

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

The integration of language with perceptual systems

Permalink

<https://escholarship.org/uc/item/290191wh>

Author

Collins, Jennifer Blair

Publication Date

2011

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

THE INTEGRATION OF LANGUAGE WITH PERCEPTUAL SYSTEMS

A Dissertation submitted in partial satisfaction of the
Requirements for the Degree of Doctor of Philosophy

in

Cognitive Science

by

Jennifer Blair Collins

Committee in charge:

Professor Seana Coulson, Chair
Professor Gedeon Deák
Professor Jeffrey Elman
Professor Karen Emmorey
Professor Victor Ferreira
Professor Judy Reilly

2011

Copyright

Jennifer Blair Collins, 2011

All rights reserved.

The Dissertation of Jennifer Blair Collins is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2011

DEDICATION

To my lovely family

TABLE OF CONTENTS

Signature Page	iii
Dedication.....	iv
Table of Contents	v
List of Figures.....	vii
List of Tables	viii
Acknowledgments	ix
Vita	xi
Abstract of the Dissertation	xiv
Chapter 1 Introduction.....	1
What is embodied cognition?	2
Perceptual symbol systems.....	3
Immersed Experiencer Framework	12
Summary.....	24
References	25
Chapter 2 Modality Switching in a Property Verification Task.....	29
Introduction	30
Materials and Methods	32
Results	33
Discussion.....	35
References	38
Acknowledgment.....	39
Chapter 3 An ERP study of the impact of modality on the conceptual modality switch effect	40
Abstract.....	41
Introduction	41
Material and Methods.....	49
Results	52
Discussion.....	56
References	65
Acknowledgment.....	67
Chapter 4 Processing specifications for the influence of vertical language on visual motion perception	68
Abstract.....	69
Introduction	70
Experiment 1	79
Experiment 2	88

Experiment 3	92
Acknowledgment.....	120
References	120
Chapter 5 The impact of pictorial and verbal motion content on representational momentum	124
Abstract.....	125
Introduction	125
Experiment 1	135
Experiment 2	145
Experiment 3	155
General Discussion.....	163
Acknowledgment.....	168
References	168
Chapter 6 Conclusion	171
Review of findings	173
Perception is part of the meaning comprehension system.....	181
Clarifying/constraining issues	183
Summary.....	191
References	193

LIST OF FIGURES

Figure 2.1. Participants saw a concept followed by a property after which they made their judgment.....	33
Figure 2.2. Relative placement of 29 scalp electrodes.	33
Figure 2.3. The N400 elicited by visual property verification targets in the switch and no-switch conditions.....	34
Figure 2.4. Topography of the switch effect for visual property verifications.....	34
Figure 2.5. The late positive complex to auditory targets in the switch relative to the no-switch conditions.....	35
Figure 2.6. Topography of the switch effect for auditory targets.....	35
Figure 2.7. Current source density maps of responses to visual and auditory properties.	35
Figure 3.1. Trial arrangement for the conceptual modality switch..	52
Figure 3.2. Tactile N400 effect.....	54
Figure 3.3. Topography of the tactile N400 effect..	55
Figure 3.4. Conceptual modality switch effect.....	56
Figure 4.1. The timing of a single trial from all three experiments.....	81
Figure 4.2. Behavioral results for Experiment 1..	85
Figure 4.3. Behavioral results for Experiment 2..	90
Figure 4.4. Reaction time results split by sensitivity.....	100
Figure 4.5. N1 plots for high sensitivity participants..	103
Figure 4.6. Participants with high sensitivity scores showed a positive ERP deflection for upward verbs at a subset of 6 posterior electrodes in the 400-600 ms interval.....	105
Figure 4.7. The coherence effect for participants with high sensitivity..	107
Figure 4.8. ERPs from the low-sensitivity group seeing coherently-moving dots... ..	108
Figure 5.1. Sample images.	137
Figure 5.2. The timing of stimulus presentation in Experiment 1.....	139
Figure 5.3. The effect of probe position on sensitivity.....	142
Figure 5.4. Effect of image prime on sensitivity to the probe.	143
Figure 5.5. The timing of stimulus presentation in Experiments 2 and 3.....	149
Figure 5.6. The interaction of language context and triangle direction.....	153
Figure 5.7. D' values for the direction by shift interaction..	159
Figure 5.8. Interaction of vertical language and vertical motion.....	160

LIST OF TABLES

Table 5.1. Sentences used in Experiments 2 and 3..... 147

ACKNOWLEDGEMENTS

I would first like to thank Seana Coulson for her guidance, patience, extensive editing, opinions, advice, and encouragement throughout my graduate school career. I have learned a lot from her and am very grateful. I would like to thank Judy Reilly for her mentorship and advice during the years I have known her. I would also like to thank Gedeon Deák, Jeff Elman, Karen Emmorey and Vic Ferreira for their time and interest in my work.

I would like to acknowledge funding from the University of California at San Diego Dean's Fellowship, University of California at San Diego Chancellor's Fellowship, and the Center for Research in Language NIH training grant. The resources available through CRL have also been invaluable in supporting my research.

I have also received professional and personal support that has enabled me to complete the work presented here. David Brang, Tristan Davenport, Marguerite McQuire, and Ying Wu from the Brain and Cognition Lab have been excellent labmates. I would not have been able to do all the experiments presented here without help from my research assistants. I owe many, many thanks to: Lauren Cardoso, Jordan Conway, Nafees Hamid, Anita Leung, Jenny Lyons, Rubén Moreno and Vanessa Williams. I would like to thank Rafael Núñez and Kensy Cooperrider whose input has helped to shape my research. I appreciate the guidance and assistance of Diane Pecher as well as her collaboration and that of René Zeelenberg on the modality switch projects.

I am thankful for the camaraderie and advice of my classmates: Nick Butko, Marguerite McQuire, Adrienne Moore, Nathaniel Smith, and Adam Tierney. Particu-

larly Adrienne and Marguerite have been wonderful friends with whom I have been lucky to share this journey. Finally, I would like to thank Tona Rodriguez-Nikl for his motivation, patience, appreciation, typesetting skills, programming help and love.

Chapter 2, in full, is a reprint of the material as it appears in *Frontiers in Psychology*; Jennifer Collins, Diane Pecher, René Zeelenberg and Seana Coulson, 2011. The dissertation author was the primary investigator and author of this paper.

Chapter 3, will be submitted for publication: Jennifer Collins, Diane Pecher, René Zeelenberg, Seana Coulson, “An ERP study of the impact of modality on the conceptual modality switch effect”. The dissertation author was the primary investigator and author of this paper.

Chapter 4, will be submitted for publication: Jennifer Collins, Seana Coulson, “Processing specifications for the influence of vertical language on visual motion perception”. The dissertation author was the primary investigator and author of this paper.

Chapter 5, will be submitted for publication: Jennifer Collins and Seana Coulson, “The impact of pictorial and verbal motion content on representational momentum”. The dissertation author was the primary investigator and author of this paper.

VITA

EDUCATION

- Ph.D., Cognitive Science University of California, San Diego, 2011
DISSERTATION: “The integration of language meaning with perceptual systems”
ADVISOR: Seana Coulson
- B.A., Cognitive Science University of California, Berkeley, 2003
B.A., Mathematics

PUBLICATIONS

Journal Publications

- Collins, J., Pecher, D., Zeelenberg, R., & Coulson, S. Modality switching in a property verification task: An ERP study of what happens when candles flicker after high-heels click. *Frontiers in Psychology*, 2 (10) 1-10.
- Teuscher, U., McQuire, M., Collins, J., & Coulson, S. (2008). Congruity effects in time and space: Behavioral and ERP measures. *Cognitive Science*, 32 (3) 563-578.

Journal Publications in Preparation

- Collins, J. & Coulson, S. (in preparation). The impact of pictorial and verbal motion content on representational momentum.
- Collins, J. & Coulson, S. (in preparation). Experimental design modulates language-perception interactions.

Conference Presentations

- Collins, J. & Coulson, S. (2011). Experimental design modulates language-perception interactions. Poster presented at the 2011 Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA.
- Collins, J. & Coulson, S. (2010). Representational momentum indexes perceptual activation by vertical and horizontal motion sentences. Poster presented at the 2010 Conceptual Structure, Discourse and Language/Embodied and Situated Language Processing Joint Conference, San Diego, CA.
- Collins, J., Pecher, D., Zeelenberg, R. & Coulson, S. (2010). ERP indices of the conceptual modality switch effect: What happens when leaves rustle after candles flicker? Poster presented at the 2010 Annual Meeting of the Cognitive Neuroscience Society, Montréal, Canada.
- Collins, J. & Coulson, S. (2009). Motion verbs affect the perceptual task of representational momentum. Poster presented at the 2009 Annual Conference on Architectures and Mechanisms for Language Processing, Barcelona, Spain.
- Collins, J., Reilly, J., & Stiles, J. (2007). An investigation of spatial language in preschool-aged children with focal brain injury. Poster presented at the 2007 Bi-

ennial Meeting of the Society for Research in Child Development, Boston, MA.

