of apes and human infants. *Human Development, 40*(1), 7–24.
Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. *Language and Conceptualization*, 162–189.
Traxler, M., & Gernsbacher, M. A. (2011). *Handbook of psycholinguistics*. New York: Academic Press, 2006.
Van der Lely, H. K. (1997). Language and cognitive development in a grammatical SLI boy: Modularity and innateness. *Journal of Neurolinguistics, 10*(2), 75–107.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vauclair, J., & Perret, P. (2003). The cognitive revolution in Europe: Taking the developmental perspective seriously. *Trends in Cognitive Sciences, 7*(7), 284–285.

Vecera, S. P., Flevaris, A. V., & Filapek, J. C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science, 15*(1), 20–26.

Vygotsky, L. S. (1978). *Mind in Society*. Cambridge: Harvard University Press.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138*(6), 1172.

Wallace, S. (1982). *Figure and ground: The interrelationships of linguistic categories* (pp. 201–223). Tense-aspect: Between semantics and pragmatics.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158.

Wilson, A. L. (1993). The promise of situated cognition. *New directions for adult and continuing education, 1993*(57), 71–79.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636.

Wittgenstein, L (2001) [1953]. *Philosophical investigations*. Oxford: Blackwell Publishing.

Wittgenstein, L. (1953/2001). *Philosophical investigations*. Oxford: Blackwell.

Wittgenstein, L. (2010). Philosophical investigations. Malden: Wiley.

Wundt, W. (1907). Ueber Ausfrageexperimente und ueber Methoden zur Psychologic des Denkens. *Psychologische Studien, 3*, 301–360.

Wundt, W. (1900). Die Sprache (2 Vols.). *Leipzig: Kroner*.

Wundt, W. (1911). *Völkerpsychologie*. Leipzig.

Wundt, W. M. (1874). *Principles of physiological psychology*. Leipzig: W. Engelman.

Zwaan, R. A. (2003). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of learning and motivation, 44*, 35–62.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*(1), 1.

Chapter 4
Attention and the Processing of Sentences

4.1 Automaticity and Sentence Parsing

We process sentences of various complexities and lengths all the time. At first this process looks very automatic and effortless. During listening one has the feeling that the whole process is going on smoothly without much demand on the central executive resources. A widely held view within cognitive sciences and psycholinguistics, following Fodor (1983), has been that sentence parsing is immune to other external influences and is therefore encapsulated (Pulvermüller et al. 2008). It is not known if general attention states have causal influences on sentence parsing. Some researchers have suggested that sentence processing is largely automatic and one need not pay excessive focal attention (Maidhof and Koelsch 2011). The projection of the syntactic structure could be automatic and is part of the innate linguistic system of humans. The question whether sentence processing is automatic or voluntary is linked to the issue of structural computation. Many psycholinguists use the word 'parsing' to refer to the process of comprehension given a sentence either in visual or spoken format. The parser's goal is to quickly assign the right structural representation to an incoming string of words and phrases, avoiding any ambiguous interpretation. The structural description of a sentence imposes certain restrictions on its computations (Frazier 1988). Computation is guided by the grammar which the native speaker has internalized. This process is intuitive and unconscious. It is, therefore, plausible to generate the hypothesis that since access to grammatical knowledge is automatic and unconscious, sentence processing seems cognitively effortless.

A widely held view within cognitive sciences is that many mental processes are automatic. Cognitive psychologists have made the distinction between 'automatic' and 'voluntary' processes around the notion of attentional involvement and effort

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_4

(Pashler et al. 2001). If a cognitive process calls for strategic utilization of attention, then it is not automatic. Automatic processes happen mandatorily without the dependence on cognitive resources such as attention (Jonides 1981). Distinct neural resources implement automatic processes at different time scales in the brain (Jeon and Friederici 2015). Although the role of attention as a causal factor in automatic processes still remains debatable (Moors 2016), attentional involvement need not be the only diagnostic of a mental process being automatic or voluntary. Nevertheless, it is important to understand if linguistic processing can happen without deliberate effort. If so, which processes are under voluntary control and which are automatic?

Listening to a string of words leads to spontaneous activation of knowledge. Listeners derive meaning from language automatically even when they are task irrelevant (Singh and Mishra 2013, 2015; Salverda and Altmann 2011). Linguistic processing of stimuli can be initiated in the brain without deliberate intention by the perceiver. Processes that require deliberate processing requires attention (Posner and Snyder 1975). Several threads of research suggest that attentional mechanisms can modulate sentence comprehension. Available evidence suggests that syntax processing is automatic but not semantics (Friederici 2011). Selectively attending to different aspects of a sentence modulates brain networks distinctively (Rogalsky and Hickok 2009). Additional mental load in the form of a dual task disturbs performance in a sentence processing task (Caplan 1996). This evidence suggests that certain aspects of linguistic processing are under strategic processing and selective attention can modulate such processing. The majority of such studies with different behavioural and neuroimaging techniques have used ‘attention as a limited processes resource’ view to examine if sentence processing is voluntary or automatic. Exogenous attention (capture of attention by external stimuli) leads to differential processing of syntax and semantic information. Gunter and Friederchi (1999) examined if syntax and semantic processing differ when endogenous attention (Chap. 2) is deployed to a specific task goal in an ERP (Event-related potential) paradigm. Participants automatically processed the word category violation even when the grammaticality judgement was task irrelevant and when they paid attention to the physical features of the font. However, the P600 effect (an effect reflecting reanalysis) almost diminished for the verb inflection violation when participants did the physical task. Interestingly, the N400 effect (an effect indicating detection of contextual anomaly) completely disappeared for the verb inflection type sentences under the physical conditions. When a participant’s attention was controlled for then a difference was observed with regard to the processing of syntactic and semantic information. These data indicate a differential and selective influence of attention on syntax and semantic processing during sentence comprehension.

Hahne and Friederici (1999) proposed that the ELAN (Early Late Anterior Negativity) effect, which is seen during violation of phrase structures very early, could reflect automaticity. If early structural computation is automatic and late reanalysis is a controlled process then different ERP effects should indicate this. The authors manipulated the proportion of correct and incorrect trials and predicted that if the ELAN effect arises automatically, then proportion manipulation should not have any effect on this effect. On the other hand, if the P600 effect reflects controlled processing, then its latency and amplitude should not be affected by proportion manipulation. The results showed appearance of ELAN under both types of proportion violations but the P600 effect appeared when the proportion of incorrect sentences was low. This showed that participants engage in selective attention to infrequently presented information. This study suggests that certain early syntactic processing could be automatic and attentional manipulation may not have much influence on this. Hahne and Friederici (2002) presented participants with three types of sentences in the spoken form. The first type had a syntactic violation (*Das Brot wurde gegessen*, The bread was eaten); the second type had semantic violation (*Der Vulkan wurde gegessen*, The volcano was eaten); and the third type had both syntactic and semantic violation (*Das Turschlop wurde im gegessen*, The door lock was in-the-eaten). Apart from expecting different types of ERP effects for different types of violations, the authors additionally manipulated attentional involvement. In one experiment, participants were asked to judge the sentences for their overall correctness whereas in another version of the same experiment, they were asked to specifically judge if the sentences were 'semantically' making sense. It was assumed that even if participants were exclusively attending to the semantic violations, some early syntactic processing would automatically take place and this should be evident in the appearance of LAN. This is to say, one would expect LAN even when attention is paid elsewhere. Selective attention in this experiment was manipulated solely through task instructions. The results showed that for sentences that had both syntactic and semantic violations, there was only ELAN but no N400 effect. This suggests that for such sentences when the parser found a phrase structure violation initially, later semantic analysis was not perused.

Can one now conclude that these ERP effects indicate controlled versus automatic processing of sentences? Recently, Steinhauer and Drury (2012), among other things, have raised issues with the claims of automaticity with reference to manipulation of the proportion of trials, as it was the case with Hahne and Friederici. The objection is that mere manipulation of the proportion of trials cannot guarantee that attentional processes are indirectly being manipulated. Further, it is possible that this way of designing the experiment might involve implicit learning and can induce statistical learning in subjects. However, Steinhauer and Drury have nothing to say against the study where attention was manipulated with task instruction. Whatever it may be, it seems likely that certain syntactic processes could indeed be automatic in the sense that attentional manipulation may not have any effect on them. The issue of attentional effects on sentence parsing is theory dependent. That is, depending on the type of parsing model one subscribes to, one would predict executive function effects on sentences (Fig. 4.1).

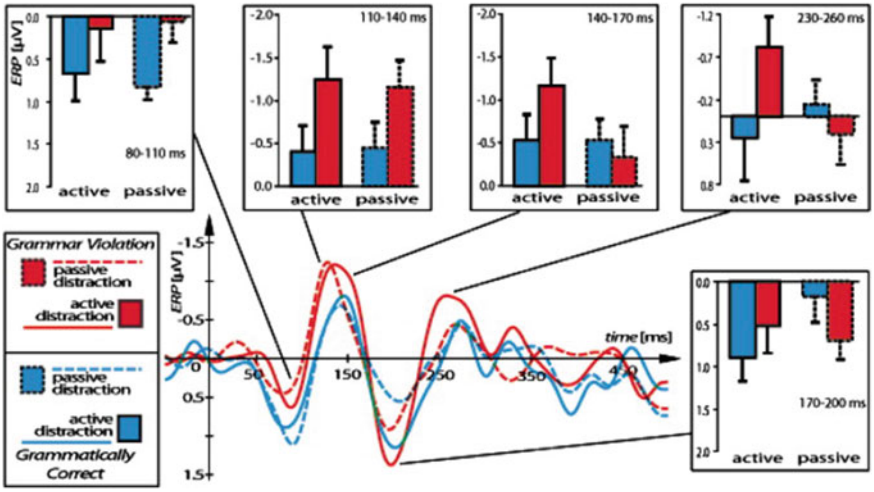

Fig. 4.1 The effect of active and passive attention on syntactic MMN. In general there is larger MMN effect for incorrect sentences as compared to correct sentences. The *upper panel* shows MMN effects after critical word onset for different time windows. The syntactic MMN peaked after 130 ms and was not modulated with the rising distraction. This pattern of results suggests that early syntactic parsing could be automatic. Attention distraction does not affect early stages of parsing. *Source* Pulvermuller et al. (2008, p. 248)

4.2 Processing Language Without Attention

Studies described above with spoken sentence processing in the ERP paradigm manipulated attention selectively. Furthermore, participants were asked to selectively process one or another aspect of the sentence. The results showed a divergent pattern of attentional involvement for syntactic and semantic information. Alternatively, it has been possible to examine if language can be processed in the absence of attention. That is when the participant's attention is not explicitly engaged with the task at hand. There is automaticity in sentence processing when attention is deployed in some other modality. The use of MMN (Mismatch Negativity) effect of ERP allowed researchers to test aspects of language processing in a rather controlled manner without summoning attention explicitly. Importantly, in this paradigm the idea has been to examine if processing sentences does not call for explicit attention.

The MMN is generated when a deviant stimuli is perceived by the sensory system among similar stimuli (Naatanen et al. 2007). The MMN reflects automatic detection of change in any sequence of stimuli. MMN has been used as an index of automatic processing in a number of different areas. Its use as a language and speech related component has allowed researchers to examine automatic nature of language processing. The most popular paradigm has been to ask participants to view a silent film and listen to spoken language stimuli while ERPs are being recorded. For example, in the domain of auditory speech perception, a MMN arises when an ‘a’ sound occurs infrequently among ‘b’ sounds. The cognitive system habituates itself when participants process the same stimuli repeatedly. However, attention is summoned if one encounters deviant stimuli among similar stimuli. This exogenous capture of attention by external stimuli can influence ongoing cognitive processing. The MMN effect can be correlated with automatic aspects of language processing while the participants’ attention is taken away from the linguistic stimuli. This is not the case with ERP studies where participants are explicitly asked to respond to the grammatical properties of sentences and judge them.

The MMN effect is seen even when participants do a concurrent task or operate under high attentional load. In an interesting study, Pulvermuller et al. (2008) examined the automaticity of syntactic parsing when participants were asked to divert their attention to some other demanding stimuli. The participants listened to grammatical or ungrammatical speech strings in one ear, while they processed some tones varying in their intensities in another ear. Critically, the tone was played when the suffixes that could be decisive for grammaticality of the phrase appeared. It was assumed that, if MMN is seen as an effect indicating early syntactic processing under this high task load, then such processing does not require attention and is automatic. It was observed that such high task load did not modulate the MMN. Similarly, in another condition participants were asked to see a silent movie strip while grammatical and ungrammatical phrases were presented in the classic oddball paradigm. In the oddball paradigm one hears deviant stimuli embedded in a train of similar stimuli. Again, the authors did not observe the amplitude or the latency of the MMN getting affected when attention was given to the visual stimuli. From this, it appears that, syntactic MMN reflects some early automatic processes of syntactic computation. For verb-subject agreement violation and word category violation one can see different topographical distribution of MMN. However, in both types of violations MMN is recorded indicating early and automatic processing of syntax (Hastings et al. 2008). This means, syntax processing is not affected when attention is summoned by any exogenous cue.

Several researchers have claimed that the MMN component is largely automatic and its greatest strength is that it can be elicited in the absence of attention. However, there are some studies in speech processing that show attention modulation of MMN (Szymanski et al. 1999). However, there is always a possibility that even though subjects have been asked to see a silent video film or read a book, they somehow manage to listen to the speech stimuli, i.e. pay attention. Speech, being such a common biological stimuli, has been shown to capture attention automatically. This means, it is nearly impossible to ignore speech stimuli that may include words, sentences or sounds. If this is so, then one can argue that the claims of the MMN paradigm may be too strong and unwarranted. It is easy to generate the impression from the above-mentioned studies that attention is not engaged in the processing of syntax at least in the earliest stages of parsing. This seems to be the case with early generation of phrase structures as evident in sentences with word category violation. In sum, current evidence suggests that sentences are processed without the engagement of explicit attention. However, if we go by current theories of attention, participants in any psychological task operate with the clear understanding of a goal which is the task instruction. The engagement of attention is contingent on this understanding of the goal in any sequence of action (Mole 2011). Therefore, attention is engaged as such participants are able to perform the task at hand with reasonable accuracy. Albeit, it can be said that participants do not need to deploy selective attention as such during normal sentence processing. It can be said that the task goals, i.e. judging the grammaticality of a sentence, engage the endogenous component of attention and exert a top-down selection strategy. Experimental manipulations such as responding to coloured fonts of words engage the exogenous attention. Therefore, it can be said that in such sentence processing tasks, automaticity seen in syntax processing reflects the maintenance of the endogenous component of attention throughout the task which overrides the exogenous capture.

4.3 Processing Sentences Under Dual Task

ERP studies have shown that certain syntactic processes are automatic since these processes do not seem to require attention. However, from this evidence one cannot be sure if processing sentences does not call for general executive resources (Logan and Gordon 2001). One important distinction between automatic and controlled processes is that automatic processes do not depend on executive resources. If sentences are processed without attention, then it is logically possible that they do not depend on general cognitive resources. One experimental strategy to examine this issue is to ask the participants to process sentences under a dual task scenario. This can be examined, for example, with a simultaneous working memory task. It is long known that maintaining something in working memory can be cognitively demanding (Engle 2002). Comprehending syntactically complex sentences while trying to keep a sequence of digits in mind can be difficult. The effect of working memory load on sentence processing in both normal and individuals with different cognitive disorders reveals different aspects of attention involvement (King and Just 1991). Engle (2002) provided evidence which suggests a close causal relationship between working memory capacity and executive attention. Participants who are better in working memory maintenance are also better in tasks that demand some form of conflict resolution, for example, the anti-saccade task or the classic Stroop task. This seems reasonable, since many conflicting situations want us to keep the cognitive goals in mind and take proper action. Brain imaging studies have shown that the fronto-paretial attention and executive control networks show activity when one engages in some working memory task (Coull et al. 1996). Further, it has been shown that spatial working memory and attention share psychological mechanisms (Awh and Jonides 2001). Therefore, engaging someone's working memory can impair a task if this task calls for central resources. Another research result shows that children with poor working memory do badly on sentence comprehension tasks. In this sense, loading the working memory should influence automaticity in the processing of sentences. In the literature, the executive control component of working memory has been known as 'attention' (Barrouillet et al. 2007) and others have called this 'controlled attention' (Engle et al. 1999).

Dual tasks require division of attentional resources. Caplan and Waters (1999) suggested that working memory functions as a temporary work pad where task specific information is kept for a short time and computation happens. Sentence processing requires such computational assistance. Importantly, their theory proposed a division in the working memory system. Their suggestion based on both normal and lesion data was that there is a specialized component of working memory which is specifically used for computation for syntactic structures. Therefore, a general deficit in working memory may not always predict a deficit in sentence processing. However, if two different cognitive processes are not tapping the same resource, then effects may be different. If sentence processing does not depend on attentional resources, then it is conceivable that their processing will not be affected under dual task conditions.

Sentence processing under concurrent working memory load results into slowness in comprehension. Some aspects of the performance will suffer since working memory will consume resources that might be required for processing the sentence. However, in order to make this assumption one has to accept the idea that processing sentences is dependent on the availability of cognitive resources, specifically, the type required by items maintained in working memory. The domain specific or domain general nature of working memory resources in sentence processing has been debated long. Caplan and Waters (1999) were the first to explicitly posit the idea that processing sentences may depend on working memory store. Individuals with low working memory span show deficits in sentence comprehension under concurrent working memory load (King and Just 1991). It has been observed that children's ability to process information in working memory positively correlates with sentence comprehension (Magimairaj and Montgomery 2012). Therefore, additional cognitive load can influence successful processing of sentences in a variety of situations.

Do complex sentences require higher attention resources? Psycholinguists often consider a sentence that has an object relative clause as a complex sentence compared to one that has a subject relative clause. For example, the sentence, 'The boy whom my friend knew is the son of my school teacher', could be considered a complex sentence. This belief stems from the assumption that complex sentences are transformed versions of simple sentences. Therefore, understanding them may require first transforming them back into simple sentences. Similarly, many consider passive sentences syntactically complex compared to active sentences. Sentences with initial objects have been shown to cause difficulties with parsing (Traxler et al. 2002). Sentences with centre embedding are structurally complex. One can consider lengthy sentences with many phrases and constituents as complex since they may call for more working memory resources for computation. The notion of symbolic computation requiring central processing resources is explicit in these theories.

Waters et al. (2003) asked university students to listen to syntactically complex and simple sentences under increasing working memory loads. Participants had to make a plausible judgement of the sentences while maintaining the digits in the working memory. The assumption was that working memory load would affect the complex sentences more, since processing them requires higher resources. However, the results showed that while there was a general increase in processing difficulty for all types of sentences with increase in the number of digits to be maintained, memory load did not affect simple and complex sentences selectively. This led the authors to suggest that there could be a domain specific working memory for sentences. This view logically entails the idea that attentional load cannot influence sentence processing. However, in contrast to this, others have produced evidence which suggests that items maintained in working memory can selectively influence processing of simple and complex sentences, showing dependence of sentences on a general purpose working memory reserve. Godern et al. (2002) asked participants to remember a list of nouns while they had to comprehend syntactically complex or simple sentences. Participants were asked to keep in working memory words like man, cat, river, boy, boot, etc. Some of these nouns later appeared at different places in a sentence. Words held in working memory interfered with other words that shared similarity. When words maintained in working memory matched words in the sentences, higher error rates were observed for complex sentences while this was not the case for simple sentences. This suggests that linguistic items in working memory created interference with processing of complex sentences and, therefore, sentence processing cannot be modular. It has been argued that increasing the number of digits in working memory may not be the right method to see effects on sentence processing; instead, it is better to have items that can create interference at a representational level with processing of sentences (Fedorenko et al. 2006). It is important to note that sentence processing performance is generally poor under both digit and noun loads. Further, a straightforward comparison between the domain general and domain independent working memory proposals is problematic since researchers have used sentences of different types and have used different load types.

Is there a specific type of executive control mechanism, i.e. memory, selective attention, which is used to understand complex sentences? In ambiguous sentences like ‘The horse raced past the barn fell’, ambiguity can arise when the word ‘barn’ is taken as the direct object of the first phrase and not as the subject of the second phrase. Reanalysis in such situations could engage executive control processes. One way to test language-specific executive control mechanisms is to seek correlations between performances in ambiguity resolution on such sentences and performance on some verbal attention task. It was recently observed that performance on a verbal Stroop task correlates with revisions in ambiguous sentences, whereas such correlations are not seen with non-verbal attention tasks (Vuong and Martin 2014). Executive control also influences the processing of figurative sentences (Columbus et al. 2014). This may suggest that those who have a superior ability to maintain goal-directed focus can do well in syntactic ambiguity resolution. An ability to quickly reanalyse a faulty sentence could stem from better working memory or even an ability to inhibit goal irrelevant information. However, from correlational studies it is often not clear if there is a specific type of attention control mechanism which is used for sorting out complex sentences; although it is clear from available evidence that understanding complex sentences and reanalysis calls for working memory resources.

Memory load can create problems for sentence processing in participants with neuropsychological conditions. Sentence processing with working memory load in patients who have cognitive deficits such as the Parkinson's disease (PD) patients reveal dissociation. Parkinson's disease is a progressive neurological condition where one sees impairments in memory and motor coordination. Patients with Parkinson's disease are also slow in delivery of speech. Grossmann, who has studied sentence processing in Parkinson's disease patients, believes that in these patients one may see a deficit in recruiting attentional resources during the computation of complex sentences. In his view, working memory resources required for sentence processing help in sorting out computational problems that come up when sentences are complex or of longer length. Therefore, it makes sense to consider working memory as playing a crucial role in the executive control system required for sentence processing (Grossman 1999). Parkinson's disease patients have executive function deficits in general and this is evident in many tasks they do. On the other hand, there is an involvement of working memory in sentence processing specifically, which shows interference. If the theory that executive functions contribute towards sentence processing is correct, then one can see errors in such patients, but not necessarily in tasks that call for working memory involvement. If the assumption that sentence processing requires the available cognitive resources is correct, then one can see some deficits in parsing in these patients, at least for the syntactically more complex sentences. fMRI studies have shown that PD patients recruit the important frontal and temporal areas that support working memory to a lesser extent while processing sentences, as compared to control subjects (Grossman et al. 2003). Since sentence comprehension—particularly for those sentences that have an object cleft—requires sorting out the long distance dependency problem, PD patients show deficits in this, as they also show working memory problems (Hochstadt et al. 2006). Huntington's disease is a neuro-degenerative disease where there is neural damage in the striatum (Vonsattel et al. 1985). Huntington disease patients may show deficits in memory retrieval, sequencing and reasoning abilities. In contrast to PD patients, patients with Huntington's disease do not show impairments in sentence processing with concurrent working memory load (Samblin et al. 2012). Interestingly, patients with Lewy body spectrum disorder (a form of dementia resulting in cognitive impairment) show sentence processing deficits when working memory load increases (Gross et al. 2012). Patients with the dementia with lewy bodies show executive control and visuo-spatial deficits (Horimoto et al. 2003).

It is important to note that dual task studies have exclusively paid attention to the processing of syntactically complex sentences compared to simple sentences while someone is holding some items in working memory. Therefore, it becomes difficult to exactly pinpoint the role attention might play in such circumstances, since it can be only indirectly inferred. Further, several researchers have used a wide range of sentence types as stimuli which makes direct comparison difficult. Nevertheless, if it is assumed that holding something in working memory is equivalent to attending to it, in this sense, working memory tasks can tell us something about attention. Lee et al. (2003) used a word detection task to see if Parkinson's disease patients were able to pay attention to certain phonetic errors embedded in sentences. This task does not call for sentence processing as such but some sensitivity to what the sentence contains. It was observed that Parkinson's disease patients are poor in noticing such errors or they are slow in reporting them. This makes sense, since noticing errors calls for alertness, which is an important aspect of the executive control system. Such evidence points towards the fact that when the general executive control capacities are poor, sentence processing could be affected. Similarly, Parkinson's disease patients do not show robust activation in the attention and control networks of the brain such as in areas like the anterior cingulate cortex (an important brain area whose functions include action control). It has been observed that such patients show limited activity in these brain regions during attentionally demanding tasks; for example, spotting errors in sentences (Grossman et al. 1992). The executive dysfunction in general creates deficits in the alerting mechanism and also in selective attention tasks.

This evidence from neuropsychology suggests that executive control mechanisms and not just attention play a critical role in language processing. Attention is a component of executive control processes that include goal-directed action, inhibition and task switching and a general level of alertness (Rubinstein et al. 2001). Therefore, if an experimental task has not explicitly manipulated selective attention, the task outcomes can be seen from the point of view of one or another executive control processes.

It would be an excellent idea to construct a theory where domain general executive resources serve both language and non-language material. This will allow predicting correlations between linguistic and non-linguistic tasks. For example, if someone has a weak attentional system or memory system, it can be predicted that s/he may show deficits in language processing. This is what we have just seen with Parkinson's disease patients who have a neurological condition that weakens the general executive control system. Children with ADHD (attention deficit hyperactivity disorder) often show problems with sentence processing as well as tasks that call for selective attention (Cohen et al. 2000). A domain general attention network seems to be supporting various aspects of language processing in different situations.

4.4 Influence of Language on Orienting of Attention

While attention can modulate language processing, language itself influences attentional allocation (Talmy 2000). Language has its own mechanisms to draw attention selectively towards some aspects and not others. For example, attention is paid to words that are mentioned first in a sentence. Spoken language itself has several mechanisms to mark 'focus' on several constituents. Speakers stress on particular word more heavily when more emphasis is required. This makes such an entity more focused and attention is summoned. 'Topic' and 'focus' in language have been the object of extensive discussions for long within psycholinguistics (Gundel and Fretheim 2004). For example, sentential subjects often occupy places of attentional significance. Similarly, certain prosodic cues help mark certain words in a sentence. Different languages use different mechanisms to mark topic and focus in sentences (see Gundel and Fretheim 2004 for an introduction). These mechanisms often show the intimate connection between pragmatic aspects of language and how they can influence structural choices. In passive sentences objects are brought into focus by placing them in the beginning of sentences.

Languages have different particles that encode focus (Lambrecht 1996). Take for example, the Hindi particle '*hī*' (ही). If I utter the sentence, '*Ram hī tha jo wahan gaya* tha', 'It was Ram who went there'. The word '*hī*' puts excessive focus and emphasis on the discourse about Ram. Depending on the discourse necessities, language users use either focusing participles or stress certain words for successful communication. An interesting proposal will be if such discourse entities recruit the attentional network during language comprehension. In an interesting demonstration dubbed as the semantic illusion effect, participants do not often catch errors in sentences such as, 'How many animals of each kind did Moses take on the Ark?' (Erickson and Mattson 1981). The correct person should have been Noah and not Moses. This slip is possible since participants do not pay attention to every anomaly in speech and, moreover, in this case Noah and Moses belong to the same semantic field. However, when the anomalous word Noah is focused with a prosodic cue or with additional particles like 'It was Noah', listeners immediately become aware of the problem. This suggests that information structure devices, i.e. prosody, particles, immediately recruit the attention of the listeners. There is a caveat here. Which means, as such, during normal conversation listeners do not pay attention and only do so when some elements are excessively focused.

Kriestensen et al. (2013) explored how manipulating prosodic structures of speech could involve attentional networks during spoken language manipulation. The idea was that when people listen to some sentence under normal circumstance, they often do not pay attention to each element. Interestingly, they also do not seem to be attentive to major semantic flaws. Kriestensen et al. presented sentences that either contained a semantic anomaly or was correct. Further, on some trials, the anomalous words were prosodically stressed to make them unique. The prediction was that participants will immediately attend to the semantically problematic elements when they are focused compared to when they are not. And this will cause the general attentional network of the brain to show activation. The above-mentioned researchers administered several spatial attention tasks to the participants to see if the same attention network is active for linguistic and non-linguistic stimuli. The results showed that important attention areas like the superior/inferior parietal cortex, superior and middle temporal cortex, as well as inferior, middle, and posterior parts of the frontal cortex are active for both types of stimuli. Participants used a domain general attention network to find out the errors in the spoken language. This result demonstrates that attention networks in the brain are summoned to sort out issues when certain elements demand so. Further, errors are easily detected when selective attention is paid to the problematic parts. The important thing here to note that language itself has many devices that recruit attention selectively.

In another closely related study, Li and Ren (2012) examined if accenting a word would engage the attentional system more compared to when words are not accented at all or even de-accented. Modern techniques in speech manipulation offer these possibilities to entertain. Li and Ren presented sentences that contained a semantically anomalous word. It is well known in the ERP literature that words that cannot be semantically integrated into the context lead to a specific component known as the N400 effect. The latter indicates the failure to integrate upcoming information into an already built context. It was predicted that when the semantically anomalous word is accented, it would engage the selective attention network. The results showed that the N400 effect had a faster latency when the word was heavily accented and the effect was almost non-significant when the word was de-accented. The findings do indicate that the prosodic quality of speech engages the attention network and this in turn facilitates language processing. Long back, Cutler and Foss (1977) had observed that items that were stressed in sentences were processed faster. Many might object that these manipulations appear too artificial and the effects could be just specific to certain experimental situations. However, it is a common observation that when the speaker emphasizes certain words, we pay higher attention. Deployment of selective attention also facilitates processing of sentences that are acoustically degraded. Many a time, we listen to sentences in noisy environments. Imagine trying to listen to someone on a railway platform when a train is passing by. In this situation, the attentional networks show higher activation, since the perceptual system can focus and retrieve some of the information that could be missed (Wild et al. 2012). Wild's study found activation of the attention network for sentences that contained noise whereas there was no attention requirement for sentences processed under normal conditions.

There are a few things to sum up in this section. Studies have shown that a general attentional network is used when certain linguistic features draw attention. This facilitates a range of processing. Further, people use attention to comprehend speech when there is noise. The general purpose attentional mechanism for both linguistic and non-linguistic stimuli can help to interpret findings from many studies. In the sentence processing studies of attention, attention has been considered both as a limited resource process as well as a mechanism that helps select the right stimuli. It has been implicitly assumed that attending to one stimulus will result in non-attendance to another. This may not be true since it is possible to attend to stimuli in parallel. Attention to some highly stressed word in a sentence might indicate bottom-up capture of attention. At the same time this attention also facilitates extraction of meaning. It is often possible to pay selective attention to some aspects of the sentence. Attending voluntarily to the structure or meaning part of the sentence affects processing of the sentence. We know that task instructions can affect how a language processor selects to attend to stimuli. Rogalsky and Hickok (2009) asked people to report for either semantic or syntactic errors that may be present in sentences. This technique is different from the traditional way of asking to make a grammatical judgement. In this way, the participants are expected to pay selective attention to either the meaning or the grammar part. The brain imagines results show that an area known as the anterior temporal lobe. was active for both types of tasks. Some others have suggested that the anterior temporal lobe is involved in syntactic computation apart from the Broca's area (Brennan et al. 2012). However, this neural structure was selectively active when participants were looking for the grammatical errors. This evidence suggests that attention to specific linguistic events leads to superior processing.

4.5 Conclusion

In this chapter we saw that many have found early syntactic processing to be automatic. Evidence using the ERP method suggests that attention is not necessary to compute the phrase structure of a sentence. However, it seems likely that for processing semantics one may need some attention. Normally, we do not feel this dissociation during every day speaking activities. It is likely that the syntactic automaticity conceptualization is an outcome of the syntax versus semantics debate in linguistics. Those who thought that syntax is rule-governed and computational as well as should happen before semantics have argued for its automaticity. It is still not clear if the methodologies used to control for attention are foolproof against many other external factors. For example, the manipulation of trials or dual tasks may induce other constraints in the experiments. Additionally, many recent studies have indicated individual differences in sentence processing and attentional allocation (Rogalsky and Hickok 2009). Additionally, eye tracking studies show a dynamic interaction between attention and linguistic processing. Automaticity, here, means processing in the absence of attention. However, many in attention research may argue that there is no such thing in the cognitive processing domain which does not require attention. It is very difficult to prove this in our context. With available knowledge it can be safely said that some aspects of sentence processing seem to proceed without explicit allocation of attention while some other aspects require attention. It is likely that different levels of language interact with attention differently. This is a good hypothesis since now we know that linguistic levels constantly interact and influence one another.

References

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5*(3), 119–126.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*(3), 570.

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language, 120*(2), 163–173.

Caplan, D. (1996). Syntactic processing in sentence comprehension under dual-task conditions in aphasic patients. *Language and Cognitive Processes, 11*(5), 525–551.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(1), 77–94.

Cohen, N. J., Vallance, D. D., Barwick, M., Im, N., Menna, R., Horodezky, N. B., & Isaacson, L. (2000). The interface between ADHD and language impairment: An examination of language, achievement, and cognitive processing. *Journal of Child Psychology and Psychiatry, 41*(03), 353–362.

Columbus, G., Sheikh, N. A., Côté-Lecaldare, M., Häuser, K., Baum, S. R., & Titone, D. (2014). Individual differences in executive control relate to metaphor processing: an eye movement study of sentence reading. *Frontiers in human neuroscience*, *8*.

Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory. *Neuropsychologia, 34*(11), 1085–1095.

Cutler, A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech, 20*(1), 1–10.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. *Models of working memory: Mechanisms of active maintenance and executive control, 4*, 102–134.

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior, 20*(5), 540–551.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language, 54*(4), 541–553.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology. MIT press*. Boston, MA: The MIT Press.

Frazier, Lyn. (1988). The study of linguistic complexity. In Alice Davidson & Georgia M. Green (Eds.), *Linguistic complexity and text comprehension: Readability isues reconsidered* (pp. 193–221). Hillsdale, NJ: Lawrence Erlbaum Associates.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews, 91*(4), 1357–1392.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13*(5), 425–430.

Gross, R. G., McMillan, C. T., Chandrasekaran, K., Dreyfuss, M., Ash, S., Avants, B., & Grossman, M. (2012). Sentence processing in Lewy body spectrum disorder: The role of working memory. *Brain and cognition*, *78*(2), 85–93.

Grossman, M. (1999). Sentence processing in Parkinson's disease. *Brain and Cognition, 40*(2), 387–413.

Grossman, M., Carvell, S., Stern, M. B., Gollomp, S., & Hurtig, H. I. (1992). Sentence comprehension in Parkinsons disease: The role of attention and memory. *Brain and Language, 42*, 347–384.

Grossman, M., Cooke, A., DeVita, C., Lee, C., Alsop, D., et al. (2003). Grammatical and resource components of sentence processing in Parkinson's disease: An fMRI study. *Neurology, 60*(5), 775–781.

Gundel, J. K., & Fretheim, T. (2004). Topic and focus. *The handbook of pragmatics, 175*, 196.

Gunter, T. C., & Friederici, A. D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology, 36*(01), 126–137.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11* (2), 194–205.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13*(3), 339–356.

Hasting, A. S., Winkler, I., & Kotz, S. A. (2008). Early differential processing of verbs and nouns in the human brain as indexed by event-related brain potentials. *European Journal of Neuroscience, 27*(6), 1561–1565.

Horimoto, Y., Matsumoto, M., Nakazawa, H., Yuasa, H., Morishita, M., Akatsu, H., & Kosaka, K. (2003). Cognitive conditions of pathologically confirmed dementia with Lewy bodies and Parkinson's disease with dementia. *Journal of the neurological sciences*, *216*(1), 105–108.

Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language, 97*(3), 243–257.

Jeon, H. A., & Friederici, A. D. (2015). Degree of automaticity and the prefrontal cortex. *Trends in cognitive sciences, 19*(5), 244–250.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. *Attention and performance IX, 9*, 187–203.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580–602.

Kristensen, L. B., Wang, L., Petersson, K. M., & Hagoort, P. (2013). The interface between language and attention: prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex, 23*(8), 1836–1848.

Lambrecht, K. (1996). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Vol. 71). Cambridge, UK: Cambridge University Press.

Lee, C., Grossman, M., Morris, J., Stern, M. B., & Hurtig, H. I. (2003). Attentional resource and processing speed limitations during sentence processing in Parkinson's disease. *Brain and Language, 85*(3), 347–356.

Li, X. Q., & Ren, G. Q. (2012). How and when accentuation influences temporally selective attention and subsequent semantic processing during on-line spoken language comprehension: An ERP study. *Neuropsychologia, 50*(8), 1882–1894.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*(2), 393.

Magimairaj, B. M., & Montgomery, J. W. (2012). Children's verbal working memory: Relative importance of storage, general processing speed, and domain-general controlled attention. *Acta Psychologica, 140*(3), 196–207.

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience*, *23*(9), 2252–2267.

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology*, *67*.

Mole, C. (2011). *Attention is cognitive unison: An essay in philosophical psychology*. New York: Oxford University Press.

Myers, J. (2009). Syntactic judgment experiments. *Language and Linguistics Compass, 3*(1), 406–423.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology, 118*(12), 2544–2590.

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52*(1), 629–651.

Posner M. I., Snyder C. R. R. (1975). Attention and cognitive control. In Information Processing and Cognition: The Loyola Symposium, (Ed.)R. L. Solso (Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 55–85.

Pulvermüller, F., Shtyrov, Y., Hasting, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language, 104*(3), 244–253.

Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex, 19*(4), 786–796.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27* (4), 763.

Salverda, A. P., & Altmann, G. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1122.

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: evidence from Hindi-English bilinguals. *Frontiers in Psychology*, 4.

Singh, N., & Mishra, R. K. (2015). Unintentional Activation of Translation Equivalents in Bilinguals Leads to Attention Capture in a Cross-Modal Visual Task. *PLoS ONE, 10*(3), e0120131.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language, 120*(2), 135–162.

Szymanski, M. D., Yund, E. W., & Woods, D. L. (1999). Phonemes, intensity and attention: differential effects on the mismatch negativity (MMN). *The Journal of the Acoustical Society of America, 106*, 3492.

Talmy, L. (2000). Towards a cognitive semantics: Vol. I. Conceptual structuring systems. Cambridge, MA: MIT Press.

Traxler, Matthew. (2011). *Introduction to Psycholinguistics: Understanding Language Science*. Wiley-Blackwell: Chicester, UK.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*(1), 69–90.

Vuong, L. C., & Martin, R. C. (2014). Domain-specific executive control and the revision of misinterpretations in sentence comprehension. *Language, Cognition and Neuroscience, 29*(3), 312–325.

Vonsattel, J. P., Myers, R. H., Stevens, T. J., Ferrante, R. J., Bird, E. D., & Richardson, E. P, Jr. (1985). Neuropathological classification of Huntington's disease. *Journal of Neuropathology and Experimental Neurology, 44*, 559–577.

Waters, G., Caplan, D., & Yampolsky, S. (2003). On-line syntactic processing under concurrent memory load. *Psychonomic Bulletin & Review, 10*(1), 88–95.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience*, *32*(40), 14010–14021.

Chapter 5
Attention in Speech

5.1 Interference, Inhibition and Naming

Most cognitive processes result from competition (Miller and Cohen 2001). Selection of the goal appropriate concept is critical to naming and speaking. Speaking is also an intentional act (Levelt 1989). That is, the speaker must have an intention to speak. Selection of concepts is influenced by both top-down and bottom-up features that influence processing in naming. It has been a issue of debate how speakers choose the right concept and words to speak amid many. Target objects are often surrounded by many objects, whose names might resemble one another. For example, during the naming of 'cat', the picture of a 'bat' can cause some subtle interference as well as the picture of a 'dog' (Schriefers et al. 1990). That is because these objects either share phonology or belong to the same semantic field. It is interesting to examine how the language production system handles such interference. Many studies have suggested that naming isolated objects often calls for some kind of attention selection where executive control is necessary. For instance, executive control abilities such as inhibition measured through a stop signal task (a task where a response has to be stopped on some trials given a cue) significantly predicted naming latencies of both objects and actions (Shao et al. 2012). Previous work on single word production and naming did not consider the role of attention and other cognitive resources very seriously since many believed that these process are automatic (Coltheart 1999). However, it is now being observed that selective attention could be important in the early stages of naming leading to formation of concept and phonological word form retrieval (Griffin and Bock 2000).

One proposal has been that lexical selection proceeds through the inhibition of competing concepts (Schriefers et al. 1990; Bock and Levelt 2002). Speakers are not consciously aware of such subtle interference during naming but latencies in picture naming shows some delay as a result of competition. The necessity of selective attention arises when one has to name a particular concept or word form

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_5

that has to be selected from many competing ones. Naming need not always be automatic and involuntary and may call for deliberate processing (Bloem and Heij 2003). Executive control is necessary to select the target concept and utter its name (Roelofs et al. 2013). It is important to note that the engagement of attention is contingent on the notion of parallel activation of concepts and interference. Naming a picture as such may not involve attention unless the name to be uttered has a large number of semantic or phonological competitors (Kan and Thompson-Schill 2004). Since activation of several related concepts with any input word or picture seems to be the norm, it is important that the executive control system helps in selecting the right response (de Zubicaray et al. 2001). Spreading activation is a very commonly observed phenomenon in most psycholinguistic studies.

The now classic picture naming task (Levelt 1989) has been used to examine the nature and locus of competition during the naming process. Given a picture to name, mental representation of objects arises through the visual perception and then one accesses the phonological form and finally articulates (Meyer et al. 2012). This follows from a serial view of name generation where conceptual access leads to retrieval of form. There is also evidence that language production is an incremental and predictive process that has strong links with the comprehension processes (Dell and Chang 2014). Thus, theoretical debates on the processes of word planning and speaking differ with regard to notions like interference and inhibition. Discrete and serial models assume that naming begins with the formation of an intention to generate a label for an object. While naming an entity, assuming that its name is known, at first the speaker would generate a message plan and retrieve a lemma (Levelt 1993). A lemma has been considered to be an abstract entity that contains information about the word's syntactic properties and also information related to gender and number (Levelt 1989). This abstract conceptualization is then used to frame the appropriate syllabification and phonological form retrieval. For example, with the picture of one cat, the phonological form is 'CAT' and not 'CATS'. This kind of specification in the speaking process evolves over time (Indefrey and Levelt 2004; also see Strijkers and Costa 2011). Therefore, from the visual perception of a picture till the production of its name, there are a series of incremental processes, largely encapsulated from one another and each process is completed in some specified time frame (Indefrey and Levelt 2004).

This serial account of language production thus consists of two discrete steps. In the first stage which is semantically driven, information related to syntactic and semantic properties is retrieved. In the second stage, phonological encoding takes place before being articulated. Therefore, this model would assume semantic and phonological interference affecting production process at two distinct stages. Schriefers et al. (1990) used a picture naming task with auditory stimuli that provided interference. They observed a semantic facilitation at an early timescale, i.e. within 150 ms, and a phonological facilitation after 150 ms. The result suggested that different types of information affect word naming at different time points. The WEAVER++ model (Word-form Encoding by Activation and VERification) of single word planning and production makes the explicit assumption that word planning may require some form of attentional involvement (Roelofs 1997). It is a new generation model of picture naming which also includes attention control and response selection. Thus, executive control processes are engaged during naming. It is important to note that attention here is part of the large executive processes.

However, there have been alternative proposals to language production that emphasize an interactive architecture and parallel processing of naming (Rapp and Goldrick 2000). The interactive two-stage model (Dell et al. 1997) assumes that during lemma selection, phonological information is active in a cascaded manner and during the processing of phonological information, feedback connections exist to the lexical level. Thus, information flow is continuous where multiple attributes of words belonging to different processing levels can affect one another. These models do not assume a step-by-step processing that progresses through clearly defined time windows, as one would get from the model of Inderfry and Levelt (2004; see Fig. 5.1). Evidence in support of the interactive models comes from data that show influence of both prior conceptualization processes on articulation as well as influence of articulatory features on conceptualization (Dell 1986). That is, information flow is truly bidirectional. Baese-Berk and Goldrick (2009) found that phonetic information affects lexical access during language production, suggesting interactivity. Phonetic properties of particular sounds affect production in a cascaded manner (Goldrick and Blumstein 2006). These examples suggest that different types of information dynamically affect language production and this process may not necessarily be a serial process. Strict encapsulation between production modules and a unidirectional flow of information is thus challenged (Vigliocco and

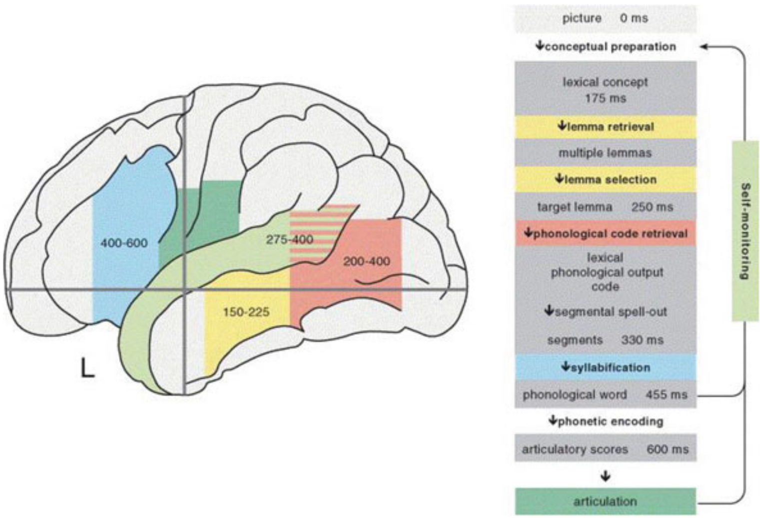

Fig. 5.1 Figure shows cortical areas involved during naming process and the time course for speaking assuming a serial process. Meta-analysis of several brain imaging studies of picture naming has shown that there is a time bound completion of distinct processes starting with an intention and ending with articulation (Reproduced from Indefrey and Levelt (2004, p. 121) with permission)

Hartsuiker 2002). Data from speech errors and experiment-induced tongue twisters suggest interactivity among levels as well as connectionist simulation studies (Harley 1993). It is likely that depending on how one views the word production process, i.e. as discrete or cascaded, the involvement of attentional mechanisms is considered.

Below I will attempt to examine the role of selective attention in naming with regard to notions of competition and interference, as discussed earlier. Psycholinguistic studies have used the picture-word interference paradigm to study this kind of difficulties in planning and production. It has been observed that there could be some delays in naming when there is some semantic or phonological competition from another concept (La Heij 1988). This sort of thing happens when, for example, one is asked to name the picture of an elephant while ignoring the word giraffe written on it (Fig. 5.2). The role of attention in the picture-word paradigm has been looked at from the lexical competition account (Levelt 1999) or the response exclusion account (Mahon et al. 2007). The selection via competition account predicts increased reaction times in naming in the presence of semantic distractors while the response exclusion account does not predict any interference from semantic distractors. Miozzo and Caramazza (2003) found that high frequency unrelated distractors led to faster naming latency compared to low frequency distractors. This effect of frequency on naming latencies cannot be explained by accounts that propose competition. While the first account would postulate a competition between lexical items and some inhibition later via the executive control system, the second account proposes that interference arises since written words gain entry into the articulatory buffer automatically. A buffer is a place where information is stored temporarily. The selection via competition account invokes inhibitory control mechanisms while the response exclusion model assumes a top-down, a priori selection of the correct word without involving much competition. Below we examine the evidence concerning these positions as both of these views have used data from the picture-word interference paradigm.

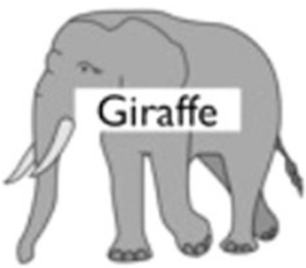

Fig. 5.2 The figure shows a classic picture-word interference stimulus. The participant is asked to name the picture while avoiding the written word. In this case both the picture and the written words belong to the same semantic field. Picture-word interference effects differ depending on the relationship between the picture and the word

In picture-word interference tasks, however unnatural they may be, one sees interference since it is the activation of semantics of the competitor that creates conflict. This task as such does not examine the involvement of selective attention in any manner but only indirectly shows the inefficiency in goal maintenance in the face of conflict. Participants are asked to name the picture while ignoring the written words, since these words are task irrelevant. Since written words automatically activate meaning, this leads to interference or facilitation with the naming of the picture. Competition is intense when the distractor picture is semantically related to the target since this can lead to difficulty in the selection of the desired lemma (Glaser and Düngelhoff 1984). The claim is that extracting the meaning 'giraffe' from the written word will directly compete with the extracted concept ' elephant' from the picture and naming will be delayed.

Speaking being an important activity related to the expression of thought, speakers must exercise extra caution unlike listeners whose activities may appear to be somewhat passive. The speaker initiates the communication process while exercising their executive control over this process. Speaking thus could be considered an important manifestation of planning and acting out our intentions. The whole action control system therefore comes into activity when one plans and articulates speech. Brain imaging studies have revealed higher activation in the frontal attentional networks and particularly in the ACC (anterior cingulate cortex) in such a situation that involves conflict in action planning (Thompson-Schill et al. 1997). ACC is largely known to bring down conflict when it arises because of two competing representations and only one correct response has to be selected. In a direct examination of the involvement of selective and non-selective inhibition in the picture-word interference paradigm, Shao et al. (2013) asked participants to name objects while ignoring superimposed words on line drawings that were either semantically related or were neutral. Participants also did a stop-signal task that measured their non-selective inhibitory mechanisms. The results showed that performance on the stop-signal task was related to overall mean reaction times of word production. The authors also observed that those participants who suffered less in the semantic interference (latency difference between related and unrelated conditions) also applied higher selective attention. This suggests that in picture-word interference tasks, both selective and non-selective inhibitory mechanisms may be at work and they may be potentially dissociable. Further, it has been shown that domain general executive control processes contribute to naming of objects and actions (Shao et al. 2012).

While a picture-word interference paradigm may be sufficient to show semantic interference, it says little about the attention–language interaction, since attention per se has rarely been manipulated. In the picture-word paradigm, written words attract attention automatically. It is likely that pictures are surveyed later in the sequence and this causes processing delay. Studies with ERPs suggest that conceptual information from pictures is extracted very fast, i.e. within the first 100 m, and this information then comes into conflict with meaning extracted from written words. An alternative method that could be to present the written word may be at variable SOAs (Stimulus Onset Asynchrony) with the onset of the picture

(Dell'Acqua et al. 2010). That is, the written word appears after some delay from the onset of the picture. Lupker (1979), after investigating various parameters that could influence picture-word interference effects, suggested that the locus of the interference lies at the response level competition. Competition during stimuli selection can happen at many levels. One way is to imagine that only one alternative is considered from the very beginning. Another way would be to assume that both alternatives are entertained till the final response is selected amid competition. The semantic or associative relationship between the picture's name and the word is not the only cause for interference. Phonological similarity between the picture and the word can also cause interference. Lupker (1979) suggested that one can see interference in this paradigm if the written word is 'relevant' to the main task. Interestingly, it has also been the case that one observes interference when the picture and the word are from the same semantic category, i.e. CAR and TRAIN, but there is facilitation if this written word is semantically associated one but not from the same category: 'BUMPER' (Costa et al. 2005). Thus, executive control processes are modulated depending upon task effects as well as the linguistic relationship between the target and the competitor.

Schnur and Martin (2012) used the psychological refractory period paradigm to study if degree of semantic interference in this paradigm is modulated by temporal gap between this task and prior decision-making task. Participants made judgment about the frequency of a tone and then named the picture while ignoring either a semantically related or unrelated word. The naming task came after a variable SOA which was either long or short. The interference effect was similar for both long and short SOAs, suggesting that semantic competition occurs at a later response selection stage. When Aristei et al. used a picture-picture paradigm (Aristei et al. 2012) in a novel manner to examine the claims of the response exclusion hypothesis, they observed interference with context pictures that were semantically similar. In this study, participants saw two pictures and named them as a compound, i.e. apple and cherry. Naming latencies for the first picture were higher when the second picture belonged to the same semantic field. It is important to note that the response exclusion hypothesis does not say anything about picture stimuli since picture names would not find a place in the buffer but only written words. Interference is observed when the participant is asked to ignore the interfering word in a picture-word paradigm. Roelofs (2008) asked participants to name pictures while ignoring another distractor picture, and found interference. However, this interference disappeared when participants were asked to name both the pictures. This suggests that when attention is sustained on the interfering picture, interference is less. This may indirectly mean that in the conventional picture-word interference task, the involvement of attention is regulated by the nature of the stimuli and task demands. Therefore, attention control in this scenario, as far as it is used as a representative paradigm to study lexical access and naming is controlled in a top-down manner. It further makes sense considering the fact that naming being an intentional process is goal directed and, therefore, top-down control should help avoid distractors and select the relevant target words for speaking. There is evidence which suggests a top-down task decision may be more helpful in getting the right word out during the naming process (Piai et al. 2011). Alternatively, it is possible that the selection does not happen beforehand and both the correct and incorrect responses are entertained for some time and then after some evaluation by a monitoring system, the wrong response is inhibited (Dhooge and Hartsuiker 2011). An emphasis on monitoring explicitly suggests a role of later response exclusion (Dhooge and Hartsuiker 2012). Whatever it is, pre-selection or post-exclusion, it seems that executive control system plays a very important role in the word naming process. It is important to note that attentional mechanisms could play causal roles differently depending on task manipulations. Both behavioural and ERP as well as brain imaging experiments show that interference happens due to word and picture similarity at some level and more certainly when they are semantically from the same field.

Selection through competition indicates that the attention system influences the word naming activity in a dynamic manner. Examples considered so far come from monolingual naming tasks. However, many interesting insights into the inhibitory mechanisms operative during picture naming and speech production have come from studies with bilinguals. Further, parallel activation of lexicons during language processing in bilinguals is now an established fact (Mishra and Singh 2014; Marian and Spivey 2003). It is observed that when a bilingual sees a picture s/he also activates the name in his/her other language and this parallel activation leads to a conflicting situation (Gollan et al. 2005). Bilinguals exercise a general-purpose executive control system to exert inhibitory control on the word node that is not tasking relevant (Green 1998). The model assumes an inhibitory control mechanism which resolves competition during word generation. Bilinguals seem to have superior inhibitory control mechanisms even in non-linguistic Stroop task (Singh and Mishra 2012). However, Costa and Caramazza (1999) found evidence that with English–Spanish and Spanish–English bilinguals there is no interference in lexical selection. These bilinguals were asked to name pictures either in Spanish or English while avoiding interference of the type described above from a written word printed on the picture. Naming latency was the same irrespective of the language of the word and facilitation was observed when this word belonged to the same semantic field as the picture. These effects were also generalized with Spanish–Catalan bilinguals who showed the same interference with the same language and different language distractor words and higher facilitation when the distractor word was from the same semantic field and from the same language (Costa et al. 1999).

In sum, studies with picture naming and with variants of the picture-word interference task suggest that lexical competition is commonly noticed during naming. Further, general executive control mechanisms are recruited during planning and language production as well as monitoring processes. Individuals, depending on whether they are bilinguals or monolinguals, entertain interference to different extents.

5.2 Speaking, Looking and Attending

Studies with the picture-word interference paradigm or the simple object naming paradigms do not tell where exactly people are looking during the preparatory process or naming; or, in other words, if selective attention to the visual stimuli is necessary for naming. The studies with picture-word interference paradigm do not clearly discuss the nature of attentional involvement per se. However, they do show competition arising out of the spreading activation of concepts. Further, the debate between the phonological and semantic nature of competitions and their timings has not been resolved in this paradigm. What it shows is that object identities enter into perceptual awareness even without our intention and start affecting behaviour. In the picture-word interference paradigm the distractor picture too receives attention because of its very close proximity to the to-be-named picture. Therefore, it is not clear what would happen if it is possible to control attentional engagement to this distractor picture. It is not clear what role executive control plays in maintaining attention only on the object to be named. Unfortunately, attention always spills over in the visual domain from one point to another. Objects that are spatially close by also receive some attention. And this spill-over attention sometimes causes conceptual activation sufficient to give rise to phonological form. The use of eye tracking as a methodology has revealed the dynamics of attention interaction during object naming.

Eye movements have long been known to reflect attentional mechanisms (Rayner 1998) and for a decade or so, several researchers have been using them to study attentional and visual processes in a range of linguistic activity such as speaking, listening and reading (Mishra 2009). Movements of the eyes through fixations and saccades could tell someone about the internal cognitive states of the viewer and also something about the focus of attention. Eye movements are also interesting since they can tell when someone has shifted his attention more explicitly from one location to another (Hoffman and Subramaniam 1995). While speaking, visual processing of objects and events provide the necessary conceptual structures. Interestingly, it has now been possible to examine how shifts of eye movements during naming can be correlated with levels of linguistic processing (Meyer et al. 1998).

Eye tracking during object naming can tell, for example, how long it takes for visual analysis, can derive conceptual information and phonologically encode the stimuli (Meyer and Lethaus 2004). The WEAVER++ model of spoken word planning assumes that it takes about 200 ms to see a picture and retrieve its phonological name for production (Indefrey and Levelt 2004). Fixation durations on to-be-named objects can thus tell how long it takes to reach the phonological stage when trying to name a visually presented picture. Since naming an object is an intentional activity, looking at a to-be-named object should engage selective attention and facilitate lexical access (Strijkers and Costa 2011) that in turn affects eye movements. Eye tracking has also been used to examine how multiple objects are named (Morgan and Meyer 2005). These issues have been recently addressed by researchers to examine attention involvement in spoken word planning (Griffin and Bock 2000).

An ability to grasp information from the parafoveal visual field (Rayner 1998) suggests parallel processing of both the stimuli to some extent. Compared to the picture-word interference type studies, eye tracking studies have controlled better for the fovea–parafoveal distinction. Participants look systematically and shift their gaze depending on several visual as well as linguistic factors. The claim is that gaze shifts occur only after phonological encoding on the foveated object has been complete. Crucially, while phonological encoding is still being getting processed, speakers seem to be already processing some semantics and features of the adjacent object. This type of parafoveal pre-processing certainly minimizes efforts later as speakers can be faster in giving a response as soon as they have looked at it more explicitly. Given two objects to name, speakers continue to look at the first picture as long as it is necessary to retrieve its name, then shift their gaze towards another location only when this process is reasonably complete. Further, studies show that psycholinguistic property of the second to-be-named picture such as its semantic similarity with the first, its length and frequency could affect the time speakers take to look at the first picture (Morgan and Meyer 2005). This can happen in an experimental situation when the second picture is within the parafoveal view and the speaker can derive much information without moving the eyes towards it. It appears that focal attention must be paid for some amount of time in order to retrieve phonological information during picture naming and eye movements can indicate that.

Interestingly, it has been observed that, given two different pictures to name one after the other, speakers first look at the first picture for about 200 ms and then shift their gaze towards the second picture (Meyer et al. 1998). The syllabic length of the word to-be-named is also important for the amount of time someone needs to look at this picture. The longer the word in terms of phonological units, it might take more time to extract this information (Meyer et al. 2003). However, highly frequent and familiar longer words are easily retrieved. There is some evidence which suggests that speakers can retrieve phonological and conceptual information from two objects in parallel. Malpass and Meyer (2010) found that when the second object was easy, speakers looked longer at the first object compared to when the second object was difficult. This suggests that allocation of attention in parallel over two objects is affected by the naming complexity linked with the objects. Do speakers shift attention to be named objects in a multiple object naming scenario only when they have fully done with the first object? Recently, Schotter et al. (2014) manipulated the available preview of a to-be-named object, making it contingent on attention shift from the first object. It was observed that speakers used preview benefit from objects that were in non-foveal regions even much before they shifted attention to this object or programmed an eye movement. This shows that attention shift between multiple objects during naming is a complicated process that is affected by visual and linguistic complexity of objects, location of objects and the task demands, i.e. name aloud or silently. The amount of time one looks at the first picture can be indicative of how much effort it takes to retrieve the phonology.

Interestingly, people sometimes already have a vague idea of what is the other picture to the left using their parafoveal vision. A variety of other experimental data points towards the fact that speakers are somehow able to deploy parafoveal attention towards the second picture even when they are processing the first picture (Meyer et al. 2003). If pictures are shown side-by-side speakers try to look at the first picture till they are sure of its phonological name, and then shift their gaze to the second one. Attentional resources, therefore, are divided depending upon the task goals, and this is a dynamic process. It has also been observed that speakers are faster in naming an object if they are primed with preview information about the location where the object is to be (Schotter et al. 2013). In the Shotter et al. study, researchers explored if intention to name any object at any location should influence the eye movements. There were line drawings on the screen positioned at all corners and subjects were asked to name only some and not others. As participants started naming one object and made a saccade towards another, this object was changed. This was a gaze contingent paradigm where during the flight of the saccade from the first object to the second object there was this change. The results showed that if participants did not wish to name an object in a particular location, then they did not pay attention and therefore somehow this object skipped their awareness. Performance on monitoring tasks like operation span tasks is found to be correlated with naming speech as well as interference during a dual task (Piai and Roelofs 2013). Thus, attention to objects as well as locations of objects helps speakers retrieve the required conceptual and phonological information for planning articulation.

While fixation durations on the foveal object and the timing of shift of attention to the second object suggest a dynamic relationship between lexical access and eye movements, it says little about if selective attention is necessary for naming visual objects. Lamers and Roelofs (2011) presented participants with a picture to be named on the left visual field and an arrow, flanked by symbols, on the right. Participants were asked to name the picture and then report the direction of the arrow, giving a manual response. Thus, the task used a dual task paradigm putting selective attention under some constraints. In order to be successful in this task, the participant would have to keep two task goals in mind and create some type of division of attentional labour. The interest was to see how long participants take to look at the picture before they explicitly shift their gaze towards the arrow, when the SOA between the appearance of the picture and the arrow is manipulated. Interestingly, speakers took more time to respond to the arrows when the time gap between the arrow and the picture was less compared to when it was say 1,000 ms. This shows that speakers need to attend to the picture for a certain amount of time in order to retrieve its name and then they can attend to other tasks and take complex decisions. Gaze shifts towards the arrow were determined by the amount of time it took to phonologically encode the picture. Such an experimental demonstration seem to suggest that given two complex tasks to be performed one after another, human subjects modulate the attentional resources required for each task in order to be highly successful. Attention helps in creating goal plans and goal directed action. When attention is directly manipulated in a dual task scenario, performance crucially depends on the availability of resources and task demands.

The question of whether central attention is absolutely a must for conceptual understanding and phonological name retrieval has also been examined with regard to automatic and strategic processes. One might assume that the stages of early conceptual information retrieval and phonological encoding may be automatic processes, something along the lines of what researchers in the sentence processing domain have claimed for early stages of parsing (Friederici 2002). If the phonological form retrieval stage is automatic and does not need central attention, then one can administer a concurrent task to test it. Ferreira and Pashler (2002) asked participants to name pictures while simultaneously doing a tone discrimination task. In one experiment participants first did a cloze task and then named pictures. Cloze tasks are used to ask participants to complete a sentence that has one or more words missing. For instance a participant is given the sentence "Every morning I like to drink my—with sugar". In this example, the participant has to supply the missing word 'coffee' using contextual knowledge. This task checks language comprehension ability. Thus, experimenters know the most likely words that language users can come up with, given a fragment. It was observed that tone discrimination as well as naming latencies went up when the picture names and the cloze words shared a frequency relationship with cloze sentences. In another experiment, participants named pictures with simultaneously presented distractor words. In both experiments, it was evident that one cannot do a concurrent task while naming pictures and, therefore, naming might require central attention. Selective attention seems also to facilitate word production when pictures are surrounded by distractors (Nozari and Dell 2012).

To summarize, it is clear that central attention is required for generating words from pictures in many cases. During visually presented picture naming, linguistic processes affect attention and eye movement measures. Evidence from eye tracking studies as well as studies with other behavioural paradigms indicates that some amount of attention is necessary for speaking out names.

5.3 Executive Control, Monitoring and Naming

Speaking is a linguistic action that involves intentional selection of a goal and timely execution. Speaking also involves resolving conflicts that may arise due to multiple activations of concepts. Many important areas in the brain linked to conflict management and action control show activity during speaking (DeLeon et al. 2007). This is because psycholinguists assume that to name an object the speaker must first retrieve the concept of this object following an intention to name and then prepare for articulation. From conceptualization till final articulation, several important brain areas such as the frontal, the language areas and the cerebellum participate (Indefrey and Levelt 2004). The question is, does intention to name an object imply attention selection and does it call for the exercise of executive control processes? It has been observed that areas that serve attentional selection, such as the frontal and parietal areas, show activation when one is asked to look at pictures and name them (Price 1998). For reading aloud words where selective attention has to be paid to the stimuli, the visual world form areas and angular gyrus show activation. Both reading and naming pictures demand selective attention and therefore the language system must work in harmony with the attentional system for these tasks (Mesulam 1990). Thus, the act of naming involves attentional resources to the extent that speakers attempt to be accurate with regard to the concepts that they want to speak and not others (Levelt 1989). Below we will see that naming and language production general also recruits general executive control areas of the brain. Hickok (2012) asserts that speech production is a process that involves both psycholinguistic computation and motor control. Psycholinguistic work linking attention control and speech generation has revealed dependency between the two systems. Speech appears to be a flawless cognitive performance since both intuitive linguistic knowledge and motor control blend dynamically. It is one thing to talk about the role of attention as a focusing and selective mechanism and another to understand it in terms of how it regulates behaviour. The issue of control during naming arises from the simple observation that given any concept, many other concepts that are related also compete for selection. Therefore, if speaking should require constant monitoring of competing information, the cognitive control processes become active as it would be in any other scenario (Botvinick et al. 2001). It is intuitive that speakers have some idea about the planning of their speech and they often get the feeling that they can control their articulation and action (Postma 2000). It has been suggested that speakers use knowledge from speech perception to monitor and edit speech during successful communication (Hartsuiker and Kolk 2001). Speech production is affected when speakers are provided a delayed feedback of their produced speech suggesting a close functional link between the production and the perception system (Hashimoto and Sakai 2003). It is also seen in many neuropsychological cases that the ability to monitor and edit speech production is impaired (Postma and Oomen 2005). Speakers pay a high level of selective attention to their inner speech during speech planning and production in a dynamic manner (Perrone-Bertolotti et al. 2014). Speakers monitor their speech production considering several discourse-related factors and ongoing communication constraints in a particular situation (Hartsuiker 2014). Therefore, the monitoring system in speech production uses top-down control which stems from the general purpose control systems of the brain that work for other action control systems. The executive control system in the brain looks after the accuracy and control of all types of cognitive behaviour. There appears to be no separate attention and executive control system for managing linguistic stimuli and their processing. Young children do not show much control in speaking and produce context inappropriate words and errors since the main cortical sites of action control, such as the dorso lateral pre-frontal cortex (DLPFC) and ACC as well as the whole frontoparietal attentional system, are not well developed.

Is monitoring during speech production contingent upon the linguistic feedback available from speech production? Recently, there have been some concerns with comprehension-based accounts of error monitoring in speech production (Nozari et al. 2011). It is quite possible that we do not often edit what we speak based on what we hear. It looks like speakers often fail to detect errors in their own speech while they can easily detect errors in another's speech. Further, speakers are rarely conscious of each word they speak and hear. Alternatively, one can consider the production system itself as monitoring its activity and not depending on the signals from the comprehension system. Conflict monitoring in the production system operates without the awareness of the speakers. Studies with ERPs have shown that a component known as error-related negativity (ERN) is commonly seen when a motor response is selected in conflict. ERNs have been found in Stroop tasks where participants give a verbal response (Masaki et al. 2001). Monitoring is a context-dependent thing, which means that we monitor our behaviour more when the situation demands it. If the situation does not demand any monitoring then the ACC may not show much activity. In an interesting study, Ganushchak and Schiller (2008) motivated participants not to commit errors during naming. Participants were given lists of objects to be named that were semantically similar, while another list had dissimilar objects. The amplitude of ERN was high in the motivated condition since presumably participants were operating under high monitoring. Monitoring therefore seems to be highly context dependent and also related to internal task goals of the participants. During speech monitoring speakers do not depend on feedback from comprehension but their own internal cognitive states affect speech production. Riès et al. (2011) measured the error-related negativity component of ERP during a production task and observed that this component peaks much before information from perception has arrived during the monitoring process—thus, indicating that general-purpose monitoring mechanisms are used by the brain during speech production.

It is also a possibility that efficient speakers resolve any conflict in speech production before production. Speech monitoring involves the ACC, which supports error monitoring in many other task situations (Christoffels et al. 2007). ACC sorts out conflict when the speech production system does not get immediate feedback from the perception system. One can see this sort of conflict when speaking while wearing ear plugs. There is a tendency to speak louder and this mismatch happens because of lack of feedback. In a direct exploration of the role of ACC during error monitoring of speech, Barch et al. (2000) asked participants to produce verbs under different constraints. ACC was selectively active when there was maximal conflict at the level of response. ACC must work in tandem with the speech comprehension and the attentional system to execute monitoring which leads to error check and correction.

It is practical to expect that bilinguals should face higher conflict during naming since they are known to activate concepts and lexical words from both the languages given any input. Acheson et al. (2012) found that Dutch–English bilinguals showed a larger ERN during naming of cognate words, indicating higher conflict in case of words that are lexically similar. Bilinguals use general-purpose inhibitory control mechanisms to monitor and resolve conflict that may arise because of dual language activation (Green 1998). Early bilingualism significantly modulates the activity of the ACC that is used for conflict monitoring during language production (Abutalebi et al. 2012). Bilinguals need to endure higher conflict during language production and therefore bilinguals show significant brain activity in some areas that monolinguals do not show, mostly in areas that deal with conflict monitoring (Jones et al. 2012). This evidence suggest that bilinguals develop a highly sophisticated mechanism of conflict management, drawing from the domain general executive control resources, and put this to use during language production.

5.4 Generating Sentences and Attention

Attention control in the above discussed single object naming was top-down. This means, speakers control and monitor what they uttered. However, there is a chance that bottom-up features of the world, including visual features, might affect the way attentional mechanisms work during speaking. In this section, it will be shown that arrangements of objects in the visual world and manner of seeing them influences conceptualization and encoding during sentence production. Attention control that is required during object naming in a laboratory is not the same as is required during speaking a sentence. Attentional control during generation of whole sentences or phrases operates under context. Much of the processes discussed like executive control and editing also happen during speaking sentences. Sentences bring in additional grammatical complexity that interacts with attentional states differently.

Speaking while seeing involves attention to objects and actions. Attending leads to visual perception of the distinctive nature of the objects and filters out irrelevant objects. This process is essential in situations where one aims to utter a sentence while looking at a scene that may consist of one or more objects, agents and some action. In this scenario, the central issue is: how we see and what we see in some order affects the choice of items to be inserted into the sentence. That is, how attentional allocation to entities affects the serial nature of speech production.

Given a simple picture that shows an action between two entities, attention selects the order in which the objects will be described. The speaker must generate a message plan that conceptually maps the perceptual events and finally a sentence with words. The selection of the agent (who is doing what to whom), and establishing a functional relationship between the agent's action and the acted upon object or theme are important. It has been observed that speakers normally make the agents the subjects of their sentences (grammatically) and then speak the rest of it while incorporating other elements. When a speaker's attention is drawn towards one of the competing entities, these entities become subjects of the sentences and they occupy prominent roles in the sentence. Therefore, focusing attention could enhance the perceptual salience of entities and affect syntactic processing in sentences. In an early study, Tomlin (1997) used an animated picture where one fish is shown gradually eating another fish. Tomlin used Posner's visual cuing paradigm (Posner 1980) to explicitly summon attention of the speakers towards one of the fish and explored how this might affect the choice of elements in the produced sentences. It was observed that English speakers normally used the fish that was on the attentional focus as the subject and created the sentences. When the actual agent, the fish that eats another fish, was given an exogenous cue, there was a good chance that speakers would speak an active sentence. On the other hand, when the fish that gets eaten was cued, speakers were more likely to produce a passive voice sentence. This indicates that attentional manipulation could powerfully affect the construction of sentences produced in a bottom-up manner. Therefore, attention capture can affect sentence planning and production (Myachykov et al. 2009).

Sentence production studies have shown that the first mentioned noun phrase is also the subject of the sentence (Gleitman et al. 2007). However, an explicit cue automatically influences the cued entities' prominence in the sentence and interferes with top-down goals. When attention is first summoned towards a location and after a brief delay a picture appears at this location, there is a good chance that speakers will make this object the first mentioned noun phrase of their utterance. It is another issue if this is also the grammatical subject of the sentence or it is just mentioned first. Languages with free word order like Russian and Finnish should behave differently on attentional manipulation. Speakers, depending on what they first focus attention on, create sentences while rearranging the other constituents of the sentences. Attentional cuing affects the choice of linguistic objects in sentence creation since they make these entities more accessible to consciousness and facilitate their selection. If sentence production is incremental then one should see an effect of attention on constituent ordering. Sentence production in the eye tracking paradigm has been also found to be influenced by attentional capacity of the individual. For example, children with autism and specific language impairment look longer at depicted events before they produce speech compared to normal children (Norbury 2011). Delays in linking the depicted visual information to frame linguistic structures suggest a weakness in basic attention capacities.

An entirely bottom-up attention capture account of sentence processing, at least the way it has been done with the Posner-type task, cannot be entirely satisfying. That is because speakers know what they need to speak given any situation. Both explicit and subliminal cues in simple sentence production tasks have shown that speakers create sentences by first mentioning objects that are cued (Nappa et al. 2004). However, an alternative possibility is that irrespective of this attention capture, speakers proceed to utter the sentence that they have already planned and constructed. This view proposes that the structure of a sentence is already fully formed before its utterance (Bock 1989). Apart from this, it also seems that the speaker's choice of sentence structures can be influenced by prior syntactic priming along with attention capture (Myachykov et al. 2012). In this paradigm, a prime sentence is first presented and then a picture for producing sentence. The choice of word order is affected by priming. This shows structural influences on planning and articulation. Myachykov et al. (2012) manipulated both attention, using exogenous cues, and structural choice, using sentence primes. Speakers were more likely to produce a passive voice sentence when the patient was cued and they also produced a passive voice sentence when the prime sentence was a passive voice sentence

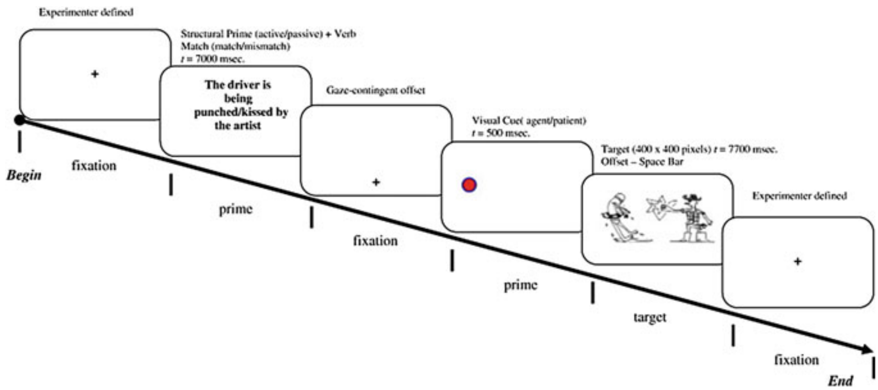

Fig. 5.3 In this experiment, an effect of structural priming as well as attention control was examined. Participants first read a prime sentence and then they had to describe a picture. Crucially an exogenous cue was used to attract their attention towards the visual referent, i.e. towards the agent or the patient. Eye tracking was used to track eye movements as participants spoke sentences. The results showed the influence of attention manipulation on the choice of sentence structure during production (Reproduced from Myachykov et al. 2012, p. 308)

itself. Attention factors seem to affect selection of sentential subjects and sentence structure in such cases (Fig. 5.3). This evidence of course suggests a highly dynamic view of language–cognition interface that goes beyond the simple structural notions of sentence processing (Jackendoff 2012).

Attentional focus on any object might make this object more salient and push it to the subject's position, depending on the language. However, speaking sentences looking at visual actions also is guided by the speaker's real-world semantic and pragmatic knowledge—that is, who can do what to whom and some knowledge about the type of action. For example, in a scene where a policeman is punching a boxer, each agent is linked to some real-world knowledge (Myachykov et al. 2012). In the real world, the act of boxing is more salient as an action with the boxer, while on the other hand policemen are also capable of exerting force on entities. These visual and pragmatic discrepancies have to be tackled by the language production system when planning sentences around them. In such situations, attention cues towards these entities can make them salient and more competitive for subjecthood, but speakers will always create sentences that follow real-world knowledge. In this case, attentional cues interact with pragmatic expectations and the result could be the production of a passive voice sentence. Therefore, attention cues could have a positional effect and also have an effect on ordering while they do not seem to distort the speaker's knowledge of the affairs of the world and actions.

What role does attention assume in these examples? Since the visual world is much cluttered and one has to find an object to begin speaking, attention narrows down the search. It is gross to suggest that low-level attentional cues always override top-down goals or alter pragmatic understanding. However, attention manipulations can create temporary interference or bias. Studies with exogenous cues do show that attending to some entity can influence a linguistic plan during production. Attention makes certain objects more prominent and they therefore occupy important roles in sentences.

Studies that have used exogenous cues to attract attention towards one or another object in a picture to see if this affects choice of syntactic construction have explicitly assumed a definition of attention where attending aids perceptual enhancement and also does filtering. However, as we have seen elsewhere, some other theories of attention (Knudsen 2007; see also Chun 2000) view attention as a combination of several processes such as top-down, goal-driven selection, ability to restrict interference from distractors and selection of relevant objects. Most importantly, these processes are part of the general executive control mechanisms that aid cognitive control and conflict resolution in real-life situations. Summoning attention exclusively towards one location affects the goal-driven aspects of speech planning.

Although attention cues may capture attention, briefly leading to higher activation of some object over another, it is possible that speakers can also disengage attention from this object, if this does not conform to their cognitive goal plan and then proceed in an altogether different manner. Attentional selection, thus, is goal driven in the face of conflict and speakers must use this strategy in order to be successful communicators. Such demonstration would require creating an experimental paradigm that examines how speakers manage conflict due to bottom-up attentional capture and still manage to speak the sentence that is context and goal relevant. Therefore, a capture theory of attention may not be the only account that explains the selection of linguistic structures during sentence generation. However, the cueing experiments certainly demonstrate the fact that attending to objects can alter speaking plans in a significant manner.

Do speakers look at visual events in the same sequence as they are integrated into the sentences they produce? Cueing studies may reveal attentional capture and its influence on sentence constituent selection, but they say little about any commonalities that visual and linguistic processing may share during speaking. In one study, children and adults were asked to look at colour photographs and generate sentences freely while their eye movements were recorded (Mishra 2013). There were three types of pictures. These pictures showed transitive, intransitive and di-transitive events with different objects. The idea was to examine how participants shift their attention between the subject and verb regions while conceptualizing and speaking. Both children and adults looked at the verb regions more than other regions. However, it was not possible to create links between regions of the image and linguistic units. The language in question was Hindi, which is a verb final language as far as canonical structures are concerned. The study also explored eye movement shifts between verb and subjects regions for this type of word order. For English, one can progress from subject to verb in a somewhat sequential manner. Verbs provide important information about other constituents of the sentence and the way they are related to one another. However, the data did not show that speakers looked at the verb region first to develop the structures for sentences in Hindi. What was clear was the distribution of attention between certain crucial regions from where speakers derive conceptual and linguistic structures. Much more can be done with such ideas to find out how attention shifts can be indicative of linguistic and conceptual processing during both planning and stages of articulation in speaking. Coco and Keller (2012) examined how similarities of sentences can lead to similar nature of scan paths during scene viewing and speaking (Fig. 5.4).

In this study participants were shown clip art scenes and were asked to produce sentences freely. However, they were also primed by a key word whose pictorial description was in the scene. The hypothesis was that scan paths across different scenes should be similar depending on the similarities of linguistic structures that are generated. The results showed that similarities of scan paths are correlated with similarities between sentences. This means that what we want to say and in which order are influenced by how we see a scene. Attention probably helps link different perceptions that arise out of visual perception and makes it possible for language to tie them all up with words and phrases. These studies are some of the most recent investigations that show a close relationship between the attention, visual and linguistic system. These data also speak against strictly modular theories of cognition.

This above demonstration suggests very close and dynamic interactions between the linguistic and conceptualizing system, and the visual system. As one looks around, linguistic labels are generated for what one sees. However, it is difficult to know if a linguistic mechanism projects a certain way of looking or a certain way of looking triggers a certain linguistic structure. Since we know that linguistic processing is incremental, once a certain thing has been encountered, it leads to predictive processing. This type of data also suggests that our perceptual world exerts a very strong influence on the type of linguistic structures one prefers to generate. Recently, it has been shown that even very young children show strong interaction between attention and grammar when speaking (Ibbotson et al. 2013). As children show increasing competence in attention control during development, they also show better control of linguistic production. This is a fine example of how language and attention could interact in crucial cognitive function like speaking.

If one has to generate a theory of attention and speaking, the endogenous component of attention should be taken seriously. Speaking is a voluntary activity. To speak, one has to generate an intentional plan. This plan itself involves selection. Therefore, endogenous attention helps us select topics and also exert editing control over our speech production system. Although we often speak with respect to some cue, i.e. someone's speech. This means, external cues are internalized and an action plan is generated for speaking. It is important to link components of attention that facilitate both object naming as well as sentence production. In sentence production, one is always aware of the larger pragmatic effects. Therefore, sometimes, we do not much attend to the smaller constituents while speaking. Attention on the topic of the discourse operates at a global level while when in need we give selective attention to some constituents. This polarization of attention during speaking happens because speaking itself involves many attentional mechanisms.

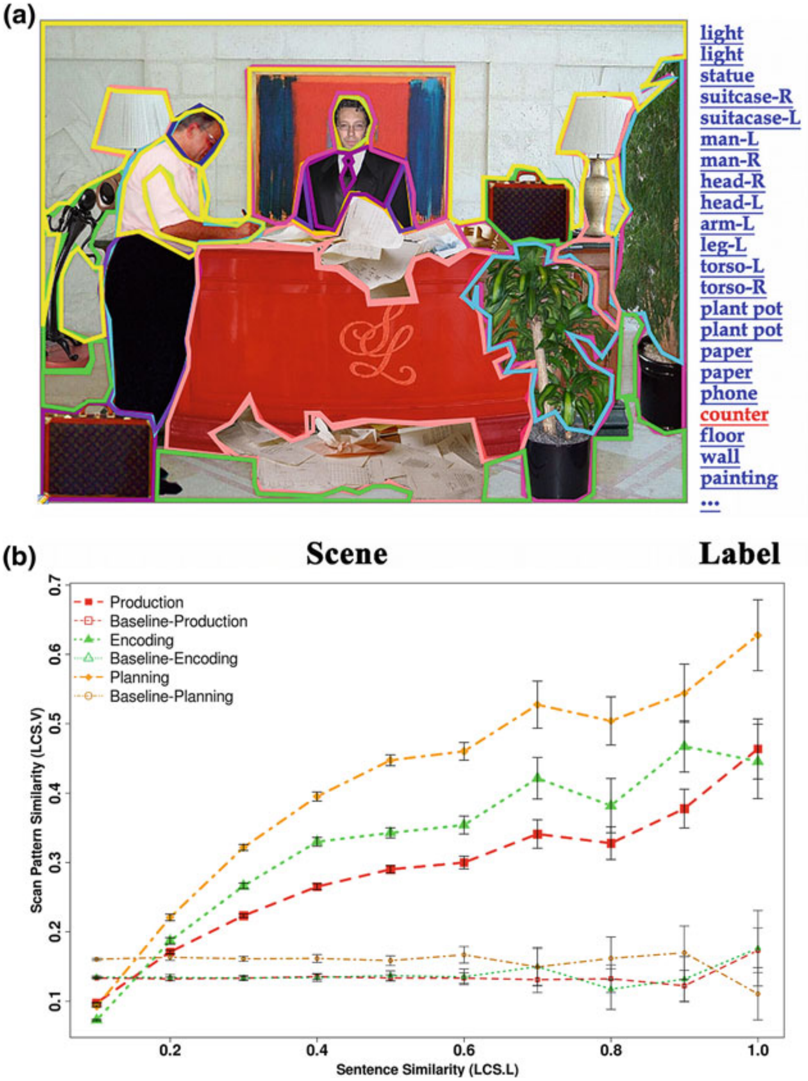

Fig. 5.4 The *top panel* shows the various areas of interest and different linguistic labels associated with them. In eye tracking studies of scene perception it is customary to designate different areas of interest to calculate fixations on that area. The *bottom plot* shows similarity in scan paths as a function of linguistic similarity. This shows speakers inspected areas in a similar order when they conceptualized and spoke similar sentences. *Source* Coco and Keller (2012, p. 5, 9)

5.5 Attention in Learning to Name

In this section I will try to examine if attention plays a key role during developmental years, particularly when it comes to speaking. Wittgenstein had observed that '… an important part of the training will consist in the teacher's pointing to the objects, directing the child's attention to them, and at the same time uttering a word; for instance, the word "slab" as he points to that shape' (1929, p. 4). So far we have seen that in adults, attention is paid to objects to be named, and this helps in retrieving conceptual and phonological information. One may wonder how attention might help acquire concepts and names during development itself. Infants have been shown to attend selectively to objects (Bulf et al. 2013). Young children grasp the perceptual and functional features of objects in the world while attending to them. Often, this mode of attention plays an important role in internalizing the features of objects. Children also look at objects selectively paying selective attention, which triggers lexical learning during naming (Smith et al. 1996). Children who learn to attend selectively to certain object properties like shape do well in naming tasks with objects that share this shape (Smith et al. 2002). This evidence suggests that early during the stages of language learning, visual attention to objects plays a critical role in naming ability and categorization. Attending to some features of objects selectively helps generate names and acquire concepts. Children have been shown to name novel objects with salient perceptual features if their attention is drawn towards them in an exogenous manner (Smith et al. 1996). Children, while focusing on some typical aspects of the object, acquire core perceptual features that help them further categorize and name them. During the early phase of word learning children must link phonological names of objects to the objects themselves.

Attentional mechanisms can be directly deployed towards objects other than features and spatial regions (Scholl 2001). For example, when someone says 'this is a cat', the child can integrate the spoken word 'cat' to the visual object cat. Early learning of object names and visual processing pathways are functionally connected (Smith 2013). Attention to object features helps retrieve associated conceptual and semantic information. Even attention to occluded portions of images helps retrieve words. Attention glues together different kinds of representations belonging to different modalities and allows the generation of a single unified concept. It is difficult to imagine how otherwise this sort of conceptual, phonological and perceptual blending could be possible (Medin and Ortony 1989) but with the involvement of attention. When children are given novel shapes to name spontaneously, they generate names based on their attention to the features and other perceptual features of these objects (Samuelson and Smith 2005).

Children also pay close attention to the mouth of a talking face when they are acquiring speech (Lewkowicz and Hansen-Tift 2012). This could have an evolutionary basis to it since we have always looked at objects that are talked about and also we look at faces when there is a speaker who speaks. Attention in this situation helps integrate audio-visual cues together which seem crucial for concept acquisition and language learning. Similarly, when children pay attention because of noise, their learning suffers (Riley and McGregor 2012). Children who show deficits in selective attention also show deficits in naming and word learning and retrieval (Traver et al. 1976). Attention then establishes a strong functional relationship between perceptual features of objects, their names and their meaning.

The developmental trajectory of attention mechanisms in children has functional and causal relationship with their language skills. Very young children show attention selectivity which allows them to focus attention on relevant objects while filtering distractors (Tipper et al. 1989). It is not very clear if children show inhibitory mechanisms when they are young. It is assumed that some form of inhibition is always part of selection. While novel objects might attract attention automatically, it is also important to inhibit distractors. Attention reinforces visual and other features of objects in memory and facilitates later recognition and naming (Vuilleumier et al. 2005).

The selecting and filtering property of attention appears to be crucial in acquiring the linguistic labelling. In fact, over a lengthy period of time children learn to deploy specific attention to certain relevant aspects of the object and not to other irrelevant features (Hagen and Hale 1973). It is crucial to first know which linguistic labels are related to particular objects. This coupling needs the involvement of attention in an important manner. Children's ability to control attention and deploy it in a goal directed manner to useful aspects of the stimuli and sustain attention for action has been known to help them learn literacy and numerical skill (Steele et al. 2012). Development of domain general cognitive skills causally affects the ability to combine information and express thoughts in language.

Even if one believes in an extreme form of nativism in child language acquisition one must explain that perceptual sensory information influences learning. Attention being a core mechanism of the cognitive system it helps in perception and conscious knowledge. The human child orients towards stimuli and learns to sustain attention on it. This type of executive control later proves to be crucial in binding perception with action. Objects and their names form a single gestalt and enter into the lexicon to be retrieved at will later. Both visual and auditory attention benefits concept acquisition formation of linguistic labels. This might already explain why as adults we look at objects before speaking, since that is why we have learnt naming in the first place.

5.6 Summary

The data acquired from different areas convincingly show that attention plays a critical role in language development and in adult language production. Further, during speaking attention and linguistic conceptualization interact dynamically. Looking at objects helps develop the linguistic knowledge that later helps in communication. The conceptualization of attention as a filtering agency has been important in developing theories of learning and evolution of complex cognitive behaviour. However, attention is much more than this and is critically linked to executive control. We saw that during spontaneous speaking we also edit and know what we should not say. Therefore, attention mechanisms help in moving along a goal-directed plan. However, we still do not know how different types of attention interact and influence different aspects of speaking, i.e. single objects versus sentences and bigger chunks. It is still not clear how the pragmatics of speaking affects attention deployments during sentence production. On the other hand, the use of eye tracking has made it possible to know the moment-by-moment nature of cognitive plans as they evolve and decision-making happens. Intuitively, one may think that naming and speaking are automatic processes that do not call for deliberate processing. However, studies discussed in this chapter indicate attentional mechanisms are at play during concept selection as well as encoding. It is not clear though, which aspect of attention, i.e. endogenous versus exogenous attention, is crucially involved during speaking. Unlike reading, where visuo-spatial attention seems to be more important, speaking may call for more endogenous attention. This is because one selects concepts and ideas wilfully for articulation. Therefore, one has to keep the modality of language use and particular tasks while referring to attentional mechanisms. While the issue of lexical competition, i.e. pre- or post-lexical access, remains debated in psycholinguistics, ample evidence shows attentional involvement in general.

References

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex, 22* (9), 2076–2086.

Acheson, D. J., Ganushchak, L. Y., Christoffels, I. K., & Hagoort, P. (2012). Conflict monitoring in speech production: Physiological evidence from bilingual picture naming. *Brain and Language, 123*(2), 131–136.

Aristei, S., Zwitserlood, P., & Rahman, R. A. (2012). Picture-induced semantic interference reflects lexical competition during object naming. *Frontiers in psychology, 3*, 28.

Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes, 24*(4), 527–554.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12*(2), 298–309.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language, 48*(3), 468–488.

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition, 31*(2), 163–186.

Bock, K., & Levelt, W. (2002). Language production. *Psycholinguistics: Critical concepts in psychology*, *5*, 405.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, *108*(3), 624–652.

Bulf, H., Valenza, E., & Turati, C. (2013). How a hat may affect 3-Month-Olds' recognition of a face: An eye-tracking study. *PLoS ONE*, *8*(12), e82839. doi:10.1371/journal.pone.0082839.

Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Human Brain Mapping*, *28*(9), 868–879.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170–178.

Coco, M. I., & Keller, F. (2012). Scan patterns predict sentence production in the cross-modal processing of visual scenes. *Cognitive Science*, *36*(7), 1204–1223.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences, 3*(3), 115–120.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition*, *2*(03), 231–244.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41*(3), 365–397.

Costa, A., Alario, F. X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin & Review, 12*(1), 125–131.

de Zubicaray, G. I., Wilson, S. J., McMahon, K. L., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: An event-related fMRI study employing overt responses. *Human Brain Mapping, 14*(4), 218–227.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., & Newhart, M., et al. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*, *130*(5), 1408–1422.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634), 20120394.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801–838.

Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., & Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture–word interference paradigm. *Frontiers in Psychology*, *1*, 177. doi:10.3389/fpsyg.2010.00177.

Dhooge, E., & Hartsuiker, R. J. (2011). The distractor frequency effect in a delayed picture-word interference task: Further evidence for a late locus of distractor exclusion. *Psychonomic Bulletin & Review, 18*(1), 116–122.

Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring: Effects of lexicality, context, and time pressure in picture-word interference. *Journal of Memory and Language, 66*(1), 163–176.

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*(6), 1187–1199.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.

Ganushchak, L. Y., & Schiller, N. O. (2008). Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. *Neuroimage, 39*(1), 395–405.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 640–654.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language, 57*(4), 544–569.

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes, 21*(6), 649–683.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220–1234.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(02), 67–81.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11* (4), 274–279.

Hagen, J. W., & Hale, G. A. (1973). The development of attention in children. *ETS Research Bulletin Series*, *1973*(1), i-37.

Harley, T. A. (1993). Connectionist approaches to language disorders. *Aphasiology*, *7*(3), 221–249.

Hartsuiker, R. (2014). Monitoring and control of the production system. *The Oxford Handbook of Language Production*, p. 417.

Hartsuiker, R. J., & Kolk, H. H. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42*(2), 113–157.

Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping, 20* (1), 22–28.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*(2), 135–145.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57*(6), 787–795.

Ibbotson, P., Lieven, E., & Tomasello, M. (2013). The attention-grammar interface: Eye-gaze cues structural choice in children and adults. *Cognitive Linguistics*, *24*(3), 457–481.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1), 101–144.

Jackendoff, R. (2012). *A user's guide to thought and meaning*. Oxford: Oxford University Press.

Jones, Ō. P., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., et al. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cerebral Cortex, 22*(4), 892–902.

Kan, I. P., & Thompson-Schill, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience, 4*(4), 466–482.

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience, 30*, 57–78.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory & Cognition, 16*(5), 400–410.

Lamers, M. J., & Roelofs, A. (2011). Attentional control adjustments in Eriksen and Stroop task performance can be independent of response conflict. *The Quarterly Journal of Experimental Psychology, 64*(6), 1056–1081.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. (1993). *Speaking: From intention to articulation* (Vol. 1). Boston, MA: The MIT Press.

Levelt, W. J. (1999). Models of word production. *Trends in cognitive sciences, 3*(6), 223–232.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences, 109*(5), 1431–1436.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485–495.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*(3), 503–535.

Malpass, D., & Meyer, A. S. (2010). The time course of name retrieval during multiple-object naming: Evidence from extrafoveal-on-foveal effects. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*(2), 523–537.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, *6*(02), 97–115.

Masaki, H., Tanaka, H., Takasawa, N., & Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *NeuroReport, 12*(9), 1851–1855.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. *Similarity and analogical reasoning, 179*, 195.

Mesulam, M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology, 28*(5), 597–613.

Meyer, A. S., & Lethaus, F. (2004). The use of eye tracking in studies of sentence generation. In *The interface of language, vision, and action: Eye movements and the visual world*, pp. 191–211.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*(2), B25–B33.

Meyer, A. S., Roelofs, A., & Levelt, W. J. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language, 48*(1), 131–147.

Meyer, A. S., Wheeldon, L., Van der Meulen, F., & Konopka, A. (2012). Effects of speech rate and practice on the allocation of visual attention in multiple object naming. *Frontiers in psychology*, *3*, 39. doi:10.3389/Ffpsyg.2012.00039.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167–202.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General, 132*(2), 228–252.

Mishra, R. K. (2009). Interaction of language and visual attention: evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K. (2013). Attention in sentence processing: A review. *Psychological Studies, 58*(3), 308–317.

Mishra, R. K., & Singh, S. (2014). Activation of shape and semantic information during ambiguous homophone processing: Eye tracking evidence from Hindi. *Cognitive Processing, 15*(4), 451–465.

Morgan, J. L., & Meyer, A. S. (2005). Processing of extrafoveal objects during multiple-object naming. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(3), 428–442.

Myachykov, A., Garrod, S., & Scheepers, C. (2009). Attention and syntax in sentence production: A critical review. In *Discours. Revue de linguistique, psycholinguistique et informatique*, Vol. 4.

Myachykov, A., Garrod, S., & Scheepers, C. (2012). Determinants of structural choice in visually situated sentence production. *Acta Psychologica, 141*(3), 304–315.

Nappa, R. L., January, D., Gleitman, L. R., & Trueswell, J. (2004). Paying attention to attention: Perceptual priming effects on word order. In *Proceedings of the 17th Annual CUNY Sentence Processing Conference*.

Norbury, C. F. (2011). Developmental language disorders: Overview. *The SAGE Handbook of Developmental Disorders*, p. 329.

Nozari, N., & Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(4), 1084–1090.

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology, 63*(1), 1–33.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Loevenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research, 261*, 220–239.

Piai, V., & Roelofs, A. (2013). Working memory capacity and dual-task interference in picture naming. *Acta Psychologica, 142*(3), 332–342.

Piai, V., Roelofs, A., & Schriefers, H. (2011). Semantic interference in immediate and delayed naming and reading: Attention and task decisions. *Journal of Memory and Language, 64*(4), 404–423.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77*, 97–131.

Postma, A., & Oomen, C. C. E. (2005). Critical issues in speech monitoring. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 157–166). Hove, New York: Psychology Press.

Price, C. J. (1998). The functional anatomy of word comprehension and production. *Trends in Cognitive Sciences, 2*(8), 281–288.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*(3), 460–499.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Riès, S., Janssen, N., Dufau, S., Alario, F. X., & Burle, B. (2011). General-purpose monitoring during speech production. *Journal of Cognitive Neuroscience, 23*(6), 1419–1436.

Riley, K. G., & McGregor, K. K. (2012). Noise hampers children's expressive word learning. *Language, Speech, and Hearing Services in Schools, 43*(3), 325–337.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*(3), 249–284.

Roelofs, A. (2008). Tracing attention and the activation flow in spoken word planning using eye movements. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(2), 353.

Roelofs, A., Piai, V., & Schriefers, H. (2013). Context effects and selective attention in picture naming and word reading: Competition versus response exclusion. *Language and Cognitive Processes, 28*(5), 655–671.

Samuelson, L. K., & Smith, L. B. (2005). They call it like they see it: Spontaneous naming and attention to shape. *Developmental Science, 8*(2), 182–198.

Schnur, T. T., & Martin, R. (2012). Semantic picture–word interference is a postperceptual effect. *Psychonomic Bulletin & Review, 19*(2), 301–308.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80*(1), 1–46.

Schotter, E. R., Ferreira, V. S., & Rayner, K. (2013). Parallel object activation and attentional gating of information: Evidence from eye movements in the multiple object naming paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(2), 365.

Schotter, E. R., Jia, A., Ferreira, V. S., & Rayner, K. (2014). Preview benefit in speaking occurs regardless of preview timing. *Psychonomic Bulletin & Review, 21*(3), 755–762.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29* (1), 86–102.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *The Quarterly Journal of Experimental Psychology, 65*(10), 1927–1944.

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition, 41*(8), 1200–1211.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(04), 771–781.

Smith, L. B. (2013). It's all connected: Pathways in visual object recognition and early noun learning. *American Psychologist, 68*(8), 618.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition, 60*(2), 143–171.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*(1), 13–19.

Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The multiple subfunctions of attention: Differential developmental gateways to literacy and numeracy. *Child Development, 83*(6), 2028–2041.

Strijkers, K., & Costa, A. (2011). Riding the lexical speedway: A critical review on the time course of lexical selection in speech production. *Frontiers in* Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language*, *65*(4), 345–362.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792–14797.

Tipper, S. P., Bourque, T. A., Anderson, S. H., & Brehaut, J. C. (1989). Mechanisms of attention: A developmental study. *Journal of Experimental Child Psychology, 48*(3), 353–378.

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In *Language and conceptualization*, pp. 162–189.

Traver, S. G., Hallahan, D. P., Kauffman, J. M., & Ball, D. M. (1976). Verbal rehearsal and selective attention in children with learning disabilities: A developmental lag. *Journal of Experimental Child Psychology, 22*, 375–385.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin, 128*(3), 442–472.