Coulson, S., Collins, J., & Teuscher U. (2006). ERPs to congruous and incongruous moving stimuli after literal and metaphoric sentence primes. Poster presented at the 2006 Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA.

HONORS & AWARDS

Superior TA Award, Department of Cognitive Science, 2009-2010

Superior TA Award, Department of Cognitive Science, 2008-2009

NIH Training Grant Recipient through the Center for Research in Language, 2007-2008

Chancellor's Fellowship Recipient, 2006-2007

Honorable Mention TA Award, Department of Cognitive Science, 2006-2007

Dean's Fellowship Recipient, 2004-2006

NSF Graduate Research Fellowship Honorable Mention, 2004

PROFESSIONAL MEMBERSHIPS

Cognitive Neuroscience Society

Society for Research in Child Development

TEACHING

Instructor, Oregon State University, Winter 2011/Spring 2011. COURSES: Cognition, Cognition Research

Teaching Associate, UC San Diego, Spring 2010. COURSE: Language Development

Teaching Assistant, UC San Diego, 2004-2010. COURSES: Neurobiology of Cognition; Learning, Memory and Attention; Language and Reasoning; Cognitive Development; Design and Analysis of Experiments; Introduction to Cognitive Science; Cognitive Ethnography

Instructor, UCSD Extension, 2006/2008. COURSES: Introduction to Cognitive Science; Cognitive Neuroscience in the Media

Instructor, Center for Talented Youth, Lancaster, PA, 2005. COURSE: Cognitive Psychology

Head Undergraduate Student Instructor, UC Berkeley, Fall 2003. COURSE: Introduction to Cognitive Science

Undergraduate Student Instructor, UC Berkeley, Spring 2003. COURSE: Introduction to Cognitive Science

SERVICE

Laboratory Mentor, Brain and Cognition Lab, 2007-2010

Alumni of the Month Co-Founder, Cognitive Science Department, 2008-2010

Embodiment and Cognition Seminar Organizer, UC San Diego, 2007-2008
“Wa” (Cognitive Science Brown Bag Series) Organizer, UC San Diego, 2007-2008
Vice President of Academic Affairs, Graduate Student Association, UCSD, 2006-2007
Graduate Student Association Representative, UC San Diego, 2005-2007
Cognitive Science Peer Advisor Co-Founder, UC Berkeley, 2003

ABSTRACT OF THE DISSERTATION

THE INTEGRATION OF LANGUAGE WITH PERCEPTUAL SYSTEMS

by

Jennifer Blair Collins

Doctor of Philosophy in Cognitive Science

University of California, San Diego, 2011

Professor Seana Coulson, Chair

Recent theories on language and concepts suggest that when comprehending the sentence, *the deafening jets soared over the UCSD campus*, we activate perceptual systems that have previously been used for hearing and seeing related events. There is a growing body of research in support of the idea that perceptual features interact with meaning comprehension, but the current research is not sufficient to describe how and why this interaction occurs. The work presented here investigates the processing mechanisms behind the integration of language and perceptual systems. The studies focus first on property terms such as *deafening*, *shiny* and *rough*, that describe experiences in different perceptual modalities, and then on motion verbs such as *soar* and *tumble*.

The results verify that information about perceptual modality and motion direction is available during comprehension. Furthermore, the results suggest that the

stages of processing at which perceptual features of language are accessible and used are variable. In the semantic domain of motion for example, the results show perceptual motion information in language is processed at a high cognitive level and is influenced by context. The conclusions that fall out of the varied results are that language meaning has perceptual components as initially suggested, but also that language meaning might be co-opted by perceptual processes. We must not confuse these two related possibilities when investigating the perceptual nature of meaning.

Chapter 1

Introduction

What is embodied cognition?

Our world is a rich, complex system, to which we easily attribute meaning. Through the use of verbal and written language we represent these meanings and use them to describe, discuss and learn, among other uses. The meanings that we can construct with language have the powerful ability to extend beyond what we experience directly, to describe events that are displaced from us in time and space, and can even represent events we have never experienced. These powerful abilities lead naturally to the question for cognitive science of how the brain and body support the representation and construction of meaning via language.

Determining how language meaning is connected referentially to the world is an issue that has been grappled by philosophers, psychologists and linguists for centuries. Recent instantiations of this problem arise in the wake of successful artificial symbol processing systems that produce intelligent behaviors based on rule-guided manipulation of formal symbols. These symbols, like words, are related to their referents arbitrarily and by convention, not based on resemblance between the symbol and referent or any other quality intrinsic to the symbol. Endlessly referring to other meaningless symbols, these symbol systems do not appear to have the capacity to acquire meanings. This is often referred to as the “grounding problem.”

One proposed solution to the grounding problem is to argue that concepts are grounded in virtue of their experiential basis. The thesis of this type of theory, often referred to as “embodied” or “grounded cognition,” is that understanding meaningful concepts takes place through the reactivation of brain systems that would be used to perceive that concept’s referent (Barsalou, 1999; Zwaan, 2004). With the emphasis on

bodily experiences *as* cognition, the representations underlying meaningful language can be thought of as equivalent to traces of the experiences themselves. For example, an action concept, such as *walking*, activates brain areas used to perceive oneself and others walking as well as to execute walking. This approach to the grounding problem limits the need for associations between a representation and the world.

Below I present two theories of grounded concepts and embodied language processing, Perceptual Symbol Systems and the Immersed Experiencer Framework. I describe behavioral and neural specifications of these perspectives and outline empirical support for each. Finally, I introduce some open issues for each of these theories, along with the approaches that will be pursued in later chapters of this dissertation.

Perceptual symbol systems

The Perceptual Symbol Systems (PSS) hypothesis is one of the most influential accounts of embodied or grounded approaches to meaning. Barsalou (1999) presents a detailed theoretical contribution for how mental symbols are used for concept representation and the related processing of broad activation patterns in perceptual and motor areas (Simmons & Barsalou, 2003). According to Barsalou, schematic representations are built from many experiences through selective attention and are encoded in perceptual areas of the brain used during these experiences. It is the relationship with the neural substrate of experiences (i.e. perceptions and actions) that allows for a natural source of symbol grounding for concepts, words, and meanings. The Perceptual Symbol Systems theory also maintains ties to more classical ideas of meaning representation. In particular, Barsalou recognizes the theoretical power of abstract

symbols that can be combined in systematic ways to create infinite combinations of richer meaning.

The representation of experiences is neither complete, nor exact, but rather is schematic and changes with an individual's experiences. Perceptual symbols are supported by neural patterns in any modality including introspection and proprioception. A symbol can represent meaning insofar as that meaning was experienced. But this view of "experience" can extend to attention directed inward for construal of a previous event, and the derivation of new symbols from existing ones. Broad patterns of activation corresponding to interaction in the world become unconscious, schematic representations that can be accessed for any number of cognitive applications in the future, independent of the initial perceptual state. However, the critical difference between this theory and previous symbol systems is that perceptual symbols are never divorced from their representation in perception and action systems of the brain (Barsalou, 1999).

The perceptual symbol systems theory states perceptual areas are important for representing these schematic symbols with the help of non-perceptual brain areas such as *convergence zones* (coined by Damasio, as discussed by Barsalou, 1999) and *conjunctive neurons* within convergence zones (Simmons & Barsalou, 2003; Barsalou, 2008). Conjunctive neurons are those which activate sets of perceptual brain regions initially co-activated. Cross-modal associations can activate multiple conjunctive neurons from different perceptual modalities resulting in multi-modal representations of events and objects located in higher-level convergence zones. Barsalou emphasizes that conjunctive neurons reactivate perceptual brain regions independently of percep-

tual stimulation and that these representations will never be dissociated or abstracted away from activation in these brain regions. He calls the form of concept representation *simulators* and the process of activating a concept *simulation* (Barsalou, 1999; Simmons & Barsalou, 2003).

This framework of abstracted symbols based on perceptual experiences has several important implications. First, perceptual symbols are abstract enough to be productively combined to make infinite different complex concepts, allowing for hierarchical relationships, type-token mappings, and falsity/negativity to be represented. Secondly, variable embodiment describes the tight dependence of the conceptual system on the embodied experiences. Because people have had different experiences – variable embodiment – their conceptual systems will be different and related to particular contexts. Finally, perceptual symbols can represent ideas that have never been experienced through the senses through states that allow us to focus attention internally such as introspection, selective attention, abstraction and reference to other symbols. This claim is important for all concepts to be based in perceptuo-motor experiences.