Vuilleumier, P., Schwartz, S., Duhous, S., et al. (2005). Selective attention modulates neural substrates of repetition priming and implicit visual memory: suppressions and enhancements revealed by fMRI. *Journal of Cognitive Neuroscience, 17*, 1245–1260.

Chapter 6
Language, Attention and Individual Differences

6.1 Unconscious Language Activation in Bilinguals

Social status (Labov 1972), bilingualism (Grosjean 1982) and gender (e.g. Coates and Pichler 1998) could affect language performance. While language performance seems to be within our voluntary control, recent studies show that listeners and speakers activate unwanted lexical items unintentionally. Therefore, it becomes necessary to examine how executive control and selective attention influence such processes. Language use in bilinguals and monolinguals differs in many ways. Learning to use two languages requires the neural structures that serve different linguistic as well as non-linguistic functions. How one becomes a bilingual is a matter of accident. Speakers are born into a bilingual society or learnt a second language at school and developed fluency in it. In either case, as a bilingual, there is something which seems common to all, to different degrees.

If Wittgenstein said that language is a game we play, it makes sense even to consider language performance a skill that we master. If it is a skill, then expertise in it should confer some advantages generally. It is like an excellent basketball or video game player who shows a good amount of executive control on many other tasks. It is like transfer of skills from one domain into another. For bilinguals, who confront the simultaneous activation of two lexicons, the advantages appear in another form. The more they shift, switch and inhibit, their executive control system gets better. We have seen already that the human attention system helps in attending goal-directed action while avoiding interference from distracters. This is what presumably happens with those bilinguals who constantly battle with two languages to make their discourse error free. This shows that one's language use patterns could enrich general attention mechanism.

Bilingual individuals show robust activation of the unused language during listening, reading and speaking of any one language (Kroll et al. 2015). Bilinguals are slower in picture naming compared to monolinguals (Costa and Caramazza 1999). That is because given any picture the bilingual brain also activates its

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_6

translation or words related to its name and this causes a delay in selection. In bilinguals, these activations are seen even when they are paying attention to one language node (Marian and Spivey 2003; Mishra and Singh 2014; Dijkstra and Van Heuven 1998). Many bilinguals who are not proficient need to translate their L2 words into L1 words to understand their meanings (Kroll and Stewart 1994). However, it has also been seen that highly proficient bilinguals translate words unconsciously (Mishra and Singh 2014; Sunderman and Priya 2012). Therefore, accessing the context-irrelevant language seems to a feature of the bilingual brain rather than linked to proficiency. Bilinguals also activate cross-language lexicons unconsciously and when the experimental context is monolingual. Chinese–English bilinguals access the first language when they read in second language (Thierry and Wu 2007).

Listening to spoken words can lead to automatic shifts in visual attention towards relevant objects in the environment. Visual world eye tracking allows tracking the automatic and unconscious nature of mental activations (Huettig and Altmann 2005; Mishra 2009; Huettig et al. 2011). In this paradigm, eye movements are tracked towards visual objects as a function of spoken language input. The most common finding is that participants immediately move their eyes towards objects that match the spoken language input (Tanenhaus et al. 1995). Such eye movements reveal the rapid integration of visual and linguistic integration during multi-modal processing (Mishra and Marmolejo-Ramos 2010). For the present context, it is important to note that these automatic eye movements with spoken language input also show that linguistic input can led to changes in the attentional states. Many researchers have explored if bilinguals activate cross-language competitors during listening using the visual world paradigm (Weber and Cutler 2004; Spivey and Marian 1999; Mishra and Singh 2014). Language non-selective activation is seen during simultaneous processing of both spoken and visual information in bilinguals (Shook and Marian 2013). In a now classic study, Spivey and Marian (1999) examined Russian–English bilinguals who had been staying since a long time in the USA. Participants were asked to listen to words either in Russian or English and look at a display containing four line drawings. One of these pictures depicted the spoken target word, and another was a cross-language phonological competitor. For example, for the target word, 'marker' in English, the display could contain '*marka*', a Russian word resembling the English word in sound. Interestingly, it was observed that these bilinguals were looking at this competitor word most of the time, indicating activation of cross-language phonology. Similarly, Weber and Cutler (2004) found that Dutch–English bilinguals looked more at pictures whose names were phonologically similar to Dutch objects. These studies show that bilinguals quickly orient their attention towards objects whose names match to the spoken input in any language direction.

Mishra and Singh (2014) presented written words on a computer screen, while participants listened to spoken words. On critical trials, one of the written words was a phonologically similar word of the translation equivalent of the spoken word. For example, if the spoken word was '*bandar*' in Hindi, the display contained the written word 'money', a word similar in sound to 'monkey', the translation of '*bandar*'. Bilingual participants immediately oriented their attention towards this particular word more often than distracters. Bilinguals activated translation equivalents in both L1–L2 and L2–L1 language directions (Fig. 6.1).

Bilinguals immediately activate cross-language semantics when they listen to words in either one of their languages. Activations of this sort with language pairs that do not share script or phonology show the very extreme nature of language non-selective activation in bilinguals. Does this involve attentional mechanisms in any manner? It does so, since activation of things that are not one's goal shows lack of inhibitory control for the organisms. In many visual search tasks, people look at distracters that are not targets. Similarly, getting primed by distracter stimuli can also divert attention in the main search task. These are the first evidence in the spoken word processing domains which show 'unintentional' activation of words which the bilingual is not actually dealing with. Bilinguals also are known to exercise inhibitory control during language selection (Green 1998). That is, when conflicting information is presented, bilinguals apply inhibitory control to suppress the task-irrelevant unit and select the appropriate one. The above-cited evidence of language non-selective activation in bilinguals suggests a lack of control. Why cannot bilinguals apply inhibition to select the relevant lexicon during listening? Indecently, recent work suggests that bilinguals who have higher inhibitory control also show less language non-selective activations during spoken word comprehension. Geizen and colleagues (2015) examined cross-language activation and inhibitory control in bimodal bilinguals (bilinguals with both spoken and sign language abilities). They found that superior ability in inhibit control (smaller

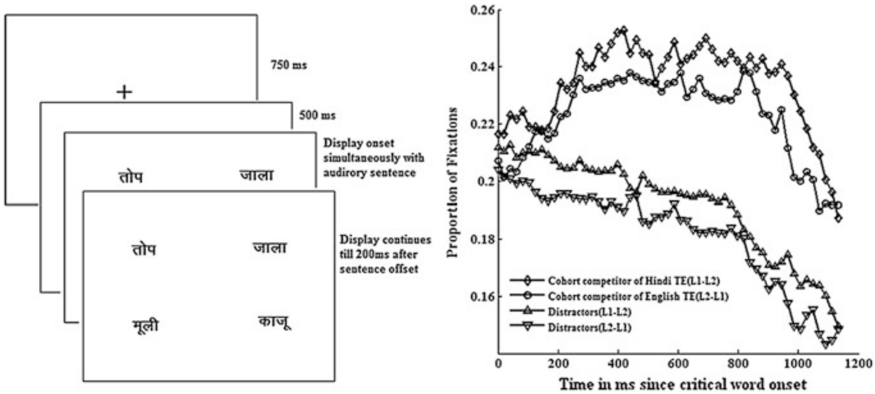

Fig. 6.1 The *left panel* shows a sample experimental trial. First, the participants fixated on a central cross. After a brief interval, there was a display containing four written words, and an auditory spoken word was presented simultaneously. Sample trial sequence in the L2–L1 direction with the auditory target word '*parrot*' was paired with a display containing '*top* (tank)' as a cohort competitor of translation equivalent '*parrot-tota*' along with three other distractors. Participants were merely asked to listen the spoken word and look at the display. There was no specific task as such. The *right figure* shows proportion of fixations to different written words from the onset of the auditory spoken word. The data show that eye movements towards the word which was a phonological competitor of the translation of the spoken word were immediate and higher compared to distracter. *Source* Mishra and Singh (2014, pp. 136,140)

effects in a Stroop task) was correlated with lower cross-language action. Similarly, in another study, Blumenfeld and Marian (2011) observed that performance on the Stroop task correlated with bilinguals ability in resolving conflict during cross-language activation. Therefore, bilinguals' executive control and attention ability influence the degree of parallel language activation.

However, not always do bilinguals manage to control the interference that may arise because of spurious activation of task-irrelevant lexicons. Interestingly, bilinguals have also been found to activate cross-linguistic information when the spoken words are irrelevant for the task at hand. It is known that spoken words are processed automatically even when one wants to ignore them (Salverda and Altmann 2011). Singh and Mishra (2015a, b, c) examined if such unwanted activation of lexicons during listening leads to interference in a visual task in Hindi–English bilinguals. High and low proficient Hindi–English bilinguals were asked to programme a saccade towards a line drawing which changed colour among other distracters. A task-irrelevant spoken word was presented along with the display (Fig. 6.2). On some trials, one of the distracter's names was phonologically similar to the translation of the spoken word. It was seen that participants were slow in

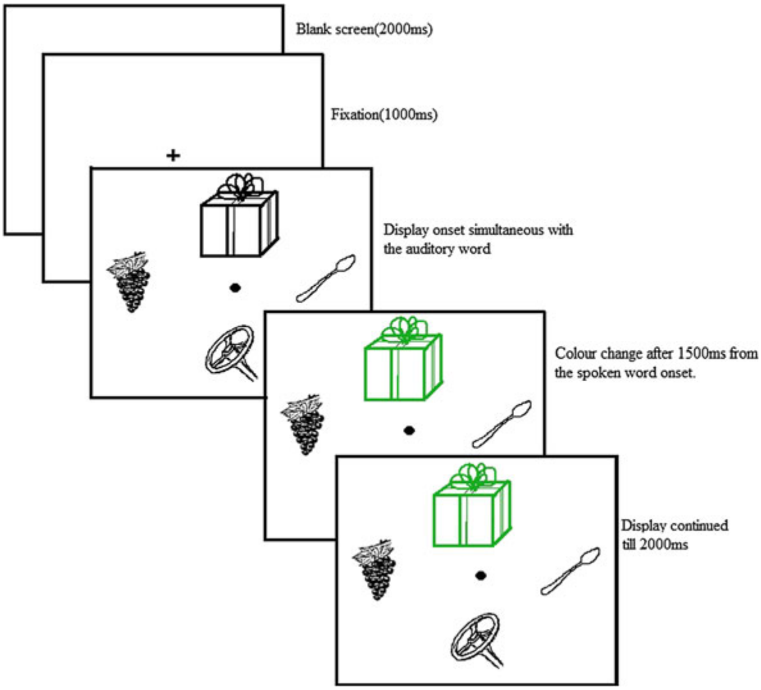

Fig. 6.2 In this trial, the display contained four line drawings. One of them changed colour. Participants listened to the English spoken word "ring" (Hindi translation is *angoothi*). The display has the picture of grapes, whose Hindi name '*angoor*' has a phonological overlap with '*angoothi*'. Participants looked at this cross-language competitor, which slowed down their saccade towards the object that changed colour. *Source* Singh and Misra (2015a, b, c, p. 8)

saccade execution when the distracter was related to the spoken word. This evidence suggests that bilinguals suffer interference because of spurious cross-language activation which happens most of the time. Such activations can affect actions in other domains.

6.2 Cognitive Control in Bilinguals

Some bilinguals grow up in bilingual societies and learn both the languages as basically first languages and develop balanced competence. Others learn a second language at school and develop competence. A third category of bilinguals learn a second language much later in life and do not often develop good fluency. The cognitive consequences of bilingualism manifest differently in these three types of subjects. When some sort of life experience or expertise brings in changes in the executive control system, it is important to show what exact components of attention are modulated. For example, attention control could mean selecting the right target among distracters, inhibiting the distracters, suppressing the interference from the distracters without inhibiting, maintaining the correct goal in working memory, etc. All these constructs are different shades of attention, and different tasks have been used to understand how bilingualism selectively enhances one mechanism and not another. Importantly, this can only be fruitful if we create a task that better mimics the bilingual's everyday life experiences. For example, if one claims that bilinguals are better in switching between two cues, then there is a good chance attention and advantages will be seen in tasks that exploit this aspect. If one thinks that bilinguals basically activate both the lexicons and inhibit the unnecessary ones, then tasks that involve conflict resolution will be useful (Green 1998). On the other hand, if bilinguals use their top-down attentional strategies to select the right language schema, then they are going to show benefits on tasks that exploit top-down aspects of attention. Invariably, most of these non-linguistic tasks are exploiter selective attention in some form or another. The current debate in the field surrounds around these detailed fractionations of the executive control system and finding how bilingualism is linked to it.

Could bilingualism then affect the general executive control system? Hilchey and Klein (2011) proposed that bilinguals need not as such have any specific inhibitory control mechanisms, but may enjoy a much more general executive control advantage and this makes them overall faster. Further, it appears that the bilingual speed advantage appears in task conditions where demand for attentional involvement is much higher (Costa et al. 2009). Costa and colleagues manipulated the percentages of congruent and incongruent trials to examine how this would modulate monitoring. The assumption was that even when congruent and incongruent trials are of equal numbers there will be high degree of uncertainty, and this would require higher monitoring. When one type of trial is maximal in number, there will not be enough demand for monitoring as a function of some habituation. Bilinguals outperformed monolinguals in a flanker's task only when the congruent and incongruent trials were in equal numbers. This could suggest that bilinguals may have higher ability to monitor conflict situations and bring in their ability to resolve the conflict when the situation is demanding.

Bilinguals are better at keeping task-related information in working memory for longer time and monitor the situation for better goal-directed action. This has given a new turn to the theorization on bilingualism and attention link. Additionally, bilinguals could be better employing top-down attention and select the right goal (Colzato et al. 2008). To most these may mean the same thing. But to cognitive psychologists they are not. It is one thing to judiciously select the right response and suppress interference and another to entertain conflict and then inhibit it. If bilinguals are better at keeping the task goals and monitor the contextual variations in task demands, they should also be better at voiding distracters. However, it is not to suggest that top-down goal directedness does not involve any form of inhibition. If bilinguals are forced to pay attention to incoming spoken stimuli in any one ear, as is done with dichotic listening tasks, they do better than monolinguals (Soveri et al. 2011). However, in normal everyday situations, there is a need to quickly suppress the non-target lexicon before it creates interference. Although studies have shown that bilinguals are slower in picture naming compared to monolinguals, not much evidence exists that suggests that they are also slower in listening. It appears that bilinguals do exercise inhibitory control during spoken word comprehension similar to their language production (Blumenfeld and Marian 2011).

Understanding the various components of cognitive control and how they interact could be a complex process. Braver (2012) makes a distinction of two types of cognitive control that may be useful in understanding the bilingual situation. Braver says that on the face of conflict, which a task like Stroop poses, participants could bring in proactive and reactive types of control. The proactive type of control manifests in better anticipation of the correct goals and also keeping the task sets in mind much before conflict occurs. This early selection could prevent conflict when it arrives. On the other hand, one could wait for the conflict to approach and then exert an active type of inhibitory control. In essence, it is preparedness versus firefighting. Braver seems to suggest that the proactive form of control could be more resource demanding since one has to keep the task goals all the time, throughout the task, while the reactive form of control could be somewhat less demanding. However, for the reactive type of control, the participant has to constantly reactivate the goals each time when there is a conflict. It is quite possible that both types of controls are triggered by the task context. For example, if a Stroop task has a very high percentage of either congruent or incongruent trials, then goal maintenance might not be useful, since the participant has to just keep more or less one response strategy in mind, whereas if the congruent and incongruent trials come randomly and are in equal number (context requires high monitoring), a mixture of both strategies could be useful. Bilinguals show superior performance on conflict tasks when the trial context is more demanding (Costa et al. 2009). Bilinguals probably bring in both types of control and it all depends on the scenario where it is more useful. Since different attention control tasks demand different types of control, it is not always possible to compare performances on them.

Do bilingual speakers exercise cognitive control to manage dual language interference? Bilinguals automatically activate cross-language words during naming, which is a goal-directed and intentional action (Levelt 1989; Strijkers et al. 2011). A bilingual often needs to translate a particular word into another lexicon for different discourse reasons. Assuming that the concept is the same that has to be named in two different languages, the bilingual speaker should select the right lexical word. For example, when a Dutch–English bilingual speaker is asked to translate an English word 'RABBIT' into Dutch 'KINIJN', they are faster in the presence of a semantically contextual picture of 'DOG' but are slower when in the presence of the written word 'DOG'. This is because the picture directly aids in concept activation, while the written word activates the lexical name and creates a bottleneck (Bloem and La Heij 2003; see also Roelofs et al. 2012). Bilinguals also seem to name pictures faster if they know its name in both the languages (Gollan et al. 2005). Currently, there are two views concerning mechanisms of bilingual naming process. One view holds that though bilinguals activate both the lexicons during naming, and they somehow intuitively know which lexicon to select for the task at hand (Costa and Caramazza 1999). Others have argued that activation of both the lexicons would require the bilingual to inhibit the unwanted word and select the right one (Abutalebi and Green 2007; Green 1998). While many assume that bilinguals need to inhibit the task-irrelevant lexicon, others have proposed a bifurcation in the inhibitory process itself. Inhibitory mechanism could be general or selective to task requirements (Shao et al. 2013). Brain imaging data show that when bilinguals entertain conflict during naming, the ACC and the fronto-parietal attentional network are active. Brain networks sub-serving Conflict monitoring show activity during bilingual language processing (Abutalebi et al. 2012; see also Van Heuven et al. 2011). The amount of conflict that the bilinguals will entertain will depend on the strength of the distracters. During both picture naming and reading aloud, the bilingual brain does extra work (Jones et al. 2011). Inhibition thus applied constitutes a part of the overall executive control system and is not specific to language. Bilinguals can select a task schema and continue to speak in one language while inhibiting the other. Recently, there have been proposals which have called for a modification of this idea in the light of recent evidence (Runnqvist et al. 2012). It appears that bilinguals still suffer interference from activation of language nodes from within and across the language. Speakers are often slower in the current trial if the previous trial involved naming in a different language. Parallel activation of concepts affects switching between languages. Bilinguals are slower when they shift from their L2 to L1 than vice versa (Meuter and Allport 1999). Unless one is a simultaneous bilingual and has acquired both the languages together, there is a good chance that L2 is weakly represented. This means, when speaking in the second language, they must resist interference from the stronger L1. Interestingly, for highly fluent bilinguals, the switch costs are similar in either direction (Costa and Santesteban 2004). Further, the bilingual brain may treat the mechanisms or translation and language switching differently (Price et al. 1999).

The evidence from bilingual naming studies both with behavioural and neuroimaging methods suggests that bilinguals' language performance is controlled by the general purpose neural architecture that deals with attention control and inhibition (Branzi et al. 2015). This inhibitory control mechanism in bilinguals changes over time with language maturity and brain development (Calabria et al. 2015). Branzi et al. (2015) administered both linguistic and non-linguistic switching tasks to bilinguals and measured neural activity using fMRI. It was observed that the left prefrontal cortex was active during response selection in linguistic and non-linguistic tasks. This indicates domain-general brain mechanisms for execution cognitive control for diverse stimuli. Importantly, brain areas known to exercise domain-general inhibition have been found active during bilingual language control (de Bruin et al. 2014). Thus, bilingualism influences how critical cognitive control measures are implemented in the brain.

In infancy, bilingual children show certain additional cognitive advantages compared to their monolingual peers. Bilingual infants show distracter suppression and task anticipatory behaviour even at the pre-verbal age (Kovács and Mehler 2009). These infants shift their gaze towards objects in advance of the cue. This could arise from their constant handling of two languages from birth and their constant effort to anticipate changes in the linguistic signals. Bilingualism early in the environment helps the infant to immediately comprehend shifts in language and use the cues productively (Sebastián-Gallés et al. 2012). And finally, the bilingual child may show superior capacity and organization for memory (Brito and Barr 2012). These data suggest that early exposure to two languages enhances a domain-general cognitive system and these advantages manifest on different attention tasks. Importantly, such infants show specific skills on tasks that require some sort of conflict resolution and goal-directed action (Poulin-Dubois et al. 2011).

What exact attentional mechanisms are affected by bilingualism during infancy? If the bilingual child learns to separate out different speech streams (Werker 2012) and focus on one, then it must develop selective attention. Attention is being now viewed as a fundamental cognitive skill, which is also a core component of executive control (Bialystok 2015). Bilingualism can also act as a potential cognitive intervention programme. For example, it has been suggested that bilingualism can cognitively benefit pre-term infants who are otherwise known to have long-term cognitive deficits (Head et al. 2015). Bilingual infants have been shown to have superior anticipatory mechanisms (Kovacs and Mehler 2009). Any infant that finds itself amid several language cues tries its best to develop sufficient attention resources to tackle the problem. This in turn rewires the working memory and the attentional networks of the brain further, resulting into superior performances on non-linguistic attention control tasks later (Petitto et al. 2012). If we leave apart the bilingualism factor, even training infants on attention control tasks can enhance their overall abilities in different executive control spheres (Wass et al. 2011). The ability to exploit conflicting environmental cues and exert control seems to be a hallmark of the bilingual brain. How does all this reflect in adulthood?

Therefore, bilingualism confers certain distinct cognitive advantages that manifest differently in a range of tasks. However, as is the case with any empirical phenomena in psychology, successive replications only provide validity. Recently, several meta-analyses (Hilchey and Klein 2011) and empirical studies (Paap 2014;

Paap et al. 2014) have shown the non-existence of any bilingual advantage on cognitive tasks. Valian (2015) argues that unless we understand the diversity that exists within cognitive tasks and the resources that they require, it is unclear how bilingualism may confer advantages to some and not to others (see also Mishra 2014). At the same time, many more studies have shown cognitive and attentional advantages to bilinguals in different domains, from better anticipatory control (Singh and Mishra 2015a, b, c) to creativity (Hommel et al. 2011).

It is likely that language use affects bilinguals' ability to maintain correct goal plans and actions. Bilinguals exert the control which their language skills have conferred them. It is important to understand how goals are activated in the first place and if attention is involved in action planning. If bilinguals take a 'conscious' decision about the language they are going to select given some discourse context, this may call for specific action planning. It is presumable that bilinguals do not necessarily stop and think at every point in the discourse about the most appropriate language to use. Many a times the language of the interlocutor, the discourse context and several other socio-linguistic fuctions provide subtle cues for the correct action plans. If a bilinguals choice of language behaviour is not 'intentional' and 'pops up' given a situation, it is important to understand how such 'unconsciously' raised goals are maintained for a long time. How do they activate the correct task schemas? Some striking insights about the nature of unconscious goal-directed behaviour is available in the theoretical proposal advanced by Dijksterhuis, in his 'theory of unconscious thought' (Dijksterhuis and Nordgren 2006). Dijksterhuis thinks that most of our goals are formed unconsciously and they influence decision making. Many a times goals are formed unconsciously and action schemas are generated with environmental trigger. Concerning the role of attention in this process, Dijksterhuis and Nordgren write,

> As a general rule, attention in the service of goals involves two interconnected faculties that usually act in close harmony: stability or focus (the ability to keep information active for action or further processing) and flexibility (the ability to be flexible enough to switch to, and take advantage of, contextual variations). The balance between focus and flexibility is crucial for goals to do their work effectively. (Dijksterhuis and Nordgren 2006, p. 471)

Most of our language behaviours are unconscious in the sense that we do not actively select each word we speak or how we speak. For a bilingual, similarly, most choices of this sort are unconscious; yet, they display superior goal maintenance and avoid conflict during action. One way to consider the bilingual behaviour is to suggest that bilinguals have better ability to generate subtle cues from the environment and these create goals in them which they later act out with top-down attention control. The unconscious though proposal also suggests that for goals to get formed and behaviour to generate unconsciously, practice is necessary. This means, one must have encountered that situation many times and a habit should have been formed.

Bilinguals regularly participate in and exercise their bilingualism in discourse and, therefore, given any context, they can form readily the correct goals. Once these are formed, correct action schemas can trigger behaviour. This is another way of explaining the behaviour than explicitly suggesting conscious selection of action. The flawlessness in behaviour may arise if most actions in selective attention are not always needed. Therefore, bilinguals' better performance in selective attention tasks does not arise from their ability to attend to each change in movement but from a robust task schema. There is evidence which suggests that an unconscious form of attention module may be at work which inhibits distractors and helps carry forward the task (Dijksterhuis and Nordgren 2006). Attention could then be merely modulating this behaviour a supervising entity. This is even important whether issue is how an unconscious behaviour like language could influence executive control abilities which are in turn modulating other non-linguistic functions that demand control. This proposal is different from other proposals where a top-down control strategy in bilinguals has been emphasized (Hilchey and Klein 2011).

6.3 Individual Differences and Bilingual Executive Control

The benefits of bilingualism lie with the *exercise* of bilingualism. That is, if one continuously uses two languages, then cognitive benefits appear. Therefore, the comparisons between monolinguals and bilinguals could be misleading since monolinguals anyway have very different neural networks than bilinguals. For instance, bilingual brains show superior functional connectivity between fronto-parietal network and default mode network. Therefore, when bilinguals and monolinguals are compared on executive control tasks in adulthood, interpretations may be problematic. One possible way is to compare two groups of bilinguals who differ in their experience and proficiency of second language (Abutalebi et al. 2012). In a series of studies, Mishra and colleagues (Singh and Mishra 2012, 2013, 2015a, b, c) who have studied high and low proficient Hindi–English bilinguals have shown that bilinguals who have higher proficiency also show superior performance on many executive control tasks compared to bilinguals with lower L2 proficiency.

Singh and Mishra (2012) examined high and low L2 fluency Hindi–English bilinguals in an experiment which was an oculomotor version of the classic Stroop task (Fig. 6.3a). Participants had to make a correct eye movement towards the colour patch which matched the colour of the written word. Participants in this task had to resist the interference produced by the written words on incongruent trials. Highly fluent participants were overall faster on different trials and also showed reduced conflict costs.

These data suggest that higher second language proficiency induces superior attention control on tasks that require keeping the task goals in mind. In another experiment (Singh and Mishra 2013), we manipulated the percentages of congruent and incongruent trials to induce differential demands on monitoring that is required. In this study, the authors used symbolic arrows in place of words (Fig. 6.4a).

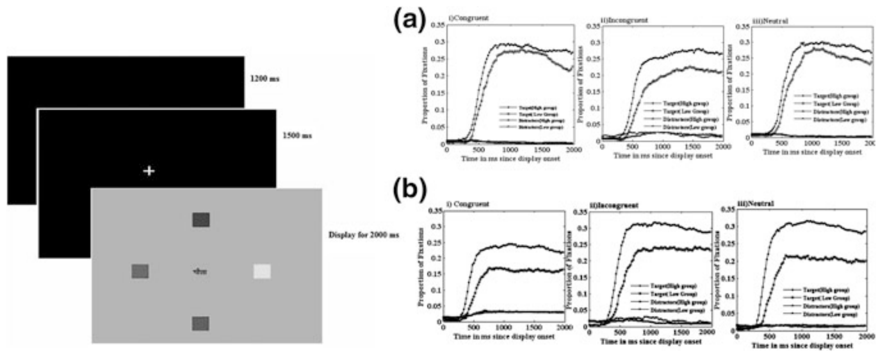

Fig. 6.3 The figure shows oculomotor version of the classic colour word Stroop task. Participants were asked to make an eye movement towards the colour patch which matched in colour with the written word's font. The results show that highly fluent bilinguals were faster and more accurate in programming a correct saccade towards the correct colour patch. The proportion of fixations shows looking behaviour towards different colour patches over time. The results thus indicate superior attention control in the oculomotor domain for highly fluent bilinguals. *Source* Singh and Mishra (2012, pp. 774, 776)

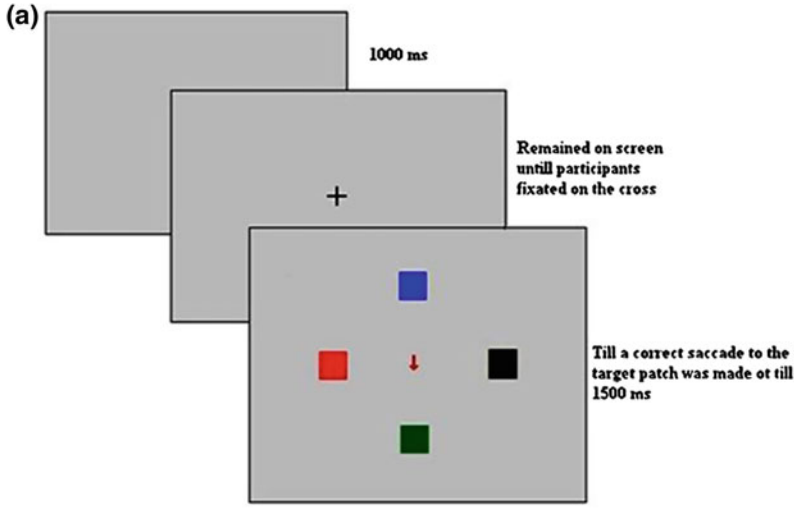

Fig. 6.4 The *top figure* shows an experimental trial. Here, the participants were asked to programme a saccade towards a colour patch which was similar to the *arrow*'s colour. The *bottom figure* shows saccadic latency for different blocks of trials. In general, highly fluent bilinguals were faster compared to low proficient bilinguals on all types of trials. *Source* Singh and Mishra (2013, pp. 5, 6)

Arrows can trigger attention shift reflexively like eye gaze. The percentage manipulation was done to induce greater attention demands for some blocks. It is assumed that the lock where both congruent and incongruent trials are in equal

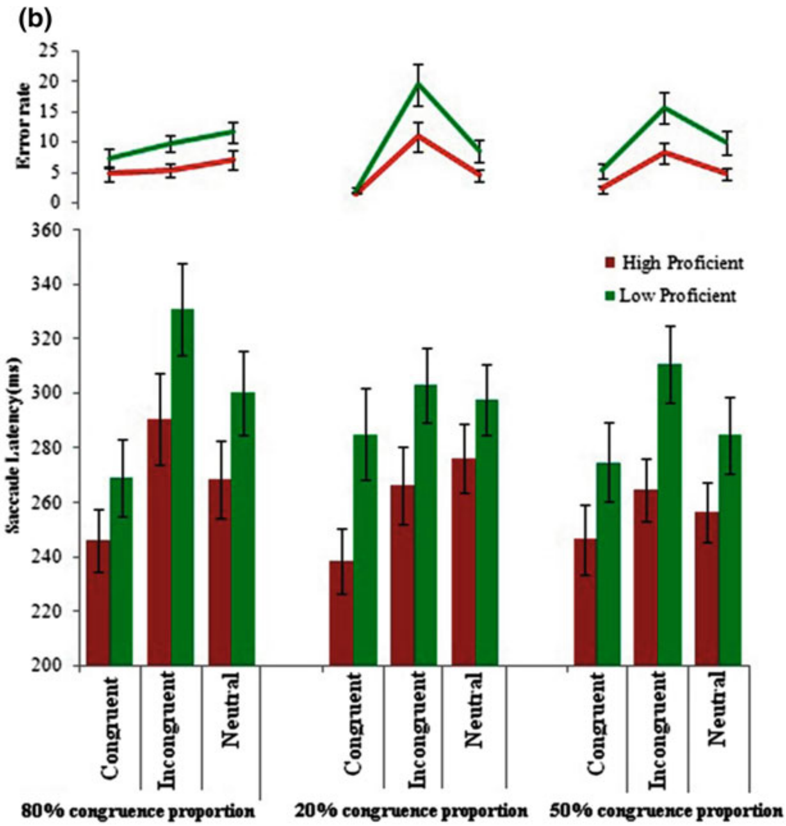

Fig. 6.4 (continued)

numbers, attention demands will be higher. This is because one will not be able to guess what could be the next trial. If in a block one type of trials is in maximum numbers, then such load will be less. We expected that highly fluent bilinguals will show attention advantage in the block which calls for higher monitoring. We also expected to see overall higher speed in such bilinguals for all the different blocks. Even in this study the highly fluent bilinguals outperformed.

The less fluent ones indicate better goal maintenance and monitoring. Many of these effects have already been shown with manual responses with monolinguals and bilinguals (Costa et al. 2009). This pattern of results is also in harmony with the thesis of executive control advantage postulated by Hilchey and Klein 2011).

Many theories on bilingualism's effects on control have now polarized around the issue of the presence or absence of inhibitory control. No particular cognitive psychological task has been found to be the right one to mimic the bilingual control mechanism. What exactly a bilingual person does when he is faced with multiple activations? One possibility could be that the bilinguals are good at disengaging attention faster from the unwanted lexical word. To mimic this with an attention

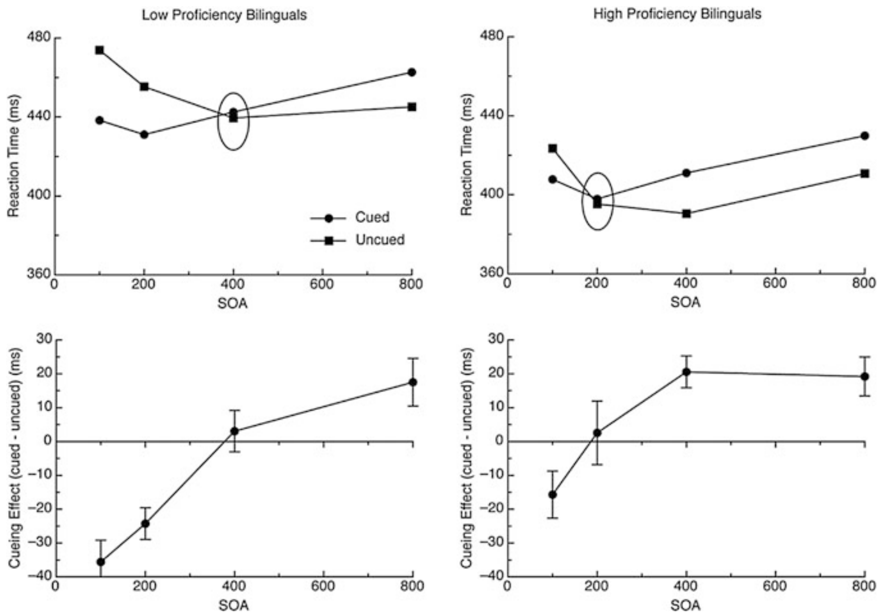

Fig. 6.5 The figure shows early appearance of inhibition of return (IOR) in highly fluent bilinguals compared to low proficient bilinguals. In this task, participant's attention was summoned by an exogenous cue towards a location. On some trials, a target appeared at this location, while on other trials there was no target. Participants had to press a key when they identified a target. The IOR effect was calculated as the delay in response time on valid trials when the cue-target delay was higher than 250 ms or so. The evidence was that for the highly fluent group we found an early appearance of IOR. Inhibition of return is the phenomenon that indicates the unwillingness of attention to return to a location immediately. *Source* Mishra et al. (2012, p. 1506)

task, we used the Posner's cuing paradigm (Mishra et al. 2012). In this paradigm, if a target appears at a cued location immediately, then there is facilitation. However, a delay causes inhibition. This means that attention does not want to move where it has been recently. We wondered if second language fluency modulates the ability to disengage attention from a task-irrelevant cue (Fig. 6.5). Participants had to detect a target which was sometimes cued and respond. However, the peripheral cue did not predict the appearance of the target. We observed that highly fluent bilinguals were faster in disengaging attention from this task-irrelevant cue when it did not predict the location of the target.

Language appears to be a skill whose extensive practice strengthens the executive control system for some. If playing action video game would boost attentional system (Green and Bavelier 2003), then when you manage the constant conflict with two languages, there has to be some sort of deep effect on cognition. The evidence cited above strongly suggests that not only attentional mechanisms contribute towards the efficient processing of language, but also in turn language affects the very system in any ways. Multiple theories concerning the influence of bilingualism on executive control have proliferated in recent times. While the idea that shifting between languages and inhibiting a wrong response strengthens the executive control system is appealing, there are many external social factors that can affect this correlation. For example, the contribution of socio-economic status, language dominance, age of learning of L2 and overall language environment can influence the link between bilingualism and control. Furthermore, there are many types of attention tasks that have their own assumptions and constraints on the interpretation of the data. For example, performance on a Stroop-type task is not similar to visual search task. These tasks demand different contributions of memory and attention for successful performance. There is no single attention task that best mimics the bilingual's situation. Assumptions based on 'inhibition' of the wrong language network differ from those that view 'shifting' to be the main issue with bilingual lexical access. Importantly, the correlations between bilingual's performance on non-linguistic executive control tasks and linguistic tasks are often not clear. In spite of these issues, majority of research shows some type of impact of bilingualism on cognitive control. This type of research is an excellent example of situation where a certain pattern of language use over a period of time leads to noticeable changes in the general cognitive architecture. This also shows how environmental factors in which language users find themselves can affect cognition in a more direct manner.

6.4 Attention in Second Language Learning

The executive system, of which attention is a part, gets strengthened in bilinguals since they have to constantly deal with cross-linguistic activations. However, this kind of thing can be seen only with those who have been bilinguals since an early age or who have developed excellent fluency (native like) later. Also, it is important to note that for the executive system to benefit, the bilinguals should constantly engage in bilingual discourse and live in bilingual communities. There is track of research in the field of second language learning which has shown that attention is a crucial element to develop good fluency and master a second language. Learning a foreign language seems to be contingent on attentional resources, even to the extent that attention to specific aspects of the linguistic material influences memory retrieval. Such findings are necessary to understand the individual difference issues in a larger context.

Why might attention be necessary for second language learning? Chomsky believed that the child's linguistic competence matures in an unconscious manner and therefore one needs no focus attention or learning (Chomsky 1988, cited in Schmidt 1990). Language learning is triggered by biology and therefore no voluntary effort is required. While this assumption may be true to an extent for native language development, it is not the same for a second language or a foreign language. It is today impossible to argue that linguistic performance is not affected by the psychological profile of the speaker. Memory, attention and several other motivational forces constantly influence the learning experience. The idea that language acquisition is a purely unconscious process has been the long-held view in many areas of applied linguistics and developmental studies. However, one finds a rather incisive discussion of the role of conscious learning and attention involvement in second language learning in Schmidt (1990). Schmidt says that much of linguistic knowledge, in the second language learning stage, happens when learners focus attention on the linguistic material. Focusing attention leads to the conscious perception of the stimuli. Further, objects of focus also enter immediately into memory and are retained longer. If this view is correct, then attention must play a causal role in mastering the new linguistic material for the second language learner. Working memory resources are also critical for maintenance of the learnt material. Attention to both form and meaning in the new language is important for internalization of the properties. It also makes sense since the typological features of the new language could be very different from the native language. Attending to such structures could lead to better acquisition and retrieval.

For example, being a native English speaker, it might be very difficult to internalize the gender system in a language like Hindi, unless they are noticed and their variations and uses are observed with the intention of reproduction. Many second language learners miss these subtle aspects of grammar of a new language. The difficulty is higher when the L2 has a very different structure compared to the L1. Japanese speakers of English do not master many phonological elements of English. There is some evidence that suggests that second language learners who pay lesser attention to learning material do not achieve perfection (Schmidt 2012). These examples suggest that attention in this sense creates the motivation for learning. If learning a second language, or language in general, is an unconscious process (Rebuschat and Williams 2012), then attention and conscious perception may be irrelevant. These two schools of thought cannot be easily reconciled, unless it is clear what aspects of second language learning may need attention and what processes can grow out of the universal grammatical system. It appears that second language learners pay selective attention to the most fundamental structures that are crucial for the development of automatic competence (Jiang 2007).

An exceptional amount of emphasis is generally added on the development of good speaking skills for the second language learner. Second language learners could have problems with accent of the new tongue or with several of the grammatical features. It has been observed that when second language learners pay attention to errors that they produce, this helps their overall language ability (Swain and Lapkin 1995). Several processes of attention such as orienting and executive control can be influential in guiding the learning process (Tomlin and Villa 1994). If the learner has paid attention to features during language intake and again attends to the productions, there is a good chance that errors will be minimized and language fluency could grow. It is important to note here that attention in this scenario is not posited as a limited resource process but as a system that helps in selecting and maintaining learnt information (see McLaughlin et al. 2006 for a different view). As Schmidt says, learners must be conscious while they participate in second language instruction and therefore the involvement of attention is logical (Schmidt 1990). When attention is drawn to errors committed explicitly, it enhances learning (Aljaafreh and Lantolf 1994).

Even during native language learning, infants seem to selectively attend to visual and linguistic cues (Lewkowicz and Hansen-Tift 2012). Infants orient attention towards speakers and gaze directly. This form of selective attention to the mouth of the speaker helps the infant to combine the spoken words with the experience of speaking. Similarly, when second language learners learn a language whose phonological system is very different from their native language, i.e. English speakers learning Korean, they seem to attend to the face of the trainer and notice how sounds are produced. Imitation of this sort helps the necessary audio-visual integration, which is important for learning the language. Apart from this, general executive processes like attention and memory for linguistic structures influence the second language learners' system (Engel de Abreu and Gathercole 2012). When the second language learners are given intensive training on the various aspects of the language in question, their ability to automatically detect errors is enhanced (White et al. 2012). This suggests that selective attention during learning a second language helps develop the kind of automaticity that fluent language use manifests. The role of attention therefore seems to be causal for amplification of newly learnt structures and later fluent use. Brain imaging data suggest that different brain networks are used during explicit and implicit types of learning (Yang and Li 2012).

Earlier, we have seen that language has an inbuilt mechanism to direct attention towards itself. For example, subjects of sentences attract attention and these elements are fore grounded. Different languages use different grammatical features such as deictic markers to point at things. Many a times case markers indicate particular grammatical functions of nouns. It seems particularly crucial that the second language learner pays attention to such subtle discourse markers during ongoing conversations. In the initial stages of learning a language, attention is generally higher towards basic structural elements, but as proficiency develops, attention is given to discourse elements. Self-monitoring during second language comprehension helps acquire these critical elements (O'Malley et al. 1989). Therefore, for learners, attention mechanisms supporting their language acquisition evolve over time, often as a function of their fluency. This argument nicely gels which has been shown earlier with the influence of bilingual fluency on attention control. There seems to be a complex interplay between the endogenous and exegenous forms of attention when we look at learners of a second language and those who have already developed fluency. For the second language learner, exogenous attention helps in grasping new stuff and maintaining them in memory. After much use, this form of attention control changes into more goal-directed and endogenous form of control. It is difficult to say at this point in time the development and maturation of these two forms of control.

6.5 Schooling, Literacy and Attention

Another important variable that influences cognition is literacy in the civilized world. Formal literacy for example transforms brain networks for acquisition of reading and writing along with advantages for visual processing (Dehaene et al. 2015). Literates perform better on cognitive tasks than illiterates (Huettig and Mishra 2014). Most psycholinguistic theories have emerged from data collected with highly educated and efficient college students. The data examined so far in different contexts have individuals who have received formal literacy of many years. Why should literacy matter to a cognitive psychologist or a psycholinguist? There is a moral reason as well as a cognitive reason. Cognitive theories of any type must be for everyone, irrespective of class, colour, gender, culture and literacy level. If these theories and the claims they make or the data used to construct these theories are about a small percentage of individuals, their validity will remain poor. Most cognitive psychological and psycholinguistic experiments have been on highly educated college students, a minor population considering the overall population of the world today. Apart from this, there are cross-cultural differences in cognition that one can see because of diversification. The second reason is that learning to read and write profoundly alters the brain and in most cases rewires the neuronal networks (Dehaene et al. 2010; Booth et al. 2004). Furthermore, learning to read and write also affects attention in different task contexts (Olivers et al. 2014). Even though our brains are not designed to read and write as such, it has become the cornerstone of modern-day civilization and brain plasticity allows us to do such complex and demanding activities effortlessly. Learning reading and writing could influence interaction of language and vision. The Russian psychologist Vygotsky was of the opinion that schooling and formal literacy could have important influence on the development of logical thought. Vygotsky considered language as a very powerful tool that modulates intellectual adaption (Vygotsky 1962). He also emphasized the importance of the social structure and the learning environment for the development of language and other cognitive skills. Vygotsky anticipated how the brain could develop higher mental functions like memory, attention and perception through societal interactions. Language then forms a part of such a shared universe of meanings in the brain (Nooteboom 2012).

Individuals who have never gone to school or have never learnt to read and write demonstrate difficulties in visual perception and attentional mechanisms. Even though learning to read and write is culturally acquired skill, continuous practice can modulate different cognitive systems. When illiterate individuals are asked to repeat non-words, their brains do not show activations in the traditional areas responsible for phonological processing (Castro-Caldas et al. 1998). In some cases, illiterate individuals also find it difficult to name objects of certain complexity (Reis et al. 2001). These deficits indicate that reading and writing influence spoken language skills.

Does reading change brain networks because it changes the attention network? It is a plausible hypothesis since reading is an attentionally demanding cognitive activity. Dehaene et al. (2010) measured brain activations in illiterates when they saw different kinds of visual stimuli, written language and processed spoken language. Higher levels of literacy enhanced top-down modulation from spoken words and also brain networks for visual perception. The left fusiform gyros showed specialization for processing letters. These changes due to literacy suggest a selective rewiring of the brain due to literacy acquisition. What is important to note here is that orthographic acquisition influences networks that directly modulate visual perception (Kolinsky et al. 1994). Illiterates seem to have problems in searching a visual target (Bramao et al. 2007). Orthographic training affects visual scanning and modulates the ability to deploy visual attention in a consistent manner (Ostrosky-Solis et al. 1991).

Several recent studies have looked at language and information processing in illiterate and low literate subjects and have used more novel behavioural methods like measuring eye movements (Huettig et al. 2011; Mishra et al. 2012; Olivers et al. 2014). These studies have looked at how formal literacy influences both predictive language processing as well as visual attention. Previous research has shown that illiterates are slower on some language tasks and even on some visual tasks (see Huettig and Mishra 2014 for a review). The use of eye movements to measure attentional orientation towards visually presented objects can be a nice way to show if literacy modulates one language component over another in any specific way. In one study, we compared low literates and highly literate adults in a look-and-listen visual world eye tracking task to examine the temporal difference between activation of semantic and phonological information. It is widely accepted that learning to read and write enhances phonological awareness (the ability to map individual sound identities in a word in an abstract manner). It was predicted that low literate adults will be slow in activating phonological information in spoken words compared to high literate adults, whereas access to semantic information may remain comparable. Hindi-speaking literate and low literate participants listened to spoken words and looked at an array of pictures. One of the pictures was a phonological competitor of the spoken word, while the other was a semantic competitor. We assumed that literacy must modulate how phonological and semantic associations are built dynamically during this cross-modal task. The results show that low literates were either very slow or did not activate the phonological information at all on the exposure to the spoken word, while they were almost good on the semantic front (Fig. 6.6). This shows that low literates may show a subtle temporal delay in activating the phonological information during spoken language processing.

One method to directly test selective attention is to use a visual search task. In visual search, a target may pop out among distracters because of its salient features. This parallel search has been argued as not involving selective attention, and is easy (Nakayama and Silverman 1986). In contrast to this, serial search takes more effort and calls for selective attention. However, set size (number of distracters) can affect search efficiency in either case. Olivers et al. (2014) examined how formal literacy level may influence attentional allocation during visual search. The study examined the effect of literacy on simple and difficult visual search tasks. In the simple search,

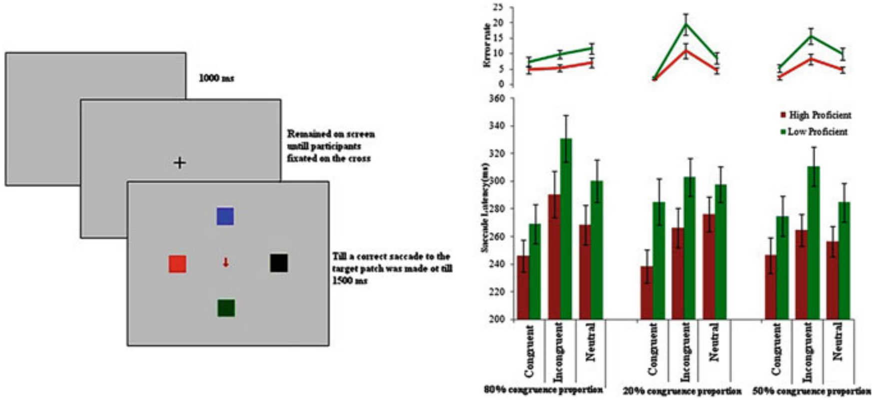

Fig. 6.6 The *top left panel* shows a sample trial which has four line drawings. Each such display is matched with a spoken word. In this case, the spoken word which the participant hears is '*magar*' (crocodile, in English). The display has the picture of 'peas' (*matar*, in Hindi) which is a phonological competitor and the picture of a tortoise, which serves as a semantic competitor. There are two other pictures that are unrelated distracters. The interest was to see the orientation of attention towards these two competitors as the spoken word unfolded over time. Participants did not have to do any task except listening and looking at the display as they liked. The *top right plot* shows fixation proportions to the phonological and semantic competitors as well as distracters for both high and low literate subjects from the onset of the spoken word. The figure shows an immediate activation of phonology of the cohort word for highly literate participants while this was absent for the low literate subjects. The *bottom panel* shows proportion of fixations for 'filler' trials. In these trials, the display contained the picture which the spoken word mentioned. The experimental trials were target absent. *Source* Huettig et al. (2011, pp. 5, 7, 10

participants searched for a coloured chicken among other chickens of a different colour. In the difficult search task, a thin chicken had to be searched among fat chickens (Fig. 6.7).

In both task conditions, set size was manipulated by increasing the number of distracters from 4 to 16. Importantly, the position of the target chicken was varied across the screen, from fovea to parafovea. We hypothesized that if reading enhances selective attention, then low literate subjects must show some deficits in search in the foveal region apart from general slowness when set size increases. Also, they should be slow in searching these target as the chicken are placed farther away from central vision towards the periphery. The results showed that illiterates in general were slower compared to highly literate subjects in this search task. Highly literate subjects were generally efficient in the search when the target appeared in the central regions and especially on the right. The low literate individuals did not show any such preferences (Fig. 6.8).

This pattern of results suggests that highly literate subjects not only deploy higher foveal attention, but they are also better in finding targets in the parafoveal regions. More interestingly, the search efficiency varied as a function of position of the target chicken with regard to the centre of the screen. The results show that reading has a direct effect on visual and attention processing.

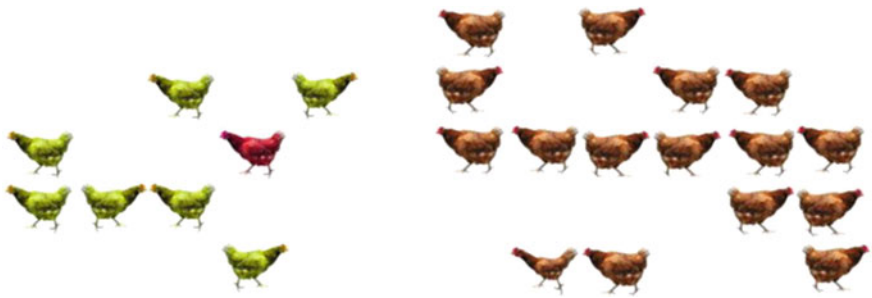

Fig. 6.7 The *left figure* shows an example of easy search where the target is a *red* chicken among *green* chickens where set size is 8. The *right figure* shows a difficult search scenario where the target is the skinny chicken among the fat chickens with set size 16. *Source* Olivers et al. (2014, p. 78)

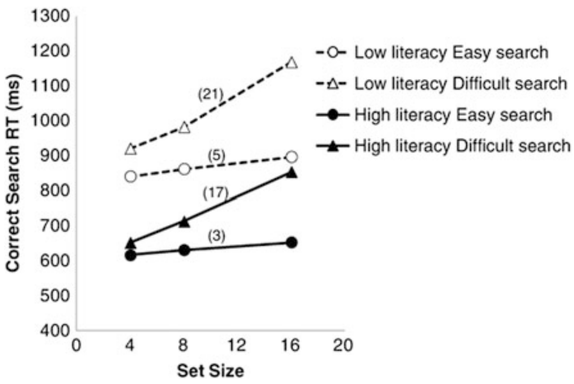

Fig. 6.8 The figure shows reaction times for easy and difficult search for literates and low literates as a function of set size (*low* and *high*). It is apparent that search times are higher as a set size is increasing. *Source* Olivers et al. (2014, p. 81)

We also examined if literacy modulates anticipatory eye movement behaviour in simple language-mediated search tasks (Mishra et al. 2012). Language users select objects that the language refers to by anticipation and prediction (Altmann and Kamide 2007). Anticipation helps in quick orientation of attention in space where the target object is most likely to be found. Low literates and high literates participated in a simple visual world eye tracking study where they had to listen to a sentence and look at a display containing four line drawings. We manipulated the gender agreement phenomena of the Hindi language. Hindi adjectives agree in gender with the nouns. Each spoken sentence had a gender marked adjective and a target noun. Highly literate subjects immediately shifted their attention towards the most appropriate target object much ahead in time compared to the low literate subjects. We argued that literacy boosts anticipatory behaviour, which in turn affects how one allocates attention to objects. Language-directed visual search is then influenced by the literacy level of the participant. This is another example of how one's linguistic profile can have substantial effects on overall attention mechanisms. The task was simple enough to extract anticipatory shifts in visual attention without any additional task.

While all these studies might have shown that illiterates perform poorly on some verbal or visual tasks, it is not quite simple to separate formal learning of reading and writing from the effects of schooling. Schooling is a broader cultural notion and has much more dimensions than mere ability to identify graphemes or knowing how they sound. Some of the studies considered illiterates who were very old when data were collected, thus confounding the effects of brain maturation due to age. There are many societies that have no formal system of literacy acquisition but they do have highly sophisticated cognitive skills. Although it is conceivable that learning to read a particular orthography could influence visual scanning in some manner, it is not clear how this could also enhance participants' general semantic knowledge, unless schooling is taken into account. There are data to suggest that schooling enhances semantic fluency (da Silva et al. 2004). This could be because by reading one also knows about things and their general uses, even things that are not experienced. Furthermore, attentional abilities linked with reading could be influenced by different orthographic styles as well as reading methods. Many studies in this area have been on illiterate populations who are old and, thus, it is difficult to separate age-related factors in cognitive decline on tasks. Many writers even consider literacy as cultural and not just not an ability to identify orthography. Illiterate individuals who learn to read and write later in life have been shown to develop the brain networks found in literate participants. There is, however, sparse research that is longitudinal in nature. Such studies are necessary to examine the precise changes that literacy acquisition brings to attentional and other cognitive skills. What is clear from recent findings is that reading and language use alter and modify cognitive networks.

Cognitive factors pertaining to individual differences show pervasive influence on speech and language processing. These effects are also found in monolinguals and even otherwise highly literate populations. For instance, attention and working memory influence how language users process auditory and visual stimuli (Ou et al. 2015). Huettig and Janse (2015) examined if working memory and processing speed capacity of individuals influenced anticipatory mechanisms, in a visual world eye tracking study. A large number of participants of different age groups were asked to click on an object as they saw visual displays containing line drawings and heard a spoken instruction. Individuals' working memory capacity measured on standardized tasks and processing speed was strongly correlated with the speed with which participants shifted visual attention towards objects that were gender congruent with an article in the spoken Dutch sentences. This indicates that cognitive capacities influence how one does predictive language processing. In another study, Rommers et al. (2015) explored if performance on a spatial cueing task predicted anticipatory eye movements in a visual world task. Dutch listeners listened to spoken instructions and looked at a display containing pictures. Contextual information in the spoken instruction predicted a highly probable noun which was in the display among unrelated distracters, i.e. 'In 1969 Neil Armstrong was the first man to set foot on the moon.' On some trials, the display contained the picture of an object whose shape matched this probable last word, i.e. picture of a tomato that is similar in shape to the moon. Participants' performance on a non-verbal spatial cueing task strongly correlated with predictive eye movements in the visual world task. This further shows that basic cognitive skills such as attention, processing speed and working memory influence predictive language processing. Therefore, a large amount of heterogeneity can be seen in such multi-modal processing even in a general population when language abilities are kept constant.

However, it is one thing to generalize from correlational studies and another to directly manipulate attentional and memory resources and see if participants' performance slows down with regard to anticipatory eye movements. It is not well established that bilingual individuals spontaneously activate cross-language competitors during spoken language processing (Mishra and Singh 2014; Singh and Mishra 2015a, b, c). In one study, Prasad and Mishra (In preparation) examined if concurrent working memory load influences eye movements in a visual world study in bilinguals. Hindi–English proficient bilinguals first saw a series of letters to be kept in working memory for a later backward counting task. Immediately following the working memory task, participants were presented with a display containing four written words, drawings and a spoken word. Spoken words were both in Hindi and English. One of the written words was a phonologically related word to the translation of the spoken word and others were distracters. Based on previous findings, it was expected that bilingual participants should immediately orient their attention towards this written word that was a phonological neighbour of the spoken word's translation compared to other unrelated distracters. However, the crucial question was whether the proportion of fixations varies as a function of working memory load (low and high load). Eye movement analysis showed that working memory load does not influence proportion of fixations to the translation competitors as such. Participants' performance was also significantly accurate on the working memory task. This suggests that a concurrent working memory load may not as such constrain automatic eye movements seen during visual world studies. This, however, may vary depending on the experimental context and task.

6.6 Conclusion

In this chapter, we saw that language ability and particular patterns of uses could influence attention mechanisms significantly. For bilinguals, handling two languages enhances their general executive functioning in many ways. However, one should be cautious in drawing strong conclusions, since bilinguals are not homogeneous communities and many differ in terms of age of acquisition, fluency, style of L2 acquisition, etc. The language prevailing in a language environment also plays a role in modulating the attention abilities depending on how bilingualism is practised in everyday life. Most data so far do indicate that fluency enhances monitoring abilities and also the ability to exert goal-directed action. Similarly, we saw that learning to read and write affects visual attentional mechanisms. In this case, the brain networks specialize and accommodate a learnt skill, and with practice also influence verbal language abilities. The discussion thus shows that language and attention could interact and influence one another in many situations. Attention is necessary during the early years of language acquisition and this in turns helps develop automaticity in language processing. This kind of research with individual differences with its theme has made psycholinguistics and cognitive psychology more interactive in the truest sense.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of neurolinguistics, 20*(3), 242–275.

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex, 22*(9), 2076–2086.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal, 78*(4), 465–483.

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502–518.

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives, 9*(2), 117–121.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language, 48*, 468–488.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition, 118*(2), 245–257.

Booth, J. R., Burman, D. D., Meyer, J., Gitelman, D. R., Parrish, T. B., & Mesulam, M. (2004). Development of brain mechanisms for processing orthographic and phonologic representations. *Cognitive Neuroscience, Journal of, 16*(7), 1234–1249.

BramAO, I., MendonCA, A., FaISca, L., Ingvar, M., Petersson, K. M., & Reis, A. (2007). The impact of reading and writing skills on a visuo-motor integration task: A comparison between illiterate and literate subjects. *Journal of the International Neuropsychological Society, 13*(2), 359–364.

Branzi, F. M., Della Rosa, P. A., Canini, M., Costa, A., & Abutalebi, J. (2015). Language control in bilinguals: Monitoring and response selection. *Cerebral Cortex,*. doi:10.1093/cercor/bhv052.

Braver, T. S. (2012). The variable nature of cognitive control: A dual-mechanisms framework. *Trends in cognitive sciences, 16*(2), 106–113.

Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental science, 15*(6), 812–816.

Calabria, M., Branzi, F. M., Marne, P., Hernández, M., & Costa, A. (2015). Age-related effects over bilingual language control and executive control. *Bilingualism: Language and Cognition*, *18*(01), 65–78.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121*(6), 1053–1063.

Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures* (Vol. 16). MIT press.

Coates, J., & Pichler, P. (Eds.). (1998). *Language and gender: A reader*. Malden: Blackwell.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(2), 302–312.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition*, *2*(03), 231–244.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*(4), 491–511.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135–149.

da Silva, C. G., Petersson, K. M., Faísca, L., Ingvar, M., & Reis, A. (2004). The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 26*(2), 266–277.

de Bruin, A., Roelofs, A., Dijkstra, T., & FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: fMRI evidence from trilingual speakers. *NeuroImage, 90*, 348–359.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*(6009), 1359–1364.

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience, 16*(4), 234–244.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science, 1*(2), 95–109.

Dijkstra, T., & Van Heuven, W. J. (1998). The BIA model and bilingual word recognition. In *Localist connectionist approaches to human cognition*, pp. 189–225.

Engel de Abreu, P. M., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology, 104*(4), 974–986.

Giezen, M. R., Blumenfeld, H. K., Shook, A., Marian, V., & Emmorey, K. (2015). Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition, 141*, 9–25.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220–1234.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(02), 67–81.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA.: Harvard University Press.

Head, L. M., Baralt, M., & Mahoney, A. E. D. (2015). Bilingualism as a potential strategy to improve executive function in preterm infants: A review. *Journal of Pediatric Health Care, 29*(2), 126–136.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*(4), 625–658.

Hommel, B., Colzato, L. S., Fischer, R., & Christoffels, I. K. (2011). Bilingualism and creativity: Benefits in convergent thinking come with losses in divergent thinking. *Frontiers in psychology, 2*, 273.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23–B32.

Huettig, F., & Janse, E. (2015). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience* (in press).

Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind–a critical examination of theories and evidence. *Language and Linguistics Compass, 8*(10), 401–427.

Huettig, F., Mishra, R. K., & Olivers, C. N. (2011a). Mechanisms and representations of language-mediated visual attention. *Frontiers in Psychology, 2*, 394.

Huettig, F., Singh, N., & Mishra, R. K. (2011b). Language-mediated visual orienting behavior in low and high literates. *Frontiers in psychology, 2*, 285.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*(1), 1–33.

Jones, Ō. P., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., et al. (2011). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cerebral Cortex, 22*(4), 892–902.

Kolinsky, R., Morais, J., & Verhaeghe, A. (1994). Visual separability: a study on unschooled adults. *PERCEPTION-LONDON, 23*, 471–471.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences, 106*(16), 6556–6560.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–149.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics, 1*(1), 377–394.

Labov, W. (1972). *Sociolinguistic patterns* (Vol. 4). Philadelphia: University of Pennsylvania Press.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: The MIT Press.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences, 109*(5), 1431–1436.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, *6*(02), 97–115.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second Language learning: An information processing perspective1. *Language Learning, 33*(2), 135–158.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*(1), 25–40.

Mishra, R. K. (2009). Interaction of language and visual attention: evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K. (2014). Lets not forget about language proficiency when link bilingualism with executive control. Comment on [Valian, in press]. *Bilingualism: Language and Cognition*, *18*(1), 39–40.

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing, 11*(4), 295–305.

Mishra, R. K., & Singh, N. (2014). Language non-selective activation of orthography during spoken word processing in Hindi-English sequential bilinguals: an eye tracking visual world study. *Reading and Writing, 27*(1), 129–151.

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012a). On the time course of exogenous cueing effects in bilinguals: Higher proficiency in a second language is associated with more rapid endogenous disengagement. *The Quarterly Journal of Experimental Psychology, 65*(8), 1502–1510.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012b). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320*(6059), 264–265.

Nooteboom, B. (2012). Embodied cognition, organization and innovation. In *Handbook of Knowledge and Economic*s, p. 339.

Olivers, C. N. L., Huettig, F., Singh, J. P., & Mishra, R. K. (2014). The influence of literacy on visual search. *Visual Cognition, 22*(1), 74–101.

O'Malley, J. M., Chamot, A. U., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics, 10*(4), 418–437.

Ostrosky-Solis, F., Efron, R., & Yund, E. W. (1991). Visual detectability gradients: Effect of illiteracy. *Brain and Cognition, 17*(1), 42–51.

Ou, J., Law, S. P., & Fung, R. (2015). Relationship between individual differences in speech processing and cognitive functions. *Psychonomic bulletin & review*, *16*, 1–8.

Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology, 26*(3), 242–255.

Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *Journal of Cognitive Psychology, 26*(6), 615–639.

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The 'Perceptual Wedge Hypothesis' as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language, 121*(2), 130–143.

Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology, 108*(3), 567–579.

Price, C. J., Green, D. W., & Von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain, 122*(12), 2221–2235.

Rebuschat, P., & Williams, J. N. (2012). Implicit learning in second language acquisition. *The Encyclopedia of Applied Linguistics,*. doi:10.1002/9781405198431.wbeal0529.

Reis, A., Petersson, K. M., Castro-Caldas, A., & Ingvar, M. (2001). Formal schooling influences two-but not three-dimensional naming skills. *Brain and Cognition, 47*(3), 397–411.

Roelofs, A., Dijkstra, T., & Gerakaki, S. (2012). Modeling of word translation: Activation flow from concepts to lexical items. *Bilingualism: Language and Cognition*, *1*(1), 1–11.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics, 77*(3), 720–730.

Runnqvist, E., Strijkers, K., Alario, F., & Costa, A. (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language, 66*(4), 850–869.

Salverda, A. P., & Altmann, G. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1122–1133.

Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning1. *Applied Linguistics, 11*(2), 129–158.

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. *Perspectives on Individual Characteristics and Foreign Language Education, 6*, 27–48.

Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science, 23*(9), 994–999.

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition, 41*(8), 1200–1211.

Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition*, *16*(2), 304–324.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(04), 771–781.

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: Evidence from Hindi-English bilinguals. *Frontiers in psychology, 4*, 322.

Singh, J. P., & Mishra, R. K. (2015a). Effect of bilingualism on anticipatory oculomotor control. *International Journal of Bilingualism*. doi: 10.1177/1367006915572398.

Singh, N., & Mishra, R. K. (2015b). Unintentional activation of translation equivalents in bilinguals leads to attention capture in a cross-modal visual task. *PLoS ONE, 10*(3), e0120131.

Singh, N., & Mishra, R. K. (2015c). Performance monitoring and response inhibition in a saccadic countermanding task in low and high proficient bilinguals. *Frontiers in Psychology*, *5*, 1481.

Soveri, A., Laine, M., Hämäläinen, H., & Hugdahl, K. (2011). Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm. *Bilingualism: Language and Cognition*, *14*(03), 371–378.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*(3), 281–284.

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language, 65*(4), 345–362.

Sunderman, G. L., & Priya, K. (2012). Translation recognition in highly proficient Hindi-English bilinguals: The influence of different scripts but connectable phonologies. *Language and Cognitive Processes, 27*(9), 1265–1285.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371–391.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*(30), 12530–12535.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in second language acquisition, 16*(02), 183–203.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(01), 3–24.

Van Heuven, W. J., Conklin, K., Coderre, E. L., Guo, T., & Dijkstra, T. (2011). The influence of cross-language similarity on within-and between-language Stroop effects in trilinguals. *Frontiers in Psychology, 2*, 374.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge MA: MIT Press. Retrieved from http://www.simplypsychology.org/vygotsky.html#sthash.LtNrwvUS.dpuf.

Wass, S., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology, 21*(18), 1543–1547.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1–25.

Werker, J. (2012). Perceptual foundations of bilingual acquisition in infancy. *Annals of the New York Academy of Sciences, 1251*(1), 50–61.

White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain Responses before and after intensive second language learning: Proficiency based changes and first language background effects in adult learners. *PLoS ONE*, *7*(12). doi:10.1371/journal.pone.0052318.

Yang, J., & Li, P. (2012). Brain networks of explicit and implicit learning. *PLoS One*, *7*(8), doi:10.1371/journal.pone.0042993.

Chapter 7
Attention, Language and Vision

7.1 Interaction of Language and Vision

The Danish linguist Otto Jespersen had pondered on the question of language and vision interaction. Jespersen (1924) provides an example that seems very contemporary with regard to the issue of how linguistic labels can refer to visual objects. Imagine a child who sees a red apple and says the word 'apple' when asked to identify the object. The same child sees a green apple and again only utters the word 'apple'. These instances involve different visual perceptions about colour, shapes and other features of the object. However, the linguistic system condenses these various perceptions and uses an average word that mostly satisfies the demands of the discourse, unless one is explicitly asked to mention the exact perceptions. The child identifies the 'category' to which the object belongs and which the child thinks is true. Similarly, we use quantifier words such as 'a few' and 'a lot of' to represent an average perception. It is very subjective indeed to know exactly how much is the spatial extent of 'a few' for different perceivers. Thus, language encodes and expresses visual perception directly. Jespersen wrote:

> [Y]ou see this definite apple, definitely red in one part and yellowish in that other part, of this definite size and shape and weight and degree of ripeness, with these definite spots and ruggedness, in one definite light and place at this definite moment of this particular day, etc. As language is totally unable to express all this in corresponding concreteness, we are obliged for the purpose of communication to ignore many of these individual and concrete characteristics: the word 'apple' is not only applied to the same apple under other circumstances, at another time and in another light, but to a great many other objects as well, which it is convenient to comprise under the same name because otherwise we should have an infinite number of individual names and should have to invent particular names for new objects at every moment of the day'. (Jespersen 1924, p. 63)

Language thus takes the visual input and transforms them to be used, keeping its own constraints into account. Ideas that Jespersen was talking about have now been empirically examined and it has been shown that language cues are powerful enough to retrieve visual features in their absence. Visual perceptions that do not

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_7

have readily available linguistic tags are not expressed. On the other hand, linguistic labels can activate perceptual features of objects that they refer to (Huettig and Altmann 2007). Lupyan and Ward (2013) showed that people can easily identify visually degraded objects when they hear a name which matches with the object. That is to say, language compresses and brings back awareness in the absence of explicit visual cues. Linguistic words lead to heightened perception of visual objects (Lupyan et al. 2010). Linguistic input can affect visuo-motor systems leading to successful action.

The above-described referential ability of language functionally links it with visual perception in the service of cognition. Thus, linguistic words and objects they refer to are automatically linked, affecting perception. Different examples suggest that listening and looking are tightly coupled both temporally and spatially, which influence learning and development. Infants learn language while they look at things around them. Children as young as two years old look preferentially at objects that spoken language predicts (Mani and Huettig 2012). Attention is shifted towards things when someone utters their name, even sometimes unintentionally (Salverda and Altmann 2011). Before speaking out the name of any object, one looks at it briefly, in order to retrieve its form and meaning (Meyer et al. 1998). Thus, visual attention in the external world often gets guided by language input. These instances suggest that language and vision interact dynamically during different types of psycholinguistic processes. In this, attention plays a crucial role, in helping the cognitive system select those items referred to by language (Mishra 2009). These examples also suggest that the interaction of language, vision and attention is highly multi-modal.

During everyday cognition, we use information from multiple sources. For example, we listen and search objects in the visual environment. One important stance in the discussion about language and vision has been around the linguistic information leading to restructuring of visual perception (Altmann and Kamide 2009). Mental representations derived from language are used in the evaluation of visual material during cognitive processing (Altmann and Kamide 2009). This can happen because language expresses our beliefs about the things and events that are to be cognized (Mishra et al. 2010). Therefore, visual perception of objects does undergo dynamic modification with linguistic input (Spivey 2007). Attention likely plays an important role in this dynamic exchange of information between linguistic and non-linguistic systems (Mishra 2009; see also Huettig et al. 2012). Likewise, visual input often affects the understanding of spoken language. Visual information is used rapidly to comprehend linguistic information (Knoeferle et al. 2005). This sort of interaction between visual and linguistic representations has revealed the very multi-modal basis of language and cognition. Language is therefore not just a mirroring device for mental computation but plays a crucial role in updating and contextualizing this activity. This is possible since linguistic and non-linguistic representations interact dynamically, affecting cognition in different situations.

7.2 Visual Analysis and Language

Ray Jackendoff has proposed a coherent theory of how language and visual information could interact (Jackendoff 1997). Jackendoff compares a theory of language to David's Marr's theory of vision. David Marr's computational theory of vision (Marr 1982) emphasized on the generation of 3D objects during perception. Jackendoff refers to Marr extensively in many of his works on the theory of language, particularly with reference to how spatial language connects to vision (Jackendoff 2011). For Marr, objects are represented at a 3D level that is independent of the perspective of the viewer. Marr's theory of 3D object representation offers a theory that builds holistic structures for perception through recursive and hierarchical processing. These object representations are then transformed as suitable input for linguistic representations. Jackendoff suggests that these representations of objects may not be specific for visual analysis only but also must be modality independent. When one wants to name an object, these representations are used by language. This means that there is an interface between visual and linguistic representations at a basic level. These representations are just 'spatial' and can be used by different modalities including language (Landau and Jackendoff 1993). It is important to note that language refers to things with names and spaces and locations with prepositions and deictic (relating to a word or expression whose meaning depends on who says it, where they are, who they are talking to, etc., for example 'you', 'me', 'here', 'next week') expressions. Spatial representations are thus used by language to encode object identities and locations. Interestingly, not all types of spatial representations are captured by language. For example, Jackendoff suggests that language does not refer to states in an analogue fashion but it filters out many things and uses terms that show continuity of perception. This means, language usage indicates both finished and on-going cognitions. The perceptual world out there is compressed by the perceptual system for language for suitable expression. This indicates a give-and-take at the representational level between language and vision. As more perceptual information arrives, they are integrated online with the existing representations (Spivey 2007). Information from multiple sources is integrated in this continuous flow of cognition.

There are similarities between vision and language that contributes to their representational interaction. Jackendoff proposes that Marr's theory of object identification and vision works recursively, similar to syntax. This is to say that one builds the overall representations by working recursively on the same structures (Jackendoff 2011). Chomsky has always maintained that recursivity seen in human syntax is the most important hallmark of the human language faculty (Chomsky 2007). Humans can produce sentences of infinite length and syntactic complexity since this is what grammar allows as a computational system. Similarly, in vision, from the moment light from objects hit our eye is a process of recursion and computation. In this manner, for Jackendoff, Marr's theory of vision is closest to one of the most influential linguistic theories (Jackendoff 2011). Recursivity is then not unique to only linguistic computation but is also seen in vision. This could be the reason why language and vision collaborate so effectively during mental processing. The other simple explanation could be that language refers to things in the world which we are able to see and perceive visually.

Concerning hierarchical representations in vision, Jackendoff provides examples of the following sort (Landua and Jackendoff 1993). Visual objects appear embedded in contexts. We always view things in terms of how they are part of some larger entity. For example, we know that the face we are seeing is part of a head which is part of a body and so on. Similarly, objects at some locations are perceived in terms of their being situated in some large visual context, i.e. the apple is on the table, the table is in the dining hall, the dining hall is in the house and so on. It is important to note that different visual features of objects perceived through senses must also be integrated for holistic perception. This feature of integration, Jackendoff says, is domain independent and is a hallmark of the human cognitive system. Such representations are then used by language for linguistic expressions. Without this type of interface at the level of representations between language and vision, it might be difficult to say things like 'The red apple is on the table'. Visual and spatial representations connect to linguistic concepts in unique ways.

Linguistic and visual representations can interact in real time, if vision is cognitively penetrable and has dynamic interface with the language and other central executive systems. Abstract mental representations arise from visual analysis of things displayed as well as what language describes. The above views of Jackendoff may suggest an interactive position on cognitive modules. Commenting on his book *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Spivey and colleagues (Spivey and Gonzalez-Marquez 2003) note that Jackendoff does not let go his allegiance to generative grammar and still believes in modules. That is, even though he talks of interactionism, he does so only vaguely and there is enough evidence to suggest that he is still protecting the legacy of the generative grammar. Jackendoff does not dismiss the idea that language is modular and syntactic computation does not require meaning. It must be noted that structure and meaning in language have always been considered to be separate computational processes with no commitments to one another. Spivey's comments also reveal that Jackendoff does not do much to link generative linguistics with other competing proposals that have come from embodied cognition and cognitive linguistics. The idea that structural computation of language is a module has been echoed by Fodor and Pylyshyn as described elsewhere in this book. Nevertheless, it is important to note that Jackendoff has at least attempted to bring some unification among ideas stretching from linguistics to neuroscience and this offers a good ground for empirical exploration. It is of course another matter that his references to Marr's levels or neurophysiology do not do enough to reveal how exactly language interacts with different systems. Jackendoff does not ever offer a full frontal attack on the notion of 'modularity' as it is the case with much of contemporary psycholinguistics, particularly the ones that have revealed a very fluid interactive scenario for language and vision (Spivey 2007).

The above discussion suggests that language and visual representations interact dynamically, affecting cognitive processing. In this chapter, our goal is to understand what role if any attentional mechanisms play in this interaction. It is important to note that neither Jackendoff nor Marr has spoken of attention in this whole business of perception or language–vision interaction. How can we look at something if we do not attend to it? Attention mechanisms help language and vision come closer in different experimental and real-life situations. Things perceived through vision are used by language in a modality-independent fashion. Language and vision have interface at the levels of representations. In sharp contrast to the Chomskyan theorization and Fodor's modularity, this sort of analysis brings language and vision much closer and makes it conducive to refer to central systems of attention in this interface. Vision, in this case, is not merely an input mechanism to identify and perceive objects; but the representations collected are used in different formats. What follows is an attempt to examine eye movements triggered by language comprehension in the presence of visual stimuli that provide robust evidence of language and vision interaction in dynamic situations.

A substantial amount of recent research has focused on the dynamic interaction between language and vision, using eye movements as a response system. That is because eye movements reveal on-going cognitive activities, as information from different sources is considered during perception (Spivey and Dale 2006). Eye movements are also intimately linked to the attention system (Rolfs et al. 2011). Conceptual knowledge derived from language affects where and when we look (Huettig and Altmann 2005). Evidence presented below comes from studies that have used the visual world eye tracking method using visual scenes and spoken words or sentences (Tanenhaus et al. 1995). These studies have had a very significant influence on current understandings of language–vision interaction and the role attention plays in it (Mishra 2009).

7.3 The Linguistic Penetration of Vision

When one says that language affects the way we see and perceives the visual world or vice versa, then this refers to top-down influences affecting vision. Language users bring in world knowledge and experiential information to perceive and recognize objects (Salverda et al. 2011). Without the assumption that visual perception is cognitively penetrable, it is not possible to appreciate how language can trigger attention shift during visual exploration, which is the core theme of this chapter. Similarly, it will not be possible to appreciate how linguistic and visual representations are processed together. Strictly, modular views of cognition gave rise to a type of psycholinguistics where the study of language processing was undertaken as a unitary process, on away from the influences of most other central process. Linguistic information would not penetrate visual modules if modules were to be encapsulated and their processing could not affect one another (Fodor 1987). Zenon Pylyshyn has long held the view that many aspects of vision are cognitively impenetrable (Pylyshyn 1999). Pylyshyn (2009) has described how language and vision work in a more modular fashion. One way to show differences between the two systems is to look at how they use some basic mechanisms. Pylyshyn (2007) proposed that the visual system is inherently constrained only to generate the most recognizable precepts, as there could be so many possible objects given the range of visual inputs one gets. Pylyshyn proposed that demonstratives like ‘this’ and ‘that’ work well for language when it comes to referencing. For vision, on the other hand, direct visual indexing in the form of ‘FINST’ work (see Chap. 2). These are tags that are linked to objects in space and when one wants to individuate them these tags are activated. Several experiments with the paradigm called multiple objects tracking (MOT) have shown the existence of FINSTs during visual attention tasks (Pylyshyn and Storm 1988). Thus, there are two different attentional mechanisms working for vision and language. If this view is true, then one would not expect much interaction between language and vision as well as language triggering attentional shifts in the visual world.

It has been long known that patterns of looking at the objects are heavily influenced by cognitive goals. Buswell (1935) observed that participants inspect the same painting very differently depending on what information they are looking for. Similarly, much recent observation suggests that top-down goals influence where we look even for simple visual tasks (Land and Hayhoe 2001). While inspecting natural images, participants do not look at blank surfaces but look at things and people (Henderson 2003). These suggest that vision is conceptually penetrable and is heavily influenced by the goals of the perceiver. Studies in visual perception have shown that people just do not look at an object because it happens to be visually salient from its surrounding (Foulsham and Underwood 2011). While these examples suggest that looking itself is not always guided by bottom-up factors, but cognitive goals, we have to now consider looking in connection with language.

Many a times, looking is coupled with language. For example, we often orient our attention towards relevant objects and locations depending on what we hear. Speech therefore guides vision in a natural, goal-directed manner. For example, while looking at the picture of a kitchen, and the spoken word ‘knife’, immediately orients the gaze towards the table where there is a good chance of finding the knife, and not towards the wall. This sort of schematic knowledge of things, places and affordances comes handy when looking for things. Top-down knowledge guides vision and perception, leading to successful action (Wolfe et al. 2004). This knowledge is of course is not dependent on linguistic representations alone. Linguistic labels activate this knowledge during multi-modal information processing.

These examples of multi-modal information processing (Henderson and Ferriera 2004) challenge the position that the human mind consists of separate and encapsulated modules that do not interact or influence one another. Most common-sense observations about how language works suggest that it constantly interacts with inputs from vision in different psycholinguistic tasks. Classical cognitive science views of purely a modal symbolic computation in encapsulated modules cannot explain how language and vision interact. This is not true and conceptual structures affect the way we perceive language and also what we see. Seeing is heavily influenced by the background knowledge of the perceiver (O’Regan and Noë 2001).

7.4 Language-Mediated Eye Movements

Are eye movements and attention shifts in the visual domain influenced by language? Some of the most important evidence of spoken language guiding attention in space comes from visual world eye tracking studies. In these studies, eye movements to visual objects are tracked in the co-presence of spoken language. Language-mediated eye movements are a class of eye movements seen during simultaneous processing of linguistic and visual material. Listening to the name of an object and looking at its direction means orienting attention which leads to activation of the oculomotor system (Cooper 1974). Language-mediated eye movements therefore manifest this interaction between linguistic and visual information. These eye movements arise as a function of processing the linguistic input from speech. Given a few line drawings and a spoken word, very fast eye movements are seen towards visual referents that match spoken words (Altmann 2011).

Visual world studies vary in many ways and have provided a wide range of data that have consequences for the study of language and vision interaction (Huettig et al. 2011). It is important to note that these studies were not designed originally to examine attention but psycholinguistic questions dealing with spoken word recognition or sentence processing. Nevertheless, their very design—which includes visual displays and spoken words, and where eye movements are measured—provides an important framework to study attention in cross-modal scenarios.

One group of studies has looked at the time course of lexical access during spoken word recognition, as fixations are time-locked to speech input (Huettig and McQueen 2007). The idea is to see if perceivers look at some objects preferentially, thus attending more to them compared to other objects, as a function of their language understanding. Visual world data provide real-time data of cognitive processing as it unfolds over time and more crucially, how the visual system considers alternatives before settling on the target objects. Cooper first observed that participants are more likely look at pictures of 'zebra' or 'snake' when hearing the word 'Africa' than towards unrelated objects (Cooper 1974). This shows that background knowledge stored in the long-term memory immediately guides visual attention towards relevant objects (see Mishra 2009; Huettig et al. 2012 for reviews). Listeners look at semantic competitors of spoken words (Yee and Sedivey 2006). It has also been observed that such shifts in visual attention happen when objects in the visual field match in shape and colour the things described in spoken language (Huettig and Altmann 2007). Dahan and Tanenhaus (2005) observed that perceivers immediately looked at the picture of a rope when they heard the word 'snake'. Apart from the quick activations of conceptually or perceptually related objects, eye movements have also been shown to covertly orient towards objects whose names partially match with the spoken words.

In the psycholinguistics of spoken word comprehension, the Cohort model (Marslen-Wilson and Zwitserlood 1989) assumes that when one begins hearing a word, at each moment, multiple word candidates are active which resemble the current input. For example, with the input word 'cat', with the arrival of the phoneme ‘ca’, words such as ‘carrot’ and ‘cab’ are active. This sort of activation happens in a rapid spreading activation style. As auditory processing progresses, activations that are not relevant die out and finally the target is perceived. During simultaneous processing of visual objects and spoken words, Allopena and colleagues (Allopena et al. 1998) showed that when one hears the word ‘Candy’,

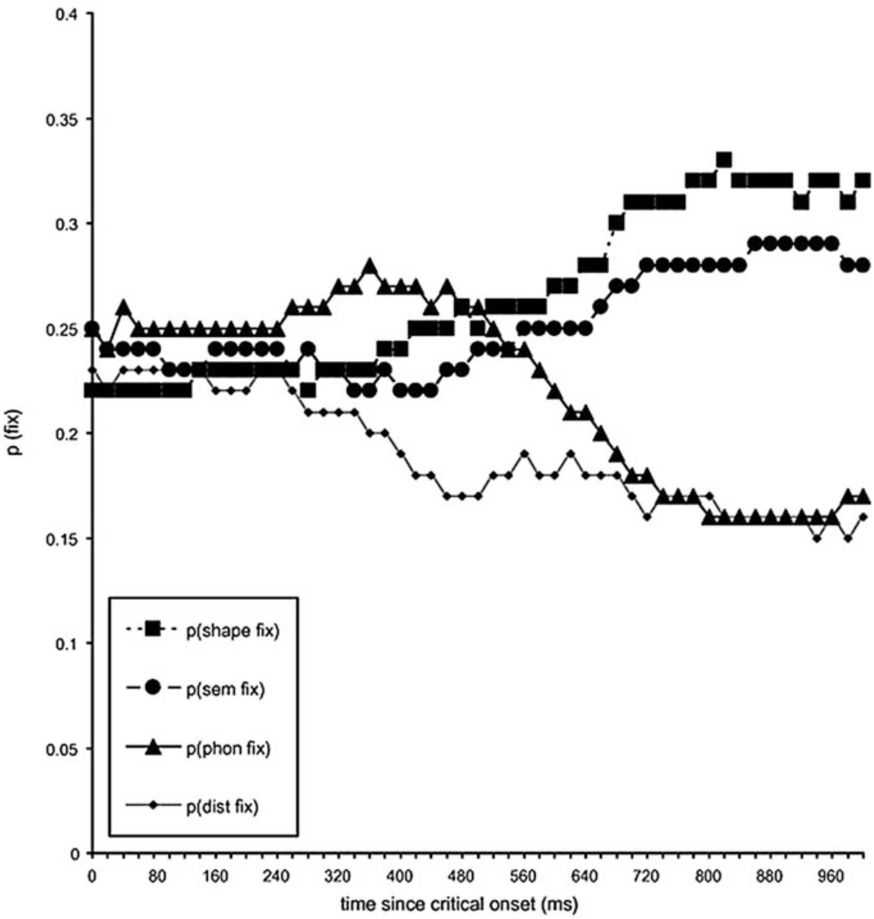

Fig. 7.1 In this study, participants looked at an array of line drawings and listened to a spoken word. The panel had a phonological competitor, a shape competitor, a semantic competitor and unrelated distractors. It was of interest to see how participants look at these various objects as soon as they hear the spoken word. The figure shows fixation proportions to different objects since onset of the spoken word. The proportion of fixation curves shows that at the onset of the visual objects, there is not any bias towards any particular object. Within 200 ms, fixations towards shape, phonological and semantic competitors start to diverge from distractors. It is also important to note that among the three competitors first looks are towards the phonological competitors. The figure shows competition and attentional bias as the spoken signal unfolds over time. *Source* Huettig and McQueen (2007, p. 466)

participants also look at the picture of a 'candle', among other distractors. Interestingly, this kind of attention shifts towards an object whose initial sounds match with the spoken targets which happen in the presence of the target itself. This suggests that language input affects attention shifts towards multiple objects as cognition progresses. As soon as spoken words are comprehended, there is an immediate competition between phonological, conceptual and other perceptual features (Huettig and McQueen 2007, Fig. 7.1).

Huettig and McQueen's study clearly shows that with the onset of any spoken word, a range of perceptual as well as conceptual features are activated. Depending on how these match to objects in the visual field, or how the objects in the visual field afford this information, attention is biased. This biasing of attention is dynamic and is controlled by both visual and linguistic information. It shows that attention system is sensitive to language input during simultaneous processing of visual and linguistic information. Recently, Huettig et al. (2011) provided further evidence that suggested that these activations are limited by one's literacy level. In this study, low literate individuals did not orient their attention towards phonological competitors of spoken words where as high literate individuals did so immediately. Language perceivers also immediately orient their attention towards printed words whose phonological representations match with the spoken input (McQueen and Viebhan 2007; Mishra and Singh 2013).

It appears that this sort of quick interaction between linguistic input and visual search has an evolutionary lineage. Even infants orient their attention towards objects whose names match partially with spoken input (Mani et al. 2012). Infants always look around when they listen to speck. Somewhere else, I had quoted Wittgenstein, who also thought that language and visual object identities merge to generate the concrete concepts. Spoken words activate latent semantic concepts and the perceiver actively searches among the available visual entities for a best match. These studies have not just shown how interactive and parallel lexical access is but how linguistic and visual representations could interact online affecting search. Many of these studies are conducted when perceivers look at an array of line or coloured objects and listen to some speech, often without any explicit task. Eye movements recorded show quick orientation of attention towards related objects within the first 100 or so milliseconds (Altmann 2011). These certainly suggest that linguistic input has an immediate effect on visual search.

One argument in this case is that the oculomotor system is sensitive to visual objects that best match with the available input from spoken words (Huettig and Altmann 2007). Objects that match best with the spoken descriptions bias attention towards them in the presence of other distractors. Interestingly, this sort of eye movements towards objects semantically or conceptually related to the spoken words is rather automatic and sometimes occurs even without the participant's conscious awareness. Language cues can also facilitate visual search by activating perceptual knowledge that is used by the visual system. Spivey and colleagues (2001) showed that when language is presented concurrently with visual displays, the search for targets with conjoined features is facilitated. Moreover, in this type of search where normally distracters influence search efficiency, language leads to facilitation. This shows that language led to efficient object search in spite of visual crowding by distracters. Language activates objects represented in the long-term memory and this guides attention in the visual world.

A second category of visual world studies has looked at anticipatory eye movements towards context appropriate objects much before the linguistic signal fully unfolds. In these studies, the main interest has been to examine if perceivers consider various aspects of the visual scene and linguistic information in order to make predictive eye movements towards objects much before language has described them. Altmann (2004) showed participants clip art images that had a man, a woman, a cake and a newspaper. Altmann's interest was to see if depending on the type of sentence participants listen; anticipatory eye movements towards the context appropriate objects will be seen. Anticipatory eye movements are seen when people already start looking at an object when it fulfils certain contextual demands. Altmann and Kamide (2007) showed that participants look at the picture of a cake much before the word cake has been mentioned in the sentence 'the boy will eat the cake'. This happens because participants use information from the verb 'eat' to look at an object that is 'eatable' among other objects. Participants also show sensitivity towards agents and what they may do in real world to look anticipatorily at visual objects. Kamide et al. (2003) observed that people look at a 'motorbike' after hearing the sentence fragment 'the man will ride…'but look at a 'girl' after the sentence fragment 'the girl will ride..'. This indicates rapid use of contextual and real word knowledge to programme eye movements towards relevant objects. Anticipatory eye movements can also been seen when there are typical linguistic elements in languages that trigger the expectation of certain types of objects. For example, Mishra et al. (2010) observed that Hindi speakers looked ahead towards an object that shared gender with the adjective.

Do linguistic and non-linguistic anticipation share underlying mechanisms? If this is so, then attentional mechanisms that drive eye movements during anticipation tasks may could be similar for different domains. Rommers et al. (2013) examined if listeners can anticipate objects that match the shape of target words in a visual world study. For example, participants listened to sentences like 'in 1969 Neil Armstrong was the first man to set foot on the moon'. The visual display contained the picture of a tomato among other distractor object. Participants looked anticipatorily at the tomato much before the critical word 'moon' arrived in the speech stream, indicating the influence of prediction on visual attention. The authors also found that these subjects were also good at target identification in a Posner-like cueing task with predictive endogenous cues. These results suggest that both linguistic and non-linguistic anticipation may share underlying mechanisms. These data suggest listeners use information from language and world knowledge to look at objects even much before language has described them. Thus, anticipatory eye movements further show how visual perception is affected by top-down knowledge. For instance, Mishra and Singh (2014a) examined if Hindi speakers activated homophones of words when the sentence context was biased towards the dominant meaning. Homophones are words that have the same spelling but mean different things for example the English word "bank". In this example, this word has a dominant meaning which is its use to refer to financial institutions. Its subordinate meaning is to refer to river bank. These dominant and subordinate meanings of homophones are linked to frequency of use. Several studies have shown that when one listen to an ambiguous word that has homophones, both the dominant and subordinate meanings are activated (Rayner and Duffy 1986), while others have suggested that depending on context of the sentence, either the dominant or the subordinate meaning is activated (Vu et al. 1998). Mishra and Singh (2014b) examined these issues with regard to the processing of Hindi homophonous words using the visual world eye tracking paradigm. The authors presented homophonous words such as “choti” both in neutral sentences and also presented them in dominant-biased contexts. The sentence context biased the word towards its dominant meaning. In this example, the dominant meaning of the Hindi word “choti” is “hair lock”, whereas the subordinate meaning is “mountain top”. Participants listened to sentences where these words were in dominant or in neutral context. They also saw that a visual display contains four pictures where in one experiment, one of the pictures was related in shape to the subordinate meaning of the word. In the other experiment, the picture was conceptually related to the word. The authors were interested if participants still activated the shape and semantic competitors of subordinate meaning of the homophone when the sentence context was biased towards the dominant meaning. Eye tracking evidence (Fig. 7.2) suggested this in spite of the contextual bias towards the dominant meaning. These data also suggest that sentential context and top-down knowledge influence activation of shape and semantic information during language–vision interaction as seen in visual world studies. These data also show that language leads to shifts in attention in an anticipatory fashion.

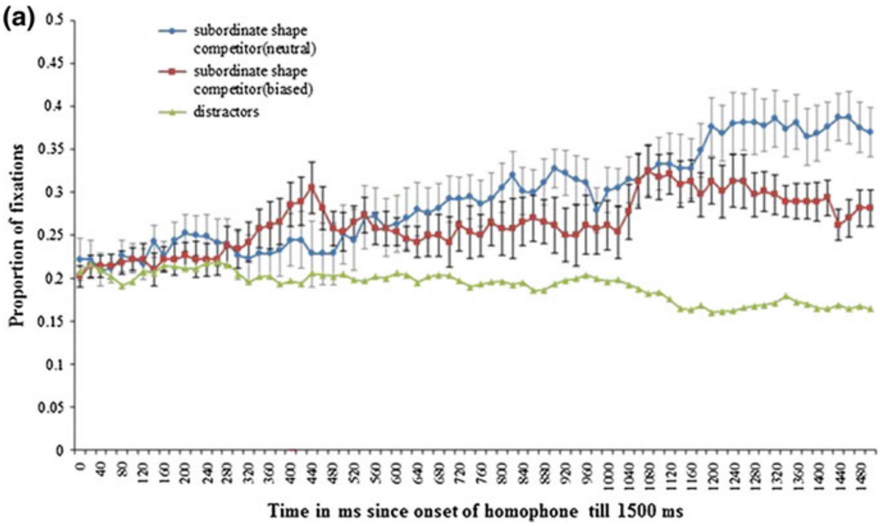

Fig. 7.2 Proportion of fixations towards **a** subordinate shape competitors **b** subordinate semantic competitors. *Source* Mishra and Singh (2014b, pp. 457–458)

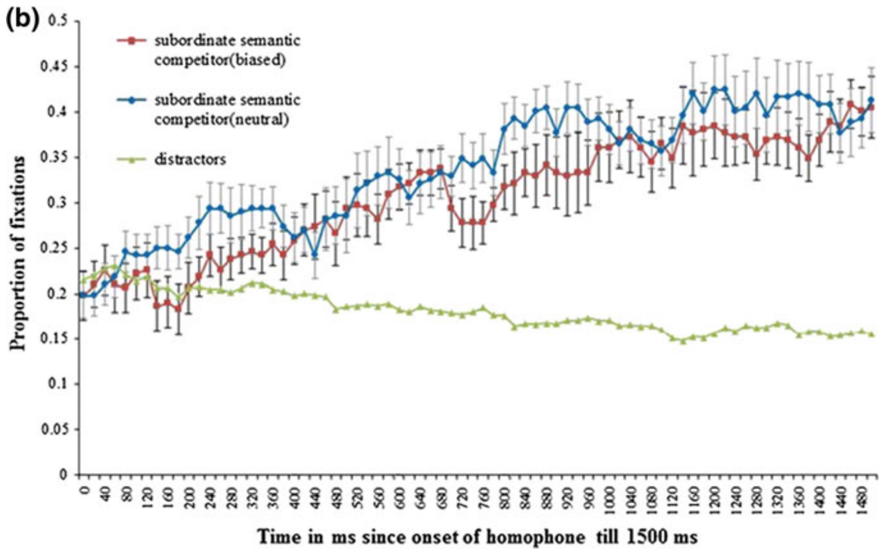

Fig. 7.2 (continued)

A third category of studies has looked at how complex visual scenes and action sequences are understood in the presence of language. Complex scenes themselves provide abstract mental representations that are visual (Knoeferle and Crocker 2006; Henderson 2003). Knoeferle and Crocker (2006) presented stereotypical actors performing actions and sentences that were temporarily ambiguous. Eye movements showed that listeners used both representations from language and depicted scenes to disambiguate comprehension. Language comprehenders seem to dynamically update their knowledge considering both visual and linguistic representations and eye movements manifest this. Listeners also seem to restrict their looks to the most pertinent object that fulfils the discourse demands. Chambers et al. (2002) observed that given the sentence ‘Put the cube inside the can’, listeners looked at the picture of a container that was large enough to accommodate the cube in question. Thus listeners rapidly integrate situational knowledge to look at the appropriate visual world. Many a times language is ambiguous and does not refer to things straight away. In an influential study involving structurally ambiguous sentences and a visual display of objects, Tanenhaus et al. (1995) found saccades towards different objects helping in disambiguating the language input. Viewers often look at objects described by the speech incrementally with the hope of comprehending the sentence. Incremental parsing of sentences lead to shifts of attention towards objects that gradually emerge as the suitable and correct referents.

Attentional selection of visual referents therefore seems to be quite dynamic and an outcome of the interaction among visual analysis, linguistic input and world knowledge (Huettig et al. 2011). Anticipatory eye movements are affected by a consideration of the immediate visual context and the language which describes it.

For example, once inside a movie theatre, one knows where the exit might be and direct saccades appropriately when needed. Spatial location of objects arrived out of schematic knowledge helps in directing vision. Altmann and colleagues studies have shown that participants immediately look at the potential actors who are best suited to do the action and also keep track of events as they may change over time. Attention, in this scenario, helps satisfy the pragmatic and contextual requirements imposed by language. Language here modifies visual representations (Mishra 2009).

Spoken language drives attention towards objects by creating some kind of attentional bottleneck in the attentional system. They cause interference in visual processing by diverting attention towards things that are mentioned by language. Further, language input immediately changes or substantially modifies our perceptual world. The referential power of language leads the attentional and visual system towards comprehension. The above-described studies have shown that processing language leads to immediate changes in the ways one looks at things. Spoken words activate conceptual information which leads to attentional orientation in space. This could be simply because language always talks about things that we see around. On the other hand, we also learn about things by looking at them and at the same time hearing their names. This give-and-take between visual and linguistic information constitutes a strong cognitive base.

7.5 Eye Movements During Looking at Nothing

What we have discussed so far suggests that spoken language leads to attention shift towards objects that it directly or indirectly refers to. Everyday, cognition involves circumstances where objects referred to by language may not be readily available at particular locations in the visual field. It has been suggested that looking at objects helps spatial indexing of their locations in memory for later retrieval and use in cognitive processing (Spivey 2004). Although it is not yet clear if attention is object or location based (see Vecera and Farah 1994), viewers typically look at locations that contain objects (Henderson 2003). In the case of visual world type studies, it is quite natural to expect eye movements to go towards objects that match the acoustic input. It is strange that many a times people look at blank locations that were previously occupied by an object. Altmann (2004) showed that viewers look at the location where an object was previously if the sentence refers to it. Altmann and Kamide (2009) showed that listeners update mental representations and look at locations likely to be occupied by discourse relevant objects. Altmann argued that this is only possible when the viewers have already constructed a mental representation which has every detail of the scene schema and information related to affordances. This mental representation then guides attention towards locations when a sentence arrives that describes some aspects of this world. It is also likely that in the absence of the scene, viewers indulge in mental imagery which eye movements reflect (Johanson et al. 2006).