Neuroimaging evidence for perceptual symbols

One source of evidence cited by Barsalou for the multimodal, perceptual bases of concepts is based in an extensive cognitive neuroscience literature showing that different conceptual categories are represented in brain areas distinguished by the roles they play in perception and action (Martin, 2007). Animals, faces and human forms activate lateral fusiform gyrus in the ventral temporal lobe important for perceiving faces and making fine visual discriminations between items (Chao, Haxby & Martin,

1999). The distribution of activity for animal concepts is more diffuse than that elicited by faces corresponding to the fact that faces are more homogeneous in their visual form than animals. These patterns are consistent for a variety of tasks, including delayed match to sample, naming, property generation, and passive viewing, suggesting that this organization is truly a basis of the conceptual organization and not specific to how the information is elicited (verbally, visually, etc.) (Chao et al., 1999).

Relative to the lateral fusiform gyrus activation for animals and faces, tool and house concepts consistently elicit more medial fusiform gyrus activity demonstrating separable feature-based representations (Chao et al., 1999). Additional activation is seen for objects also associated with motion perception in posterior superior temporal sulcus (pSTS) for animals and posterior medial temporal gyrus (pMTG) for tools (Beauchamp, Lee, Haxby & Martin, 2002; Chao et al., 1999; Noppeney, Josephs, Kiebel, Friston & Prince, 2005). These areas are also used for perception of motion of the different visual forms – human bodies and tools, respectively (Beauchamp et al., 2002). Tools and other manipulable objects, such as clothes, fruits and vegetables, show extended cortical activation into left ventral premotor cortex (PMv) and intraparietal sulcus (Gerlach et al., 2002). These areas are important for judgments regarding how to use objects and are also modulated by experience with use (Weisberg, van Turrenout & Martin, 2007; James & Gauthier, 2003). Action concepts elicited through verbs classified as typically performed with the feet, hands or mouth (e.g. *kick*, *grab*, *bite*, respectively) result in differentiable brain activation organized somatotopically (Hauk & Pulvermüller, 2004; Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006; Buccino et al, 2001; Buccino et al. 2005).

The study of color concepts offers mixed evidence for the specific claim that concepts recruit precisely the same neural regions as perception. (This is not necessarily a requirement in Barsalou's formulation of perceptual symbol systems as he suggests convergence zones will re-activate perceptual brain areas during concept simulation but does not require precise overlap). Only some conceptual tasks have revealed occipital activation for color word concepts (Chao & Martin, 1999), but they more consistently show ventral temporal activation. Ventral temporal cortex, while not critical for color perception, is involved in color imagery and challenging color perception tasks (Goldberg, Perfetti & Schneider, 2006).

Depending on the task, the same stimulus (e.g. a line drawing of a child's wagon) can elicit patterns of brain activation more similar to color perception areas (with a color generation task) or motor areas (with an action word generation task), even though neither feature was explicitly represented by the picture (Martin et al., 1995). Similarly, a single word (*comb*) can represent an object (*a comb*) or an action (*combing*), but the use of the word as a verb elicits more pMTG brain activity than the same word used as a noun (Tranel, Martin, Damasio, Grabowski & Hichwa, 2005). Presumably this is because verbs elicit action (and motion) concepts while nouns prototypically elicit object concepts. The particular instantiation of the word is processed relative to the appropriate conceptual and visual activation.

These effects support the claim that there is a predictable, organized similarity between the representations of brain areas activated for interaction with the world and those activated by meaningful language and concepts. The findings support Barsalou's

claim that concepts are represented multimodally and the meaning will re-activate the currently relevant modal representation(s) and features thereof.

Behavioral evidence for perceptual symbols

Behavioral evidence is important as a counterpoint to neuroimaging studies. Well-designed behavioral studies can demonstrate that perceptual and motoric features recruited during conceptual activation actually contribute qualitatively to the concept activation. For example, *watermelon* is associated more readily with *green* and *rind* while *half watermelon* elicits features including *red* rather than *green*, and *seeds* rather than *rind* (Wu & Barsalou, 2009). This type of finding demonstrates that when a concept is recruited with the purpose of making a judgment, perceptual qualities of the particular experience are also recruited. Barsalou argues that this is not the same thing as compositionality claimed by theories that would suggest that there are particular representations of each concept *half* and *watermelon* and the combined representation is a combination of the two (e.g. Smith, Osherson, Rips & Keane, 1988). Many theories of concepts predict them to be invariant across multiple instantiations such as occlusion, perspective, and size. The fact that concept combinations such as *glass car* elicit features that are typically occluded, such as *engine*, shows adaptability rather than invariance and is evidence that simulation is occurring for novel combinations of concepts (Wu & Barsalou, 2009).

Property verifications are used in classic studies of semantic relatedness under the assumption that faster responses indicate a feature has a closer semantic association with the object to which it is applied (e.g. *birds have wings* versus *birds have lungs*). This method has been used to show that perceptual modalities that dominate an

instantiation of a concept play a role in online responses even when the modality is irrelevant to the explicit task. Pecher and colleagues (Pecher, Zeelenberg & Barsalou, 2003) used property verification for perceptual properties (e.g. *buttermilk is sour*) and organized trials of the experiment into pairs. Some pairs of trials recruited the same perceptual domain such as *buttermilk is sour* (taste) and *cucumbers are bland* (taste). Other pairs of trials recruited different perceptual domains like *buttermilk is sour* paired with *bird eggs are speckled* (vision). This trial organization was based on effects found initially in studies of cross-modal and multimodal perception. When stimuli relying on different perceptual systems such as a beep and a flash are detected in succession, the response to the latter is slowed relative to two subsequent decisions made in the same perceptual domain (Spence, Nicholls & Driver, 2001). The prediction generated from the theory of perceptual symbols was if concepts activate perceptual representations then pairs of property verification trials should evoke analogous results as seen using auditory and visual stimuli. Indeed, longer reaction times were found for the second of a pair of property verification trials describing different perceptual modalities than the second of a pair describing the same modality.

The conceptual modality switch effect has been replicated extensively to show that timing of presentation does not matter (Pecher et al., 2003), and that the activated modality stays primed for several seconds (Pecher, Zeelenberg & Barsalou, 2004). Research suggests this effect is independent of word association (Pecher et al., 2003), category membership (Marques, 2006), and participants' mental imagery ability (Pecher, van Dantzig & Schifferstein, 2009). Further, an extension of the paradigm in which perceptual stimuli (flashes, beeps, vibrations) alternate with conceptual stimuli

(property verifications with visual, auditory and tactile items), yielded the same outcome – a conceptual response to *blenders are loud* is faster following a beep than when it follows a flash or vibration (van Dantzig, Pecher, Zeelenberg & Barsalou, 2008). Similarly, verifications of visual properties speed the detection of later picture identification relative to verifications of auditory properties (Pecher, Zanolie & Zeelenberg, 2007), suggesting the mere thought of a visual feature can influence visual detection. These last two studies demonstrate that concepts actually interact with perception in modality-specific ways, and thus must activate similar neural networks.

Outstanding issues – processing level and modality

The studies of conceptual modality switching provide a coherent body of work supporting the claims that concepts activate relevant perceptual systems in the course of activating simulations. An important outstanding issue about these simulations is when during conceptual processing they occur. On one hand, Barsalou has proposed that perceptual systems form the basis for conceptual representations, and therefore the activation of perceptual systems should arise as part of the meaning interpretation process itself. On the other hand, the locus of the conceptual modality switch effect might arise at any stage of the property verification task, including after participants' process the meaning of the property terms. In fact, the timing of perceptual recruitment is relatively long because it arises over pairs of trials (even showing access after several minutes, Pecher et al., 2004; Pecher et al., 2007) and could suggest that in certain cases simulation is a top-down process relying on decision-making processes.

The questions of the processing stages involved in the activation of perceptual symbols will be addressed primarily in Chapter 2 and also in Chapter 3. In these chap-

ters, we describe two studies in which participants' electroencephalogram (EEG) was recorded as they performed a property verification task in the conceptual modality switch paradigm. EEG allows for good temporal resolution of the processing required for a cognitive activity. In this case, EEG will index the switch between property verification trials and their related perceptual modalities. The primary question asked in Chapter 2 is, when during processing do we observe an impact of the conceptual modality switch manipulation? Examining electrophysiological correlates of the conceptual modality switch effect may enable us to determine whether this effect is driven primarily by perceptual, semantic, or decision-making processes (Rugg & Coles, 1995).

Another open issue regarding the perceptual nature of concepts is the impact played on these conceptual tasks by different modalities of interest. For example, the conceptual modality switch effect implies that a decision about a *shiny apple* is different from a decision about a *crunchy apple*. Specifically, one is about a visual property of apples while the other refers to the tactile modality. If the particular perceptual modality plays a role in the following decision and different patterns of localization are found for property verifications from different modalities, is the task of verifying visual properties qualitatively different from the task of verifying tactile properties? A related direction of exploration is to determine potential differences that arise in the patterns of effects when a property is associated with more than one sensory modality. For example, *smooth* is both a tactile property and a visual one. Chapter 3 will present an ERP study of conceptual modality switching similar to that presented in Chapter 2 using different modalities of stimuli. Comparison of the results in Chapters 2 and 3

will allow for a discussion of how properties recruiting different modalities are processed differently, how multimodal concepts are represented, and the implications of these results for the perceptual symbol systems hypothesis.