Mishra and Singh (2010) used this technique to explore if viewers simulate abstract motion while listening to fictive motion sentences and look at blank locations of the display when the visual object is not there. Fictive motion sentences are about inanimate objects incapable of biological motion, but with a motion verb there is illusory simulation of motion. For example, in the sentence 'The valley runs across the forest', participants treat the object valley as if it is literally capable of running (Richardson and Matlock 2007). Interestingly, the eye movements scanning patterns showed evidence of motion simulation in the absence of the scene. Participants looked across the length of a space previously occupied by the picture of a bridge when they heard a fictive motion sentence about it. Such evidence does suggest that visual attention towards locations is triggered by enactment of sensorimotor events and also sometimes by the construction of a rich mental representation. It has been argued that looking at nothing could indicate the role of eye movements in memory retrieval (Spivey and Geng 2001). Fixations and saccades help extract the spatial indices that have been tagged with certain objects or mental representations. There have been suggestions that looking at nothing could reflect a merging of both linguistic and visual representations (Ferreira et al. 2008). Sensorimotor experiences associated with the visual input are activated as soon as the linguistic representations match with visual information (Mishra et al. 2010). In this debate about what may drive eyes towards vacant locations, different proposals have been offered. Richardson et al. (2009) propose that 'spatial indexing' is the mechanism which helps eyes to go back to locations that were occupied by objects. This account considers eye movements working as links between external memory systems which is used in offloading cognition into the environment. When one looks at an object, its location is maintained in an internal representation. When that object's name is again processed, this location helps access many other perceptual features related to that object and this helps in dynamic cognition. It is not quite clear if one can propose both an internal and an external memory system for object representation and locations, but Richardson et al. (2009) proposed that a dynamic cognitive system may use both for optimal efficiency. Linguistic input thus helps in the immediate activation of such representations from memory and eye movements and shifts in attention reflect it. This evidence also suggests that language leads to attention shifts towards locations where objects are found.

7.6 Attention in Language Vision Interaction

The visual world studies show visual and linguistic interactions during spoken word recognition and real-time conceptual activation. Fundamentally, in this paradigm, linguistic input affects the attention mechanisms so that visual search is facilitated. One can always say that attention shifts towards the objects that matches the spoken input is an outcome of some strategy (Trueswell and Tanenhaus 2005). Truswell and Tanenhaus (2005) have dubbed it as the 'closed set' problem. Since there are only four objects, it is quite likely that participants infer that one of the objects is linked to the spoken word in some manner and then they preferentially look at it. Given the fact that attention shifts happen within 100 ms of the spoken word input (Altmann 2011), it is quite unlikely that some strategy is at work. Altmann (2011) suggests that spoken language can guide eye movements within 100 ms, regardless of what the visual world contains or the task demands are. Several psychophysical studies on the oculomotor programming have shown that it might take around 200 ms to programme a saccade towards an object (Saslow 1967). This very short latency of language-mediated eye movements can only indicate that they are non-strategic and automatic (Mishra et al. 2012). It is also not likely that participants have time to sub-vocally name all the pictures and then match the linguistic input. These fast saccades also show that top-down contextual knowledge affects visual processing immediately as reflected through eye movements. Nevertheless, automaticity eye movements does not mean that such eye movements are not sensitive to the nature of the task and cognitive goals of the participant (see Salverda et al. 2011; as also, Mishra et al. 2013 for recent reviews).

In visual search studies, it has been shown that selective attention is influenced by top-down bias which leads to selection of objects (Desimone and Duncan 1995). Many visual search studies have shown that top-down knowledge biases oculomotor programming (see Noudoost et al. 2010). Visual search mechanisms during simultaneous processing of spoken and visual material are an outcome of both top-down and bottom-up factors. In visual world studies, speech signal describes some of the actions or one of the objects in the scene. As long as there is no speech, most objects will be competing for selection with equal probabilities. There is a good chance that some objects might attract attention because of their saliency or lexical frequency. Language input accentuates the probability of some objects for selection (Wolfe 1994). Speech input modifies attentional states, leading to bias towards one object (Spivey 2007). Attention selection in this case begins with immediate orientation of attention towards that object which speech either directly describes or indirectly refers to. Since in many visual world type studies there is no strict instruction given to the participants, visual search should progress quite naturally. If there is an instruction to look at a particular object, then also one sees covert activation orientation of attention towards another object (Dahan and Tanenhaus 2005). Competition for selection determines sustenance of attention as speech unfolds. It is important to note that most models of selective attention have taken only visual search tasks and not cross-modal stimuli. There is thus danger in seeking applications of these models to visual world type data (see Huettig et al. 2011).

Another way to understand how attention is involved in the visual world type studies is to consider the attention state as a landscape (Spivey 2007). This landscape changes continuously with several attractors (Kukona and Tabor 2011). For example, the display has four objects and one object is referred to by the speech signal. As soon as the speech is processed, one of the objects becomes a strong attractor and biases attention. This object competes for attention selection more strongly compared to distractors. Distractors also simultaneously compete for selection depending on how they are related to the target. Therefore, attention landscape is in a constant state of dynamic change and this is determined by the speech input. Given a visual world with objects, perceivers also actively try to find suitable referents as the linguistic input suggests (Mayberry et al. 2009). Attention in this case serves as a glue that sustains the interaction of both linguistic and visual representations. Representations derived from the visual analysis are sustained in memory and also connect to knowledge stored in the long-term memory system. Working memory plays an important role in the triggering of language-mediated eye movements and cognitive processing (Huettig et al. 2012).

Huettig et al. (2011) have proposed a general model to account for eye movements and attention shift during language-mediated visual search. In this model (Fig. 7.3), there is an explicit characterization of the role of working memory as a binding factor that helps in the interface between emerging linguistic and visual representations. The authors suggest that the key to understand how language drives eye movements and attention in the visual world lies in the long-term memory. Long-term memory contains information about all the concepts we know as well as semantic information about them. If one sees a shape and it looks like an apple, then it is because this visual analysis triggers the conceptual representation of 'apple' stored in the long-term memory. According to the model, when one hears a spoken word, the word's phonological structure is activated which in turn activates the semantics (meaning). In a cascaded fashion, this activation leads to the further activation of the visual features of the object. This might explain why attention is biased towards one object that is the best match with the spoken word input compared to other distractors. Similarly, when one sees the picture of an object, the visual features lead to the activation of the semantic features (Huettig and Altmanneye movements in an active 2005). These activated information about the

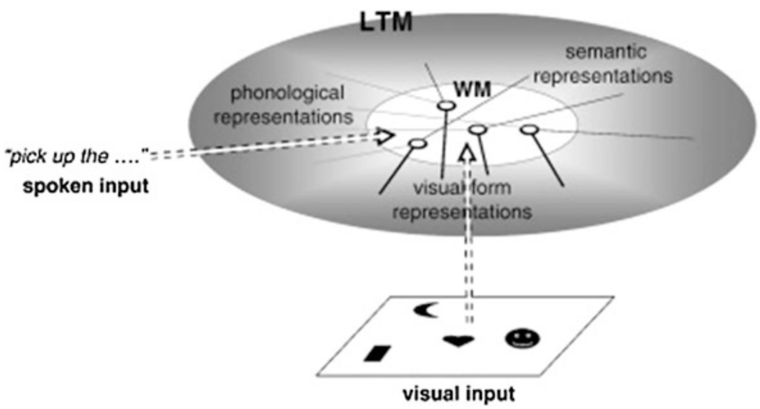

Fig. 7.3 The model considers long-term memory as a store house of visual, phonological and conceptual representations. Both visual and linguistic input activates semantics and this is temporarily kept in the working memory. Objects in the visual space are kept in the visuo-spatial working memory. When one hears a name of one of them, this is tagged to the location of the object. Therefore, the model assumes that linguistic information and visuo-spatial information interact dynamically affecting attention shift and eye movements. *Source* Huettig et al. (2011, p. 141)

object are temporarily stored in the working memory for further processing. On hearing a word in the presence of some familiar objects, there is stimulation of representations and the most pertinent object and its location are selected for attention shift and eventual saccade. This model has made an explicit attempt to link findings from visual search and use it to explain language-mediated eye movements. It is important to note that 'spreading activation' of concepts and a dynamic flow of information between phonological and semantic levels play a crucial role in this mechanism.

Do we look at objects based on what attributes they may have? Attention has been known to guide eye movements in an active manner (Findlay and Gilchrist 2001). Humans are foraging creatures, and attention shifts and eye movements are used to bring new information for further processing and allocate foveal vision (Klein 2000). Looking around the environment and shifting attention towards relevant things help exert goal-directed action on them. The visual world paradigm fixations and saccades therefore reveal a dynamic competitive state rather than certainty (Spivey 2007). When one listens to spoken language, one starts actively sampling the environment to see if the language makes sense or if the language can lead to some satisfactory action. On listening to the word 'fire', one must find out the escape route and this means shifting attention in a much more active manner towards the desired location. Much recent theorization on attention is also aimed at explaining its social and experiential role (Posner 2012). Posner (2012) does not offer a social theory of attention as such. Posner shows how orienting, engaging and controlling attention is crucial for a range of cognitive functions and these skills develop during infancy. Attention in this case is not merely illuminating things or locations that the spoken language refers to. Attention orients towards things that are 'afforded' by the scene as well as the language (Altmann and Kamide 2007). Attention shifts therefore are an outcome of sensorimotor knowledge and affordances of objects and situations (Meier et al. 2012; Mishra et al. 2010). Human perceivers do not just see a scene but also actively follow changes that agents and objects in the scene might go through (Knoeferle and Crocker 2006). Imagine you are seeing two people talking. One person, for example, uses filthy language, and your attention automatically and anticipatorily moves towards the face of the perceiver to see the consequences of this kind of language use. Human observers keep track of such changes in the environment and systematically shift attention. Anticipatory shifting of attention or covertly attending towards a location forms the active sampling of the environment.

Visual world studies and their findings in language and vision interaction have certain limitations. So far, most studies have used a static set of few isolated objects or at best ensemble of clip art pictures. When spoken language refers to one of the images, attention shifts towards it. Often, spoken language does not refer to only one object but to a collection objects. Further, everyday life experience tells that objects and actors could be in perceptual motion, dynamically moving and executing actions in the visual field. How does language-mediated attention work in such cases? Imagine yourself watching two groups of children playing football. When one utters the phrase, 'the children in blue', you need to keep track of all or most children who have put on the blue clothing. This situation is similar to the multiple objects tracking experiment that Pylyshyn has done over the years. These studies have shown that attention can be spread over a few objects at the same time. Although Pylyshyn's experiments have involved only visual objects, it is possible that spoken language input could additionally amplify the situation and help attentional sustenance. Depending on what exactly language refers to, i.e. a single object or a group of objects, attentional mechanisms show dynamic modulation. Language helps attention 'split' either among many discrete objects that dynamically shift positions or treats them together as a single object. For example, single nouns and quantifier terms in language affect language-mediated eye movements differently. Huang and Snedeker (2011) found that listeners interpret quantifiers like 'some' differently as time elapsed. Initially, listeners interpreted 'some' with the understanding that it also included the meaning of 'all'; however, after 800 ms, listeners excluded all referents that were compatible with 'all'. It might as well be the case that language-mediated attention is contingent on grammatical and other facts of individual languages and their typological features.

7.7 Automaticity of Language-Mediated Eye Movements

Does language drive attention automatically towards objects and locations? It appears that spoken words do drive attention in an automatic fashion even when the perceiver wishes to exercise some type of control. The evolutionary coupling of looking and listening may account for the quick integration of visual and linguistic stimuli, even automatic eye movements with regard to spoken stimuli. Spoken language comprehension not only leads to semantic access but also creates bias in attention states. Speech even if irrelevant can have distracting effect on the main task (Elliott and Briganti 2012). When subjects are given a visual task and at the same time some type of meaningless speech stimuli is played, it has been found that participants cannot avoid paying some attention to this irrelevant speech and therefore become slow in their main response (Ellermeier and Zimmer 1997). Similarly, emotional content of speech captures attention (Bertels et al. 2011). Speech, when meaningful, influences attention in a more serious manner. Speech can drive attention unintentionally during visual search. This view contrasts with what visual world studies have suggested so far. Visual world studies seem to suggest that only when a visual representation matches with a linguistic input, attention is biased towards it (Huettig and McQueen 2007). That is, attention shift is contingent upon 'meaningfulness' of a spoken word. But it seems even pure tones and other task-irrelevant speech fragments capture attention (Jones and Macken 1993).

It is one thing to say that language leads to automatic shifts of attention towards speech relevant visual objects, and another to say that language users look at certain objects guided by specific goals or tasks. In the second case, language-mediated eye movements are not automatically driven by lexical access, but what language comprehenders prefer to look at for performing a certain task. In order to understand these distinctions, one has to understand the various manipulations that are used in a visual world task. Salverda et al. (2011) proposed a language comprehension-based and a task-based view of language-mediated fixations seen during visual world studies. In the comprehension view, listeners are not given any task and they see some objects and a spoken word appears (Huettig and McQueen 2007; Huettig et al. 2011; Mishra and Singh 2014a, b; Huettig and Altmann 2007). In this passive listening version, with the onset of speech signal, eyes automatically move towards the object that matches with the speech. As the speech signal is processed incrementally, eyes move towards that object that is the eventual target. Huettig and Altmann (2007) found that when a sentence context was biased towards one of the meanings of an ambiguous homophone words, participants looked more frequently towards the shape competitor that matched this representation in comparison to unrelated distractors. But as soon as the key critical word arrived in the speech signal, participants started to look at the object that was related to this critical word. This indicated that looking in the visual world studies is driven by lexical information and is automatic. This view may suggest that goals of perceivers and task demands are ineffective in controlling fixations in a visual world study. There have been many studies where participants are given a clear task that may involve reaching towards a particular object or shifting one object from one location to another (Allopena et al. 1998). In these studies, fixations towards potential objects are partially influenced by the task demands. Language users also look at objects keeping track of both visually derived information and linguistic information (Chambers et al. 2002). Therefore, when a task is given, looking towards visual objects is a result of information derived from both bottom-up and top-down goals. Salverda et al. (2011) thus provide a linking hypothesis that links both a task-based approach and a language comprehension-based approach to understand attention shifts triggered in visual world studies. Therefore, in visual world studies, language-mediated eye movements are sometimes automatic and sometimes are under the control of perceiver's goals.

In the visual world eye tracking paradigm, participants generally shift their attention towards relevant objects with the onset of the speech. It is important to know if listeners can prevent such eye movements when the speech is not relevant for the main task. Salverda and Altmann (2011) presented participants with line drawings and asked them to make an eye movement towards one target object that underwent a colour change. Just before this colour change, participants heard a spoken word which either referred to a distractor or referred to an object that was not in the display. Salverda and Altmann (2011) observed that saccades towards the object that changed colour were slow when the spoken word referred to one of the other distractor objects present, than when it was not so. Listeners were slow in detecting a location change of a target visual object in the presence of a spoken word which was not directly relevant for the task. Participants were faster in their response when the spoken words directly mentioned the object that underwent the colour or the location change. These data indicate both facilitation as well as interference with visual search in the presence of spoken word. Most importantly, spoken words that referred to a distractor caused disruption with the primary visual task. The authors interpreted their findings suggesting that, like Strop interference seen with written words, spoken words are processed automatically and this leads to attention capture. It might be difficult to deploy selective attention to targets in the visual modality while ignoring spoken words. This research implies that the automatic activation of semantics of spoken words creates some kind of attention bottleneck and reorients saccadic programming, which is goal driven. Thus, spoken words act more or less like exogenous cues that capture attention and disrupt goal-driven processing in a cross-modal scenario. The results suggest automatic processing of spoken words that affected oculomotor programming. Recently, Mishra et al. (2012) reviewed studies from visual world and visual search paradigms and observed that in order to establish automaticity of language-mediated eye movements, one has to either control attention or increase search difficulty with higher set sizes, as has been done traditionally with many visual search tasks. Importantly, most visual world studies have at best presented three or four line drawings.

Other research has shown that processing spoken words can unintentionally facilitate visual search. Lupyan and Swingley (2012) asked participants to name objects during a visual search task. Search was found to be more efficient when the spoken verbal label matched the search target and there was higher interference when the name was different from the search target. This indicates that spoken words processed unintentionally narrow down search when the extracted semantics match the search target. Visual referents that are related to the spoken words compete with the target for selection and bias attention, causing interference. Verbal labels have even been shown to guide attention automatically during visual search in a top-down manner (Soto and Humphreys 2007). Cleland et al. (2012) asked participants to respond to a visual stimulus (a word, face or a shape) in the presence of a spoken word. Response times were higher when the spoken word's uniqueness point was closer to the appearance of the visual stimulus and also effects were stronger as the gap between the spoken word and the visual stimulus decreased. Further, response interference was higher when the visual stimulus was a word compared to a shape and a picture. Previous eye tracking studies have shown that bilinguals covertly activate cross-language information rapidly and automatically (Marian and Spivey 2003a, b; Spivey and Marian 1999; Weber and Cutler 2004; Mishra and Singh 2013). Bilinguals have been shown to activate translation equivalents of words automatically and also words that are similar in phonology (see Sundermann and Priya 2012; Guo et al. 2012). It can safely be assumed that spoken words are difficult to ignore when one wants to focus on a visual task. One immediately activates semantics of the task-irrelevant spoken word and this causes some sort of attentional bottleneck which causes delay in the visual processing. It is also important to note that when the spoken word directly refers to the visual target, processing is sometimes facilitated. These data further strengthen the theoretical stand that vision is cognitively penetrable and audition and vision interact dynamically at many levels.

These data suggest that spoken words automatically receive rich phonological as well as semantic processing and thus create attentional bottlenecks during search for visual targets. This can happen when both spoken words and visual targets draw on the same attention mechanisms for processing. Even unintentional processing of spoken words can lead to shifts in attentional states which in turn may lead to changes in goal planning and action selection (Spivey 2007). Previous studies have shown that visual and spoken words that are semantically related to the targets or distractors influence saccadic behaviour by driving eye movements towards these task-irrelevant referents (Moores et al. 2003; Telling et al. 2010). It is quite possible that spoken words enter into short-term memory and thus capture attention (Klatte et al. 2010). Attention capture due to spoken words can happen even when subjects are not explicitly instructed to keep verbal information in working memory (Downing 2000). Neurobiological studies on speech processing suggest that spoken words can activate important brain circuits even when listeners do not pay any attention to them. Shtyrov et al. (2010) found that despite explicit instruction to ignore the spoken words, the results showed that within 120 ms of the spoken word onset, the negative going ERP effect was modulated by the frequency and lexicality of the words. This suggests that words can affect neuronal processing without attentional engagement. While these data could be interpreted as suggesting strategy-free processing of spoken words, they also suggest attention capture.

Although I have invoked the notion of automaticity quite a lot in the preceding paragraphs, the notion of automaticity still remains contested in cognitive psychology (Logan 1978; Jonides et al. 1985). Cognitive processes are automatic when they are not under one's voluntary control and occur unintentionally. The most crucial test perhaps is if they still occur when attention is controlled for another task. Moors and De Houwer (2006) have argued that it might not be enough just to suggest that some mental process is automatic unless we identify what component of automaticity that process recruits. For example, under this compositional view of automaticity, while activation of semantics in a colour word Stroop task could be unintentional, it also demands selective attention. If attention is not paid to the written word, one may probably not see Stroop interference. Further, it is crucial to examine if automaticity holds under task demands or when attention is strictly controlled for. The load theory of attention (Lavie 2005) argues that one may not see any automatic interference with a visual task caused by auditory stimuli when the visual task demands excessive amount of selective attention. Thus, it remains to be seen if spoken words in visual world paradigms capture attention when there is an intense primary visual task to be performed. Another possibility may be to examine language mediated eye movements with working memory load.

7.8 Conclusion

The evidence discussed in this chapter shows that linguistic input can influence visual processing significantly. More importantly, both visual and linguistic representations affect attentional states during multi-modal processing. Further, individual differences affect this interaction. Attentional orientation is often contingent upon incremental processing of speech. The visual world paradigm has been very useful in sorting several psycholinguistic issues related to speech processing and sentence comprehension. More significantly, data from visual world paradigm also demonstrate the dynamic patterns of interactions seen between linguistic and visual information. This chapter has shown that successful cognition results when information pertaining to different systems interacts dynamically. Furthermore, both visual and linguistic representations influence perception (see Mishra and Marmalejo-Ramos 2010 for a review).

Top-down knowledge along with representations derived from visual context facilitate language-driven search. Further, language seems to move attention in the real world in an automatic fashion. This only shows the co-patterning of both representations and the everyday use by the cognitive machinery. It is also important to note that prevailing theories of attention have no paid attention to such multi-modal processing and there lies the weakness of many such models. While they can account for how we merely select visual objects, they cannot explain what it is that we see and understand in the presence of language is. Language too does not express everything that the visual system generates.

References

Altmann, G. (2004). Language-mediated eye movements in the absence of a visual world: The blank screen paradigm. *Cognition, 93*(2), B79–B87.

Altmann, G. T. M. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta Psychol. (Amst.), 137*, 190–200.

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502–518.

Altmann, G., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition, 111*(1), 55–71.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419–439.

Bertels, J., Kolinsky, R., Bernaerts, A., & Morais, J. (2011). Effects of emotional spoken words on exogenous attentional orienting. *Journal of Cognitive Psychology, 23*(4), 435–452.

Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art*. Oxford, England: University of Chicago Press.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language, 47*(1), 30–49.

Chomsky, N. (2007). Approaching UG from below. *Interfaces + recursion = language*, 1–29.

Cleland, A. A., Tamminen, J., Quinlan, P. T., & Gaskell, M. G. (2012). Spoken word processing creates a lexical bottleneck. *Language and Cognitive Processes, 27*(4), 572–593.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*(1), 84–107.

Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychological Bulletin & Review, 12*, 455–459.

De Houwer, J., Smith, C. (in press). Go with your gut! Effects of Affect Misattribution Procedures become stronger when participants are encouraged to rely on their gut feelings. Social Psychology.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science, 11*(6), 467–473.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the irrelevant speech effect. *The Journal of the Acoustical Society of America, 102*(4), 2191–2199.

Elliott, E. M., & Briganti, A. M. (2012). Investigating the role of attentional processes in the irrelevant speech effect. *Acta Psychologica, 140*, 64–74.

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in cognitive sciences, 12*(11), 405–410.

Findlay, J. M., Brown, V., & Gilchrist, I. D. (2001). Saccade target selection in visual search: the effect of information from the previous fixation. *Vision Research, 41*(1), 87–95.

Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge: The MIT Press.

Foulsham, T., & Underwood, G. (2011). If visual saliency predicts search, then why? evidence from normal and gaze-contingent search tasks in natural scenes. *Cognitive Computation, 3*(1), 48–63.

Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. (2012). On the time course of accessing meaning in a second language: An electrophysiological investigation of translation recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*, 1165–1186.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in cognitive sciences, 7*(11), 498–504.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Huang, Y. T., & Snedeker, J. (2011). Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes, 26*(8), 1161–1172.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23–B32.

Huettig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*(8), 985–1018. doi:10.1080/13506280601130875.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*(4), 460–482. doi:10.1016/j.jml.2007.02.001.

Huettig, F., Mishra, R. K., & Olivers, C. N. (2012). On the mechanisms and representations of language-mediated visual attention. *Frontiers in Cognition, 2*, 394.

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011a). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica, 137*, 138–150. doi:10.1016/j.actpsy.2010.07.013.

Huettig, F., Rommers, J., & Meyer, A. S. (2011b). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151–171.

Huettig, F., Singh, N., & Mishra, R. K. (2011c). Language-mediated visual orienting behavior in low and high literates. *Frontiers in Psychology, 2*, 285. doi:10.3389/fpsyg.2011.00285.

Jackendoff, R. (1997). *The architecture of the language faculty* (Vol. 28). Cambridge: The MIT Press.

Jackendoff, R. (2011). What is the human language faculty?: Two views. *Language, 87*(3), 586–624.

Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science, 30*(6), 1053–1079.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: implications for phonological coding in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*(2), 369–381.

Jesperson, O. (1924). The Philosophy of Grammar. University of Chicago Press.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133–156.

Klatte, M., Lachmann, T., Schlittmeier, S., & Hellbrück, J. (2010). The irrelevant sound effect in short-term memory: Is there developmental change? *European Journal of Cognitive Psychology, 22*(8), 1168–1191.

Klein, R. M. (2000). Inhibition of return. *Trends in cognitive sciences, 4*(4), 138–147.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95*(1), 95–127.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science, 30*(3), 481–529.

Kukona, A., Tabor, W. (2011). Multi-word bottom-up effects in the visual world paradigm. Poster presented at the 17th Annual Architectures and Mechanisms for Language Processing Conference, Paris, France.

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences, 9*, 75–82.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25), 3559–3565.

Landau, B., Jackendoff, R. (1993). What and where in spatial language and spatial cognition. *Behavioral and brain sciences*, *16*, 217–217.

Logan, G. D. (1978). Attention in character classification: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107*, 32–63.

Lupyan, G., & Swingley, D. (2012). Self-directed speech affects visual search performance. *The Quarterly Journal of Experimental Psychology, 65*(6), 1068–1085.

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science, 21*, 682–691.

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences, 110*(35), 14196–14201.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847.

Mani, N., Durrant, S., & Floccia, C. (2012). Activation of phonological and semantic codes in toddlers. *Journal of Memory and Language, 66*(4), 612–622.

Marian, V., & Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24*, 173–193.

Marian, V., Spivey, M. (2003b). Competing activation in bilingual language processing: Within and between-language competition. *Bilingualism: Language and Cognition*, *6*, 97–115.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word Knoeferle, P., Crocker, M. W., Scheepers, C., Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. Cognition, 95(1), 95-127.onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 576–585.

Marr, D. (1982). *Vision*. New York: W. H. Freeman and Co.

Mayberry, M. R., Crocker, M. W., & Knoeferle, P. (2009). Learning to attend: A connectionist model of situated language comprehension. *Cognitive science, 33*(3), 449–496.

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology, 60*(5), 661–671. doi:10.1080/17470210601183890.

Meier, B. P., Schnall, S., Schwarz, N., & Bargh, J. A. (2012). Embodiment in social psychology. *Topics in cognitive science, 4*(4), 705–716.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*(2), B25–B33.

Mishra, R. K. (2009). Interface of language and visual attention: Evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing, 11*(4), 295–305.

Mishra, R. K., & Singh, N. (2010). Online fictive motion understanding: An eye-movement study with Hindi. *Metaphor & Symbol, 25*(3), 144–161.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012a). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Mishra, R. K., Olivers, C. N., Huettig, F. (2012). Spoken language and the decision to move the eyes: To what extent are language-mediated eye movements automatic?. In *Progress in Brain Research, Vol. 202, Decision Making: Neural and Behavioural Approaches*.

Mishra, R. K., Huettig, F., & Olivers, C. N. (2013). Automaticity and conscious decisions during language-mediated eye gaze in the visual world. In N. Srinivasan & V. S. C. Pammi (Eds.), *Progress in Brain Research: Decision Making: Neural and Behavioral Approaches*. Amsterdam: Elsevier.

Mishra, R. K., & Singh, N. (2014a). Language non-selective activation of orthography during spoken word processing in Hindi-English sequential bilinguals: an eye tracking visual world study. *Reading and Writing, 27*(1), 129–151. doi:10.1007/s11145-013-9436-5.

Mishra, R. K., & Singh, S. (2014b). Activation of shape and semantic information during ambiguous homophone processing: eye tracking evidence from Hindi. *Cognitive Processing, 15*(4), 451–465.

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience, 6*, 182–189.

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297–326.

Noudoost, B., Chang, M. H., Steinmetz, N. A., & Moore, T. (2010). Top-down control of visual attention. *Current Opinion in Neurobiology, 20*(2), 183–190.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–972.

Posner, M. I. (2012). *Attention in a social world*. New York: Oxford University Press.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism*. *Spatial Vision, 3*(3), 179–197.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioural and brain sciences, 22*(3), 341–365.

Pylyshyn, Z. (2007). *Things and places: How the Mind connects with the World*. Cambridge: MIT Press.

Pylyshyn, Z. (2009). The empirical case for bare demonstratives in vision. Compositionality, Context and Semantic Values. Studies in Linguistics and Philosophy Volume 85, pp 254–274.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191–201.

Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition, 102*(1), 129–138.

Richardson, D. C., Altmann, G. T., Spivey, M. J., Hoover, M. A. (2009). Much ado about eye movements to nothing: a response to Ferreira et al.: Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *13*(6), 235–236.

Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature Neuroscience, 14*(2), 252–256.

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia, 51*(3), 437–447.
Salverda, A. P., & Altmann, G. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1122–1133.
Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica, 137*(2), 172–180.
Saslow, M. G. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *JOSA, 57*(8), 1024–1029.
Shtyrov, Y., Kimppa, L., Pulvermüller, F., Kujala, T.(2010). Event-related potentials reflecting the frequency of unattended spoken words: a neuronal index of connection strength in lexical memory circuits?. Neuroimage, *15*:*55*(2), 658–68.
Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 730–737.
Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial Activation of an Irrelevant Lexicon. *Psychological Science, 10*(3), 281–284.
Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*(4), 235–241.
Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science, 12*(4), 282–286.
Spivey, M. J., & Gonzalez-Marquez, M. (2003). Rescuing generative linguistics: Too little, too late? *Behavioral and Brain Sciences, 26*(06), 690–691.
Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science, 15*(5), 207–211.
Spivey, M. J. (2007). *The continuity of mind*. Oxford, England: Oxford University Press.
Sunderman, G. L., & Priya, K. (2012). Translation recognition in highly proficient Hindi-English bilinguals: The influence of different scripts but connectable phonologies. *Language and Cognitive Processes, 27*(9), 1265–1285.
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.
Telling, A. L., Kumar, S., Meyer, A. S., & Humphreys, G. W. (2010). Electrophysiological evidence of semantic interference in visual search. *Journal of Cognitive Neuroscience, 22*(10), 2212–2225. doi:10.1162/jocn.2009.21348.
Trueswell, J. C., & Tanenhaus, M. K. (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge: The MIT Press.
Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*(2), 146–160.
Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & cognition, 26*(5), 979–1001.
Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken- word recognition. *Journal of Memory and Language, 50*, 1–25.
Wolfe, J. M. (1994). Guided Search 2.0: A Revised Model of Visual Search. *Psychonomic Bulletin & Review, 1*(2), 202–238.
Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*(12), 1411–1426.
Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*(1), 1–14.

Chapter 8
Attention in Reading

8.1 Visual Word Recognition

Reading is in some sense artificial since it is not what the language system has evolved for. Even today we have many societies that are illiterate but who use verbal language for communication. However, reading and writing are finest manifestations of human cognitive evolution and remain cornerstones of modern civilization. Reading is not like listening or speaking in a fundamental sense. Reading is a learnt cognitive activity and depends on one's knowledge of sounds and how they connect to different symbols generating meaning. Reading may differ with regard to scripts and phonological structures in different languages (Katz and Frost 1992). Children learn to read and write, after they have mastered the sound systems of their language. Therefore, reading is contingent on phonological analysis as well as visual analysis of symbols (Goswami 1999). A critical awareness of one's sound systems and their links to visual symbols play a causal role in reading acquisition (Wagner and Torgesen 1987). During reading, a linear sequence of symbols is visually inspected, which results in a phonological representation leading to lexical access. In sum, although reading requires a great deal of linguistic process, it is also fundamentally based on visual and attentional processing resources (Rayner 1998).

Neuroimaging research indicates that the visual word form area specializes in identification of visual words. The invention of reading and its habitual practice has gradually created this specialised brain network (Dehaene 2009). This type of specialization often arises when certain skills are regularly practised and transferred as cultural products. The visual word form area is claimed to be specifically active during letter recognition compared to other tasks (McCandliss et al. 2003). Dehaene and Cohen (2011) suggest that the left lateral occipito-temporal sulcus or more commonly known as the visual word form area shows activation only for letter strings and not for other types of visual objects like faces or houses. This area is also found to be active when readers read different orthographies. The authors also suggest that learning to read modifies the existing visual object recognition

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_8

networks for letter recognition. However, since not many neuroscientists believe today in the theory of one cortical area and one function, there is contradictory evidence about the specialization of the visual word form area (see Pricea and Devlinb 2003). There is evidence that brain areas that are active during reading are also engaged in other cognitive tasks (Vogel et al. 2013). However, it is a fact that, to some extent, continuous exposure to print reorganizes brain cortical networks.

Cognitive models of reading have long assumed that reading a word must first begin with some type of visual analysis of letters (Frith 1980). Since orthographic symbols are first and foremost visual objects, there must be an initial analysis to generate an abstract mental representation which then leads to lexical access. This preliminary analysis leads to access of orthographic information, generating phonology and finally meaning (Besner et al. 1984). However, there has been considerable disagreement over the exact interactions or chronology of events in this process (Rastle 2007). When one looks at a printed word, presumably there is activation of some abstract letter representation (Rapp et al. 2001). These abstract visual entities do not resemble any specific orthography. However, these abstract visual objects must then correspond to particular letters of the language concerned. The orthographic lexicon supposedly contains information about letters. Lexicons are storehouses where all the alphabets of one's orthography are kept. Many psycholinguists assume the presence of both a phonological and an orthographic lexicon (Ferrand and Grainger 1993).

A central question in word recognition is how readers map visual forms to meaning. Words are made of letters which are visual objects that must be connected to their linguistic representations. In his *Consciousness and the Computational Mind*, Jackendoff (1987) considers visual word recognition as a translation from a 3D level representation into a linguistic representation. Jackendoff proposes David Marr's theory of visual object recognition (Marr 1976) and the generation of the 3D percept as a basic step in the visual analysis of objects. Jackendoff writes 'above all, reading involves a recognition process-whether one is recognizing individual letters or larger units'. These considerations suggest that the visual part of reading is a function of the 3D model level, where size and shape invariance can be specified. From the point of view of vision, then letter and word recognition is a learned object discrimination task, not unlike, say 'automobile recognition' and for example discriminating between a B and an R is qualitatively similar to telling a Ford from Buick (p. 210). Jackendoff does not consider that for reading there is any specialized language faculty, but instead merely depends on the interface between the letter recognition system and semantics with phonology linking the two. It is important to explore what role attention plays in this transformation of visual objects into meaningful units of language. Furthermore, for languages like English, the same letter might involve different 'sounding out' (grapheme to phoneme conversion processes); since this orthography has irregular grapheme to phoneme conversion rules (Seidenberg et al. 1984). Other languages like Hindi are transparent in this regard.

Top-down influences on reading have led to the development of interactive models of visual word recognition. As in theories of visual object recognition, it is critical to examine this issue of top-down and bottom-up interaction in order to understand what role attention may play. The influential interactive activation model of word recognition (McClelland and Elman 1986) proposed that identifying a sequence of letters as a word happens through an interactive process. First, there is activation of feature-level information, letter-level information and then orthographic information (Norris 2013). Interestingly, information from letter level and word level flow both ways. When one visually inspects a written word, the abstract features activate different letters and they in turn activate whole words. It is important to note that in many languages letters are visually similar and at least share common design features. For example, in English, the letters "b" and "d" or even M and N share similarities. It is crucial to note that these are perceived as separate entities as there are many words the reader may know that contain these letters. Rumelhart and McClelland (1982) showed that letter identification was improved when they occurred in letter strings that were real words or were non-words that resembled real words. Readers detect letters faster when the letter is embedded inside a word than appears alone or in a nonword (Baron and Thurston 1973). Thus, readers can predict letter identities using contextual information. The interactive activation model uses a top-down mechanism where activation of unrelated words is inhibited. The inhibition works laterally till the correct target is identified. Lateral inhibition means that inhibition works within the word level itself. Letter strings that share features inhibit each other and this helps in the emergence of the most appropriate target. The TRACE model (McClelland and Elman 1986) is an interactive model, which considers both top-down and bottom-up information during perception. TRACE is an interactive model in the sense that sentence context has an immediate effect on speech processing. TRACE accounts for the fact that listeners use their top-down lexical knowledge to understand variations in speech. For example, even if a sound is unclear, or missing in the speech fragment, listeners can make out using their knowledge of the word. One important aspect of these models is the role of inhibition. It is widely accepted that because of spreading activation, related words are activated at each level. For example, the visual system first processes the letter "C" in a sequence. It is possible that other words that begin with this letter are also activated (Norris 2013). This unintentional activation must be inhibited so that only the correct target is identified. As processing proceeds temporally, the search becomes narrower and narrower and finally the target is retrieved. Another influential model of visual word recognition considers reading in terms of print to meaning through several distinct routes (Coltheart et al. 2001). The DRC model (Dual route cascaded model of visual word recognition) of visual word recognition is a computational model, which consists of a lexical and a non-lexical route to meaning comprehension. The model proposes that through the non-lexical pathway one sounds out serially all the letters into their corresponding phonemes, a process known as grapheme to phoneme conversion. The lexical route is active when phonology is directly retrieved from the assembled orthography. The model considers that both phonological and semantic information influence word recognition but they are not both necessary. The DRC model has been shown to be useful in explaining the different sub-types of developmental dyslexia as well as acquired disorders of reading (Coltheart et al. 2001).

Much of this evidence suggests that for skilled readers word identification may not be an attentionally demanding task. Further evidence of top-down influences on reading comes from studies that have shown how both orthographic and phonological factors influence word identification. Readers have tacit knowledge about letter shapes and also orthographic shapes of words (Treisman and Souther 1986). For example, when a prime consisting of letters of the target word is presented, then targets are accessed faster. However, for this effect to occur, the letters must maintain their position (Peressotti and Grainger 1999). For example, the letter sequence “grdn” will facilitate the access of “garden” but not “gdrn” (Grainger and Whitney 2004). This indicates a very fast mapping of top-down orthographic information on visual word recognition. Top-down influences on the initial stage of abstract letter perceptions lead to transposed letter effects. For example, if the written non-word “knfie” is encountered, it is often read as “knife”, even though the letter sequence is incorrect. Priming studies have shown that the letter sequence ‘garden’ facilitates recognition of ‘garden’. When most letters are shared between the primes and the targets, transposing one letter does not affect recognition (Perea and Lupker 2003).

Skilled readers do not necessarily visual process all the letters and sound them out one after the other, but instead use their “top-down” knowledge of words that they already know. The use of top-down information about the visual object, in this case a word or nonword, facilitates accurate identification without complete analysis. Many studies have shown that we can identify objects that are occluded or partially hidden (Bhanu and Ming 1987). Even infants can identify objects that are not fully visible to them (Kellman and Spelke 1983). Similarly, one can identify common written words even if many portions of the word are not totally visible. This shows how stored top-down information leads to rapid identification of visual objects. This leads to the following question: To what extent is focal attention needed during visual word recognition? If readers can access letter knowledge using top-down information, then it is not clear how focal attention is involved in this process. However, there is always an immediate influence of bottom-up information on cognitive processing related to word recognition. Readers also show versatility in recognizing visual words in different formats and dimensions. For example, readers can easily recognize a word if it appears in capital or small letters or even as a mixture of the two. One can tamper with the font size and colour and still word recognition does not suffer to any great extent. This can only happen if the visual system is able to extract some kind of an invariant representation which is then connected to meaning. Dehaene et al. (2004) used a priming paradigm with brain imaging to examine this issue. Each target word was preceded by a related prime word or an anagram made of the same letters. The primes were masked in order to make them totally invisible. In this way, the authors could manipulate whole word representations from letter-based representations. The results showed that in spite of these different primes, participants were able to recognise words. The evidence suggests that mental representation related to orthography is invariant.

The above discussion indicates three things. First, there is some sort of brain specialization for reading compared to other similar perceptual processes. Second, identification of visual words involves interaction of both top-down and bottom-up information. Third, identification of letters involves linguistic processing which produces meaning. The visual word recognition system is therefore an interactive system. It must be noted that, although many of these models consider visual analysis as the first step, none of them explicitly refer to attention. Visual word recognition is considered by as many as an automatic process without the need of focal attention. These models include components of visual processing, top-down knowledge but do not explicitly refer to attention as a mediating agent. Data from the time course of visual word recognition have further revealed the fine distinctions between unconscious and conscious processes that have implications for attention.

8.2 Timeline of Visual Word Recognition

How much time does it take to look at a word and access its meaning? This question is important since different types of information such as orthographic, phonological and semantic area accessed at different time scales (Balota et al. 2004). Furthermore, some type of information may be acquired very early, almost pre-consciously, while some other type of information may call for attentional involvement. The cognitive system reacts to a word's grammatical class, the number of syllables, letter frequency and familiarity, which also affects gaze behaviour. More comprehensive data regarding the timeline of visual word recognition come from studies with eye movements. It appears that for most words, word-specific information is acquired within the first 200 ms of stimulus onset (Sereno and Rayner 2003). Brain imaging data from MEG studies (magnetoencephalography) clearly show the timeline of visual word processing (Fig. 8.1).

In order to understand how attention mechanisms work in visual word recognition, it is crucial to know the time line of visual word recognition. Information

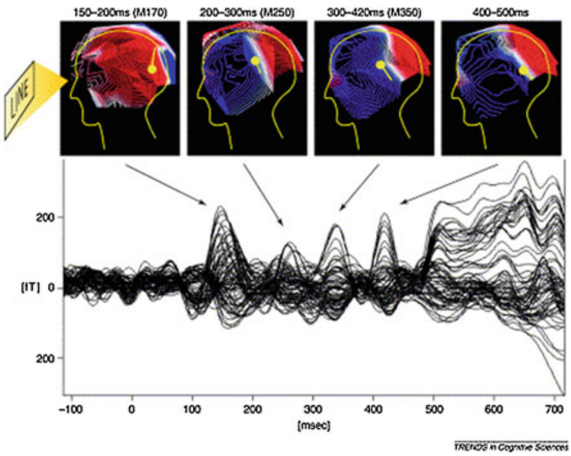

Fig. 8.1 MEG data showing the time course of visual word processing. An activity in the left hemisphere between 150 and 200 ms is indicative of letter sequence processing. Lexical access takes place between 300 and 400 ms. *Source* Pylkkänen and Marantz (2003, p. 188)

from visual words is processed over different time scales till successful recognition (Laszlo and Federmeier 2014). Orthographic information is quickly accessed and is converted into phonological information which then dominates word recognition and semantic access (Zeguers et al. 2014). A substantial amount of visual processing of letters takes place before the word is fully recognised. The initial analysis of the visual objects unconsciously occurs within the first 100 ms after which more conscious processing is experienced. The brain can detect a word from a pseudo-word within 160 ms after word onset (Hauk et al. 2006). Petit et al. (2006) presented masked letter primes that were either similar to target letters in form or in case or in identity. ERPs (event related potentials) revealed an influence of prime-target visual similarity between 120 and 180 ms. The case identity between primes and targets was evident around 180–220 ms. These results suggest that letter identification is a hierarchical mechanism where multiple types of information affect detections at different time scales. The important thing to note is that these processes are operating below the thresholds of conscious awareness which nevertheless have significant effect on recognition.

Figure 8.2 shows the amount of time it takes for the system to recognize visual words. The visual word first has to be identified as an object. It takes about 100 ms for initial visual processing to take place. Normally, attention shifts to the next word as soon as the meaning of the current word has been attained. Eye movements

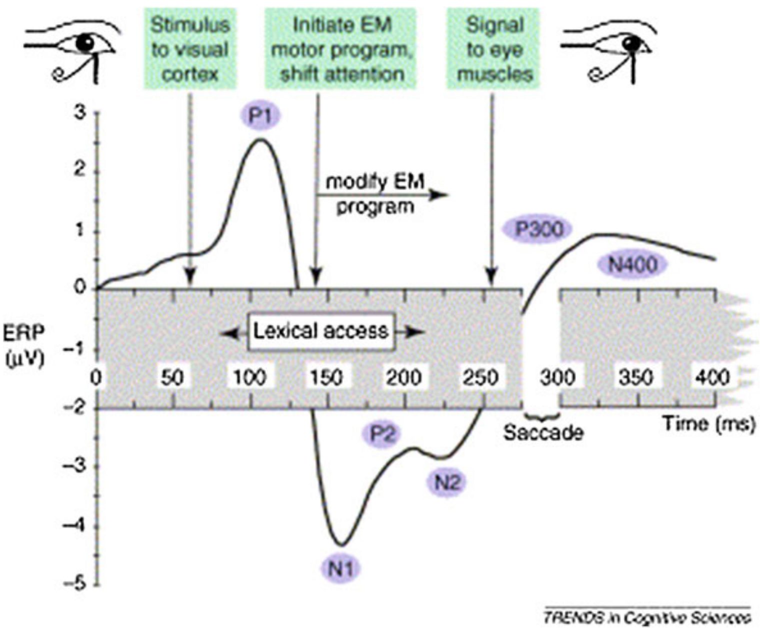

Fig. 8.2 The figure shows the time taken for lexical access during single word recognition. The plot shows both ERP and eye movement data. It shows that it takes around 200 ms or so for lexical access. Attention shifts happen before eyes move to the next word. This timeline shows specific modulations of ERP components. *Source* Sereno and Rayner (2003, p. 490) with permission

during reading text seem to be highly synchronised to attentional shifts (Rayner 1998). Since much of early visual processing happens below the threshold of conscious awareness, it is important to note that the output of this processing is used by the linguistic system to generate meaning. Word identification therefore is an outcome of very rapid visual processing and attention shifts punctuated by linguistic processing. The P300 ERP component has been considered as a signal of attention processing (Holcomb 1988). This is a positive-going ERP effect that peaks after 300 ms of critical stimuli onset. Critical stimuli could be a visual word in a sentence. This means that at this point in time one is able to report the identity of the word/visual object in question. In a comprehensive study of several psycholinguistic variables and their effects on the duration of written word recognition, Hauk et al. (2004) observed that within the first 150 ms, the brain is aware of a word compared to a nonword. Within the first 100 ms, an area in the left inferior temporal cortex is able to make decisions about the surface structures of a word. A preliminary analysis of the number of syllables and their types and other visual aspects occurs. Similarly, within the first 150 ms, readers are able to access the semantics of the word. Schendedan et al. (2003) found that ERPs to written words and faces diverge within the first 90 ms, suggesting a very fast categorization of written words compared to other visual objects. Does this mean that most important processes related to visual word recognition are automatic?

Other researchers have used the priming paradigm to examine how the unconscious perception of a prime word can influence different aspects of the target word. Holcomb and Grainger (2006) asked participants to monitor for words that were animal names in a masked priming design. Target words were full repetitions, partial repetitions or unrelated to the primes. The prime target relationship modulated the N400 effect as well as early effects such as P150 and N250. The N400 effect often indicates difficulty in contextual integration in any kind (Van Berkum et al. 2003). Similarly, effects such as P150 and N250 indicate early a perceptual processing of a stimuli. ERP effects thus can indicate a temporal, continuum of perceptual and cognitive processing of any stimuli. These effects indicate a very rapid access of both orthographic and semantic factors in visual word processing. Ferrand and Grainger (2003) observed the influence of pseudohomophones as well as non-word primes on target word recognition suggesting that an early influence of phonological and orthographic information on word recognition evidenced from MEG and from ERP studies shows that roughly the reading centres of the brain are aware of the meaning of the word within the first 250 ms of the stimulus onset (Bentin 1989). It seems believable that much early visual processing is unconscious in nature and within this time frame visual and phonological information about the written word is already computed. For highly frequent words, this time window is even shorter and similarly for words whose meanings we know.

Early visual processes are unconscious and non-conceptual. One can, however, be aware of the presence of a stimulus with attention (Lamme 2003). Awareness about these computational processes is incomplete and people can only report little after much of these processes are over (Pammer et al. 2004). The P300 ERP effect is commonly considered to be related to the initial stage of conscious attentional processing. This component peaks after 300 ms of critical stimuli onset. After 300 ms of the presentation of a visual stimulus, the subject is in a position to detect changes and report the experience. If one divides visual perception into preconscious and conscious stages (as is presented in the framework by Lamme 2003), it is easy to see how this occurs for written words. The first 250 or 300 ms of processing gives the reader a very vague idea of the word's identity in question. The content of this processing is not yet complete and is short lived. For full-blown identification, several brain areas must be coordinated which this takes time.

Lamme (2003, 2006) offers an extensive account of visual processing with regard to preconscious and conscious stages of processing. According to his model based on neurophysiological data recorded from primates during visual object processing, there are two main processes or events. The first stage is called "feedforward sweep", in which visual information reaches the primary visual cortex. The second stage is called local recurrent processing. The third and final stage are called "global" processing. In this stage, higher order cognitive processes are at work and several brain areas show activity. Lamme has linked his model to the discussion of visual perception and the role of attention. Lamme observes that all visual processes are attentional in some sense. Attention selects and facilitates the processing of some object over others. Attention basically does the work of selection of an item for further processing. Visual processing prior to 300 ms could lead to what is known as some type of "phenomenal" awareness. In this stage, the subject is vaguely aware of the stimulus but cannot report it. After 300 ms, attentional processes begin and the subject reports knowledge of the object. For visual word recognition, one can report the lexical nature of the word after this time window. See Raftopoulos (2011) for a more elaborate discussion on these issues in vision and attention from a philosophical perspective.

The data discussed so far suggest that visual word recognition evolves over time and different types of information such as visual, phonological and semantic affect recognition. Furthermore, some aspects of this processing such as early access of visual information are unconscious while later knowledge of semantics operates within attention constraints. While these studies have dealt with the time course of visual word recognition and have revealed how different types of information interact, they do not directly speak about attention requirement in visual word recognition. The discussion that follows looks critically at studies that have directly examined how attention affects visual word recognition.

8.3 Written Word Recognition and Visual Attention

With fluent reading, the issue of attention is not linked with effort. Rather, the interesting question has been how linguistic information affects oculomotor programming. Recent brain imaging evidence suggests that the visual word form area facilitates the interaction of top-down contextual knowledge with bottom-up features (Song et al. 2012). As previously stated, the visual word recognition process is very robust in the face of bottom-up variations, for example with different types of fonts or colours. Even, non-words that look like real words can be read using this top-down contextual information. How immediate is the influence of top-down information on visual word recognition? It appears that within the first 300 ms, information from the left inferior frontal gyrus activates the visual areas. This activity is higher for words than non-words or words written with false fonts (Woodhead et al. 2014). It should be noted that the left frontal areas activate semantic representations for linguistic stimuli. In this case, it appears that semantic representation from the symbol is quite fast and this in turn constrains visual processing.

There is evidence indicating that attention 'modulates' the experience of reading. Cognitive psychologists use the word 'modulate' in a variety of senses to refer to selective enhancement of some effect or processing. It refers to some kind of causal or correlational relationship between two variables. However, visual processing and selection of the sequence of letters as a single 'visual object' with some kind of conceptual representation must occur. This is crucial since we do not perceive individual letters but the whole word in a holistic sense (Perea and Rosa 2002). The letters are perceived as part of a word, and as a unified visual object. How does such a holistic perception arise when we look at the word in snapshots? Eye tracking studies have shown that for many words, readers fixate only once and then move on to the next word. The word 'fixate' here refers to looking at something with foveal focus (Rayner and Bertera 1979). Words that are fixated receive maximal processing. Some evidence suggests that readers first fixate briefly between the gap between the second and the third letter during reading (Rayner 1979). Fixations on words during reading vary as a function of orthography type as well as length of the word (Rayner 1986). Generally, with a single fixation, readers are able to grasp maximal information about the word which can perceive the whole word as a single undivided object (Balota et al. 2004).

Attention does also help 'glue' perceptions that may be distinct (see Treisman 1996 for an account of the binding problem in object recognition). Attention seems to spread across an object from one location to another (Moore et al. 1998). This movement enhances the perception for a whole object. These theories have given prominence to an object-based view of attention (see Pylyshyn 2007). Pylyshyn considers attention being object-based and rejects a location-based account. According to him, attention is glued to objects and travels with objects. Much of this evidence comes from his studies of multiple object tracking. Studies with inhibition of return with mobile objects also show that attention is tied to object representations along with spatial representations (Tipper et al. 1994). When one fixates on one letter of a word, attention spreads without additional eye movements and the word is recognized. Attention can neither glue things nor spread if things are perceived as different objects (Shomstein and Yantis 2002).

Is visual word recognition an automatic process? A mental process is automatic if it does not require voluntary processing and attention. Based on results from the Stroop task, it is widely believed that visual word processing is automatic. In this task, a certain colour word is written in another colour, for example, the word

'BLUE' is written in a red colour. Reading such a word takes time compared to when the name of the colour and the font colour match. The colour-word Stroop has been replicated millions of times and has been observed in young children. Interference occurs when there is a mismatch between the name of the colour and its font colour. Interestingly, highly fluent bilinguals show better abilities to resist this interference (Singh and Mishra 2012). Some researchers have demonstrated that the Stroop task actually requires attention (Besner and Stolz 1999). Besner et al. (1997) changed the colour of a single letter rather than the whole word and made it incompatible with the name of the colour. The Stroop effect was significantly reduced. This shows that automaticity seen in the Stroop task emerges only in certain circumstances. That is because even to identify the word and take a decision on its colour, selective attention is necessary. In a very early study, LaBerge and Samuels (1974) suggested that most aspects of fluent word decoding could be automatic and may not require attention. In their opinion when one wants to read accurately, attention is called for. Otherwise, word decoding is automatic.

The importance of attention during word recognition can be studied using a dual-task paradigm. Visual word recognition is affected when a secondary task is presented. One example is making a lexical judgement about a written sequence of symbols while at the same time doing a working memory task. Becker (1976) observed that performance on visual word recognition was relatively low under a working memory task. Furthermore, when participants did a secondary task, their ability to recognise low frequency words was especially poor. This means, making lexical decision does demand central attention to some extent, but it also depends on the type of word.

Several researchers have invoked the 'capacity limitation' attribute of attention in examining if central attention is required for visual word identification. If word identification is automatic, then attentional manipulation should not affect performance. If word identification calls for the engagement of focal attention, then different types of words, i.e. high-frequency words, low-frequency words and non-words should be differently affected by attentional manipulation. One standard experimental paradigm in this investigation has been the use of Posner's spatial cuing paradigm (Posner 1980). In this paradigm, if a target appears in the location of a briefly presented exogenous cue, then its processing is facilitated as compared to when the target appears elsewhere. Montani et al. (2014) presented high- or low-frequency words along with pseudo-words at cued or uncued locations in the parafovea. There was facilitation in the recognition of pseudo-words when they appeared in the cued location, thus indicating that central attention facilitated processing. However, no such facilitation was found for the high-frequency words suggesting that these words did not require much of attention. Auclair and Sieroff (2002) manipulated lengths of words and pseudo-words, i.e. letters ranging from 6, 8 or 10 letters and used a spatial cueing paradigm to examine the effect of attention on visual word recognition. Spatial attention affected recognition of pseudo-words regardless of length but it had some effect on the very long real words. This suggests that spatial attentional allocation directly facilitates the processing of pseudo-words, whereas it has any effect on real words only when the words are either very long demanding attention or complex. The redistribution theory of spatial attention during word recognition explains why words and pseudo-words require different levels of attention. Since words already provide top-down information regarding the meaning, attention spreads the entire length of the real word from the cued position, whereas attention operates sequentially for a pseudo-word and does not spread across the letters easily. This means that lexicality of a letter string can affect attentional allocation dynamically. However, the claim that spatial attention is differentially engaged for familiar and unfamiliar words has been challenged by others using the same cueing paradigm. McCann et al. (1992) observed facilitation in the recognition of letter strings at cued positions. However, when letter strings were sometimes presented with irrelevant stimuli on either side of the fixation, recognition times were longer. This indicates attentional resources similarly engaged for both familiar and unfamiliar words. Regardless of this controversy, it is clear that spatial attentional mechanism interacts with processes of word recognition.

One can utilise Stroop interference signifying that word recognition is automatic (Brown et al. 2002). Evidence from the Stroop task suggests that word recognition is automatic (Neely and Kahan 2001). Stolz and Besner (1999) observed that there is no automatic semantic access in a Stroop task when participants were asked to focus attention on the letter strings. In contrast to this, when Heil et al. (2004) looked at modulation of the N400 ERP component in a priming paradigm (a component indicating semantic a contextual influence on processing), semantics of primes affected the component. This happened even when participants were given a letter search paradigm. Therefore, it cannot be conclusively said that familiar words call for less attention.

If attention modulates word recognition, then one can claim that it interacts with the word recognition process. As words are primarily visual objects, cueing can boost their detection in space. When visual targets appear at the locations that have been cued, their processing is facilitated compared to those appearing at uncued locations. Similarly, when written words appear at cued locations, their identification may be facilitated (Sieroff and Posner 1988). Stolz and McCann (2000) first presented a prime word and then presented a brief cue to attract spatial attention. A target word appeared next which the participants had to judge as a word or non-word. The prime and target words were either semantically related or unrelated. Spatial attention in this case interacted with the prime target relationship affecting word recognition times. The authors argued that in this particular case, spatial attention worked as ‘glue’ in facilitating intake of orthographic information. This is another important property of attention apart from being a filtering agency. Posner and Petersen (1989) observed that recognizing a visual word may not require attention. However, attention is required if someone wants to find out a particular letter or count the number of letters.

Readers generate orthographic information from a letter string shortly after presentation. Therefore, once this basic information has been acquired, it can still be used to decide if a string is a word or not, a classic lexical decision task. Kellas et al. (1988) examined if the temporal distance between an auditory probe detection task affects word recognition. Auditory probes are basically tones of different frequencies. The participant is asked to identify them and give a response, which may require selective attention. The authors presented an auditory probe detection task after 90, 180 or 270 ms of the lexical decision task. Interestingly, the authors also manipulated the type of word which was either an ambiguous or unambiguous word. It was observed that participants were faster in the lexical decision task as well as the probe detection task for the ambiguous compared to the unambiguous words. It seems that finding out if a word was ambiguous required less attention. Does this mean that finding a visual word which is not a word or less like a word is easy? This could be so, since such words do not have meanings and there is no semantics to be activated. Both lexical decision and naming could demand attention under a dual task situation. Herdman (1992) observed that efficiency on auditory probe detection was low for both naming aloud and lexical decision on visual words, when the words were of low frequency. Participants have to pay more attention to make a decision towards a primary task which causes problems for a secondary demanding task. Many studies of this sort with visual and auditory stimuli have found interference suggesting attention involvement.

8.4 Attention Shifts During Sentence Reading

Most models have been developed around naming or reading a single word. However, normally one reads sentences and paragraphs for comprehension. There have been developments that have revealed how we read sentences. Most evidence concerning this has come from eye tracking studies during reading a sentence. The eyes movements from one word to the other are affected by specific attention and linguistic mechanisms. Eye movements during reading reveal the interaction between linguistic information processing, attention shifts and comprehension (Rayner 1998). This process is repeated continuously throughout the reading process. Thus, shifts in attention are functionally linked to lexical access in reading. Research has led to the development of various reading models that have emphasised the important role of attention (Reichle et al. 2003).

Several fixations in certain intervals are needed to process new information, which is a limitation of our visual system. For reading, since the words are rather closely aligned, one shifts the eyes only a short distance. The fovea is a mere two degree of the visual field, which is the region we get maximum information from. Beyond this region lie the para fovea and the periphery. It is natural that information acquisition from the parafovea is limited and almost negligible from the periphery. Rayner and Bertera (1979) obstructed foveal vision creating a foveal mask which moved in synchrony with the eyes. This led to serious difficulties in word recognition. Rayner et al. (1981) masked both foveal and parafoveal vision and found that word recognition requires information acquisition from both sources. Readers can acquire relevant information from few letters to the left and right of the fixation and this is called the perceptual span (Rayner 1975). Perceptual span could arise as a result of script direction. For example, when reading Hebrew, the perceptual span is extended towards the left, while reading English is it more towards the right (Pollatsek et al. 1981). The size of the perceptual span increases with reading experience from 11 letters to 14 letters to the right of fixation (Rayner 1986). Studies with different scripts and how they influence perceptual span show how reading a script affects attentional span broadly. Readers have an awareness of certain letters to the left and right of the current fixation which is known as the perceptual span.[1] The perceptual span is shorter to the left of fixation than to the right in left to right orthographies. If this is so, then there is good chance that readers have some 'preview' benefit from words that are immediately around the current fixation. Also, there is a good chance that some of the processing from the current fixation is spilled over to the next word, since attention moves before the eyes move. Readers often know without looking at the next word its linguistic identity. Sometimes, this can happen based on top-down linguistic contextual knowledge. Readers know if the currently fixated word is a noun and the first element of the sentence, the next word has to be some sort of an auxiliary or helping verb. This sort of knowledge can guide the eyes during reading. This can also explain why so many words are skipped during reading. It must be remembered that although during saccades no new information is acquired, cognitive processing continues.

Reading a sentence involves interplay between complex linguistic and visual analyses. Since reading a sentence involves making eye movements from one word to another, attention shift is mandatory (Rayner 1986). Much controversy in the field is around the nature of such attention shift and how they are linked to meaning generation. This has led to the development of cognitive models of reading, which assume both parallel and serial in processing. Particularly, the question if attention moves serially or deployed in parallel has been the centre of discussion (Rayner 2009). If attention shift operates serially, one word at a time, then it is not possible to acquire meaning from several words simultaneously. Attending to one word and accessing meaning, and moving to the next, attention shift has to be serial. On the other hand, the assumption of parallel processing can also be made, since normally eyes do not scan every word.

Without an attention shift towards a location, there will not be a corresponding saccade (Kowler et al. 1995). There is also evidence which indicates that different mechanisms are behind attention shift and saccade generation (Remington 1980). With regard to reading, since it is already shown that readers acquire information from letters and words that fall inside their perceptual spans, it is important to know if this involves saccadic eye movements. If readers can acquire information from multiple words without making saccadic eye movements, then this may indicate a parallel distribution of attention over several zones. If it is shown that attention shift to the next word is contingent on the linguistic information, then this would indicate that attention shift in reading is controlled by upcoming linguistic information.

[1]It is the region of effective vision where visual processing is at its best.

While reading a sentence, readers normally spend around 250–300 ms fixating on a word (Rayner 1986). Interestingly, eye movements are both progressive and regressive in reading. When the sentence is difficult to parse, readers may regress or return back to previous text. Saccade lengths during reading are from 1 to 15 letters. This depends on what type of text is being read and who is reading. This influences attention allocation during reading. Furthermore, the type of orthography, i.e. written from left to right or right to left, affects eye movement control during reading. People often skip words that are predictable from preceding context, or are high frequency words such as articles or prepositions as well as words that are familiar (Inhoff and Rayner 1986). Similarly, readers have longer fixation durations with longer words and low-frequency words. This means, attentional shift during sentence reading is sensitive to various linguistic as well as cognitive factors. Readers sometimes look at particular words repeatedly and this could be because of comprehension difficulties.

The above discussion points towards the fact that eye movement control during reading a sentence is dependent on the linguistic material being read as well as attentional factors. The precise functional connections between attention shift, saccades and reading comprehension are currently subject of great controversies. Below the proposals of two prominent reading models have been discussed that make contradictory claims regarding eye movement control.

8.5 Serial Versus Parallel Allocation of Attention

Modelling a cognitive process involves mimicking that process utilizing artificial conditions. As reading is a highly personal and idiosyncratic process, modelling has not been an easy endeavour. However, the use of modern eye tracking measurements has made it possible to study certain subtle aspects of attention–language interactions during reading. One reads a sentence or two on a computer screen and eye movements are recorded. It has been possible to present only selective material to the reader by controlling the amount of information available visually. It has been possible to control the display of words through a technique called the moving window paradigm, so as to accurately measure saccades from one word to another during reading (Just et al. 1982). It is also possible to change certain words or make them disappear when certain saccades cross an invisible boundary (Rayner et al. 1986). Furthermore, it has also been possible to block central vision using an artificial scotoma to examine how this may affect comprehension (Fine and Rubin 1998). These technical developments have allowed examining the issue of serial vs. parallel movement of attention during reading.

When can one say that attention shift during reading is serial? Most researchers have been curious to know if attention shift during reading is a serial process. If this assumption is correct, then one would observe saccades, i.e. at least the start of their programming after substantial processing of the word currently being fixated is over (Inhoff et al. 2000). Most readers quickly make a saccade to the word next to the foveal word as soon as they have acquired some minimum information from the current word (Rayner et al. 1983). There have been many demonstrations which suggest the simultaneous acquisition of information from at least two words to the left of the fixated word. This parafoveal processing has been found to affect fixation duration on the foveal word. In any case, such studies have taken the attention–language debate to an altogether new height. The ultimate aim of a general oculomotor model of reading comprehension has to account for the many different orthographies and reading habits of individuals that are encountered.

A crucial test for the serial view is that readers should not show the influence of the parafoveal information on fovea reading. Kennedy and Pynte (2005) examined reading corpus and found that when the currently fixated word is short, the lexical frequency of a parafoveal word affects fixation duration of the foveated word. When the foveated word is long, the parafoveal word frequency does not have much effect. Angele et al. (2008) manipulated the preview of word n + 1 and n + 2 to examine their effect on reading times on word n. There was no preview benefit of word n + 2 when the n + 1 word was a high-frequency word. This evidence points to a strictly serial model of attention allocation which is triggered by lexical access. Any effect of parafoveal information on foveal processing also reflects the distribution of attentional resources between the words. If foveated word is easy to process, then attention quickly moves towards the parafovea. When the parafoveal word is easily predictable, it is skipped (Balota et al. 1985). The E-Z model predicts some preview benefit from word n + 1 but not from word n + 2 in normal situations. This is consistent with the model's serial processing assumptions (Rayner et al. 2007). E-Z model assumes serial attention shift as reading progresses. Attention is shifted to the next word only after the processing of the currently attended word is completed (Reichle 2003).

It is crucial to know how attention shift occurs during reading a sentence under normal conditions. The E-Z model of eye movement control assumes a serial shift of attention from one word to another (Reichle et al. 1998). Readers will only shift attention to the next word when they are done with the processing of the currently fixated word and not before. The reader fixates and comprehends a word, and then moves to another word to the right. The model makes sense if one assumes a spotlight model of attention where attention can illuminate only one object at a time. The E-Z model of eye movement control is perhaps the most explicit model which links attention with language processing. One of the basic assumptions of the model is about serial shift of attention and word processing (Reichle 2003). The model assumes that letters within a word are processed in parallel. This means, attentional focus is on the whole word at any given point in time and not across multiple words. Crucially, the model also assumes that attention shifts from word to word are related to different linguistic processing stages. The E-Z model assumes that while someone is fixated on word n, his/her attention could be deployed on word n + 1. This decoupling of attention and eye movements allows this model to incorporate 'spill over effects' and also to some extent the effects of the parafoveal word on fovea processing. Many vision scientists believe that point of regard need not be the focus of attention. Looking at a point does not necessarily mean that attention per se is there. However, this is counterintuitive.

The E-Z model proposes two states of word identification. First, the visual system does a familiarity check on a word which provides the system with gross information. The second stage includes a very specific and attention processing stage (Rayner 2007a, b). According to the model, attention shifts to the next word when the second processing stage is over, provided that any saccade has not been already triggered. Therefore, the decision to shift the eyes to another word depends on the demands on the familiarity check and also to an extent on the complexity of the next word. Pre-attentive visual processing leads to the access of semantic information. Both covert and overt attention shift are controlled by linguistic processing. Attention in this model shifts from one word to another before lexical processing at the currently fixated word is completed. Readers also seem to extract information from a neighbouring word in parallel. The preview effects also show that covert attention can move to the neighbouring word prior to the subsequent saccade (Reichle 2003).

It is important to note that the E-Z model assumes separation of attentional shifts and eye movements. It assumes covert shift of attention occurring serially as one begins to read one word at a time and integrates meaning into an overall semantic representation (Reichle 2003). The model does not mean 'spatial' orientation of attention when it refers to 'attention' but the process of integrating upcoming linguistic information into the existing structure. Thus, even though some low level visual information is acquired while the eyes are on word n, saccades are only generated when word n is fully processed. This, this model does not assume that attentional orienting will automatically move eyes to the next word as it is normally assumed in the eye movement literature (Hoffman and Subramaniam 1995). Therefore, eye movements are linked to linguistic processing in a serial fashion. Lexical processing in the E-Z model drives eye movements' not attentional shifts.

In contrast to this, the SWIFT (Autonomous Saccade Generation with Inhibition by Foveal Targets) model considers attention as a gradient where one can process several words simultaneously. The reader is currently fixated on a word but is also processing information from words to the left and right. The sharpness of attentional focus dims from the fovea outwards. The SWIFT model (Engbert et al. 2005) explicitly assumes that attention operates like a gradient over many words. The model derives its assumptions from the zoom lens model of attention (LaBerge and Brown 1989). The zoom lens model assumes that attentional focus is sharpest at the fovea and gradually falls towards the periphery. Thus, words that are currently fixated are in the focus of attention. The SWIFT model bases its assumptions regarding oculomotor control in reading on the foveal load theory (Henderson and Ferreira 1990). The foveal load theory proposes that perceptual span during reading is dependent on the foveal load. That is if the currently fixated word is difficult to process (low in frequency or long) then perceptual span decreases. Consistent with this claim, it has been observed that readers grasp only very little information from a parafoveal word if the foveal word is difficult to process (Schroyens et al. 1999). Attentional scope during is modulated by foveal load (All these words receive different amounts of processing. Those words that are under foveal inspection receive the highest amount of processing. Visual acuity drops off as one move from the fovea towards the periphery. Interestingly, the SWIFT model proposes an independent saccade generator (Engbert et al. 2005). In the SWIFT model saccades are generated autonomously, that is after a random variable time interval. However, saccades are delayed if the word that is currently being processed is difficult to integrate. Unlike the E-Z model, in SWIFT, eye movements are not dependent on lexical processing as such. Saccades are regularly programmed but at modulated by difficulty level in lexical processing. In this sense it is not a serial mechanism since it does not prohibit eye movements till lexical processing at a word is fully complete. A critical prediction of the SWIFT model is that the characteristics of word n + 2 will affect processing on word n. That is because the model assumes parallel processing of words that are in the attention gradient. Kliegl et al. (2007) found parafoveal on foveal effect of word n + 1 and word n + 2 on word n suggesting extended perceptual span that may affect word processing. Parafoveal effect of this type has also been observed in Chinese language (Yang et al. 2009).

In sum, the E-Z model explicitly links eye movements to linguistic processing while the SWIFT model assumes an independent saccade generator. Much of the debate so far has not been on 'attention' per se in reading but on how certain linguistic features guide eye movements. Shifts in attention are not only guided by bottom-up factors but also by top-down factors involving prediction and contextual knowledge. Currently available data show changes in fixation behaviour with regard to the psycholinguistic properties of neighbouring words. Overall, comprehension develops as a function of attention shift. Attention does play a causal role in the reading process by linking perceptual processes with linguistic conceptualization. The shifts of attention during reading are controlled by a range of linguistic and extra-linguistic factors including predictive processes. It is important to note here that both models assume different properties of attention and link them differently to language processing. While the E-Z model assumes a spotlight model of attention with a serial processing mechanism, the SWIFT model assumes a zoom lens model that operates like a gradient. Importantly, both models relate attention shift and oculomotor programming differently to linguistic processing.

8.6 Attention in Dyslexia

A widely studied condition is dyslexia or specific reading disorder. Much discussion on this topic has been on the deficits in 'phonological' processing (Harm and Seidenberg 1999). That is, if someone lacks the knowledge of subtleties of sounds then reading is disrupted. This linguistic view of reading emphasises the establishment of a relationship between sounds and symbols. Whereas some researchers view dyslexia as primarily being a language-based disorder, others have taken a more cognitive view linking it with attention, memory and executive functioning deficits. The diagnosis and manifestation of dyslexia in different writing systems remains controversial at best. Reading involves visual and attentional mechanisms, and a dynamic interaction with language processing. Dyslexia is a good testing ground for understanding language–vision interaction. Information processing in reading involves cooperation of both attention and language, mediated through eye movements. Goswami (2002) has proposed that reading disability primarily arises because of a deficit in representing speech sounds. Further, these deficits may manifest differently in different orthographies.

Is there an attentional basis to reading disorder? Reading is a highly interactive visual process, even though linguistic comprehension is its aim. Several researchers have started to examine dyslexia from a language–attention perspective. What exactly breaks down in dyslexia? Is it the linguistic machinery or the visual-attentional system or their networking? (Cornelissen and Singleton 2005). There are several sub-types of dyslexia which are seen in both children and adults, such as developmental dyslexia, deep dyslexia and other, acquired dyslexia. Developmental dyslexia is again divided into phonological and surface types (Frith 1985). Acquired dyslexia results because of brain damage and mostly seen in adults who have strokes. In deep dyslexia readers substitute semantically similar words while reading and are quite rare (Coltheart et al. 1987). Most dyslexics show an inability to decode written words. This inability can arise because of both linguistic as well as visual factors. There has been the suggestion that developmental reading disorders could result from an inability to process several visual symbols in parallel (Bosse et al. 2007). From the SWIFT model, we know that reading occurs when attention operates like a gradient and several words fall under this attentional scope receive simultaneous processing. If one fails to receive information from several letters or words in parallel then reading could be slow. This is not to contest that the phonological factors are not responsible for reading. However, if there is limited attention span or an inability to process information in parallel, then linguistic processing could suffer. Attention shift during reading is linked with lexical access.

Dyslexics typically take longer to process written words than non-dyslexics. Studies with MEG have shown that while normal subject's brain activation shows word specific neural responses within the 200 ms, for dyslexics this is delayed (Salmelin et al. 1996). Dyslexics take time to figure out that a collection of letters represent a meaningful word. This explanation is not purely visual but also attentional. Others have found visual processing deficits in dyslexics indicating a deficit during the earliest time window of visual object processing (Iles et al. 2000). However, it is not easy to figure out if dyslexics' deficits are purely verbal or basically visual resulting in verbal deficit. There is evidence that suggests that those who are better at maintaining a visual attentional span, are also better at reading or on other non-linguistic visual tasks (Lobier et al. 2012). If children have low visuo-spatial attention then reading suffers (Gabrieli and Norton 2012).

Visuo-spatial attention plays a crucial role in selecting the meaningful visual targets for further processing. This also helps in binding the perception which helps detecting the target as a meaningful object? Dyslexics do show deficits in awareness of speech sounds and problems in phonological awareness tasks. These are the tasks that have been classically used to test metalinguistic knowledge. Many have shown that performance on these tasks predict reading ability (Goswami 2000). Apart from this non-word reading tests are also used and rapid picture naming. However, if there is a problem with visual processing and object identification, then linguistic processing may also suffer. Children who have ADHD (attention deficit hyperactivity disorder) often also experience reading deficits (Gilger et al. 1992). However, it is not to suggest that all children who show a specific reading disorder also show symptoms of ADHD or other executive dysfunctions (Pennington et al. 1993). Dyslexic children have been shown to be slow in spatial attention tasks. For example, in a Posner cueing task, dyslexics do not show a quick orientation of attention towards the cued location. This results in a non-facilitation of the targets that appear in that location (Facoetti et al. 2010). Recently, there has been support for the view that dyslexia can be seen as a disorder of visuo-spatial attention (Vidyasagar and Pammer 2010).

The relationship between dyslexia and attention is at best correlational and not causal. Since most researchers who have studied attentional deficits in dyslexia have tried to correlate problems in attentional tasks with a linguistic task like non-word reading. Early measures of attention and executive functioning can predict if someone is going to have a reading or learning disorder. Dyslexics do show some deficits in their visuo-spatial attention. Some dyslexics do not have symmetry between their right and left visual fields. This has led to the argument that dyslexics may have an uneven spread of attentional resources in their visual field (Facoetti and Molteni 2001). During reading, we acquire parafoveal information from words both to the left and to the right of the currently fixated word. If there is an abnormal distribution of attentional resources, then acquisition of information from such words may be impaired. Dyslexics do not seem to take advantage from spatial cues? In shifting their attention and target identification (Roach and Hogben 2004). They also seem not to be able to filter out visual information that is not task and goal related (Roach and Hogben 2007). This may indicate that an attentional deficit has led to less goal-directed action. Second, it may lead to visual target identification and information acquisition. Finally, it may lead to failure of executive functioning in many ways.

Eye movements of dyslexics typically exhibit repeated fixations and higher saccades when presented with longer words (De Luca et al. 2002). They fixate on a word repeatedly since they may be converting grapheme to sounds sequentially. Previously, Pavlidis (1981) observed that dyslexics show poor eye movement control in non-linguistic sequential tasks. Dyslexic individuals show problematic eye movements on simple non-linguistic tasks where either they show repeated fixations, higher amplitude or poor fixation maintenance (Eden et al. 1994). It has been suggested that their oculomotor problems are independent of their language disorder. However, oculomotor problems of this nature could lead to poor performance on visual word recognition and comprehension. However, Olson et al. (1983) did not find any difference between dyslexic and control subjects on a task which required making saccade towards a visual target. Therefore, they argued that an oculomotor deficit could not be at the roots of dyslexia. Visual and attentional issues in dyslexia are now beginning to be explored more. Bucci et al. (2012) investigated eye movement behaviour in dyslexics and normal readers in a simple reading and visual search task. In the reading task, the participants were asked to read a short story. In the visual search task, they were asked to find the letter ‘r’ and its occurrence in the text. Eye movement patterns from dyslexic subjects show that dyslexics do exhibit some erratic movement of saccades. The authors found that dyslexics have longer fixation durations and also higher number of saccades, indicating greater task difficulty.

As already discussed, reading does require some sort of allocation of attention. However, it is still not clear if one needs selective attention for reading comprehension. It is also not clear how fluency affects attentional allocation during reading. There have been some other demonstrations that indicate a role of attention in visual word recognition. Written words have letters that are very close to one another in a visual space. Observations from a phenomenon called 'visual crowding' demonstrate how one can be aware of an object without attending to it. Visual crowding occurs when many different objects are in close proximity to one another. It is difficult to individuate each object when there is crowding. Chakravarthi and Cavanagh (2009) have suggested that we can be aware of a group of crowded visual symbols or objects without recognizing or differentiating between individual items. It is like seeing the forest without the trees. Crowding can impair identification of the individual items but not the detection of the texture (Whitney and Levi 2011). There are claims that one does not need attention to recognise the visual object as a whole? We do not necessarily attend to the trees but to the forest. This is to say, we do not need to attend to each letter in the word. This can be considered to be a constraint of crowding in the visual system.

Crowding can arise when an object of interest is flanked by other objects. Consider the following sequence of letters: XXHHX and HHHHH. Imagine, further, that you are seeing it from a distance. Many visual psychophysical tasks ask participants to identity an object to the right of the fixation while they keep fixating on a cross to the left. The above two examples of letter crowding can differently affect the identification of the middle H. In this example, participants may not be able to report the identity of the central H but can report it if this letter is replaced by something else. This can occur without attention (Block 2013). It has been suggested that word length creates a situation of crowding for reading (Bouma 1973). It has been observed that developmental dyslexic children read more slowly when there is crowding (Martelli et al. 2009). In one study, dyslexics found it very difficult to read aloud words when they appeared embedded among other words (Spinelli et al. 2002). This extra sensitivity of dyslexics to the flanking words leads to slow and error-prone word identification. If this is correct, then it is likely that dyslexics must find it visually overwhelming to look at words when reading a sentence. Crowding could offer a purely visual explanation of reading. This state of affairs can be compared to distinctions between local and global states of attention. Earlier it was shown that one can voluntarily attend to a visual object locally or globally. The local state requires focused attention while in the global state; attention operates in a more dispersed fashion. It is possible that dyslexia affects these two states of attention allocation strategy differently. Reading failure due to crowding could be related to an inability to deploy global attention. Actually, most reading studies with eye tracking have not considered these two aspects of attentional operation during word recognition. It remains to be seen if a pure attention failure can account for reading dysfunction in different sub-types of dyslexics.

8.7 Conclusion

The evidence and theories discussed in this chapter indicate that although several processes associated with word recognition appear to be unconscious driven by top-down influences, the role of visual and attentional processing is paramount. One's capacity to acquire information and process from text is linked to attentional span. However, in reading, attention and language are linked bi-directionally. That is, linguistic properties like orthography and phonology affect how attention will be deployed during reading. Attentional capacities of individuals influence their reading as is evident from several cognitive disorders. Eye movement data from reading show how properties of words shift attention further. It is apparent from this discussion that reading behaviour involves a complex and dynamic interaction between attention and linguistic processing. Contemporary oculomotor models of reading have shown how attentional shifts are triggered by distinct stages of linguistic processing. Data from dyslexia also illustrates how an attentional and visual account could inform us better about the aetiology of this complex disorder. Reading studies have demonstrated an attention–language interaction, albeit one that can be argued is relatively artificial to the language system. There are some problems which future research efforts can illuminate. Most reading researchers do not pin point the type of attention that is involved during lexical access in reading. Attention is not a unitary entity and different readers using different aspects of it during reading. Second, reading is a highly individual process. These individual differences need to be considered in model building. Thirdly, reading is a highly cultural and orthography-specific process.

References

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n + 1 and n + 2 previews simultaneously. *Visual Cognition, 16*(6), 697–707.

Auclair, L., & Sieroff, E. (2002). Attentional cueing effect in the identification of words and pseudowords of different length. *Quarterly Journal of Experimental Psychology, 55A*, 445–463.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133* (2), 283–316.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*(3), 364–390.

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology, 4*(2), 207–228.

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*(4), 556–566.

Bentin, S. (1989). Electrophysiological studies of visual word perception, lexical organization, and semantic processing: A tutorial review. *Language and Speech, 32*(3), 205–220.

Besner, D., & Stolz, J. A. (1999). Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin & Review, 6*(3), 449–455.

Besner, D., Coltheart, M., & Davelaar, E. (1984). Basic processes in reading: Computation of abstract letter identities. *Canadian Journal of Psychology/Revue canadienne de psychologie, 38*(1), 126–134.

Besner, D., Stolz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review, 4*(2), 221–225.

Bhanu, B., & Ming, J. C. (1987). Recognition of occluded objects: A cluster-structure algorithm. *Pattern Recognition, 20*(2), 199–211.

Block, N. (2013). Seeing and Windows of Integration. *Thought: A Journal of Philosophy, 2*(1), 29–39.

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104*(2), 198–230.

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research, 13*(4), 767–782.

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic?". *Journal of Experimental Psychology: General, 131*(2), 220–240.

Bucci, M. P., Nassibi, N., Gerard, C. L., Bui-Quoc, E., & Seassau, M. (2012). Immaturity of the oculomotor saccade and vergence interaction in dyslexic children: Evidence from a reading and visual search study. *PLoS One*, *7*(3), DOI: 10.1371/journal.pone.0033458.

Chakravarthi, R., & Cavanagh, P. (2009). Bilateral field advantage in visual crowding. *Vision Research, 49*(13), 1638–1646.

Coltheart, M., Patterson, K., & Marshall, J. C. (1987). *Deep dyslexia since 1980*. New York: Routledge.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256.

Cornelissen, P., & Singleton, C. (Eds.). (2005). *Visual Factors in Reading*. England: Blackwell.

De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language, 80*(3), 617–626.

Dehaene, S. (2009). Origins of mathematical intuitions. *Annals of the New York Academy of Sciences, 1156*(1), 232–259.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences, 15*(6), 254–262.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words behavioral and neuroimaging evidence. *Psychological Science, 15*(5), 307–313.

Eden, G. F., Stein, J. F., Wood, H. M., & Wood, F. B. (1994). Differences in eye movements and reading problems in dyslexic and normal children. *Vision Research, 34*(10), 1345–1358.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777–813.

Facoetti, A., & Molteni, M. (2001). The gradient of visual attention in developmental dyslexia. *Neuropsychologia, 39*(4), 352–357.

Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., & Zorzi, M. (2010). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *Journal of Cognitive Neuroscience*, *22*(5), 1011–1025.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the psychonomic Society, 31*(2), 119–122.

Ferrand, L., & Grainger, J. (2003). Homophone interference effects in visual word recognition. *Quarterly Journal of Experimental Psychology Series A Human Experimental Psychology, 56*(3), 403–420.

Fine, E. M., & Rubin, G. S. (1998). Reading with central field loss: number of letters masked is more important than the size of the mask in degrees. *Vision Research, 39*(4), 747–756.

Frith, U. (1985). Beneath the surface of developmental dyslexia. *Surface dyslexia*, *32*.

Frith, U. (Ed.). (1980). *Cognitive processes in spelling*. London: Academic Press.

Gabrieli, J. D., & Norton, E. S. (2012). Reading abilities: importance of visual-spatial attention. *Current Biology, 22*(9), R298–R299.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research, 16*(2), 123–144.

Gilger, J. W., Pennington, B. F., & DeFRIES, J. O. H. N. (1992). A twin study of the etiology of comorbidity: attention-deficit hyperactivity disorder and dyslexia. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*(2), 343–348.

Goswami, U. (1999). The relationship between phonological awareness and orthographic representation in different orthographies. *Learning to read and write: A cross-linguistic perspective*, 134–156.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6*(2), 133–151.

Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia, 52*(1), 139–163.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences, 8*(2), 58–59.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review, 106*(3), 491–528.

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*(4), 1383–1400.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 301–307.

Heil, M., Rolke, B., & Pecchinenda, A. (2004). Automatic Semantic Activation Is No Myth Semantic Context Effects on the N400 in the Letter-Search Task in the Absence of Response Time Effects. *Psychological Science, 15*(12), 852–857.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*(3), 417–429.

Herdman, C. M. (1992). Attentional resource demands of visual word recognition in naming and lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 460–470.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57*(6), 787–795.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*(1), 66–85.

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience, 18*(10), 1631–1643.

Iles, J., Walsh, V., & Richardson, A. (2000). Visual search performance in dyslexia. *Dyslexia, 6* (3), 163–177.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40*(6), 431–439.

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception and Psychophysics, 62*(7), 1474–1484.

Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge: The MIT Press.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*(2), 228–238.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology-Amsterdam, 94*, 67.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology, 15*(4), 483–524.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*(2), 153–168.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the time course of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 601.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n + 2. *Journal of Experimental Psychology: Human Perception and Performance, 33*(5), 1250–1255.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96*(1), 101–124.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*(1), 12–18.

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10*(11), 494–501.

Laszlo, S., & Federmeier, K. D. (2014). Never seem to find the time: evaluating the physiological time course of visual word recognition with regression analysis of single-item event-related potentials. *Language, Cognition and Neuroscience*, (ahead-of-print), 1–20.

Lobier, M., Zoubrinetzky, R., & Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. *Cortex, 48*(6), 768–773.

Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *275*(942), 483–519.

Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision, 9*(4), 14.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*(7), 293–299.

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1015–1029.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1–86.

Montani, V., Facoetti, A., & Zorzi, M. (2014). The effect of decreased interletter spacing on orthographic processing. *Psychonomic Bulletin & Review, 22*(3), 824–832.

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9*(2), 104–110.

Neely, J. H., & Kahan, T. A. (2001). *Is semantic activation automatic? A critical re-evaluation*. In Roediger, H.L., III, Nairne, J.S., Neath, I., Surprenant, A.M. (Eds.) The nature of remembering: Essays in honor of Robert G. Crowder. Science conference series., (pp. 69–93). Washington, DC, US: American Psychological Association, xix, 396 pp.

Norris, D. (2013). Models of visual word recognition. *Trends in Cognitive Sciences, 17*(10), 517–524.

Olson, R. K., Kliegl, R., & Davidson, B. J. (1983). Dyslexic and normal readers' eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 9*(5), 816–825.

Pammer, K., Hansen, P. C., Kringelbach, M. L., Holliday, I., Barnes, G., Hillebrand, A., & Cornelissen, P. L. (2004). Visual word recognition: the first half second. *Neuroimage, 22*(4), 1819–1825.

Pavlidis, G. T. (1981). Do eye movements hold the key to dyslexia? *Neuropsychologia, 19*(1), 57–64.

Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology, 29*(3), 511–523.

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition, 31*(6), 829–841.

Perea, M., & Rosa, E. (2002). Does 'whole-word shape' play a role in visual word recognition? *Perception and Psychophysics, 64*(5), 785–794.

Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception and Psychophysics, 61*(4), 691–706.

Petit, J. P., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2006). On the time course of letter perception: A masked priming ERP investigation. *Psychonomic Bulletin & Review*, *13*(4), 674–681.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*(1), 174–180.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Posner, M. I., & Petersen, S. E. (1989). *The attention system of the human brain* (No. TR-89-1). Washington Univ St Louis Mo Dept of Neurology.

Pricea, C. J., & Devlinb, J. T. (2003). The myth of the visual word form area. *Neuroimage, 19*, 473–481.

Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences, 7*(5), 187–189.

Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world*. Cambridge: MIT press.

Raftopoulos, A. (2011). Late vision: processes and epistemic status. *Frontiers in Psychology, 2*, 382.

Rapp, B., Folk, J. R., & Tainturier, M. J. (2001). Word reading. *The handbook of cognitive neuropsychology: What deficits reveal about the human mind*, 233–262.

Rastle, K. (2007). Visual word recognition. *The Oxford handbook of psycholinguistics*, 71–87.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7* (1), 65–81.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*(1), 21–30.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*(2), 211–236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*(4417), 468–469.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 40*(4), 473–483.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7*(1), 167–179.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007a). Do readers obtain preview benefit from word n + 2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 230–245.

Rayner, K., Li, X., & Pollatsek, A. (2007b). Extending the E-Z reader model of eye movement control to Chinese readers. *Cognitive Science, 31*(6), 1021–1033.

Rayner, K., Slowiaczek, M. L., Clifton, C., & Bertera, J. H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance, 9*(6), 912–922.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445–476.

Remington, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 6*(4), 726–744.

Roach, N. W., & Hogben, J. H. (2004). Attentional modulation of visual processing in adult dyslexia: A spatial-cuing deficit. *Psychological Science, 15*(10), 650–654.

Roach, N. W., & Hogben, J. H. (2007). Impaired filtering of behaviourally irrelevant visual information in dyslexia. *Brain, 130*(3), 771–785.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*(1), 60–94.

Salmelin, R., Kiesilä, P., Uutela, K., Service, E., Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology, 40*(2), 157–162.

Schroyens W., Vitu F., Brysbaert M., d'Ydewalle G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 52A(4):1021–1046.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*(3), 383–404.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*(11), 489–493.

Shomstein, S., & Yantis, S. (2002). Object-based attention: Sensory modulation or priority setting? *Perception and Psychophysics, 64*(1), 41–51.

Sieroff, E., & Posner, M. I. (1988). Cueing spatial attention during processing of words and letter strings in normals. *Cognitive Neuropsychology, 5*(4), 451–472.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(04), 771–781.

Song, Y., Tian, M., & Liu, J. (2012). Top-down processing of symbolic meanings modulates the visual word form area. *The Journal of Neuroscience, 32*(35), 12277–12283.

Spinelli, D., De Luca, M., Judica, A., & Zoccolotti, P. (2002). Crowding effects on word identification in developmental dyslexia. *Cortex, 38*(2), 179–200.

Stolz, J. A., & Besner, D. (1999). On the myth of automatic semantic activation in reading. *Current Directions in Psychological Science, 8*(2), 61–65.

Stolz, J. A., & McCann, R. S. (2000). Visual word recognition: Reattending to the role of spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*(4), 1320–1331.

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 20*(3), 478–499.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178.

Treisman, A., & Souther, J. (1986). Illusory words: The roles of attention and of top–down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception and Performance, 12*(1), 3–17.

Van Berkum, J. J., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17*(3), 701–718.

Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences, 14*(2), 57–63.

Vogel, S. E., Grabner, R. H., Schneider, M., Siegler, R. S., & Ansari, D. (2013). Overlapping and distinct brain regions involved in estimating the spatial position of numerical and non-numerical magnitudes: An fMRI study. *Neuropsychologia, 51*(5), 979–989.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192–212.

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168.

Woodhead, Z. V. J., Barnes, G. R., Penny, W., Moran, R., Teki, S., et al. (2014). Reading front to back: MEG evidence for early feedback effects during word recognition. *Cerebral Cortex, 24* (3), 817–825.

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 1192–1204.

Zeguers, M. H. T., Snellings, P., Huizenga, H. M., & van der Molen, M. W. (2014). Time course analyses of orthographic and phonological priming effects during word recognition in a transparent orthography. *The Quarterly Journal of Experimental Psychology, 67*(10), 1925–1943.

Chapter 9
Cultural and Sensorimotor Events During Language–Attention Interaction

9.1 Speaking and Linking Things

Language is a symbolic structure that captures spatio-temporal relationships in many situations. Language is used to bind the visual perception of distinct objects in space and map them into a structural form. Spatial cognition and linguistic structures have been shown to interact (Herskovits 1987; Bloom 1999). Referential terms like prepositions and deictic terms are used to refer to spatial structures to map space. Referring to an object which is away from us requires us to map the space and use a particular term available in the language. Now, the question is, do all speakers who use similar terms also quantify space in similar fashion? Or do speakers of different languages who use different terms map space differently? Much evidence for these questions has come from cross-linguistic studies that show a clear effect of linguistic and cultural habits on the perception of space. Attention is automatically involved when language refer to objects or events in space.

How does language refer to objects that are located in space? While the attentional mechanism may select the object in space, language if often used to describe its position. Therefore, language and attention interact more dynamically during mapping of the spatial world. Different spatial words are used to refer to lotion of objects and also spatial interrelationship of objects by speakers. Many have studied (Levinson 2006; Li and Gleitman 2002) these from a cultural perspective to see if this is universal or shows cultural specificity. Interestingly, language has its own mechanisms to represent space through different structural constituents. A sentence like "The cat is on the table near the cake" indicates the spatial and relational position of different objects in space. One can talk about the cat's exact position using natural language with respect to other objects. If different speakers use different linguistic constructs to refer to this relationship differently, then this may entail that these speakers are also allocating attention differently. Attention plays an important role in this process, since without paying attention it is not possible to generate spatial relations (Logan 1994). This includes important components of

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_9

attention like selection, orientation and engagement when speakers select objects for referencing and description. Specific spatial term in one's language constrains, selection and orientation of attention (Talmy 1983).

A central question in this context has been which 'perspective' the speaker takes while perceiving and referring to the spatial relationship. Perspective refers to the vantage point from which the speaker is viewing the objects in the space. Since this perspective influences which particular spatial word the speaker uses to describe the location or the spatial relationship. Specific terms like 'under' and 'above' reflect different points in space which the speaker has perceived and ha preferred to use in a given context. As the location of the speaker changes in pace with regard to the object or the object changes its location, different spatial terms are used. Therefore, speaker's mapping of space and use of particular spatial structures is contingent upon the spatial relationship between the speaker and the object itself. If this is true then this will also influence how speakers are channelizing attention towards the object in this process. For example, in the sentence 'the cup is on the table' an object is at a particular location with regard to another object. How the viewer takes a perspective and cognizes the visual scene depends on available grammatical resources as well as observable spatial relations.

Do speakers use different attentional strategies while describing spatial relationship between objects in different cultures? Cross-linguistic data shows that speakers of different languages are constrained by their linguistic environments as far as mapping space is concerned (Levinson 2006). Speakers of specific languages employ certain fixed terms to refer to situations of similar spatial contour (Majid et al. 2004). Therefore, variations in the use of linguistic terms affect how speakers deploy attentional strategies interact. English speakers use terms like 'besides' or 'adjacent to' while describing spatial location while these may not be commonly found in other languages. Cognitive linguistic studies show that these terms code very specific spatial relations for English speakers (Gentner et al. 2013). As soon as the dimensions change, speakers employ a different word to refer to the new situations. It is like watching a kite flying and going up in the sky. At different levels of height, speakers of different languages will use different terms. However, it is very difficult to say, if two different terms used by two different speakers have cut the space with similar dimensions. For example, the use of 'on' and 'above' differs at a particular threshold for speakers of one language. If this is the case, then it is possible to argue that different words that refer to spatial relations actually affect attention mechanisms and visual processing. If speakers of particular languages cut space using particular words, then it may be implied that absence of certain linguistic terms should manifest in not being able to grasp spatial relations (Munnich et al. 2001).

Language users often link things and events with particular perspectives. These perspectives are often consistent across speakers of a language. Frames of references are used to describe the spatial relation between objects. For example, imagine a cup and a pen placed horizontally on the table. If asked to describe where the cup is, one may say the cup is to the left of the table. In this case, the speaker is expressing the position of the cup from his own vantage point. He describes the spatial relations from an egocentric perspective. In this case, the body of the viewer in space works as the reference. However, from the perspective of the pen, the cup could be to its right. The first manner of saying has been called the relative frame of reference while the second is called intrinsic. Research has shown that speakers of languages with different typological and grammatical structures use different frames of references for establishing spatial relations. For instance, if one is speaker of a language called Guugu Yimithirr (an Australian aboriginal language), then the description to our above example will be ‘The cup is to the south of the pen’ (Majid et al. 2004). This has been called the absolute frame of reference where speakers use exact geometric frames of references to describe spatial position. Does this difference in use of spatial terms indicate a difference in attentional allocation?

Majid et al. (2004) explored if Dutch and Tzeltal (language spoken in Mexico) speakers refer to change in spatial location of things in space differently. Dutch uses the relative kind of frame of reference with the use of words such as ‘to the left of’ or ‘to the right of’, etc. These words reflect the assessment of the spatial layout with regard to the positioning of the viewer egocentrically. Whereas, Tzeltal speakers use the absolute type of reference where descriptions include ‘to the north of’ or ‘to the south of’. These descriptions reflect the viewers’ quantification of the spatial layout independent of him and more in terms of absolute geometry of space. In the experiment (Majid et al. 2004) the participants were shown a card which had a larger and a smaller dot. Then the table on which the card was rotated 180°. The participants were then asked to identify the card among four alternatives. It was of interest if these two language speakers will code in their memory the change depending on the frame of reference they use. It was assumed that since the Dutch speakers used relative frame of reference, they would rotate the spatial layout egocentrically but the Tzeltal speakers will rotate geocentrically. Speakers were asked to describe the spatial layout using terms in their language. The responses showed that Dutch speakers used the relative references most of the time. Whereas for the Tzeltal, it was the absolute frame. That is, Dutch speakers rotated the spatial layout with regard to their own position where as the Tzeltal speakers did not use egocentric reference. These demonstrations indicate that different people cognize spatial relationships differently if they speak different languages (Fig. 9.1).