Immersed Experiencer Framework

The Immersed Experiencer Framework (IEF) established by Zwaan and colleagues (Zwaan, 2004; Zwaan & Madden, 2004) addresses the embodied nature of meaning in language comprehension, rather than in the conceptual system in general. The IEF suggests some explicit characteristics of the integration process involved in sentence and discourse comprehension leading to experienced-based representations. Language is a naturalistic context in which concepts are activated and refined. The particular instantiation of a conceptual representation is frequently specified by using more linguistic detail. The concepts of *apple* or *car* can be made more specific with even the simple addition of an adjective, changing them to *tart apple*, or *glass car*, for example.

Zwaan's theory builds from many of the principles established by prior psycholinguistic frameworks (e.g. Kintsch, 1988). Kintsch first described the comprehension experience as having the goal of settling on a meaningful representation by integrating the relevant parts of the sentence. The view of both Zwaan and Kintsch is that the goal of language is not for the speaker to represent the state of the world in which a given sentence is true, but rather, to have a rich mental representation of the situation described by the sentence, along with its implications. Such a representation is called a *situation model*. The creation of a situation model begins with diffuse activation of possible meanings that are subsequently constrained when they are integrated with

more linguistic context to arrive at the relevant representation of the full situation (Kintsch, 1988; Zwaan, 2004). The divergence of Zwaan's theory from the previous theoretical instantiation established by Kintsch is that rather than background information being encoded in a network of propositional knowledge, the fundamental form of meaning is physical principles derived from experience.

The functional processes involved in online language comprehension, as proposed in Zwaan's Immersed Experience Framework (Zwaan, 2004), are *activation*, *construal* and *integration*. These three processes act at the word, phrase, and sentence levels, respectively. They mutually influence each other and should be considered continuous parallel processes rather than a 3-step linear one. Activation is the initial representation of a word, and involves widespread activation of various perceptual and action areas. This distributed representation reflects all experiences with that word weighted by frequency. As such, this process is similar to Barsalou's characterization of perceptual symbols. But Zwaan's activation stage is restricted to word and morpheme representations while clausal meaning and discourse meaning rely on further processing.

The construal level integrates the perceptual networks activated by words and morphemes within a phrase thereby constraining, specifying, or in Zwaan's terms, *articulating* the perceptually-based representation in the context of the utterance. The construal phase results in the word sense and other experiential elements such as a contextually modulated perspective, timing, spatial information, entities, and features. The IEF diverges from more traditional situation model theories in the importance it places on these features and their perceptual and motoric nature. The mental represen-

tation created during construal tends to be processed as an event. For example, hearing *The car was blue* would result in a representation of seeing a blue car (Zwaan & Madden, 2004). A construal is a schematic representation because of processing limitations for initial experiences that require attentional resources directed to limited elements in the world as well as processing limitations in online language comprehension. During the construal process an experience and its spatio-temporal perspective (e.g. attention on the color of the car, viewed from some perspective with little or no attention to the background) are derived from word meaning, closed-class words, morphemes, and grammatical constructions. For example, these structures can direct attention to old/new information (definite/indefinite articles), to the focal element of the event (passive/active constructions), or the process focus (aspect) (Zwaan & Madden, 2004).

Finally, the integration process is that by which multiple construals co-constrain each other to build a more articulated representation. Integration brings together construals in working memory and corresponds to related transitioning processes in human experience such as attentional modulation. During natural interaction with the world, Zwaan suggests that humans shift attention from one element of a landscape to another, or switch focus between sights, sounds, thoughts, desires, etc. Despite these shifts, human experience is perceived as continuous. Linguistic cues guide the listener in understanding what to predict in subsequent language and how to integrate this information coherently. The articulated mental model matters too; the more a linguistic representation characterizes continuous experiences the easier the integration phase and overall processing. The extent to which an experiential trace is activated is a matter of its similarity with the current eliciting input (the language) and

elements that are already activated (Zwaan, 2008). This degree of similarity and its ability to result in linguistic predictions are critical features of the comprehension process.

Empirical evidence for the Immersed Experiencer Framework

The evidence presented in support of the Perceptual Symbol Systems hypotheses can also be considered evidence for Zwaan's general claim that widespread perceptuo-motor activation underlies word meaning and drives the articulation of physical principles. Evidence for the construal phase of the Immersed Experiencer Framework (Zwaan, 2004) initially came from two experiments in which participants were presented with a sentence followed by a picture, and were asked to determine if the object in the picture had been mentioned in the previous sentence. In one experiment, sentences in two different conditions described the same objects but in different locations e.g. *the pencil was in the cup* versus *the pencil was in the drawer*. Participants responded faster to verify that a picture of the object (the pencil) was mentioned in the sentence when the orientation of the object corresponded with the orientation implied in the sentence (vertical orientation for a pencil in a cup versus a horizontal orientation for a pencil in a drawer; Stanfield & Zwaan, 2001). Similarly, participants responded faster if the shape of a picture, e.g. a fried egg with exposed yolk, was consistent with details of implied shape constructed by the sentence (*The egg was in the frying pan*) versus when the sentence described the object in a different shape (*The egg was in the refrigerator*; Zwaan, Stanfield & Yaxley, 2002).

Directionality and motion are also features activated during construal and have been demonstrated to impact both visual processing and motor processing. For exam-

ple, sentences such as, *for the first pitch of the softball game, the pitcher hurled the ball towards you* versus *for the first pitch of the softball game, you hurled the ball toward the plate* are very similar sentences except for the perspective of the speaker and the subsequent direction of ball movement. When listening to this type of sentence, participants were asked whether two successive objects presented were the same or different shape (Zwaan, 2004). The size of the objects (the objects were balls for critical trials) was also subtly manipulated as a dimension irrelevant to the task. The fast presentation implied movement toward the listener when the second ball was larger than the first and movement away from the listener when the second ball was smaller. This manipulation created a percept of directed motion that was used to test the details in participants' visual simulations of the previous sentences. Indeed, reaction times were faster when the direction of the ball movement was consistent with the direction of movement within the accompanying sentence (e.g. a small then large ball accompanied by ...*the pitcher hurled the ball toward you*).

A similar finding has been discovered in the motor domain with sentences describing transfer and using another creative methodology to extract information about directionality (Glenberg & Kaschak, 2002). Participants started with their hand at button in a neutral middle location and response buttons were arranged so that hand motions were either away from or toward the participant's body. When a sentence described motion toward the participant, (*you opened the drawer*), button presses requiring participants to reach toward their body were faster than movements away from the body; the opposite was true for sentences expressing motion away from the body (*you closed the drawer*). This effect also held for sentences describing abstract transfer

away from the listener, *You radioed the policeman*, and toward the listener, *Liz told you the story* (Glenberg & Kaschak, 2002). Similarly, an effect with rotational actions has shown faster reading times when the direction of motion in the sentence (*You turned up the volume on the radio*) was consistent with the knob rotation (clockwise versus counterclockwise) required to progress through the sentence (Zwaan & Taylor, 2006).

In general these findings demonstrate that object identity, shape, orientation, motion direction and specific motoric features are activated by people during the course of language comprehension. These physical details are not mentioned explicitly in the sentences described above yet they speed actions and visual detections. In order to have such an effect, the argumentation goes, these features must be part of the mental model that was constructed during comprehension, and they must also have access to motor and vision brain areas. Hence, these studies are used as support for the claim that motor and perception brain regions are involved in the meaning creation process during language comprehension.

Specificity of represented information

Zwaan posits widespread activation of perceptual networks (related to a proposal by Pulvermüller, 2001) that are more constrained with linguistic context. The specificity of the representation in this view only arises with sufficient linguistic context. Barsalou, on the other hand, suggests that broad regions of sensory areas will be activated by concepts insofar as these sensory areas have been important for previous experiences. All concepts will have sensorimotor bases, some of which will be quite specific, while others might represent something more generally. Even with the elabo-

rations of these two theories, it is difficult to form clear predictions about how specifically perceptual the representations evoked by language might be. The empirical work offers mixed results that suggest the interaction of language and perception occurs at any number of different perceptual processing stages, depending on the nature of the task used.

One common way to look at effects of perceptual systems on language is by pairing particular types of statements with pictures. As described already, a picture of vertical pencil elicits a faster response after the phrase *the pencil in the cup* than a picture of a horizontally oriented pencil following the same sentence. The specific feature of orientation of the pencil is thus articulated by the linguistic context and the resulting visual image must be rather specific. Participants respond “true” to both horizontal and vertical pencils so the balanced stimuli allow the authors to suggest that orientation or shape are the particularly important features in our mental representations. However, with such an experimental design it is difficult to assess how specific the mental representation is independent of the experimental context. For example, would the details of the ridges in the pencil, its color, or the existence of an eraser be part of the mental representation before seeing the image? When we probe the mental representation by using an image we cannot know which of these features were part of the representation before participants saw the image. A recent follow-up study demonstrated effects of abstract orientation, in the form of vertical or horizontal line gratings, on the judgment of the same sentence as used to test the vertical versus horizontal orientations described above (“Mental simulation in language comprehension: new findings and new considerations”, oral presentation by Rolf Zwaan at the joint meeting of Concep-

tual Structure, Discourse and Language/Embodied and Situated Language Processing, 2010). These findings suggest that a more abstract notion of visual orientation is part of the meaning. Still, on the other hand, these effects could be derived from a rather specific image of a vertically- or horizontally-oriented pencil. The question of specificity in terms of the mental representations evoked by language is a challenging issue in developing theories of grounded cognition.

The way the specificity of conceptual representations has typically been addressed in the literature is by looking for evidence of activity by different brain systems. For language describing motor actions, the cognitive neuroscience method, transcranial magnetic stimulation (TMS), has been applied at the scalp to evoke motor potentials (MEPs) recorded at distal muscles. In response to sentences of transfer, larger MEPs have been recorded at hand muscles suggesting that sentence comprehension can involve primary motor activity independent of an actual hand motion (Glenberg, Sato, Cattaneo, Riggio, Palumbo & Buccino, 2008). Studies of perceptually-related language address these questions in terms of “low-level” or “high-level” perceptual effects. Some studies show that language can impact perceptual processes that are rather simple visual processes (Meteyard, Bahrami & Vigliocco, 2007). Other studies have used performance criteria – threshold vs. supra-threshold stimuli – as a measure suggesting that top-down effects can take place in certain situations (with supra-threshold stimuli) but in other cases (with threshold stimuli) low-level visual processing is recruited (Meteyard, Zokaei, Bahrami & Vigliocco, 2008).

Another angle used to suggest early activation of perceptual systems by language is with timing measures. The timing of perceptual and motor effects can be ob-

served within 200 ms post lexical onset (Boulenger et al., 2006) and can begin within the first two syllables of word processing (Revill, Aslin, Tanenhaus & Bavelier, 2008). However, the motoric access observed during comprehension has also been demonstrated to be unavailable in the middle of sentence processing (Kaschak & Borreggine, 2008). In many of the behavioral studies designed to illustrate the features of perception and action that are accessible in language comprehension, sentences are presented followed by an additional task (e.g. Stanfield & Zwaan, 2001; Zwaan et al., 2002). In such cases the perceptual features recruited during language are accessible for visual detection 1-2 sec after sentence processing. In the discussion of low-level and high-level processes, the faster that language can be observed to interact with visual and motoric systems the more intrinsically integrated the systems are supposed to be. The experimental work on these issues is limited but the theme available from varied sources and methods suggests that language interacts with perceptual and motoric systems at many time scales.

Visual systems that are activated for perceiving low-level luminance changes can be modulated with concurrent language processing (Meteyard et al., 2007). Meteyard et al. (2007) had participants make perceptual discriminations about whether or not they perceived vertical motion while listening to verbs that conveyed motion in either upward, downward, or horizontal directions. Their particular perceptual stimuli were random-dot kinematograms (RDKs) consisting of 1,000 moving dots presented for only 150 ms. When a percentage of the 1,000 dots moved either upward or downward while the rest of the dots moved randomly, the RDK was classified as a vertical motion stimulus. Random motion consisted of all the dots moving randomly. The find-

ings showed sensitivity for detecting the vertical motion signal was worse when participants heard words conveying motion in the opposite direction. Decision criterion measures were also lower (i.e. more liberal responses for detecting coherent motion) when motion perception took place in the context of verbs conveying motion in the same direction. These findings were used to suggest that low-level motion perception can be affected by the direction of motion conveyed by incidental sets of individual words. However, the fact that decision criterion measures were affected as well as perceptual sensitivity measures could also be interpreted as showing a top-down strategic process driving both effects.

Other perceptual processes that potentially interface with language comprehension are feedback circuits for generating a motion after effect (Dils & Boroditsky, 2010), eye movement planning and control mechanisms (Spivey & Geng, 2001) or spatial attention systems (Bergen, Lindsay, Matlock & Narayanan, 2007; Richardson, Spivey, Barsalou, & McRae, 2003). For example, in one study participants were instructed to determine the shape of a figure (circle or square) presented in the upper, lower, right or left side of the screen, preceded by short sentences like, *the mule climbed* (Bergen et al., 2007). Vertical verbs affected judgments in upper or lower quadrants depending on the specific verb (Bergen et al., 2007). Unlike the conclusions of previous studies using low-level visual tasks, Bergen and colleagues suggested that their effect was due to active mental imagery of full sentences rather than particular lexical items. Similarly, Dils and Boroditsky (2010) found that only sentences presented toward the end of a full story context elicited effects of vertical motion imagery even though early sentences of the paragraphs also contained vertical motion lan-

guage. Their findings also suggested that effects of language on perception do not take place at the lexical level but extended over unified discourse contexts.

Plans to test specificity

The research presented in Chapters 4 and 5 is designed to address the issues of low-level versus high-level processing involved in the interaction of language and perception. One goal of the experiments presented in Chapter 4 is to investigate the extent to which procedural differences can influence language-perception interactions. The experimental methods in Chapter 4 are variants of those employed by Meteyard et al. (2007). The design differs from prior work in that each trial involved a single word followed by the RDK to ensure that any observed effects would be due to the semantic processing required for that word. This contrasts with the previously reported results for which blocks of trials were presented with a continuous stream of either upward or downward verbs with no direct relation to individual RDK judgments.

Chapter 4 will address whether perceptual sensitivity, decision criterion, or the reaction time modulations are observed in different procedural circumstances: one experiment will have motion trials within a block all moving in the same direction, the other experiment will have upward and downward motion randomized. Although the measures are all derived from simple behavioral responses, they presumably arise at different stages of processing, and consequently have different implications for grounded theories of meaning. Variations of the experimental design and the corresponding level of effect (d' , C , reaction time, accuracy) can lead us to a better understanding of the circumstances under which high- versus low-level interactions are observed. Similarly, if the blocked presentation is critical for observing behavioral dif-

ferences, then claims about the low-level or high-level effects need to be framed to consider contextual factors engendered by the procedure itself.

Chapter 4 will also describe the results of an event-related potential (ERP) experiment intended to investigate the processing level at which language and perception interact, as well providing a window into the timing of this process. This study paired the recording of EEG with the paradigm described above using vertical verbs and RDKs to supplement the inferences derived from the behavioral measures. ERPs can reveal differences in how the perception, attention, decision making, and semantic processing systems react to different classes of stimuli. ERP modulations observed for the perception of RDKs as a function of the type of preceding verb will be compared with previously-described ERP effects and can provide another window into the level at which visual and language processing come together.

As noted above, there is extensive evidence that language can interact with motion perception. The key questions at this point are why this interaction takes place and what the process is. Chapter 5, will propose that language recruits neuro-cognitive processes for high level motion inference and introduce the representational momentum paradigm as a test of this proposal. Along with presenting a new method for testing the interaction of language and perception, Chapter 5 will compare effects elicited by motion language and motion images. By using the same experimental paradigm (representational momentum: a high-level motion inference task) with more than one type of meaningful motion stimulus (language and images), we will have a broader perspective to better understand the systems involved.

Summary

The Perceptual Symbol Systems and Immersed Experiencer Framework are two related theories that address the question of meaning in the human cognitive system as being grounded in life experiences. There is a growing body of literature demonstrating a relationship between meaning representations and perception. The work presented here will attempt to clarify some open issues regarding the processes underlying the connections between language and perception. Chapters 2 and 3 will focus on the conceptual modality switch paradigm as a way to investigate the cognitive processes involved in comprehending words about visual, auditory and tactile experiences. Chapter 4 will focus on motion language with a low-level motion task to understand particular circumstances under which language influences motion perception. Chapter 5 will also address the comprehension of motion language but with a high-level motion perception task to test the role of motion inference processes in the comprehension of motion verbs.

References

- Aziz-Zadeh, L., Wilson, S., Rizzolatti, G., Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*, 1818-1823.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-660.
- Barsalou, L. W. (2008). Grounding symbolic operations in the brain's modal systems. In G. R. Semin, & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 9-42). New York, NY, US: Cambridge University Press.
- Beauchamp, M.S., Lee, K.E., Haxby, J.V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, *34*, 149-159.
- Bergen, B.K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*, 733-764.
- Boulenger, V., Roy, A.C., Paulignan, Y., Deprez, V., Jeannerod, M. & Nazir, T.A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of Cognitive Neuroscience*, *18*, 1607-1615.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., & Freund, H.-J (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, *13*, 400-404.
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355-363.
- Chao, L.L. & Martin, A. (1999) Cortical regions associated with perceiving, naming and knowing about colors. *Journal of Cognitive Neuroscience*, *11*, 25-35.
- Chao, L.L., Haxby, J.V. & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *11*, 1, 913-919.
- Dils, A.T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences*, *107* (37), 16396-16400,
- Gerlach, C., Law, I. & Paulson, O.B. (2002). When action turns into words. Activation of motor-based knowledge during categorization of manipulable objects. *Journal of Cognitive Neuroscience*, *14* (8), 1230-1239.