Explaining their results, Majid and colleagues write that, ‘Experience, including experience with language, can influence the perceptual system such that it is more or less attuned to particular features in the environment’ (Majid et al. 2004, p. 2004). This means the type of language one uses; its grammatical and typological structures can influence cognition in non-linguistic domains. Cultural and anthropological features of the environment can induce specific types of spatial cognition. Elsewhere the authors say, ‘rather than cognitive categories being universal and giving rise to universal semantic categories, as is typically supposed, it seems that cognitive categories are variable and they align with cross-linguistically variable semantic categories’ (Majid et al. 2004, p. 113). Thus, language’s semantics, the words and the constructions influences perception (see also Levinson 1996 for an extensive treatment). It has further been shown that frames of references used by a speaker of a language are same for both linguistic and non-linguistic domains. For

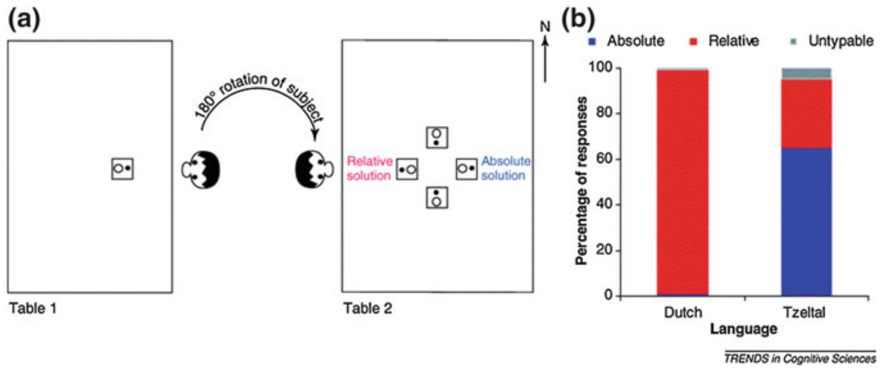

Fig. 9.1 In this task, Dutch and Tzeltal speakers were given a card that had a large and a small dot. This card was arranged in different spatial directions. After a delay the table on which the card was rotated and the speakers were asked to identify the card they had seen from four alternatives (Figure from Majid et al. 2004, p. 110)

example, if the language uses an absolute frame of reference, then one will also divide space in this manner for a non-linguistic spatial work (Levinson 1996). Therefore, the frames of references are cross-modal and are generalized across action systems (Coello and Bonnotte 2013). Speakers are not constrained by the frames of references and show flexibility in use depending on the conversational situations.

Do linguistic structures constrain perception of spatial relationship? Bowerman (1989) describes an interesting example of how Korean speakers map space when describing spatial relationship between two objects. Korean speakers use a particular verb when describing an event where an object is in loose contact with another, i.e. the paper was put into the basket. Whereas they use another verb to describe a situation where two objects are in close contact, i.e. the pencil was put into the pencil sharpener (Choi and Bowerman 1991). This indicates that Korean speakers have two different words for two different actions and they fundamentally perceive such actions differently. Does this mean English or Hindi speakers may not even perceive such perceptual differences since these languages do not have distinct words to refer to such events? Choi and Bowerman (1991) claim that certain languages allow their speakers to map spatial relationship of some types and not others. It further appears that if one does not have access to spatial language, then there could be problem in mapping space in other similar non-linguistic spatial tasks. Gentner et al. (2013) studied deaf children from Turkey who did not know any sign language. They compared their performance on some spatial tasks with deaf children who knew sign languages and normal children with verbal language. Many studies have shown that sign language has similar complexity as verbal language and syntax. Learning a sign language leads to development of specific neural networks in the brain (Bellugi 2014). The results showed that these children, who had learnt to use sign language at home, did not use any gestures even to express spatial relations. Spatial cognition in children suffers when they have an attentional disorder (Brown et al. 2003). This evidence suggests that in the absence of any kind of linguistic system, these children were unable to map space and to describe it. Spatial language and its expressions offer the tool to create co-ordinates in space and establish relationship among objects.

Other experimental investigations suggest that the manner in which viewers perceive spatial locations affect how they refer to them in language (Hayward and Tarr 1995).

Others have argued based on experimental data that refutes the claim that specific languages encode space differently (Li and Gleitman 2002). Li and Gleitman (2002) explored how native English speakers use spatial terms on a single arrangement of objects. It was observed that English speakers behaved like the Tzeltal speaking subjects when the blinds of the experimental room are raised, offering a view of the outside world. When the blinds are down, they used references like Dutch speakers. Tzeltal and Dutch have been shown to use different frames of references in the description of spatial relations. Based on these findings, the authors have argued that use of a particular spatial term is related to how the environment affects perception. Further, they suggest that one should not ascribe an excessive role of language in this behaviour.

The above examples at times indicate that our visual perceptions and conceptualizations are influenced by linguistic analysis. The symbolic structure of language thus has intricate interface with the perceptual-cognitive system. Perceptions are fed into the linguistic system and depending on the context and several other factors, speakers sue suitable linguistic terms. Work in the domain of language production has shown the influence of language variation on perception of events and actions. Language has been suggested to compress the amount of information to be verbalized and 'schematize' them in a very abstract manner (Talmy 2000). Different language speakers may use therefore different forms of schematizations. However, it is possible that the influence of language on visual perception is dependent on the task at hand. That is, speakers inspect a visual scene differently if they are asked to speak about it compared to when they need not. Papafragou et al. (2008) examined eye movements and allocation of visual attention in Greek and English speakers as they perceived scenes depicting motion events. Greek speakers are known to use more 'path' verbs while English speakers more 'manner' verbs when describing motion events. In the linguistic task, participants were asked to prepare a description of the event while for the non-linguistic condition they were asked to freely inspect the scene. Eye movements showed that Greek and English speakers were different in their visual inspection pattern when they prepared to speak a sentence on the scene. Greek speakers focused more on the path trajectory compared to the English speakers consistent with their linguistic use. However, eye movements on the scene were similar for both the groups when the task did not involve sentence planning. These data show that linguistic constrains can influence perception of visual events only when the task involved some linguistic action. Otherwise, linguistic differences do not seem to influence looking behaviour for tasks that do not call for linguistic action. Therefore, the influence of language structure on perception is constrained by task. Other evidence suggests that speakers do not necessarily show limitations of their languages while perceiving and describing events. Papafragou et al. (2006) tested Greek and English speakers to see if perceptual difficulty led Greek speakers to use more of manner verbs during event perception. Greek speakers used more of manner verbs when the scene was opaque and did not clearly indicate manner. English speakers on the other hand habitually produced manner descriptions whether the scene explicitly indicated manner or not. This shows that language users do perceive different aspects of events and this is not constrained by their languages. Depending on the demand of the task, language users insert suitable descriptions.

Recently, it has been shown that speakers of different languages allocate attention differently during perception of motion. Flecken et al. (2015) examined if German and English speakers deployed attention differently during perception of motion. The authors used event-related potentials (ERPs) to explore if brain waves were different for these speakers when they perceived motion. English language is known to emphasise attention to both trajectory and end points equally while German only focuses on end points. However, the authors explored of these linguistic differences would influence actual attention deployment during a visual task. Participants saw a prime animation where an object moved towards a geometric figure with a specific trajectory followed by test picture. In a majority of trials (75 %), the test picture did match the prime animation either in trajectory or end points. In 10 % of the trials the test and prime pictures had similar trajectory and in another 10 % they had similar end points. Only in 5 % of the trials, both trajectory and end points matched. The authors used an oddball paradigm where ERPs to deviant stimuli (unexpected or mismatching) are recorded when they appear amid other stimuli. ERPs showed that the German speakers had a larger P3 wave when the prime and the test pictures matched in end points than in trajectory. However, for the English speakers no such difference in P3 amplitude was observed. P3 ERP effect has been linked to attentional deployment in the literature (Polich 2007). These results suggest that speakers of different languages attend to motion events differently (Fig. 9.2).

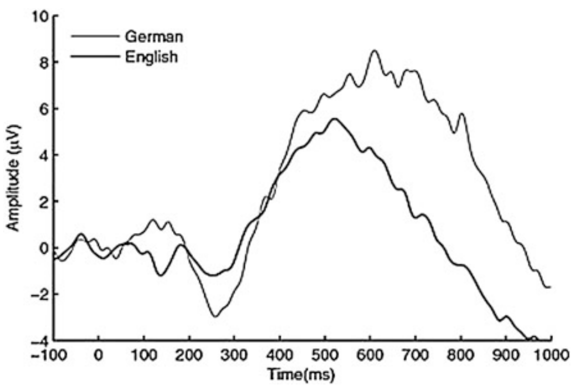

Fig. 9.2 ERP P3 wave form difference between German and English speaker (Figure from Flecken et al. 2015, p. 46)

These data indicate that certain intrinsic properties of language both at the level of word and grammar influence how speakers look at things and describe the spatial relationship among them. While language is an arbitrary collection of signs and symbols, these signs and symbols refer to the space in a very consistent manner across speakers of a language. These data also show the influence language structure has on visual cognition and attention allocation. Attention orients towards objects with regard to the way language users refer them and employ terms. Anthropological and cultural studies compliment laboratory based studies showing that the interface of language and attention can be very broad and one has to consider several other variables to fully understand this.

9.2 Sensorimotor Experience, Spatial Language and Visual Attention

Do language users use experiential knowledge in comprehending spatial language? While the previous discussion suggests a possible role of language in perception of events and attention allocation, they speak little on the issue of how language users experience these events in the first place. In other words, if language is used to code and express our varied experiences with objects and events, then these experiences should influence language comprehension itself. Theories of Embodiment have a long history in cognitive linguistics (Lakoff and Johnson 1999). Amid various definitions, these theories emphasize on the primacy of experiential knowledge in linguistic and non-linguistic cognition (Zwann 2014). Several authors have shown that listeners mentally simulate what they hear through language which plays a significant role in language comprehension (Clark 2006). For instance, listeners mentally simulate or construct mental imagery of events as they comprehend to language. If that is so, then it is natural to expect that listeners use experiential knowledge to comprehend spatial language. Language users have been shown to mentally simulate events as objects undergo change of state (Altmann and Kamide 2007).

Sensorimotor information about objects and contextual information can influence spatial language comprehension. For example, the knowledge about the canonical location of objects in space influences the allocation of attention. For instance, we expect an umbrella to be open during rain and not closed. Similarly, if someone says a sentence such as ‘the glass fell from the table…’ we expect glass pieces scattered on the floor. This knowledge results from our sensorimotor experience with objects. Much of this is achieved by mental simulation of experiential situations. Comprehension of spatial language then can be influenced by this sort of sensorimotor knowledge about objects and their location in space.

The sentence ‘the bottle is *on the* glass’ (Coventry and Garrod 2004) does not make practical sense since it does not refer to the canonical position of the bottle. That is because listeners know what kind of a surface works better for bottles. Listeners arrive at this conclusion immediately as they can mentally simulate the description of the sentence (Zwann et al. 2006). Experience suggests that bottles stand vertically on some kinds of surfaces only. Of course it is also possible that a bottle can be found lying down on glass. However, one goes by the most typical situation with specific objects with regard to their positions and also actions performed with them. Sensorimotor and embodied approaches to cognition postulate how we know about objects and mentally simulate descriptions about them during language comprehension (Zwann et al. 2004). If there is mental simulation involved in spatial language processing, then this must correspondingly affect how listeners deploy attention.

Language refers to objects that are found in both the non-canonical and canonical positions. Speakers can visually perceive non-canonical arrangements, but may face difficulty in generating appropriate linguistic descriptions. The terms one knows to refer to the canonical situations, affordances of objects may constrain the allocation of spatial attention.

Coventry et al. (2010) examined how language comprehenders consider the affordances and functionality of objects when listening to spatial language. The assumption was that depending on the type of spatial expression, perceivers will deploy specific type of attention to events depicted in a visual scene. Figure 9.3 shows a man carrying an umbrella in the face of rain to protect himself. From experience, we know that one has to keep the umbrella at a particular inclination and hold in a particular manner depending on the direction of the falling rain for good protection. Participants were presented with sentences like 'The umbrella is over the man, the man is under the umbrella, the umbrella is above the man, and the man is below the umbrella'. In one part of the experiment, participants were asked to judge the acceptability of the sentences with regard to the pictures. A close inspection of the figure would show that in some sections that umbrella is made

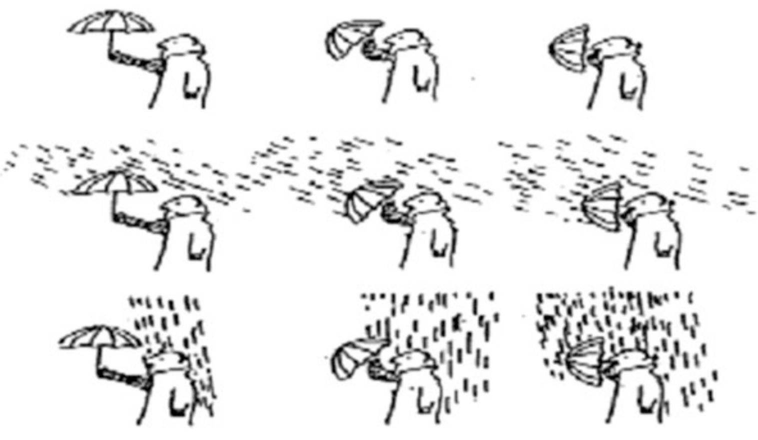

Fig. 9.3 Figure shows orientation of the umbrella with regard to the rain. The aim was to test how language users react to spatial words like 'above', 'below', 'over' and 'under' for these different situations. It was observed that geometrical orientation had an effect on how people comprehended these terms (Figure from Coventry et al. 2010, p. 204)

non-functional for protection. In some other cases, its direction is such that it will not offer protection from the rain. The acceptability test showed that participants used their experiential knowledge to judge the suitability of the descriptions. In the second part of the experiment, the authors showed various permutations and combinations of images that showed corn flakes falling from box and landing on a container. Coventry et al. (2010) used eye tracking to examine if participants were considering the knowledge of objects and changes that they are undergoing in focusing attention. The results showed that participants made eye movements to places that are likely to indicate the location where the corn flakes would fall. Perceivers used the sensorimotor system to anticipate the changes and attention then moves in the direction of this changed state. Earlier, Altmann and Kamide (2009) had shown that language comprehenders move their eyes with regard to how things change state dynamically. Thus, language users can anticipate the location of attention deployment using their experiential knowledge of the situation. Attentional shifts are regulated by what language refers to and what one anticipates in the environment as a function of sensorimotor knowledge. The understanding of spatial language with regard to visual scenes involves mental simulation of the actions. Participants not only show awareness of the current states of objects but also the states they will be after some change.

Experiential knowledge about objects and affordable actions lead to mental simulation of a scenario. Given both a visual scene and a linguistic description, listeners match both representations to examine their compatibility. Listeners explore if what they are hearing is afforded by the visual scene. Visual attention in this situation helps create a bridge between linguistic and visual representations (Mishra and Marmalojo-Ramos 2010). However, every word in a sentence may not be there in a picture. Language sometimes significantly selects and filters out things that are otherwise visually perceived. For example, a word like ‘across’ does not represent a point in space (Talmy 2000). But this word refers to a whole plane. On hearing such a word listener’s scan through the whole surface. Listeners do their best to extract information from language and see if what they see if reasonably represented by this. As for simple sentences containing spatial relationship, viewers use their background knowledge to see if what they hear is depicted, or even possible. For example, no one says a sentence like, ‘The bottle is across the table’. That is because an object like a bottle cannot afford the spatial word ‘across’ whereas it is okay for an object like ‘rope’. In this sense, sensorimotor and experiential information helps in the correct interpretation of spatial language. One may wonder if there are innumerable ways to refer to such objects and situations in different languages. Words like ‘throughout’, ‘across’ and ‘everywhere’ indicate different attentional states. These words can refer to both abstract and real objects in space (across the field versus across lifespan).

Spatial expressions in languages are usually grammaticized as prepositions, like ‘on’, ‘under’, ‘into’, etc. Jackendoff (2012b) makes an excellent point regarding the variable nature of such words in different contexts as they refer to spatial locations. Jackendoff says that the preposition ‘into’ represents different spatial relations in proportions, such as, run into the room, run into the wall. It is obvious that listeners do not deploy attention similarly for these two contexts. It is an excellent example to demonstrate that language captures some global attributes of visual perception at one go, in a minimalistic manner. With regard to the preposition ‘in’, he writes:

> The preposition *in* picks out a region within the boundary of its reference object. The reference object can have a two-dimensional interior, as in *the dots in the circle*, or it can have a three dimensional interior, as in *the beetle in the box*. The use of in also extends to some cases in which the figural object is not entirely surrounded by the reference object, e.g. the knife in the cheese, where only the end of the blade needs to penetrate the interior (Jackendoff 2012b, p. 1144, italics in original).

These examples show that the same word can force one to look at objects differently in different situations. Does sensorimotor experience play any role in this? It is quite likely, since our visual system gathers different types of perceptions from different visual contexts. Even if the same preposition is used to refer to different spatial relations, it orients the aperture of attention differently. Many times, the affordances of the concerned object, and in what way it is visually and experientially related to another, decide our looking patterns. There is probably no one-to-one match between spatial terms and the space they refer to in different situations.

9.3 Words, Mental Simulation and Attention Shift

Automatic activation of semantic and perceptual features from words can influence visual attention. Listeners activate features of word like shape, texture, colour and their canonical locations as they process them. Imagine you are sitting in your living room. When someone utters the word ‘tree’, it is likely that your attention will move outside the window, into the landscape, because that is where you find trees. If this is true, then one can assume that words drive visual attention in space (Cooper 1974). In previous sections, we have seen that prepositions direct attention in space. It is not unusual since prepositions and such other words explicitly code space. But what about a word like ‘cloud’ or ‘shoe’ and ‘respect’? Do these words direct attention in some manner? Dudschig et al. (2012) examined if words like ‘cloud’ affect attentional mechanisms. The assumption was that such words would activate sensorimotor information related to space, i.e. where does one find clouds? This sensorimotor activation should influence shifts in visual attention towards locations that are consonant with the object, i.e. above. Similarly, for a word like ‘shoe’ one may look down, since all kinds of shoes are found on or towards the ground. The authors first presented these non-spatial words and the asked participants to detect a visual target. The assumption was that after reading the word ‘cloud’ detecting a visual target in the upper part of the visual field would be much faster than when the object is down. Listeners were indeed faster in visual target detection at spatial locations when the words features were congruent with such locations, indicating an influence of sensorimotor experiences on attention. This suggests that such words directed attention automatically towards spatial locations.

Do abstract words cause attentional shift? Chasteen et al. (2010) examined if abstract words like ‘god’ and ‘devil’ have spatial qualities and affect attention. The authors assumed that listening to a word like ‘god’ might activate directions like ‘up’ and ‘right’ while a word like ‘devil’ might activate ‘left’ and below’. If this is true then target detections in these subsequent locations should be facilitated. The authors found that participants shifted their attention to locations congruent with the words’ features and were faster in target detection at such locations. This suggests that even abstract words orient attention in space. This evidence further demonstrates the close link between language comprehension and visuo-spatial attention.

On similar lines, Marmalejo-Ramos et al. (2013) examined if direction words and certain emotion words are linked to spatial locations. Speakers of different languages rated direction words like ‘left’, ‘right’, ‘up’ and ‘down’. The assumption was that humans may have the tendency to map space, i.e. vertical and horizontal, differently and may assign different valence to direction words. Additionally, participants also placed different emotion words on the vertical and horizontal grids. The results indicated that direction words like ‘up’ and ‘right’ received consistently higher ratings, whereas words like ‘left’ and ‘down’ received lower ratings. This indicates some sensorimotor relationship between direction words and emotion. It is quite possible that eye movement, i.e. visual attention, may as well be in a particular direction as a function of emotional words. Zwann and Yaxley (2003) examined if readers simulate direction when they read word pairs like ‘ATTIC-BASEMENT’—a semantic relatedness task. Participants were faster in judging the relationship when the word pairs were also represented in their iconic relationship. That is, participants were faster in judgement when the word ‘ATTIC’ appeared above the word ‘BASEMENT’ than when ‘BASEMENT’ appeared above ‘ATTIC’. This shows that participants mentally simulated the location of words and when they were in canonical relationship the response was faster. In other studies, it has been observed that listeners activate spatial information when they process verbs. Verbs such as ‘push’ or ‘respect’ can simulate ‘horizontal’ or ‘vertical’ image schemas in language users. Richardson et al. (2003) presented spoken sentences that had a critical verb, which indicated a canonical spatial image schema, i.e. the verb ‘argued with’ was linked to a horizontal representation whereas ‘respect’ was linked with a vertical schema. Participants were asked to make visual discrimination judgement after listening to these sentences. Participants were slower when the visual object appeared in the location consonance with the image schema evoked by the verb in the sentence. In this case, spatial representations activated by the verb’s image schema interfered with the visual discrimination task. In another study, the authors examined if such spatial representations influence retrieval from memory. Participants heard spoken sentences that had critical verbs linked to vertical or horizontal image schemas. They were also presented with two cartoons and they had to remember their locations later in a memory probe task. Participants were faster in correctly identifying the locations when these matched with the spatial representation of the verb. Thus, in this case there was facilitation in a memory retrieval task. Others have also observed interference in spatial compatibility tasks during language comprehension. Estes et al. (2008) first allowed participants to read words like ‘foot’ and ‘head’. These words represent ‘below’ and ‘up’ locations generally. The authors found that when participants perceived the word ‘foot’, they were slower in identification of a visual target which occurred at the ‘lower’ visual field. This means, there was no facilitation but interference in this case. This evidence further supports the idea that language users mentally simulate sensorimotor events associated with the language which then influences later cognitive processing. This influence can be seen in motor actions as well as memory performances (see Dove 2015 for a review).

Language modulates spatial perception and a range of other cognitive activities. Language is not merely used for communicating what is out there, but straight away affects what is it that we are going to say, including locating objects in the visual field (Lupyan et al. 2010). The Label Feedback Hypothesis suggests that language affects ongoing perceptual activities in different modalities online (Lupyan 2012). According to this hypothesis when we learn the name of any object, its label gets associated with distinctive perceptual features of the object. Thus when the object’s name is activated, perceptual features (such as visual features) also get activated. Furthermore, there is also a possibility that objects that share perceptual features also become active. That is, if one hears the word ‘dog’, then the chances of finding a dog or another animal of similar attributes in the visual environment are much higher. Spoken language accentuates visual search as well as semantic categorization. Recent research suggests that if you want to find where you have left your car key, just say it aloud (Lupyan 2012). By naming the object to yourself, your visual search system will be active and you will find the object. Why does language affect visual search in this manner?

Since words are sensorimotor signs, they activate motor plans, even when there is no actual plan for action. According to the level feedback hypothesis (Lupyan 2012), ‘A note of caution is in order’: Viewing language as a part of an inherently interactive system with the capacity to augment processing in a range of non-linguistic tasks does not mean that performance on every task or representations of every concept are under linguistic control. Rather, the argument is that learning and using a system as ubiquitous as language has the potential to affect performance on a very wide range of tasks. A fruitful research strategy may be therefore to investigate what classes of seemingly non-verbal tasks are influenced by language (and which are not), and on what classes of tasks cross-linguistic differences yield consistent differences in performance’. This means that language only affects some cognitive performances while a lot others are unaffected. Examining the influence of linguistic labels on visual search, Lupyan and Spivey (2010) found that when participants heard the name of a letter their visual detection of the letter increased, while this was not the case with any visual cue.

If linguistic labels could bias visual attention in a top-down manner, then the argument that vision is cognitively impenetrable may be wrong (Pylyshyn 1999). Current evidence suggests an influential role of language on visual processing. Concepts activated by language penetrate into other types of non-linguistic representation. Even intentions that have not yet been actualized in actions penetrate vision (Wu 2013). If language could modulate visual search and attention mechanisms, then there must be a dynamic interface between the two representational systems. This hypothesis has also been articulated by Jackendoff in his theory of multi-modal interface where language interacts dynamically with the visual system (Jackendoff 2012a). Language does not merely put perception into some syntactic structure; it alters and modifies it, and affects perception and action dynamically.

9.4 Simulating Motion and Visual Scanning

So far we have considered how literal use of language influences visual attentional mechanism. However, language users produce a range of non-literal language. This reflects the creative aspect of human linguistic competence, e.g. the use of metaphors and idioms. It is important to know how non-literal use of language connects to attentional mechanisms. Many a times we hear language which describes inanimate things as if they are agents capable of biological motion. For example, one can say, ‘The road runs through the valley’. Is the road actually running? This is figurative use of language. Anyone who uses language knows that language can be used literally or figuratively. In the above example, the road is an inanimate object and we all know that it cannot run. However, when we hear this sentence, it seems as if the road is capable of biological motion. Use and feasibility of such metaphoric extensions vary from one language to another. This particular use is known as fictive motion. This motion is an illusory motion in the sense that there is no true motion. How does it happen? It happens because of mental simulation of motion. When a motion verb is associated with an inanimate object, there is mental simulation of motion. Another example could be ‘the tattoo runs through his arm’. The comprehension of such figurative language has been shown to affect visual attention. Examples of fictive motion sentences show that it is not only prepositions that affect spatial and visual attention but figurative language as well.

A static display depicting a road through a valley does not indicate motion of any type. However, speakers of some languages use fictive motion constructs to describe the road and the valley. While doing so, speakers ascribe agenthood to such inanimate objects. Talmy (2000) suggested that during such scene perception, the language user is trying refer to the road as the ‘figure’ and the valley as the ‘ground’ and then spatially map motion, making the road as the agent. During perception, the language user imagines this figure changing place as a result of illusory motion. These sentences can be made with any static object, which is then often related to some background.

Does this type of sentence affect visual scanning of space? Richardson and Matlock (2007) used eye tracking and experimentally investigated how fictive and non-fictive motion language affects one’s visual perception. They presented various visual scenes which showed things like a road running across a valley. Here are some example sentences:

Fictive motion sentences

1. The road goes through the desert.
2. The fence follows a coastline.

Non-fictive motion sentences

1. The road is in the desert.
2. The fence is next to the coastline.

The difference between the two types of sentences is the use of motion verbs in the fictive motion sentences. The authors proposed that listeners will simulate illusory motion when they hear the fictive motion sentences. Participants looked at the scenes and listened to these various types of descriptions as their eye movements were recorded. The results showed that during the perception of the fictive motion sentences, participants looked longer at the scene and also their scanning patterns indicated some kind of motion simulation. This was as if they were imagining the road literally 'running'. This evidence shows that comprehension of figurative language affects visual perception. Mishra and Singh (2010) extended these findings with Hindi language speakers and also wanted to see if language users simulate the motion when the visual scene is not in front of them (Fig. 9.4).

The authors used similar visual scenes and spoken sentences which showed some objects against backgrounds. Their main aim was to replicate the findings of Richardson and Matlock (2007) in Hindi while also testing if fictive motion simulation could happen when the visual scene is not present. Participants saw the scenes on a computer screen and heard sentences that either contained the fictive motion constructions, or were literal. The free word order of Hindi language allowed the presentation of the sentences in different word orders. In Hindi, one can create six types of permutations and combinations of sentences by differently placing the subjects, verbs and objects, unlike English which has only one word order. Earlier in another study, Mishra et al. (2011) showed the effect of such scrambling of Hindi sentences on identification of semantic anomaly, using a self-paced reading task. Syntacticians use the word 'scrambling' to refer to the possibility to change word order in sentences. The use of 'topicalized' forms of sentences along with the conventional ones can rule out the fact that visual attention towards the scene is not affected by the first mention of the concerned noun in any manner. For instance, in the sentence 'The dog ran after the cat', the 'dog' is in attention focus since it is the subject of the sentence and also it performs action. However, when the sentence is paraphrased into 'It is the cat after which the dog ran' then the cat becomes the centre of attention. In this way, 'cat' is topicalized for more focus. In the first experiment, it was observed that participants allocated higher visual attention towards the scenes when they heard the fictive motion sentences. Interestingly, this was even higher when the 'trajector' was the subject of the sentences compared to when it was the object, as it was the case with the topicalized sentences. For example, in the sentence 'The road runs through the valley', the noun phrase 'The road' is called the trajector (Fig. 9.5).

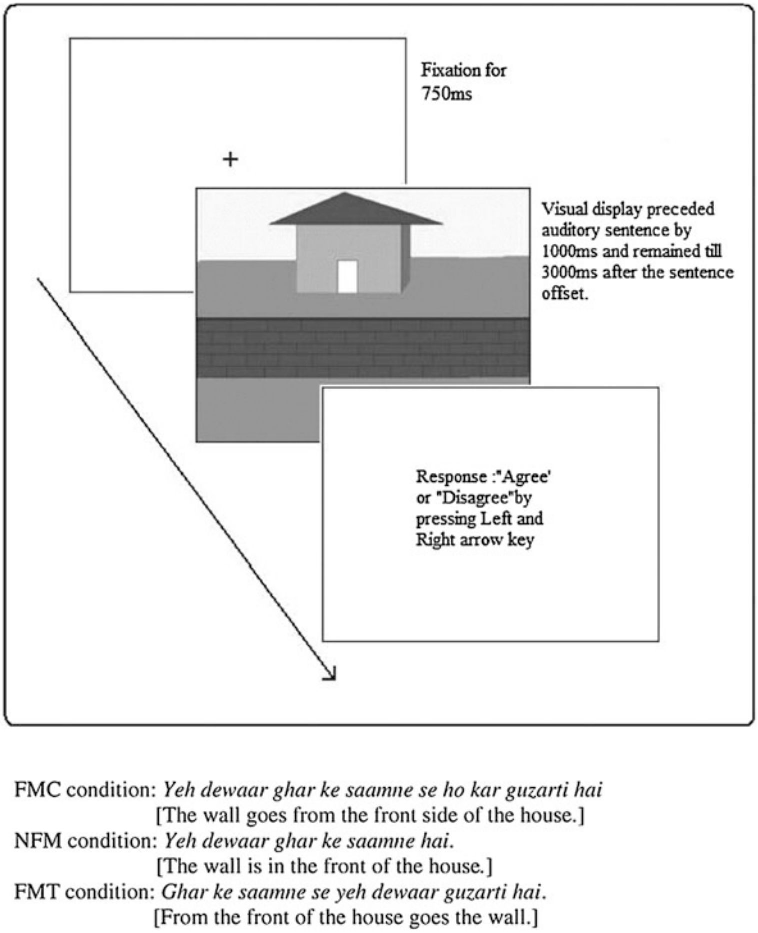

Fig. 9.4 Participants saw a clip art image that showed an inanimate object against a background. The main aim was to create objects that are regularly used as subjects of fictive motion sentences. Participants saw this picture and also heard three types of spoken sentences. One type of the sentence was a literal description of the scene; the second type carried a fictive motion component. Participants' eye movements were measured throughout. *Source* Mishra and Singh (2010), p. 147

In the second study, participants first saw the visual scenes and then after a brief moment the scene was withdrawn from the computer screen. In the absence of the scene they listened to the same set of sentences that included both literal and fictive motion sentences. This time, in the absence of the scene, motion simulation during the fictive motion sentences was low. This pattern of results suggested that during mental imagery motion simulation was difficult. However, both the experiments reveal that there could be an effect of syntactic structure on the amount of motion simulation, and it also matters if viewers are still seeing the scene when they hear the sentences.

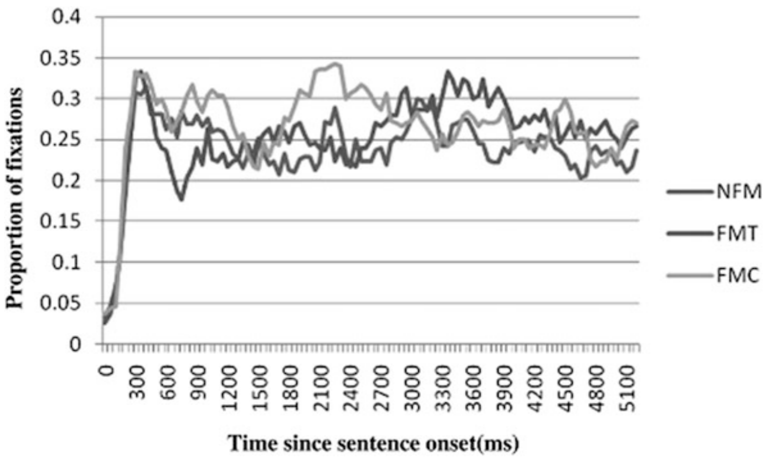

Fig. 9.5 Plot shows proportion of fixations to the trajectory region as a function of time. It is apparent that participants paid higher visual attention when they heard the fictive motion sentence compared to the literal sentence. That is because during the fictive motion sentence there was mental simulation leading to a different strategy of attention allocation. *Source* Mishra and Singh (2010), p. 151

Why did the participants allocate higher visual attention when the scene was no longer present? Mishra and Singh (2010) argued that it could have been because of limitations in the working memory for the scenes (Melcher and Kowler 2001). With a brief preview of the scenes, viewers could only remember the broad spatial layout and the gist of the scene and not detailed knowledge. When they heard sentences describing the scenes, this vague visual outline was not sufficient for mapping language onto vision. Interestingly, with single line drawings of objects, others have shown that even in the absence of scenes, viewers make correct eye movements towards blank locations when they hear names of these objects. Overall, these studies do suggest that some properties of language affect how we perceive space and look at objects. Along with linguistic and visual structures, one has to also seriously consider the various roles played by attention, memory and the perceptual mechanisms in this scenario.

Why must language lead to mental simulation? That is because language is used to describe perceptions that are both possible and impossible. The perceptions that are impossible also trigger the motor system. One can virtually imagine anything and put it forth in language. Language maps this possible perception and gives a symbolic structure for expression. Therefore, language creates the whole experience of perception in the absence of things, even when they lack clear attributes. It is interesting to see that visual scanning of objects that do not move changes with regard to certain forms of language use. This means that language can give rise to experiences which are not necessarily observable otherwise. Much of this evidence has been used to support the embodiment theory of language.

9.5 Conclusion

The chapter showed that language at all levels has an interface with attention and vision. Language has an immediate effect on our looking and attending behaviour. Language can lead to attention shift in a top-down manner. That is, listeners look at specific locations depending on the meaning of words and the sensorimotor experiences associated with them. Words can also capture attention in a bottom-up manner as it was with the studies with abstracts words. Furthermore, language influences how one maps space and describes spatial situations in everyday life. As for attention, it mediates this interface between language and cognition broadly. Language use and its grammar constrain and how attention is deployed over space. For example, in the case of fictive motion sentences, it is the linguistic use which affects visual scanning. This sort of sensorimotor response triggered by language suggests the very close and functional link between language and perception. This is what has been recently argued by Lupyan and Ward (2013) where he showed it is the language which makes absent objects again appear back. Naming and describing things make them more visually prominent. Language, then, is not merely a collection of a modal and meaningless symbols. These symbols stand on our perception and experience. These symbols trigger action in a range of situations. Eye tracking evidence certainly suggests a close link between language–attention and vision.

Language and space are part of the environment. Attention helps brains to link spatial aspects of objects to organisms. As organisms orient in space themselves, their changed perspectives are taken care of by the attention system. Additionally, cultural forces shape how certain expressions are used in some communities. Studies from anthropological linguistics and cognitive psychology show this tight link between language, cognition, brains and environment. It is perhaps very difficult to pin point the individual roles played by language and attention during spatial cognition. It is likely that language codes certain attention schemas. These schemas are activated when certain words are comprehended. That way language economizes perception and helps select the correct attentional strategy during search.

References

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502–518.

Altmann, G., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition, 111*(1), 55–71.

Bellugi, U. (2014, January). Language: perspectives from another modality. In *Modularity and the Motor Theory of Speech Perception: Proceedings of a Conference to Honor Alvin M. Liberman* (p. 145). Psychology Press.

Bloom, P. (Ed.). (1999). *Language and space*. Cambridge: MIT press.

Bowerman, M. (1989). Learning a semantic system: What role do cognitive predispositions play? *The teachability of language* (pp. 133–169). Paul H. Brookes: Baltimore.

Brown, J. H., Johnson, M. H., Paterson, S. J., Gilmore, R., Longhi, E., & Karmiloff-Smith, A. (2003). Spatial representation and attention in toddlers with Williams syndrome and Down syndrome. *Neuropsychologia, 41*(8), 1037–1046.

Chasteen, A. L., Burdzy, D. C., & Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia, 48*(2), 627–630.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*(1), 83–121.

Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends in cognitive sciences, 10* (8), 370–374.

Coello, Y., & Bonnotte, I. (2013). The mutual roles of action representations and spatial deictics in French language. *The Quarterly Journal of Experimental Psychology, 66*(11), 2187–2203.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*(1), 84–107.

Coventry, K. R., & Garrod, S. C. (2004). *Saying, seeing and acting: The psychological semantics of spatial prepositions*. Hove, UK: Psychology Press.

Coventry, K. R., Lynott, D., Cangelosi, A., Monrouxe, L., Joyce, D., & Richardson, D. C. (2010). Spatial language, visual attention, and perceptual simulation. *Brain and Language, 112*(3), 202–213.

Dove, G. (2015). Three symbol ungrounding problems: Abstract concepts and the future of embodied cognition. *Psychonomic bulletin & review* (pp. 1–13).

Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). From top to bottom: Spatial shifts of attention caused by linguistic stimuli. *Cognitive Processing, 13*(1), 151–154.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down object words orient attention to the objects' typical location. *Psychological Science, 19*(2), 93–97.

Flecken, M., Athanasopoulos, P., Kuipers, J. R., & Thierry, G. (2015). On the road to somewhere: Brain potentials reflect language effects on motion event perception. *Cognition, 141*, 41–51.

Gentner, D., Özyürek, A., Gürcanli, Ö., & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition, 127*(3), 318–330.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition, 55* (1), 39–84.

Herskovits, A. (1987). *Language and spatial cognition*. Cambridge: Cambridge University Press.

Jackendoff, R. (2012a). *A user's guide to thought and meaning*. Oxford: Oxford University Press.

Jackendoff, R. (2012b). Language as a source of evidence for theories of spatial representation. *Perception-London, 41*(9), 1128.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic books.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. *Language and space* (pp. 109–169).

Levinson, S. C. (2006). On the human "interaction engine". In *Wenner-Gren Foundation for Anthropological Research, Symposium 134* (pp. 39–69). Oxford : Berg.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83* (3), 265–294.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1015–1036.

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in psychology, 3*, 54.

Lupyan, G., & Spivey, M. J. (2010). Redundant spoken labels facilitate perception of multiple items. *Attention, Perception, & Psychophysics, 72*(8), 2236–2253.

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences, 110*(35), 14196–14201.

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science, 21*(5), 682–691.

Majid, A., Bowerman, M., Kita, S., Haun, D., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in cognitive sciences, 8*(3), 108–114.

Marmolejo-Ramos, F., Elosúa, M. R., Yamada, Y., Hamm, N. F., & Noguchi, K. (2013). Appraisal of space words and allocation of emotion words in bodily space. *PLoS ONE, 8*(12), e81688. doi:10.1371/journal.pone.0081688.

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research, 41*(25), 3597–3611.

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing, 11*(4), 295–305.

Mishra, R. K., & Singh, N. (2010). Online fictive motion understanding: An eye-movement study with Hindi. *Metaphor and Symbol, 25*(3), 144–161.

Mishra, R. K., Pandey, A., & Srinivasan, N. (2011). Revisiting the scrambling complexity hypothesis in sentence processing: A self-paced reading study on anomaly detection and scrambling in Hindi. *Reading and Writing, 24*(6), 709–727.

Munnich, E., Landau, B., & Dosher, B. A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition, 81*(3), 171–208.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition, 98*(3), B75–B87.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155–184.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128–2148.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(03), 341–365.

Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition, 102*(1), 129–138.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive science, 27*(5), 767–780.

Talmy, L. (1983). *How language structures space* (pp. 225–282). US: Springer.

Talmy, L. (2000). *Toward a cognitive semantics, Vol. 1: Concept structuring systems*. Cambridge, MA: The MIT Press.

Wu, W. (2013). Visual spatial constancy and modularity: Does intention penetrate vision? *Philosophical Studies, 165*(2), 647–669.

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in cognitive sciences, 18*(5), 229–234.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*(1), 1–11.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review, 10*(4), 954–958.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science, 28*(4), 611–619.

Chapter 10
Attention and Language: A Linking Proposal

10.1 Attention in Different Language Tasks: Goals and Intentions

Of course, there is as yet no agreement among cognitive psychological researchers on the true nature and dimensions of attention. Chapters in this book have shown that language and attention dynamically interact, affecting behaviour and cognition in many everyday acts. There are general mechanisms that are common to most domains discussed in this book. Attention selection or control has been used in explaining psycholinguistic processing. In spite of surface variations, most researchers have considered only a few facets of attention in interpreting their results, as far as language behaviour is concerned. Often, the use of a particular tool or the study of a particular phenomenon has constrained this or even debates internal to an area. It appears that researchers interested in a particular psycholinguistic process have looked at only one or another aspect of attention and not more broadly. Further, as argued before, the goals of language users with regard to task have not been fully looked at. This leads to a theorization where only bottom-up effects of some linguistic stimuli on attention mechanisms have been explored.

In Chap. 3, it was observed that historically there have been issues with the use of cognitive psychological concepts in understanding of psycholinguistic phenomena. That was because language was considered abstract enough to be examined like a simple visual object. Interestingly, this has changed in recent times and more and more researchers are seen using attention as a conceptual tool to look at language behaviour (Kurland 2011). However, one can understand how attention and language interact as far as we have good understanding of the basic mechanisms themselves. Currently, in both cognitive and experimental psychology (Anderson 2011) as well as in philosophy of mind (Mole 2011) there is considerable disagreement on the notion of attention itself and the mechanisms that this notion relates to. For example, Anderson (2011) has argued that attention is an effect of mental processing rather than a cause of mental events, as many think of it,

R.K. Mishra, *Interaction Between Attention and Language Systems in Humans*, DOI 10.1007/978-81-322-2592-8_10

and use it to explain cognitive processing. Anderson asserts that attentional effects are distributed throughout cognitive events and states and are not to be thought of as merely present or absent. While the view could be debatable, since almost all research in cognitive psychology has treated attention as a causal mechanism which the agent pays to achieve successful cognition, it nevertheless points towards new ways of talking about attention. There is also now controversy regarding the age-old dichotomy between top-down and bottom-up attention. Awh et al. (2012) have challenged this dichotomy suggesting that current experimental effects do not equivocally show that agents use top-down goals to achieve selection. Rather, often bottom-up features attract attention even when such features are not part of the agent's goal. Therefore, it is not easy to explain which mechanisms underlie attention selection in a given task. All these views have implications on our understanding of attention and language interaction looked at from a cognitive processing perspective.

Chun et al. (2011) proposed that attention is not a single, unified mechanism but a cluster of different systems that correlate with specific neural activations. According to Chun, although there are many types of attentional mechanisms that one can see in different tasks, all such mechanisms share three fundamental properties: Attention is primarily (a) a filtering mechanism, (b) it is selective and (c) it modulates the agent's behaviour during tasks. We have already seen that these three aspects of attention have been well implicated in different psycholinguistic studies. However, it is not clear if all psycholinguistic processes manifest all of these properties of attention or some of them. Different linguistic tasks involve attention differently. Further, language users recruit attention depending on their goals in a particular context. For example, as shown in the chapter on sentence processing (Chap. 4), attention to grammatical or syntactic elements of sentences can differ depending on what the participant has been asked to do (Hahne and Friederici 1999). Similarly, depending on what kind of utterances the speaker has to produce he may deploy attention selectively on a visual scene (Coco et al. 2014). Research in the domain of reading has shown how bottom-up features of text can influence comprehension as well as eye movements (Rayner 1998). In visual world studies of language-mediated eye movements, we see automatic shift of attention with regard to linguistic input (Mishra 2009).

More recently, Mole (2011) has articulated a view of attention where he proposes attention as an overall unifying mechanism for action. Mole argues that primarily attention is used by the agent for realizing goal-directed action. Without real goals to achieve, attention is not called for. Thus, in this view, attention flows from intentions of agents as they chase a goal and exert action. It is important to note that not always do agents exert actions even though they are attentive. According to Mole, we can only say that the agent is deploying selective attention in a task when we can know for sure that the agent is fully conscious and knowledgeable of his/her task goals and is not displaying the attentive behaviour as a reflex. Thus, to be attentive or to engage attention means one should be fully aware of the task and consequences of action. How does this view relate to language processing?

Language users are not often conscious of language processes themselves. For example, language users are not in a position to report the underlying linguistic computations that lead to sentence production. Similarly, listeners cannot explain what mechanisms lead to comprehension of speech. However, language users are conscious of the effects of such linguistic actions. One knows what effect a sentence will have on the listener. In this sense, following Mole's argument, while linguistic behaviours happen unconsciously inside the nervous systems of agents, agents are knowledgeable only of the effects of such processes (Wharton 2013). Thus, depending on the effect a linguistic action will have on communication and other agents, attention is involved. If Mole is correct in his theorization, then one can say that language users only involve attention when they are fully conscious and are in control of the linguistic task. Mole of course does not examine empirical research from psycholinguistics but concentrates mostly on findings from the domain of visual cognition, as it has been happened classically. In any case, for our purpose, it can be concluded that linguistic actions being goal oriented, reflect the agent's awareness of actions and, therefore, attention is involved. Without attention, such goal-directed behaviour will not result. Now, the question arises, which linguistic actions or tasks reflect such typical goal-oriented behaviour? Traditionally, language production has been considered as intentional and goal-directed performance (Levelt 1992), while listening is passive, from an agent's point of view. Similarly, reading is also passive but writing is intentional. Therefore, it is critical to examine and summarize how these various activities reflect attentional involvement as they relate to communication and linguistic cognition.

The question if attention is required for language functions is not a good question to ask. That is because it implies a causal view of attention which is currently being critiqued (Anderson 2011; Mole 2011). It also inevitably leads one towards some kind of modular thinking. The question should be how different language mechanisms involve attentional states and the role played by the agent in them. The problem of linking mechanisms of attention to subtleties of language is manifold. On the one hand, we have different modes of language performance. On the other hand, we have different levels at which attention mechanisms may seem to interact with language. Although many psycholinguists who have worked on linguistic performance in these various areas talk of attention, it is difficult to find out any particular shade of attention which is common to all of them. Interestingly, not many even explicitly refer to basic components of attention such as endogenous and exogenous. What seems reasonable is to select some linguistic functions and see if attention behaves similarly in such cases. One can also include neuropsychological cases where a very direct functional relationship between executive functions and language exists (Mesulam 1990). One way to understand the interaction of attentional mechanisms and language is to look at the demands of the task. Every task that the language user performs is controlled by goals and is affected by bottom-up features. Therefore, attentional involvement with language can be both top-down and bottom-up.

Language, though it is a symbolic system with its own unique properties, as argued in this chapter and elsewhere, manifests as an action system. Language use is always goal directed and thus, it expresses intentionality (Searle 1997). For example, when a participant is asked to name pictures in a particular language in an experiment, s/he has a goal and s/he is performing an action. Is the goal intentional? It *is* in some sense and *isn't* in some others. The subject in the psycholinguistic experiment has been given a goal which is not his/her true goal. However, once s/he understands the task well, s/he is able to focus attention (Mole 2011). Visual search studies have shown that intention for action of subjects influences how attention is deployed during any task. Visual attention to targets has been shown to be selectively enhanced depending on whether observers had the intention of merely looking or to perform some action (Bekkering and Neggers 2002). Similarly, preparing a manual grasping or pointing movement has been shown to selectively affect visual feature detection (Wykowska et al. 2009). Thus, when agents have specific actions in plan, perceptual processing is biased towards such goals (Baldauf and Deubel 2010). This fits well with the general predictions of the biased competition account of selective attention which predicts goal-related enhancement of target detection (Desimone and Duncan 1995). Therefore, unless one knows the goals and intentions of agents in any task, theorization of attentional involvement will remain incomplete. Attentional involvement during visual search has been viewed as a selection for action mechanism (Hannus et al. 2005). All the different experiments described in this book in the context of different linguistic processes are of this sort. Subjects perform the task, presumably with full attention since they have been instructed to do so. Studying attention and language in the experimental set-up thus has its own weaknesses. One is never sure if subjects are paying attention and if they are fully aware of the tasks and its constraints.

In Chap. 4, we saw that Fodor's views of modularity and particularly the notion of autonomous encapsulated processing of certain systems like syntax have led to a re-emergence of such views recently in the experimental literature (Hahne and Friederici 1999). Studies with mismatch negativity and sentence processing support the notion that focused attention need not be necessary for comprehension of sentences (Pulvermüller et al. 2008. This view stems from the divisions inside theoretical linguistics related to semantic and syntactic processing. Others have looked at attention as a limited processing resource in sentence processing. For example, the effect of working memory load on sentence processing has been taken as evidence suggesting language requiring executive processes (Gathercole and Baddeley 2014). If sentence processing is automatic, then it does not call for focused attention. On the other hand, those who have examined sentence processing under some kind of memory or perceptual load have considered the 'capacity'-limited perspective of attention. Thus, sentence processing research has vacillated between notions of 'automaticity' and 'executive control' accounts involving attention. However, an automatic processing of syntax should not mean that attention is not necessary for syntactic computation. Data from Parkinson's disease patients show that lesions in the anterior cingulate cortex leads to comprehension deficits in complex sentence processing (Grossmann et al. 1992a, b). Similarly, studies with children with different developmental language disorders like specific language impairment suggest that attentional deficit leads to linguistic deficit

(Finneran et al. 2009). If these children lack sustained attention then they show deficits in processing sentences.

Attention is required for sentence processing in general. If syntax processing is automatic and does not call for attention, then how speakers are conscious of what they are speaking? How do they know if the to-be-produced sentence is situationally correct? It is likely that overall cognitive goals exert top-down influences on selection on linguistic items as it happens in case of visual object identification (Desimone and Duncan 1995). Even though some aspects of syntax processing are automatic, speakers know if the sentences are suitable for their goals. While it is still debatable if syntax needs attention, it is likely that top-down attention influences sentence processing, making communication successful. It is another matter how attentional resources are divided among several linguistic levels like syntax and semantics. Chapters 5 and 7 have shown that even visual contexts affect how one processes language. Thus, attention interaction with language as far as sentence processing is concerned should not be viewed from a theoretical vantage point that subscribes to modularity.