- Glenberg, A.M. & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, 9 (3), 558-565.
- Glenberg, A.M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D. & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology*, 61 (6), 903-919.
- Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *The Journal of Neuroscience*, 26, 4917-4921.
- Hauk, O. & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the front-central cortex. *Human Brain Mapping*, 21, 191-201.
- James, T.W., Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology*, 13, 1792-1796.
- Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, 20, 525-540.
- Kaschak, M.P. & Borreggine, K.L. (2008). Temporal dynamics of the action-sentence compatibility effect. *The Quarterly Journal of Experimental Psychology*, 61, 883-895.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95 (2), 163-182.
- Marques, J. F. (2006). Specialization and semantic organization: Evidence for multiple semantics linked to sensory modalities. *Memory & Cognition*, 34, 60-67.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25-45.
- Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270, 102–105.
- Meteyard, L., Bahrami, B., Vigliocco, G. (2007). Motion detection and motion verbs: language affects low-level visual perception. *Psychological Science*, 18 (11), 1007-1013.
- Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, 18, R732-733.
- Noppeny, U., Josephs, O., Kiebel, S., Friston, K.J. & Price, C.J. (2005). Action selectivity in parietal and temporal cortex. *Cognitive Brain Research*, 25, 641-649.
- Pecher, D., van Dantzig, S., and Schifferstein, H. N. J. (2009). Concepts are not represented by imagery. *Psychonomic Bulletin & Review*, 16, 914-919.

- Pecher, D., Zanolie, K. & Zeelenberg, R. (2007). Verifying visual properties in sentence verification facilitates picture recognition memory. *Experimental Psychology*, 54 (3), 173-179.
- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying properties from different modalities for concepts produces switching costs. *Psychological Science*, 14 (2), 119-124.
- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, 11, 164-167.
- Pulvermüller, F. (2001). Brain reflections of words and their meanings. *Trends in Cognitive Science*, 5 (12), 517-524.
- Revill, K.P., Aslin, R.N., Tanenhaus, M.K., & Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences*, 105 (35), 13111-13115.
- Richardson, D.C., Spivey, M.J., Barsalou, L.W. & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, 27, 767-780.
- Rugg, M.D. & Coles, M. (1995). Event-related brain potentials: An introduction. In Rugg, Michael D. & Coles, Michael (Eds.), *Electrophysiology of Mind: Event-Related Potentials and Cognition*. Oxford: Oxford University Press.
- Simmons, W.K. & Barsalou, L. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, 20 (3-6), 451-486.
- Smith, E.E., Osherson, D.N., Rips, L.J. & Keane, M. (1988). Combining prototypes: a selective modification model. *Cognitive Science*, 12, 485-527.
- Spence, C., Nicholls, M.E.R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, 63, 330-336.
- Spivey, M. & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235-241.
- Stanfield, R. A. & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12 (2), 153-156.
- Tranel, D., Martin, C., Damasio, H., Grabowski, T.J. & Hichwa, R. (2005). Effects of noun-verb homonymy on the neural correlates of naming concrete entities and actions. *Brain and Language*, 92, 288-299.
- van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, 32, 579-590.
- Weisberg, J., van Turrenout, M., Martin, A. (2007). A neural system for learning about object function. *Cerebral Cortex*. 17, 513-521

- Wu, L.L. & Barsalou, L.W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica, 132*, 173-189.
- 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. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In *The Psychology of Learning and Motivation, Vol. 44*, ed. B.H. Ross, New York: Academic Press.
- Zwaan, R.A. (2008). Experiential traces and mental simulations in language comprehension. In de Vega, M., Glenberg, A.M., & Graesser, A.C. (Eds.) *Symbols, Embodiment, and Meaning* (pp. 165-180). Oxford University Press, Oxford, UK.
- Zwaan, R.A., & Madden, C.J. (2004). Updating situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 283-288.
- Zwaan, R.A., Madden, C.J., Yaxley, R.H., & Aveyard, M.E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science, 28*, 611-619.
- Zwaan, R.A., Stanfield, R.A., & Yaxley, R.H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13* (2), 168-171.

Chapter 2

Modality Switching in a Property

Verification Task



Modality switching in a property verification task: an ERP study of what happens when candles flicker after high heels click

Jennifer Collins^{1*}, Diane Pecher², René Zeelenberg² and Seana Coulson¹

¹ Brain and Cognition Lab, Department of Cognitive Science, University of California San Diego, San Diego, CA, USA

² Memory Lab, Department of Psychology, Faculty of Social Sciences, Erasmus University Rotterdam, Rotterdam, Netherlands

Edited by:

Anna M. Borghi, University of Bologna and Institute of Cognitive Sciences and Technologies, Italy

Reviewed by:

Lea Hald, Canterbury Christ Church University, UK

Frederico Marques, Lisbon University, Portugal

*Correspondence:

Jennifer Collins, Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515, USA.
e-mail: jcollins@cogsci.ucsd.edu

The perceptual modalities associated with property words, such as *flicker* or *click*, have previously been demonstrated to affect subsequent property verification judgments (Pecher et al., 2003). Known as the *conceptual modality switch effect*, this finding supports the claim that brain systems for perception and action help subservise the representation of concepts. The present study addressed the cognitive and neural substrate of this effect by recording event-related potentials (ERPs) as participants performed a property verification task with visual or auditory properties in key trials. We found that for visual property verifications, modality switching was associated with an increased amplitude N400. For auditory verifications, switching led to a larger late positive complex. Observed ERP effects of modality switching suggest property words access perceptual brain systems. Moreover, the timing and pattern of the effects suggest perceptual systems impact the decision-making stage in the verification of auditory properties, and the semantic stage in the verification of visual properties.

Keywords: property verification, modality switch, ERPs, grounded cognition

INTRODUCTION

Over the past decade, cognitive scientists have gradually moved away from the assumption that concepts are *symbolic*, that is, arbitrarily related to the things they represent, and *amodal*, or independent of any sensory modality (see Murphy, 2002 for a review of traditional models), and have increasingly come to embrace an embodied or grounded approach. These more recent accounts have focused on how concepts are grounded in our perception of, and interaction with, the physical and social world, and stressed their modal characteristics (see Barsalou, 2008 for a review). The perceptual symbol system hypothesis, for example, is that conceptual knowledge involves schematized perceptual and motor representations involved in one's prior experience with the concept's referent (Barsalou, 1999). On this account, a concept is a sensorimotor simulation involving the partial reactivation of brain regions that participated in the acquisition of that concept. For example, the concept of a dog is a simulation involving brain areas that represent one's visual, auditory, tactile, olfactory, affective, and motoric experiences with dogs. Importantly, simulations are not holistic records of experience, but can be flexibly adapted to the current context and task (Barsalou et al., 2003).

The use of visual mental images for ostensibly conceptual tasks has been demonstrated with the property verification task, in which participants are asked whether or not a particular property (e.g., *has-a-head*) is true for a given concept (e.g., *CAT*). The perceptual symbol system hypothesis suggests that accessing conceptual knowledge involves the activation of associated visual images, and thus predicts a systematic relationship between the difficulty of property verification and that of activating the relevant visual image. Consistent with this prediction, Solomon and Barsalou (2004) found that participants took less time to verify visually

large properties of a concept (e.g., that a *CAT* has a *head*) than visually smaller properties of the same concept (e.g., that a *CAT* has a *paw*). The fact that performance on this conceptual task was modulated in a similar way as performance on a visual imagery task was argued to implicate the importance of visual processes in conceptual representations.

Moreover, a functional magnetic resonance imaging (fMRI) study in which participants performed the property verification task employed by Solomon and Barsalou (2004) revealed activation in the left fusiform gyrus, an area important for object recognition and visual imagery (Kan et al., 2003). The recruitment of perceptual brain areas for the conceptual task of property verification is consistent with the perceptual symbol system hypothesis, and is also in keeping with other fMRI studies in which conceptual tasks have activated brain regions used to perceive the concepts' referents (Goldberg et al., 2006; Martin, 2007; Simmons et al., 2007).

MODALITY SWITCH EFFECTS

Although the bulk of empirical support for the perceptual symbol system hypothesis concerns the involvement of specifically visual representations, the hypothesis is, in fact, farther ranging, extending to the full multimodal characteristic of human experience. The concept of a lemon, for example, should not only represent its color, but also its taste, its smell, and its texture. Moreover, because simulations involve the coordination of information from multiple perceptual modalities, the perceptual symbol system hypothesis predicts that conceptual operations will display many of the same properties as complex perceptual operations, and be subject to similar constraints. Accordingly, Pecher et al. (2003) tested whether a property verification task using properties

from several modalities, including vision, audition, and touch, was modulated by factors known to affect perceptual detection tasks with stimuli from multiple modalities.