While the involvement of attention in the debate of 'automaticity' of syntactic processing compared to semantic processing is important, it is also crucial to refer to the proficiency of speakers when it comes to sentence parsing. In Chap. 6, we observed that recently several studies on bilinguals and illiterates have shown that the variable nature of proficiencies and strategies used directly influence sentence processing. Since most studies on parsing that have examined issues related to attention have been on monolinguals and college students, the individual difference angle is missing from the analyses. For example, bilinguals who have good proficiency in a second language and who are late learners find parsing a second language difficult (Maclntyre and Gardner 1994). Individual abilities in working memory and attention are linked to successful parsing of complex sentences in a second language (Hopp 2013). Thus, in such cases, the involvement of attention in sentence processing is linked to the concept of 'effort' and 'motivation'. Without the requisite cognitive skills parsing becomes difficult. Similarly, illiterates have been shown to be slow in online sentence processing (Mishra et al. 2012). This also nicely corroborates with the additional finding that illiterates show deficits in selective attention tasks such as visual search (Olivers et al. 2014). Therefore, it is likely that attention and memory resources aid sentence processing. However, it is still debatable if there is a selective executive control system tackling sentence processing, or if there is a single system that works for both linguistic and non-linguistic stimuli and their processing. Recently, it has been observed that domain-general enhancing of cognitive control facilitates the processing of ambiguous sentences (Novic et al. 2014). This also fits well with observation that those children who lack attentional capacity also do poorly with sentence processing tasks. For example, children with specific language impairment who show visual attentional deficits also are poor in online sentence processing (Dispaldro et al. 2013). Therefore, attentional involvement in sentence processing should be viewed as function of overall cognitive profile of individuals.

Apart from tasks that just involve language, we also saw that language input in the context of visual scenes affects attentional orienting towards relevant objects (Chap. 7), which different visual world studies have demonstrated (Mishra et al. 2013). Studies that have examined eye movements during spoken sentence comprehension in the co-presence of visual scenes do not refer to attention as a limited resource system or its role as a selective agent (Ferreira et al. 2013). Attention to visual objects during real-time language processing has been shown to reflect anticipatory and predictive processing (Altmann and Kamide 2007). Such shifts of attention as a function of linguistic input to relevant objects in the visual world also have been taken to reflect goal directedness of the comprehension system (Salverda et al. 2011). This comes very close to Mole's characterization of attention that is linked to an agent's understanding of the task and goals. Language comprehenders show very good and dynamic situational awareness during such online sentence processing, which considers the interaction of both linguistic and pictorial representations, including world knowledge (Knoeferle and Crocker 2006). In these studies, then, attention helps the cognitive system orient towards task relevant objects while guided by top-down knowledge and the semantic system. Further, such visual world studies also demonstrate the fact that linguistic processing does not operate in insulation but in a rich, multi-modal environment where both linguistic and non-linguistic information influence attention (Spivey 2007). Thus, while some have looked at attentional involvement in sentence processing from an executive control and limited resource point of view, others have looked at it more as a linking system that connects different kinds of representations (Henderson and Ferreira 2004).

Psycholinguistic studies require participants to do a task. It could be a grammatical judgement task administered through self-paced reading or spoken sentence comprehension or even object naming. Very few researchers have examined how variations in such task goals could influence attention processing. This research has not yet looked at attention as a mechanism that the agent brings to a task of which he is fully conscious. Tasks commonly used in such research have not paid attention to individual indifference that is linked to performance. While the focus has been on computation, it is not clear what role a central mechanism like attention plays in it. In order to examine how goal-directed behaviour of agents influence language performance, new paradigms should be used. For example, this has been done in visual cognition where participants choose a certain action and whose influence is seen on performance (Haggard 2005). This theorization is critical since the involvement of top-down attention is linked to the conscious intention of the agent, be it any task or stimulus (Corbetta and Shulman 2002). It is critical to know how the language user uses top-down attention in the task, since the involvement of attention is contingent on this. Of course, it has been shown that bottom-up features can influence attentional engagement, but this is not going to reveal how agents deploy goal-directed attention. One needs to design experiments where agents freely select their tasks and then one can see how such goals influence the engagement of attention. While I am conscious of the current problems with the top-down versus bottom-up division of attention (Awh et al. 2012), there is also merit in the argument that agents decide on the tasks and deploy attention depending on the scenario (Mole 2011).

There is no doubt that language reflects the intentionality of speakers. In Chap. 5, we saw that major theories of object naming consider it as a top-down process where competition between lexical items leads to a final selection (Belke et al. 2005). Selective attention towards targets helps resolve the competition. However, not everyone agrees that naming invariably involves some competition (Finkbeiner and Caramazza 2006). Attention to an object is necessary for retrieving its phonological and semantic properties as has been shown in different eye tracking studies (Malpass and Meyer 2010). Eye tracking studies with naming of multiple objects show how a certain amount of selective attention to one object should be paid for reaching a certain level in name retrieval. Attention to the next object only moves after speakers have completed phonological form retrieval of the currently fixated object (Meyer et al. 2012). Thus, visual attention and its shift across several objects during naming are contingent upon the degree of linguistic processing achieved on particular objects. Selective attention, as measured through fixation durations and saccades in eye tracking studies, is related to the retrieval of information from visual objects. It is possible that this theorization about the involvement of attention during visual object naming is an outcome of the paradigm itself, an artefact. Since we do not name single objects generally and second, objects do not appear well arranged in our visual field as they are on a computer screen during an eye tracking experiment. Finally, we also name objects when they are not present visually, retrieving their conceptual and phonological form from semantic memory, as it happens commonly during a verbal fluency task. It has been shown that prefrontal cortex shows higher activity during generation of words (Phelps et al. 1997). Thus, selection being top-down in this case, guides the generation of words. In contrast to visual object naming, where bottom-up features influence attentional engagement, verbal fluency tasks provide good data on goal-directed action during naming.

If attention is necessary for selective processing of some over others, then there should be some cost and benefits associated with such a mechanism. The biased competition account of attention proposes that attention is biased towards potential targets in a top-down manner, thus leading to inhibition of distractors. How does this work for language production? Nozaria and Dell (2012) have shown that selective attention to one visual object slows down the naming of others. This is similar to results in the field of visual cognition where selective attention to targets leads to inhibition of distractors. It has also been shown that stimulation of the prefrontal cortex facilitates naming (Nozari and Thompson-Schill 2013). It is likely that whether selective attention facilitates or induces a cost in performance depends on the goals of the agent. We also saw that when attention is cued towards particular pictures, it affects choice of linguistic structures (Tomlin and Villa 1994), although it is not clear how systematic changes in attentional allocation are linked to generation of sentential structures. What is known is that attending to something brings this object to consciousness and also linguistic structures code this object, allotting it a prominent position (Myachykov et al. 2013; see also Ibbotson et al.

2013). Attention thus enables the selection of the correct linguistic structures. It appears that one's capacity of executive control is linked to one's ability in action control in speech production (Shao et al. 2012). That is, speakers who are better in an executive control task are also better in control in naming and other language production tasks. Executive control process can predict fluency in speaking (Engeldarth et al. 2013). The direct influence of executive control on naming also suggests that naming is more top-down.
Production studies have also implicated executive control. Being able to maintain selective attention is an aspect of executive control. A general purpose monitoring mechanism subserves speech productions keeping it largely error free (Riès et al. 2011). Speakers use top-down goals to monitor speech production. Speakers can modulate monitoring depending on speaking context and other variables (Dhooge and Hartsuiker 2012). However, it should be noted that speakers do not show conscious awareness of such monitoring mechanisms during speaking. Speaking is first and foremost an intentional activity. It is always about something. Choosing to speak about something and not others itself involves attention. Attention mechanisms need not be only involved with overt orienting and eye movements. Therefore, any need to speak is an act of attention selection. Language, thus viewed, reflects intentional selection of some perceptions over others. Naming tasks that are commonly employed in psycholinguistic tasks do not capture the intention of the speaker. Does the intention to speak arise merely looking at a picture? Why intentional selection matters to me here is, if the speaker has the intention to generate a name, only then will attention be meaningfully engaged. It has been shown that phonological forms of pictures are activated even when the viewer does not have any intention of naming them. Strijkers et al. (2011a) have shown that an explicit intention to speak can influence the speed of lexical access. Therefore, attention is channelized differently depending on intention. Speaking in this sense is fundamentally different from listening since speakers are aware of their intentions for possible action and therefore for attentional engagement.

What is currently missing from the psycholinguistic investigations of naming is the link between intention and action. This link is important in understanding how attention is involved in naming and how soon attentional selection starts to affect lexical access (Strijkers et al. 2011b). It should also be noted that when the subject is asked explicitly to give a particular response, i.e. it does not involve free choice. Therefore, these studies may not directly reveal how free choice of an action should involve attention. For example, if I am a bilingual and I wish to name, it is quite possible that I choose to name in a particular language and not another. This free choice is though subject to linguistic context, attention engagement will be directly affected by it. As already indicated, studies in visual perception often involve paradigms where participants make a free choice and execute a task (Salvaris and Haggard 2014). If the task and goal based view of attention is to be realized in linguistic task, then such paradigms should be used that allow the language user the freedom to choose the action.

Reading as a linguistic cognitive activity differs from speaking and listening in a several fundamental manner (Chap. 8). Reading is parasitic on spoken language knowledge and a large percentage of world's population is still illiterates (Huettig and Mishra 2014). As noted above, concepts like 'intention' or 'control' have rarely been explored in reading. Although, like in sentence processing studies, many have taken visual word recognition as a process to be highly automatic. For fluent readers visual word recognition is effortless and takes only few hundred milliseconds (Sereno and Rayner 2003). Thus, reading does not call for selective attention as such. However, as noted in Chap. 8, contemporary eye movement models of reading have implicated shifts of attention in a serious way. Reading models certainly have brought work in cognitive psychology much closer to psycholinguistics. Many others who work with spoken language or language production view reading as an artificial situation which cannot be generalized much to other linguistic activities. However, shifts in attention during scene viewing as well as reading seem to suggest one underlying mechanism specific to the participant (Reichle et al. 2012). Oculomotor control models of reading show a direct and functional relationship between lexical access on the one hand and saccades in another. However, I am hesitant to recommend these models to someone who wants to understand how speaking and listening interact with attention at this point in time (see Rayner 2009). The question has not been if attention is required for reading per se, but if the movement of attention across the printed words is serial (Rayner 1998, 2009) or if it operates in parallel (Engbert et al. 2005). While the E-Z model proposes (Chap. 8) attentional shift linked to completion of lexical processing at the current word and serial, the Saccade-Generation With Inhibition by Foveal Targets (SWIFT) model assumes a large visual span where meaning from multiple words are acquired in parallel. Both models have struggled to explain the parafoveal word's effect on foveal word reading, a prediction linked to the serial versus parallel nature of attentional deployment (Kennedy and Pynte 2005). Eye movement research on reading have used very sophisticated paradigms such as the gaze contingent moving window paradigm that allows researchers to examine precisely the covert movement of attention and its functional role in saccades. Word features such as frequency has been shown to influence both duration of fixations and also timing of saccades towards next targets (Ghahghaei et al. 2013). It has also been shown that attentional scope automatically varies as a function of type of word, i.e. familiar or unfamiliar (Montani et al. 2014). Therefore, spatial attention and its movement during reading critically depend on features of written words and their psycholinguistic properties.

In reading research, the link between attention and lexical activation has been established in a more robust manner (Rayner 2009). That is because reading lends itself to good experimental manipulation. Most innovative paradigms with eye tracking have been developed fruitfully in reading (Reingold and Stampe 2000). Apart from the debates concerning the differences in orthography and phonological properties, most reading researchers believe that word recognition is a visuo-linguistic process. However, that does not mean that there is consensus about attention being prerequisite for reading or how attention modulates reading and oculomotor behaviour. The idea that reading is a highly automatic skill at least for the fluent reader is widely accepted among cognitive psychologists and psycholinguists (LeBerge and Samuels 1974). Looking at a word activates its semantic and phonological code automatically. Shifts in attention across words in a sentence are linked to sequential activation of meaning (Rayner 1998). The oculomotor models of reading however incorporate attention in a prominent way in modelling and theorizing. Various patterns of gaze durations indicate different aspects of cognitive processing during reading and attentional selection. Both the 'capacity limitation' view of attention and 'automaticity' are used in the interpretation of experimental data. Current data are able to model the principles on which attention shifts and eye movements happen as one begins to read. Visual information acquired from the parafovea influences the attentional mechanism during reading. Reading and eye movements show a strong connection between attention shift, eye movements and lexical access. However, the debate of whether attention shift is serial or parallel is internal to reading. In sum, reading researchers look at attention involvement from a spatial perspective and link this to language comprehension.

Reading is not automatic for dyslexic children as they may have some visual or oculomotor problem (Eden et al. 1994). Reading-disabled children show higher fixation durations and also greater regressions during reading as measured in eye movement studies. We also noted that reading in such children suffers when there is visual crowding (Chap. 8). It appears that some of the reading-related deficits can be explained by attention, just as children with specific language impairment have been found to have attention deficits. Dyslexic children seem to have a deficit in spatial and temporal attentions (Ruffino et al. 2014). Additionally, such children may have poor attention span which restricts information acquisition from multiple words in parallel (Visser 2014). Therefore, attention deficit in dyslexia leads to linguistic deficits (Castles and Friedmann 2014). This evidence suggests that attentional capacity can predict performance in higher language functions.

Is reading for comprehension then an action? It is certainly so, if one looks at reading as a cognitive activity that readers perform for extracting knowledge as they perform any other task. Psycholinguistic or cognitive psychology studies have not carefully studied this aspect, since at best they study attentional mechanisms during reading a single word or a sentence. Mostly readers are asked to judge the grammaticality or semantic acceptability of such sentences, which is not what fluent reading for comprehension is about. How does top-down and bottom-up attentions influence reading? While discussing attentional involvement in speaking, I pointed out that one's intention of speaking can facilitate lexical access during naming. That is because naming is an action. As for reading, current discussions do not focus on goal directedness or intention of the reader. If attentional engagement is linked to goal-directed intentional activity, then reading should be viewed as such. Reading is not just a visual world recognition mechanism where attention helps to move the eyes across print. Just and Carpenter (1980) had observed that eye fixations seen during reading are an outcome of both top-down and bottom-up processes. Top-down processes in this case can also refer to predictive processes that allow readers to anticipate information in advance depending on their background knowledge and fluency. Good readers do not fixate on all the words of a sentence and use their ability to predict upcoming words (Cutting and Scarborough 2006).

Goals and intentions of the reader influence attention mechanisms and ultimately eye fixations seen during reading. It has been long observed that motivation plays a key role in reading comprehension (Guthrie and Wigfield 1999). Researchers who have studied second language reading also have observed a key role of motivation (Smith 2012). Therefore, beyond helping move the eyes across print for information acquisition purposes, attention is deployed by the reader as a function of his goals and motivations. Unfortunately, not many empirical studies have investigated this angle in understanding reading. The example of word superiority effect during visual word recognition may suggest the influence of background knowledge and indicate an overlearned response, but it is not a good example of a top-down goal. It is a good example of reflex-like retrieval of stored information with brief exposure to some stimuli.

10.2 Attention, Language and Multi-modal Interaction

Chapter 7 discussed various visual world studies on spoken language comprehension where language has been shown to dynamically trigger attention shifts. How is attention to be conceptualized in such a scenario where language and visual information interact dynamically? This has given a powerful paradigm, which can be used to understand the temporal and cognitive effects of spoken word comprehension at the same time; it has also opened up debates around language–vision interaction, a broad theme. Many who have looked at cross-modal processing have considered the effect of spoken language on attention. On the one hand, one sees the movement of attention as a function of spoken word comprehension and, on the other hand, one sees automaticity in such eye movements (Mishra et al. 2013). Spoken words automatically move attention. Studies have shown that one's attention is affected even if one wants to ignore the spoken stimuli (Salverda and Altmann 2011). This kind of data points towards some kind of automaticity arising out of practiced behaviour. Listening to words makes the attention system search the referred object or even objects that are related to it. Such demonstrations indicate a functional link between comprehension and attention shift in the visual world. Studies with sentences and visual scenes also show anticipatory shifts in visual attention much before one has fully comprehended the linguistic event. Attention shifts occur as a function of prediction in language processing during comprehension.

Beginning with the studies of syntactic ambiguity resolution (Tanenhaus et al. 1995), this method has been extensively used to map time course and activation-related issues during spoken word processing (Huettig and Altmann 2007). Eye movements towards objects whose names even partially match the spoken word indicate transient activation of competitors. Saccades towards competitor objects, i.e. eye movements towards a *beaker* when spoken word is *speaker*, indicate that cognition evolves over time (Spivey 2007). Attention towards visual objects described by the language indicates an aspect of this competition through which final cognition emerges. These language-mediated eye movements and the role attention plays in them have also been linked to the sensorimotor systems (Mishra and Marmolejo-Ramos 2010). These studies indicate the very non-modular aspect of linguistic and visual cognition (Henderson and Ferreira 2004), while some still hold the view that visual processes are cognitively impenetrable (Pylyshyn 1999), language-mediated visual search suggests a complete penetration of visual experience by world knowledge (Altmann and Kamide 2007). That is, the semantic knowledge associated with language influences the oculomotor programme that orients attention towards specific objects and not towards distractors. Using background knowledge of actors and actions listeners anticipatorily look at visual scenes much before they have heard the descriptions through language (Kamide et al. 2003). Visual world studies thus provide solid evidence of the link between attention, language and vision (Mishra et al. 2012). Compared to studies on reading, these studies are more ecologically valid and natural, since listening and looking are natural. Even infants and children seem to look around at appropriate objects with slight input of spoken language (Mani and Huettig 2012). If attention as conceptualized here manifests itself in goal-directed behaviour, the eye movements seen during language-mediated visual search should be related to top-down goals of the perceivers (Salverda et al. 2011). Several investigators have given explicit tasks to participants, i.e. *move the apple*, *select the triangles*, thus making sure that looks towards objects reflect action-oriented processing (Dahan and Tanenhaus 2005; Chambers et al. 2004). Like in visual search tasks, in these studies listeners actively process the spoken instructions to do some task. Thus, attention shifts towards target and competitors reflect attention mechanisms linked to tasks. This mode of attention could be regarded as top-down, as it is used for action. Any bottom-up capture of attention by distractor objects are quickly overridden by the attentional system. However, often language input triggers attentional shifts towards objects even though they are not part of the goal or task set. Such automatic looks towards matching objects have been observed in many visual world studies (Huettig and Altmann 2005; Singh and Mishra 2015; Yee and Sidvey 2006). Bilinguals when they listen to words to one language immediately look towards objects whose names match the translation equivalents in the non-used language (Mishra and Singh 2014). In these studies participants are not given explicit tasks to perform except the instruction that they should listen and look. Huettig et al. (2011) compared language-mediated visual world studies with visual search tasks and proposed that looks towards objects with the onset of speech suggest a tight coupling of working memory, visual information and linguistic information. Even though the perceiver does not have any tasks to perform, language input drives the attentional mechanism towards matching visual objects automatically. This automatic attention shift (Mishra et al. 2012) can only arise when linguistic labels of objects and their visual forms are simultaneously activated. Quick retrieval from long-term memory influences the oculomotor system which triggers saccades. These studies thus provide evidence for bottom-up capture of attention during language-mediated visual search as opposed to task-based studies. Very similar to visual search studies where distractors matching in feature with the targets capture attention (Wolfe 1994). While Huettig et al. (2011) invoke working memory as a major player in language-mediated visual search, they do not talk about attention. From one point of view, keeping things in working memory is equivalent to attending to them (Engle 2002). If language-mediated eye movements are automatic, then they are not strategic. Thus, they do not require selective attention as such, except for cases where the participant has an explicit task to perform.

In order to understand how language could move attention such powerfully, it is important to know what language does. As we saw in Chap. 3, theorists in cognitive linguistics (Talmy 2000) and also in interactive models of language (Jackendoff 2012a, b), have taken language to have its own attentional properties. Language as a symbolic system narrows down attention to the relevant objects and allows us to perceive them in a certain manner (Talmy 2000). However, visual world researchers do not subscribe to such theories of language and attention but look at language-triggered saccades arising from memory representations (Spivey and Dale 2004). When one retrieves the linguistic label of an object one also retrieves its location and other features. That is why one can see eye movements towards blank spaces when object names are spoken even when there is no object (Spivey and Geng 2001). Thus, visuospatial attention is linked to object representation in the long-term memory which then controls eye movements during search. These studies in their various diversities have provided solid evidence for the functional links that exist between the attentional, visual and the linguistic systems.

Finally, as argued before, language-mediated visual search data also provide evidence for the cognitive penetration of vision itself (Pylyshyn 1999). Language input has been shown to drive eye movements within 100 ms at least (Altmann 2011). Such eye movements are triggered from conceptual knowledge of the language user and are automatic to the most extent. Experimental results from cognitive psychology suggest visual perception taking place in two distinct temporal time windows. It takes about 300 ms for the visual system to grasp the identity of the object and during this time participants are not able to report their experience (Dehaene et al. 2006). Attentional processing begins after 300 ms and this is indicated by the p300 ERP component. However, recent evidence suggests that viewers are in a position to consciously perceive and report very briefly presented stimuli after 22 ms of stimuli presentation (Genetti et al. 2010). Under normal circumstances it takes around 200 ms for saccades to be programmed (Hutton 2008). Therefore, if language-mediated eye movements are occurring within 100–200 ms range after spoken word onset, these must be guided by top-down knowledge. Top-down and contextual information have been shown to influence visual perception at very early time scales and can affect eye movements (Gilbert and Li 2013). Thus, it is likely that language can influence preconscious vision and saccadic eye movements. Given the fact that many visual world studies have unrelated distractors apart from the target, such saccades demonstrate the cognitive system's ability to resolve competition dynamically. Unfortunately, not much advances have been made on the neural mechanisms of such fast, language-mediated eye movements. In any case, data from visual world studies show the non-modular architecture of the cognitive system and also the fact that visual processing is influenced by linguistic information at the earliest stages.

10.3 Language, Attention, Space and Culture

Most of my attention so far has been on the diagnosis of attention as a mechanism which agents use for goal-directed linguistic action. It has been shown that the structure of language has in it attentional properties. That is because words and other elements of language are used to draw the listener's attention towards discourse relevant objects or even towards specific spatial locations. Language as a cognitive system is embedded in the environment and culture. Chapters 6 and 9 attempted to reflect this view by taking examples from studies that have explored how language affects cognition. Influential studies by Steven Levinson and colleagues on lesser studied languages have revealed how particular language speakers uniquely map space (Burenhult and Levinson 2008). Language itself has its own pointing devices in the form of deictic markers that channelizes attention towards particular entities in space. Words in the language motivate the listener to shift attention towards different objects and events. For example, when listening to a noun, one has to look for an object and an event in the case of a verb (Bloom and Keil 2001). Other evidence in this genre of research also suggests that language affects how one perceives temporality (Klein et al. 2001) and colour (Winawer et al. 2007). Thus, a language user's basic perceptual categories are influenced by language structure to an extent. It is debatable to what extent such findings are universal or if every language leads its speakers to react to stimuli differently. Therefore, while we could be satisfied with some laboratory-based experimental results that show language and attention interaction, it is impossible to generate a holistic understanding unless one considers cultural attributes and other habits of language users. Attention being a focusing device must therefore be under cultural control (Chua et al. 2005). Nisbet and colleagues have shown that East Asian and Western populations differ with regard to the styles of allocating attention, i.e. background versus foreground. Huettig et al. (2010) showed in a visual world study that language-specific attributes influence gaze behaviour during conceptual processing. Others have claimed that basic categories in particular languages, such as classifiers, influence cognition (Saalbach and Imai 2012). It is important to note that language is just one of the many ways through which cultural forces manifest their influence on cognition. Nevertheless, for a complete understanding of language–attention interaction, one should include not just the agent and his/her intentional goals, but also the cultural background as well as the context in which the linguistic performance happens.

Spatial language, we saw, directly refers to locations of objects in space. Jackendoff (2012a, b) suggests that language works like a transduction system for vision. Examples from fictive motion processing (Chap. 9) also suggest that language structure affects visual perception differently. How does spatial language influence the orientation of attention? Coventry et al. (2010) show that listeners mentally stimulate space and locations as they listen to spatial descriptions. Importantly, spatial language drives attention to precise locations. For example, a spatial preposition like 'above' is mapped very precisely in space and is used as such. This suggests that language codes attention independent of the human agent in some cases. As discussed in Chap. 3, in cognitive linguistics, Talmy (2000) also has proposed that languages have inbuilt mechanism to narrow the scope of attention. That is, a linguistic description only selects some and not others for attention processing. Entities mentioned first attract attention more than entities mentioned later in a sentence (Gordon et al. 1993). Thus, it would be wrong to assume that speakers always control attention elements in speech using the top-down goals. Language structure itself can attract attention and bias processing in an exogenous manner in listeners. Attention-grabbing devices are used by speakers as a part of conversational structure. A close analysis of linguistic structure therefore can reveal how their uses in language can affect attention.

Language and attention interactions happen in the co-presence of agents and objects in space. Human cultures differ with regard to mechanisms of shared attention, joint attention and also the way hands and eyes are oriented as a function of communicating intent. It is important to understand the channelization of attention during human linguistic interaction. The interaction of language and attention in linguistic communication is grounded in the nuances of pragmatics (Levinson 1991). Speakers in a communicative context use multi-modal signals while using linguistic structures (Levinson and Holler 2014). These signals could include eye gaze of speakers, hand movements, gestures and also facial expressions. This might explain why language even orients attention in space towards objects in the first place, as we saw in Chap. 7. Because evolutionarily, language performance has always been enacted among its users embedded in a rich canopy of multi-modal and multi-sensory signals. Meaning in language emerges through such multi-modal and multi-sensory interaction (Gardenfors 2014). Attention as a core cognitive mechanism binds different perceptions that arise during such interactions (Arbib 2013). While I have emphasized the link between language and attention from a goal-directed point of view, it is important to note that such goals arise in users of language as communicative intent. Psycholinguistic tasks that often measure single responses from isolated participants overlook this aspect. If language is there to share mental states using symbolic codes, then this is possible through joint attention and also through an understanding of others' intentions (Dominey 2013). While the mode of attention, i.e. exogenous or endogenous, may be dependent on the agent and stimulus, with regard to language use, shared intentions influence how such a mode is selected.

10.4 Attention, Individual Differences and Language Use

Whatever may be the conceptualization of the interaction of attention and language, individual cognitive profiles affect such interactions. In Chap. 6 we surveyed the literature which suggests that depending on how one uses a language or languages, executive and attentional skills may be boosted. Such evidence has come from bilinguals who negotiate two languages for communication. Constant shifting between languages and inhibiting the discourse inappropriate language can have positive influence on core systems such as selective attention (Friesen et al. 2014), inhibitory control (Green 1998), and also performance monitoring (Singh and Mishra 2014). Highly proficient bilinguals are better at monitoring goal-directed action (Singh and Mishra 2013) and conflict resolution (Singh and Mishra 2012) compared to low-proficient bilinguals. It still remains unclear what precise mechanisms lead to such cognitive advantages in bilinguals (see Valian 2015). Since bilingualism is highly culture specific and individual and depends on several other linguistic and non-linguistic variables, many studies have not observed any advantages in bilinguals compared to monolinguals (Hilchey and Klein 2011). Similarly, studies with illiterate adults show that literacy boosts selective attention (Olivers et al. 2014). Illiterate adults also are poor in predicting upcoming linguistic words during simultaneous processing of language and visual information (Mishra et al. 2012). Brain imaging studies have also shown distinct neural effects of literacy (Huettig and Mishra 2015 for review). This evidence suggests that depending on the user's cognitive profile, the language–attention interaction is modulated.

Why should learning to read and write or using two languages affect attention? The Chomskyan concept of performance is useful in understanding this mechanism here (Chomsky 1988). Chomsky considered performance to be what the native speaker does while he or she exploits the abstract and unconscious linguistic knowledge. It is likely that performance should then vary with individual cognitive profiles. The data from bilinguals and illiterates suggest that as one becomes skilful in the use of language, there is some cognitive benefit which includes sharpening of the attentional system. Illiterates, for example, have been found to be poor in visual target search tasks (Olivers et al. 2014). Not just language, recent evidence suggests that playing demanding video games can also boost attention (Green and Bavelier 2003). However, there is a difference between video game playing and constant use of language in challenging environments. Language is an intentional system and has links with propositional thinking (Mishra 2014). Language is a skill where the speaker anticipates the intentional states of the listener and dynamically modifies ongoing cognitive activity. Therefore, skilful use of language is much more an intellectual activity than playing any game. Thus, language as a cognitive skill used by the agent strengthens the attention and other executive control systems.

Focusing on the 'performance' aspect of language is important for theorizing about the attention–language interaction, as has been exemplified above. Most psycholinguistic studies described in this book measure performance of language users in various tasks. Such performance is then used to explain the competence of users. An important argument presented in this book relates to the fact that language performance is dependent upon several important cognitive capacities of the brain. Bilinguals who have better inhibitory control have been shown to be also faster in resolving dual language conflict during listening. As we already covered, children who are better in attention tasks also show superior performance in language tasks. These correlations need to be causal, yet they certainly suggest that the use of language depends on other important cognitive structures. It should also be noted that use of language is highly contextualized. For example, bilinguals who live in an L1 dominant country show different performance compared to those who live in an L2 dominant country (Sundermann and Priya 2012). With regard to acquisition of reading and writing, scripts and other such variables affect performance. Therefore, while individual differences are important in the explanation of language behaviour and attention mechanisms, we should also consider the context where such differences manifest. Since most studies have been done with highly fluent college students, it is often difficult to extrapolate results to a wider and diversified population.

10.5 Final Remarks

It is impossible to remain aloof from the current discussions on the very nature of attention that one sees in cognitive psychology (Anderson 2011), philosophy of psychology (Mole 2011) and related areas. While there is more of less agreement on the nature of language, there seems to be little agreement on the nature of attention and what it constitutes of. Anderson (2011) offers a critique of contemporary research on attention in both psychology and neuroscience and raises several interesting questions. Since this book explicitly or implicitly assumed attention as a mechanism or as causal factor in cognitive processes, it is important to evaluate if such evaluations are correct. Mole (2011) would suggest that this is not the case but some cognitive processes are either attentional or not. In this view, we can say that some language processes are attentional while at other times they are not. Therefore, investigators should explore when a language process is attentional and when it is not. For example, object naming under specific experimental constrains is attentional but speaking as such may not be. Similarly, reading carefully for later comprehension is attentional but otherwise it may not be. Anderson (2011) observes that currently used dichotomies, i.e. top-down versus bottom-up, local versus global, exogenous versus endogenous and such others are incorrect. Attention is not a thing, nor a causal agent, which makes events happen in a particular manner. He prefers attention to be an effect of a certain mode of processing governed by the current goals of the organism. For example, according to Anderson, it is incorrect to say that because of attention to certain objects, processing was facilitated. However, we have seen that most psycholinguists have used this causal view of attention in their analysis of attention–language interaction. As an alternative to the use of the causal view and use of spurious dichotomies, Anderson proposes a Bayesian approach. This approach assumes generalizations and predictions that users use (heuristics) in performing certain tasks. For example, Anderson cites visual search tasks, where faster responses are generated towards stimuli that appear at highly predictable locations. This analysis does not require a concept like attention to be invoked. While the line of reasoning proposed by Anderson is certainly open to debate, there seems to be some truth in it. A very commonly observable phenomena is different researchers use the term attention to denote a wide range of observations. Sometimes, these processes could be very different from one another. Second, as Anderson rightly observes, depending on the dichotomies assumed, researchers design experiments and interpret results. In spite of this, the mechanism of attention has remained elusive and misunderstood. In contrast to this, Mole (2011) offers an adverbial view of attention. As I have pointed out before, for Mole, agents are said to be attentive when they have complete knowledge of their tasks and goals. It is a unifying mechanism for the mind which allows agents to pursue their goals. While Mole does not seem to have a problem with the various dichotomies used in the explanation of attention, he too seems to believe that attention is not a cause but an effect. That is, similar to the conceptualization offered by Anderson, one can only say if a certain mode of behaviour is attentive or not. One cannot say because of attention one is seeing certain particular behavioural output. It is possible that Anderson and Mole's views are too metaphysical and they do not consider all kinds of empirical facts before dismissing theories of attention or advancing their own. Anderson does not consider attention research beyond visual search and Mole similarly argues around facts related to visual cognition. However, it is also important to be aware of these changing formulations in relevant fields when one attempts to link attention with language and behaviour at large.

How can one say that certain linguistic processes are attention while others are not? How does an attention state set in during the task? Is this state influenced by the goals of the agent or stimuli features? These questions are vital while we consider attention–language interaction in different situations. We saw in chapters on sentence processing and reading that these processes are automatic, including language-driven eye movements. Similarly, activation of translations during spoken word listening in bilinguals is automatic (Mishra and Singh 2014). It is also a fact that detection of grammatical or other errors during sentence processing is automatic. Therefore, when a cognitive process is automatic, no attentional effect is to be expected (Logan 1980). Since most linguistic processes appear to be automatic, it is difficult to say when attention is involved. However, it has also been argued that automatic processes need not be thought of as not needing any attention. Many so-called automatic processes are initiated by intentions of the agent (Bargh 1994). For example, sentence generation or listening may look automatic and effortless, but its origin lies in the generation of some goal or intention. Depending on such goals, attention processes set in. Therefore, the locus of the attention–language interaction should lie with intentions of the agent, which in turn modulate attentional resources. This falls in line with what Ristic and Landry (2015) have suggested with regard to attention in complex behaviour. Ristic and Landry propose that in a complex behaviour as in the case of locating objects with direction cues, both top-down goals and bottom-up signals are used simultaneously for optimal performance. Attention, thus, is used to maintain such task goals and selective processing. Most commonly, task goals are given to participants via instructions in a traditional psycholinguistic experiment. I agree with Anderson (2011) that whether attention is involved in a task should be evaluated by looking at what the subject is trying to achieve. There is a pitfall if experimenters think that their experimental manipulation is sufficient to engage attention in a task in a causal manner.

From the above discussion it is a likely conclusion that language behaviour is controlled by the goals of the participants and therefore such goals explain if the particular language behaviour has been attentional. Thus, some have viewed language as a reflex system. In the domain of syntax processing, Pulvermüller et al. (2008) demonstrated that even when subjects are distracted or are given additional attention tasks, neural responses to syntactic violation are observed within 150 ms of onset of the critical word. This suggests an extremely automatic process which occurs without attentional engagement. However, with regard to semantic violations, such automaticity is not seen. Spoken word processing has also been taken to be automatic (Salverda and Altmann 2011; Singh and Mishra 2014). While many studies have shown working memory load influencing syntactic parsing (Caplan and Waters 1999), it is not clear what its effect is on semantic processing. On the contrary, semantics in visual word recognition as in the Stroop task are viewed as automatic (Cohen et al. 1990). We have already see in Chap. 9 that words drive spatial attention unintentionally. While it is apparent that language use is a reflex, it is not a suitable argument to suggest that language processing requires no voluntary control. It seems, then, that some levels of language are like a reflex and others require effortful processing. Therefore, it would not suffice to consider language as a whole but rather to consider different dimensions within language for the construction of a theory of language–attention interaction. One can get out of this circularity with the following assumptions. Although language processing is unconscious and seems effortless, language users set up top-down task schemas that call for attention at a macro level. Without this kind of overall control, speakers would lose control of their intentions and language production may suffer.

It is strange and at the same time fascinating to see that since James till now, there has been no unanimity on the nature of attention. Louis Armstrong once said when asked what jazz is, that ‘Do not ask what it is. You will never know’ (cited in Jackendoff 2012a). Thus, the word attention is a label for a collection of myriad mental processes that have different functions—just as language is a label to a collection of meaningless signs that humans use to connect sounds with meanings and express intentions and thoughts. The same goes for words like emotion, memory and cognition. Attention is what attention does. We have seen that it is the mechanism that allows us to select, to maintain goal-directed action and to focus on relevant things. Its absence can cause vast arrays of cognitive and intellectual deficits as is evident from studies on developmental and acquired disorders. One might have issues with dichotomies such as top-down versus bottom-up but for agents these modes of acting are real. Cues do capture our attention unintentionally and we do orient towards task relevant stimuli wilfully. It is evident from current discussions on the nature of consciousness that attention mechanisms could play a major role in it. Thus, terminological impurity should not hinder research progress as long as we can experimentally isolate the cause from the effect in the study of attention and other intellectual functions. With regard to the question of its connections with language, as we have reviewed, it is subject to several constraints.

These constraints include identifying the goals and intentions of the language user in different tasks, and further, isolating aspects of such tasks where attention seems to be involved compared to others that seem automatic. As it has been described in this book at several places, contemporary investigations of attention basis of language or linguistic basis of attention are in their infancy. There are disagreements related to the question of autonomy of linguistic processing with regard to linguistic sublevels. Experimental procedures remain indifferent to the intentions of the participants but emphasize performance to stimulus. What is taken as top-down is merely the participant's willingness to cooperate with the task instructions. A more valid and ecological approach may include studying attentional mechanisms in real-life situations.

The central question to solve is how language users exert control over linguistic action in real-life communication. In answering this question, one would look for specific aspects of such behaviour that include orienting, selecting, remaining engaged, disengaging at will, and stopping the action when needed. If linguistic stimuli unintentionally catch our attention, that is because evolutionarily such signs have been perceived as useful stimuli. It is a fact that a meaningless noise also catches attention. However, with meaningful stimuli, agents can exert sustained action. Therefore, future investigators should move away from questions that ask if linguistic processes require attention. The emphasis should be rather on understanding how attention is involved in our goal-directed linguistic cognition. If intentionality of the mind reflects its 'aboutness' then language more or less manifests this. This book has presented evidence which shows that attention plays a critical role in this important aspect of human cognition. Future studies should also move away from studying isolated subjects doing unrealistic tasks to studies of group level shared behaviour. After all, language operates within groups of speakers and listeners who jointly attend to signs. How does such consensual linguistic behaviour emerge? These questions would call for developments in techniques and methodologies. And more so, in the right conceptual structures.

References

Altmann, G. T. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta Psychologica, 137*(2), 190–200.

Altmann, G., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502–518.

Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology, 2*, 246.

Arbib, M. A. (Ed.). (2013). *Language, music, and the brain: A mysterious relationship*. Cambridge, MA: MIT Press.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437–443.

Baldauf, D., & Deubel, H. (2010). Attentional landscapes in reaching and grasping. *Vision Research, 50*(11), 999–1013.

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues.

Bekkering, H., & Neggers, S. F. (2002). Visual search is modulated by action intentions. *Psychological Science, 13*(4), 370–374.

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition, 96*(2), B45–B54.

Bloom, P., & Keil, F. C. (2001). Thinking through language. *Mind and Language, 16*(4), 351–367.

Burenhult, N., & Levinson, S. C. (2008). Language and landscape: A cross-linguistic perspective. *Language Sciences, 30*(2), 135–150.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(01), 77–94.

Castles, A., & Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind and Language, 29*(3), 270–285.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30*(3), 687–696.

Chomsky, N. (1988). *Aspects of the theory of syntax* (Vol. 11). Cambridge: MIT press.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the United States of America, 102*(35), 12629–12633.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101.

Coco, M. I., Malcolm, G. L., & Keller, F. (2014). The interplay of bottom-up and top-down mechanisms in visual guidance during object naming. *The Quarterly Journal of Experimental Psychology, 67*(6), 1096–1120.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*(3), 332–361.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Coventry, K. R., Lynott, D., Cangelosi, A., Monrouxe, L., Joyce, D., & Richardson, D. C. (2010). Spatial language, visual attention, and perceptual simulation. *Brain and Language, 112*(3), 202–213.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277–299.

Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin and Review, 12*(3), 453–459.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204–211.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Dhooge, E., & Hartsuiker, R. J. (2012). Lexical selection and verbal self-monitoring: Effects of lexicality, context, and time pressure in picture-word interference. *Journal of Memory and Language, 66*(1), 163–176.

Dispaldro, M., Leonard, L. B., Corradi, N., Ruffino, M., Bronte, T., & Facoetti, A. (2013). Visual attentional engagement deficits in children with specific language impairment and their role in real-time language processing. *Cortex, 49*(8), 2126–2139.

Dominey, P. F. (2013). The tip of the language iceberg. *Language and Cognition, 5*(2–3), 177–188.

Eden, G. F., Stein, J. F., Wood, H. M., & Wood, F. B. (1994). Differences in eye movements and reading problems in dyslexic and normal children. *Vision Research, 34*(10), 1345–1358.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777–813.

Engelhardt, P. E., Nigg, J. T., & Ferreira, F. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? *Actapsychologica, 144*(2), 424–432.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language, 69* (3), 165–182.

Finkbeiner, M., & Caramazza, A. (2006). Lexical selection is not a competitive process: A reply to La Heij et al. (2006). *Cortex*, *42*(7), 1032–1036.

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 52* (4), 915–929.

Friesen, D. C., Latman, V., Calvo, A., & Bialystok, E. (2014). Attention during visual search: The benefit of bilingualism. *International Journal of Bilingualism*, 1367006914534331.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge: MIT Press.

Gathercole, S. E., & Baddeley, A. D. (2014). *Working memory and language processing*. Hove: Psychology Press.

Genetti, M., Britz, J., Michel, C. M., & Pegna, A. J. (2010). An electrophysiological study of conscious visual perception using progressively degraded stimuli. *Journal of Vision, 10*(14), 10.

Ghahghaei, S., Linnell, K. J., Fischer, M. H., Dubey, A., & Davis, R. (2013). Effects of load on the time course of attentional engagement, disengagement, and orienting in reading. *The Quarterly Journal of Experimental Psychology, 66*(3), 453–470.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience, 14*(5), 350–363.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*(3), 311–347.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(02), 67–81.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

Grossman, M., Carvell, S., Stern, M. B., Gollomp, S., & Hurtig, H. I. (1992a). Sentence comprehension in Parkinson's disease: The role of attention and memory. *Brain and Language, 42*(4), 347–384.

Grossman, M., Crino, P., Reivich, M., Stem, M. B., & Hurtig, H. I. (1992b). Attention and sentence processing deficits in Parkinson's disease: The role of anterior cingulate cortex. *Cerebral Cortex, 2*(6), 513–525.

Guthrie, J. T., & Wigfield, A. (1999). How motivation fits into a science of reading. *Scientific Studies of Reading, 3*(3), 199–205.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences, 9*(6), 290–295.

Hahne, A., & Friederici, A. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11* (2), 194–205.

Hannus, A., Cornelissen, F. W., Lindemann, O., & Bekkering, H. (2005). Selection-for-action in visual search. *Acta Psychologica, 118*(1), 171–191.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin and Review, 18*(4), 625–658.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*(1), 33–56.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23–B32.

Huettig, F., & Altmann, G. T. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*(8), 985–1018.

Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind–a critical examination of theories and evidence. *Language and Linguistics Compass, 8*(10), 401–427.

Huettig, F., Chen, J., Bowerman, M., & Majid, A. (2010). Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture, 10*(1), 39–58.

Huettig, F., Olivers, C. N., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Actapsychologica, 137*(2), 138–150.

Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain and Cognition, 68*(3), 327–340.

Ibbotson, P., Lieven, E. V., & Tomasello, M. (2013). The attention-grammar interface: Eye-gaze cues structural choice in children and adults. *Cognitive Linguistics, 24*(3), 457–481.

Jackendoff, R. (2012a). *A user's guide to thought and meaning*. Oxford, UK: Oxford University Press.

Jackendoff, R. (2012b). Language as a source of evidence for theories of spatial representation. *Perception-London, 41*(9), 1128.

Jongman, S. R., Roelofs, A., & Meyer, A. S. (2015). Sustained attention in language production: An individual differences investigation. *The Quarterly Journal of Experimental Psychology, 68*(4), 710–730.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kamide, Y., Altmann, G., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133–156.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*(2), 153–168.

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage, 13*(4), 646–653.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science, 30*(3), 481–529.

Kurland, J. (2011). The role that attention plays in language processing. *SIG 2 Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, *21*(2), 47–54.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323.

Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42*(1), 1–22.

Levinson, S. C. (1991). Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics, 27*(01), 107–161.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity* (Vol. 5). Cambridge, MA: Cambridge University Press.

Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130302.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12*(4), 523–553.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283–305.

Malpass, D., & Meyer, A. S. (2010). The time course of name retrieval during multiple-object naming: Evidence from extrafoveal-on-foveal effects. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*(2), 523–537.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847.

Mesulam, M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology, 28*(5), 597–613.

Meyer, A. S., Wheeldon, L., van der Meulen, F., & Konopka, A. (2012). Effects of speech rate and practice on the allocation of visual attention in multiple object naming. *Frontiers in Psychology, 3*, 39.

Mishra, R. K. (2009). Interaction of language and visual attention: Evidence from production and comprehension. *Progress in Brain Research, 176*, 277–292.

Mishra, R. K. (2014). Listening through the native tongue: A review essay on Cutler's Native listening: Language experience and the recognition of spoken words. *Philosophical Psychology*, 1–15. (ahead-of-print).

Mishra, R. K., & Marmolejo-Ramos, F. (2010). On the mental representations originating during the interaction between language and vision. *Cognitive Processing, 11*(4), 295–305.

Mishra, R. K., & Singh, N. (2014). Language non-selective activation of orthography during spoken word processing in Hindi-English sequential bilinguals: An eye tracking visual world study. *Reading and Writing, 27*(1), 129–151.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Mishra, R. K., Olivers, C. N., & Huettig, F. (2013). Spoken language and the decision to move the eyes-To what extent are language-mediated eye movements automatic? *Progress in Brain Research, 202*, 135.

Mole, C. (2011). *Attention is cognitive unison: An essay in philosophical psychology*. Oxford University Press.

Montani, V., Facoetti, A., & Zorzi, M. (2014). Spatial attention in written word perception. *Frontiers in Human Neuroscience, 8*, 42.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: The case of English and Russian. *The Quarterly Journal of Experimental Psychology, 66*(8), 1601–1619.

Novick, J. M., Hussey, E., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. F. (2014). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience, 29*(2), 186–217.

Nozari, N., & Dell, G. S. (2012). Feature migration in time: Reflection of selective attention on speech errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(4), 1084–1090.

Nozari, N., & Thompson-Schill, S. L. (2013). More attention when speaking: Does it help or does it hurt? *Neuropsychologia, 51*(13), 2770–2780.

Olivers, C. N. L., Huettig, F., Singh, J. P., & Mishra, R. K. (2014). The influence of literacy on visual search. *Visual Cognition, 22*(1), 74–101.

Phelps, E. A., Hyder, F., Blamire, A. M., & Shulman, R. G. (1997). FMRI of the prefrontal cortex during overt verbal fluency. *NeuroReport, 8*(2), 561–565.

Pulvermüller, F., Shtyrov, Y., Hasting, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language, 104*(3), 244–253.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(03), 341–365.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using EZ Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye–mind link. *Psychologicalreview, 119*(1), 155–185.

Reingold, E. M., & Stampe, D. M. (2000). *Saccadic inhibition and gaze contingent research paradigms* (pp. 119–145). Amsterdam: Elsevier.

Riès, S., Janssen, N., Dufau, S., Alario, F. X., & Burle, B. (2011). General-purpose monitoring during speech production. *Journal of Cognitive Neuroscience*, *23*(6), 1419–1436.

Ristic, J., & Landry, M. (2015). Combining attention: A novel way of conceptualizing the links between attention, sensory processing, and behavior. *Attention, Perception, & Psychophysics*, *77*(1), 36–49.

Ruffino, M., Gori, S., Boccardi, D., Molteni, M., & Facoetti, A. (2014). Spatial and temporal attention in developmental dyslexia. *Frontiers in Human Neuroscience, 8*, 331.

Saalbach, H., & Imai, M. (2012). The relation between linguistic categories and cognition: The case of numeral classifiers. *Language and Cognitive Processes, 27*(3), 381–428.

Salvaris, M., & Haggard, P. (2014). Decoding intention at sensorimotor timescales. *PLoS ONE, 9* (2), e85100.

Salverda, A. P., & Altmann, G. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1122–1133.

Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica, 137*(2), 172–180.

Searle, J. R. (1997). Consciousness & the philosophers. *New York review of Books*, 44(4).

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*(11), 489–493.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *The Quarterly Journal of Experimental Psychology, 65*(10), 1927–1944.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15*(04), 771–781.

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: evidence from Hindi-English bilinguals. *Frontiers in Psychology*, 4.

Singh, N., & Mishra, R. K. (2014). The modulatory role of second language proficiency on performance monitoring: Evidence from a saccadic countermanding task in high and low proficient bilinguals. *Frontiers in Psychology, 5*, 1481.

Singh, N., & Mishra, R. K. (2015). Unintentional activation of translation equivalents in bilinguals leads to attention capture in a cross-modal visual task. *PLoS ONE, 10*(3), e0120131.

Smith, F. (2012). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Routledge.

Spivey, M. (2007). *The continuity of mind*. Oxford, UK: Oxford University Press.

Spivey, M. J., & Dale, R. (2004). On the continuity of mind: Toward a dynamical account of cognition. *Psychology of Learning and Motivation, 45*, 87–142.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*(4), 235–241.

Strijkers, K., Holcomb, P. J., & Costa, A. (2011a). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language, 65*(4), 345–362.

Strijkers, K., Yum, Y. N., Grainger, J., & Holcomb, P. J. (2011b). Early goal-directed top-down influences in the production of speech. *Frontiers in Psychology, 2*, 371.

Sunderman, G. L., & Priya, K. (2012). Translation recognition in highly proficient Hindi-English bilinguals: The influence of different scripts but connectable phonologies. *Language and Cognitive Processes, 27*(9), 1265–1285.

Talmy, L. (2000). *Toward a cognitive semantics*: *Concept structuring systems* (Vol. 1). Cambridge: The MIT Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(02), 183–203.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(01), 3–24.

Visser, T. A. (2014). Evidence for deficits in the temporal attention span of poor readers. *PLoS ONE, 9*(3), e91278.

Wharton, T. (2013). 13 Linguistic action theories of communication. *Theories and Models of Communication, 1*, 241.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*(19), 7780–7785.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin and Review, 1*(2), 202–238.

Wykowska, A., Schubö, A., & Hommel, B. (2009). How you move is what you see: Action planning biases selection in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1755–1769.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*(1), 1.