In particular, Pecher et al. (2003) focused on the modality switch effect, a phenomenon observed in the literature on perceptual processing. In a study designed to assess cross-modal effects of spatial attention, Spence et al. (2001) asked participants to detect brief auditory, visual, or tactile targets at peripheral locations. The *modality switch effect* is the finding that reaction times were longer for all stimulus types when they were preceded by a stimulus from a different modality than from the same modality, and has been interpreted as an exogenously driven attentional cost for the switch trials (Spence et al., 2001; Rodway, 2005).

Pecher et al. (2003) reasoned that if conceptual processing relies on perceptual systems, the well-known cost for successive trials from different modalities in perceptual tasks might also be expected to occur on a property verification task employing properties from multiple modalities. In their conceptual analog to the modality switch studies, Pecher et al. (2003) asked participants to determine whether a property (e.g., *yellow* or *sour*) applied to the preceding concept (e.g., *LEMON* or *MOUSE*). The manipulation of interest was whether a pair of trials was from the same modality (*LEAVES–rustling* followed by *BLENDER–loud*) or different modalities (*CRANBERRIES–tart* followed by *BLENDER–loud*). As predicted by the perceptual symbol system hypothesis, Pecher et al. (2003) found longer reaction times for the second trial in a pair of different modality (switch) trials than for the second trial in a pair of same modality (no-switch) trials, the conceptual modality switch effect.

Variations on the conceptual modality switch paradigm have shown that results cannot be attributed to alternative explanations, such as word association (Pecher et al., 2003), or category overlap (Marques, 2006). The generality of the effect is supported by the demonstration of a similar switch effect on a property verification task using perceptual and emotional attributes (Vermeulen et al., 2007). Importantly, property verification has also been shown to be speeded by the presentation of a perceptual stimulus from the same modality relative to one from a different modality (van Dantzig et al., 2008). The finding that the verification of visual features of a concept is faster after the perceptual detection of visual than auditory or tactile stimuli provides strong support for the suggestion that the conceptual task of property verification recruits perceptual processing resources, as opposed to an amodal re-representation of perceptual information.

Another direction this research has taken has been to investigate the neural substrate of modality specific concepts using cognitive neuroscience methods. Goldberg et al. (2006) recorded participants' brain activity using fMRI while they engaged in a property verification task. The experiment used a design in which different blocks required participants to make decisions about properties referring to different modalities – visual, auditory, tactile, and gustatory. The brain regions uniquely activated for each property category were regions related to the perception of stimuli in the different domains. These results are particularly important given that reaction time results for similar conceptual tasks have not distinguished between responses to properties of different modalities (Pecher et al., 2009).

Neuroimaging data thus provide compelling evidence that conceptual tasks are associated with the activation of perceptual brain regions. At issue, however, is whether perceptual systems play a central or a peripheral role in cognition (Barsalou, 2008). Perceptual activations might, for example, be an artifact of the blocked design used by Goldberg et al. (2006). Alternatively, perceptual activations might reflect top-down processing initiated only after the meaning of the property words has been activated.

THE PRESENT STUDY

The present study addressed the cognitive and neural basis of the conceptual modality switch effect by recording event-related potentials (ERPs) as participants made property verification judgments about the visual and auditory properties of objects. ERPs are patterned voltage changes in the on-going electroencephalogram (EEG) that are time-locked to classes of specific processing events. As a continuous, real-time measure of brain activity, ERPs are well-suited for investigating the neural processes relevant to the conceptual modality switch effect allowing us to better understand when a perceptual system is accessed by a related concept. In particular, the present study was designed to address whether the modality manipulation affected ERP components associated with the visual processing of property terms, such as the N1 and P2, semantic processing of property terms, such as the N400, or their task-relevant categorization as typical properties of the relevant concept, indexed by the P3 or late positive complex (LPC).

We used stimuli similar to those employed by Pecher et al. (2003), but included only visual (*CANDLES–flicker*) and auditory (*NEWSPAPERS–rustle*) trials in our critical conditions. This reduction in variation was important in order to have enough trials in critical conditions for averaging ERPs. Participants' task was to determine whether or not the property applied to the concept. The correct response on all experimental trials was "true," and a large number of filler trials requiring a "false" response (e.g., *COCKROACHES–ablaze*) were included to discourage the development of a particular response bias. A subset of false filler trials included properties and concepts that were lexically associated (e.g., *STRAWBERRIES–cream*) and were intended to discourage the use of word association strategies (Solomon and Barsalou, 2004). The critical manipulation concerned whether the target concept–property trial (e.g., *NEWSPAPERS–rustle*) was preceded by a prime concept–property trial from the same modality (e.g., *HIGH HEELS–click*), or a different modality (e.g., *CHERRIES–ruby*). Half of the experimental trials involved visual and half auditory properties, and were equally likely to follow a concept–property trial from the same modality (a visual property following a visual property, or an auditory property following an auditory property, viz. no-switch trials) as one from a different modality (visual–auditory or auditory–visual, viz. switch trials).

The primary goal of the study was thus to identify electrophysiological correlates of the conceptual modality switch effect in order to determine which stage or stages of processing the switch manipulation would modulate. If concepts automatically engage early sensory processing, then the mention of a visual property such as "flicker" could modulate the actual perception of visual word forms presented shortly afterward. The converse of this type of effect was found behaviorally by van Dantzig et al. (2008).

Low-level perceptual engagement of this sort would be indexed by modulation of visual ERP components to the word form such as the N1, and P2.

Alternatively, perceptual access might be part of an extended, standard semantic network that subserves the representation of concepts. The N400, a negative-going wave evident between 200 and 700 ms after the visual presentation of a word, was of particular interest due to its association with the processing of meaningful events. The N400 is elicited by words in all modalities, whether written, spoken, or signed (Holcomb and Neville, 1990). Words preceded by semantically related words elicit smaller amplitude N400 than do words preceded by unrelated words, the N400 priming effect (Bentin, 1987; Holcomb, 1988; Smith and Halgren, 1989). The N400 is also sensitive to contextual factors related to meaning at the sentence and text level. In general, N400 amplitude varies inversely with the predictability of the target word: N400s are large for unexpected items, smaller for words of intermediate predictability, and are barely detectable for highly predictable words (Kutas and Hillyard, 1984; see Kutas and Federmeier, 2011 for a review).

Yet another possibility is that the conceptual modality switch effect is attributable to decision processes specifically induced by the property verification task. If this is the case, we would expect the conceptual modality switch paradigm to modulate later, decision-related components such as the P3, or LPC. This family of ERP components is generally thought to index the updating of mental representations modulated by processes such as allocation of attention and task-dependent target classification (Polich, 2007).

A secondary goal of the study was to test whether property terms from different modalities (viz. visual versus auditory) would activate different modality specific brain areas as found in related fMRI studies (e.g., Goldberg et al., 2006). Although the spatial resolution of the EEG is limited, such differences might be detectable as subtle differences in the scalp topography of ERPs to visual versus auditory properties. An interaction between the modality factor in our analysis and electrode site would suggest that non-overlapping neural generators underlie the brain response to auditory and visual properties, viz. that the exact same brain regions do not subservise the processing of visual and auditory properties (Urbach and Kutas, 2002). More generally, differences between the modality switch process in the visual and auditory domains would connect this paradigm with Pecher et al.'s (2003) claim that the conceptual modality switch effect results from switching between different perceptual networks.

As a time-sensitive measure of online cognitive processing, ERPs can provide more information about whether the real-time processing of property terms involves the recruitment of perceptual brain areas during early perceptual processing, during semantic processing, or whether the switch effect would be evident only later, during decision-related stimulus processing. Given Barsalou's (1999) claim that sensorimotor simulations comprise an intrinsic component of concept meaning, we hypothesized that the facilitative impact of a same modality prime would involve the semantic processing of the target trial, and thus would modulate the amplitude of the N400 component of the ERP. In particular, we predicted that no-switch trials would elicit reduced amplitude N400 relative to switch trials.

MATERIALS AND METHODS

The protocol for this study was approved by the University of California, San Diego Social and Behavioral Science Institutional Review Board. As such, informed consent was obtained from all participants prior to their enrollment.

PARTICIPANTS

Twenty undergraduates from the UCSD community (13 women) participated as part of a course requirement. Data from six additional participants were not included in the analysis due to the presence of an excessive number of artifacts (greater than 30% of trials in a critical condition). All participants were between the ages of 18 and 40 years old. As reported in a screening questionnaire, all participants had normal vision, and none had any history of neurological or psychiatric disorders within the previous 10 years.

MATERIALS

Each trial in the study consisted of a concept–property combination such as *HIGH HEELS* (concept) and *click* (property). Experimental trials involved 48 visual properties (such as *flicker*), and 48 auditory properties (such as *click*). Each property was presented with two different concepts for a total of 192 experimental trials; all properties were repeated once over the course of the experiment, while all concepts were unique. Half (96) of the concept–property combinations served as prime trials (48 involving auditory properties, and 48 involving visual properties), and half (96) served as target trials (48 involving auditory properties, and 48 involving visual ones). Experimental trials were presented in pairs, so that a prime trial was immediately followed by a target trial that was either from the same modality (no-switch condition), or the other modality (switch condition). Materials were thus comprised of 96 trial pairs in which the modality of the probe property was crossed with the modality switch dimension (24 auditory prime/auditory target, 24 visual prime/auditory target, 24 visual prime/visual target, and 24 auditory prime/visual target pairs). Apart from the modality manipulation the prime–target pairs were unrelated. All properties in experimental trials were valid for their concept so that the correct response on the property verification task was always “true.”

Materials also included 384 filler trials, 96 of which involved auditory properties that did not pertain to their concept (e.g., *LOBSTERS–bark*) and 96 of which involved visual properties also eliciting *false* responses (e.g., *LAWNS–scarlet*). These two sets were included so that participants could not strategically respond *true* to any trial involving an auditory or visual property. Another 96 filler trials involved tactile properties, half of which were valid for their concept (e.g., *CAVES–damp*), and half of which were not (e.g., *TOASTERS–damp*; one response for each property repetition). The final 96 filler trials were lexical associates (e.g., *BUFFALOS–winged*), included to discourage participants from shallow processing strategies relying on word association (as in Solomon and Barsalou, 2004). Half of the associated trials were *true* trials, and half were *false* trials. Of the 384 filler trials, the correct response on the property verification task was *true* for 96, and *false* for 288. When including the 192 experimental trials as well, the correct response on the task was thus *true* for half of the total trials, and *false* for the other. Moreover, even though the experimental trials always involved two *true* responses in a row (viz. one for the prime, and

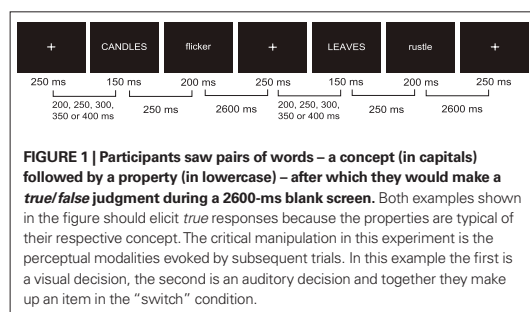
one for the probe), the inclusion of filler trials guaranteed that a correct *true* response was equally likely to be followed by a correct *false* response as by another correct *true*.

Two lists were employed so that any given target property occurred once in a switch trial (that is, following a prime from the other modality), and once in a no-switch trial (that is, following a prime from the same modality). Two variants of each were created by swapping the first and second half of each list. In this way, each concept–property combination was presented equally often in the first and second half of the experiment.

PROCEDURE

Participants were seated in a dim, sound attenuating chamber approximately 50 inches from a 17-inch computer monitor. They read a standard set of instructions telling them to “read the entity (such as objects, people, animals, etc.) and property words, . . . and respond *true* if the property was typical or often possible for the entity, and *false* if the property was highly unusual for the entity.” They read several examples and were presented with practice trials on which they received feedback. Participants were told, “after you read the property, decide as quickly and accurately as possible whether the property is true or false,” but no explicit feedback was given on either of these dimensions during the course of the experiment.

The timing of events in the experimental paradigm is presented in **Figure 1**. Each trial began with the presentation of a white fixation cross for 250 ms. The inter-stimulus interval (ISI) between the fixation cross and the concept was randomly varied with 50 ms steps between 200 and 400 ms. The concept appeared on the center of the screen in capital letters for 150 ms followed by a 250 ms ISI and the property word in lowercase letters for 200 ms. In order to limit the potential for eye-movement artifacts in our EEG signal we chose to centrally present both concepts and properties and eliminate the phrase “can be” from the original paradigm which is not a necessary aspect of the conceptual modality switching procedure (e.g., Pecher et al., 2004). All type was in white font presented on a black background. Participants had 2600 ms to make their decision and prepare for the next trial. Responses were made via a button press in which a right hand response indicated *true* and a left hand response indicated *false*. Trials were presented in ten blocks, each lasting about 3.5 min with time in between for participants to rest. The first block began with eight practice trials that were not included in the analysis. All blocks had 60 trials except for the last block which had 44 trials.



EEG RECORDING AND ANALYSIS

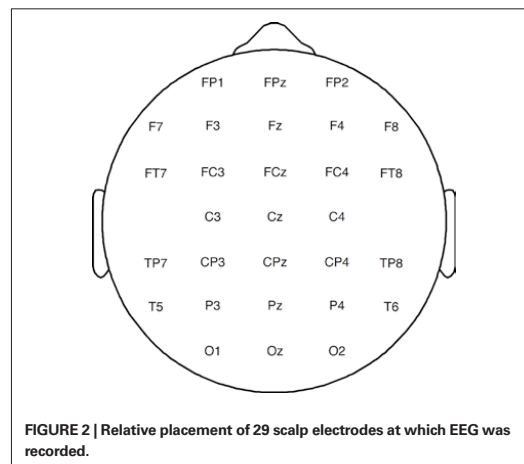
Participants’ EEG was recorded with tin electrodes mounted in an electrode cap with 29 scalp sites (see **Figure 2**). Scalp electrodes were referenced online to the left mastoid, and subsequently re-referenced to the average of the left and right mastoid electrodes. Blinks were monitored with an electrode below the right eye. Horizontal eye movements were monitored via a bipolar derivation of electrodes placed over the outer canthi. EEG was recorded and amplified with an SA Instruments isolated bioelectric amplifier at a bandpass of 0.1 and 100 Hz, digitized online at 250 Hz, and stored on a hard drive for subsequent averaging. The EEG was later monitored offline for blinks and eye movements which were rejected manually. ERPs were time-locked to the onset of property words on probe trials.

For each time interval of interest we performed a $2 \times 2 \times 29$ repeated measures ANOVA with the factors switch (switch/no-switch), target property modality (visual/auditory), and electrode site (29 levels). The dependent measure was the mean amplitude within the time intervals of interest. In cases where the overall analysis revealed a significant interaction between modality switch and property modality, follow-up analyses were conducted separately for the visual and auditory properties. Follow-up analyses thus involved factors switch (switch/no-switch) and electrode site (29 levels). The Huynh–Feldt correction was applied where relevant. For clarity, however, we report the original degrees of freedom.

RESULTS

BEHAVIORAL RESULTS

Analysis of reaction times failed to reveal any statistically significant effects in a 2×2 ANOVA testing switch (switch/no-switch) and modality (visual/auditory; all $F_s < 2$). Given that behavioral studies of this phenomenon typically do not test the modalities separately and employ data from at least 60 participants (cf. the 20 employed in the present study), these null results are likely due to a lack of power. The pattern of reaction times was in the expected direction for the visual properties (switch = 902 ms, SD = 152 ms;



no-switch = 891 ms, SD = 155 ms) but not for the auditory properties (switch = 908 ms, SD = 148 ms; no-switch = 917 ms, SD = 163 ms).

Analysis of accuracy rates revealed a main effect of modality type with auditory properties showing worse accuracy than visual properties [$F(1,19) = 13.81, p < 0.01$]. There were no significant effects of switch condition for the visual (switch = 0.92, SD = 0.07; no-switch = 0.94, SD = 0.05) nor auditory properties (switch = 0.86, SD = 0.09; no-switch = 0.87, SD = 0.09; $F_s < 1$) but both modality types showed slightly worse performance for the switch condition.

ERP RESULTS

Probe properties elicited ERPs typical of visually presented words, an N1–P2 complex followed by the N400 and a LPC. The switch manipulation did not affect ERP waveforms in the early 100–200 ms interval. The switch manipulation modulated the amplitude of the N400 (measured 200–500 ms post-stimulus) and the LPC (measured 500–800 ms), but did so differently for visual and auditory properties. Whereas visual properties elicited a larger N400 for switch than no-switch trials, auditory properties elicited a larger LPC for the same comparison.

100–200 ms

Analysis of ERPs measured 100–200 ms after stimulus onset did not show any differences for analyses of switch effects (all $F_s < 1$). Nor did it reveal differences based on the modality elicited by the properties (all $F_s < 1.4$).

200–500 ms

Overall analysis of ERPs measured 200–500 ms after stimulus onset revealed a significant interaction between the switch and the modality factors [$F(1,19) = 4.61, p < 0.05, MSE = 147.25$]. Follow-up analyses of each individual modality revealed no effects in the auditory modality ($F_s < 1$; auditory switch = 5.08 μV , auditory no-switch = 4.76 μV), but a reliable switch effect in the visual one [$F(1,19) = 4.93, p < 0.05, MSE = 135.52$]. The latter reflects the slightly more negative (0.7 μV) ERPs elicited in the visual switch (4.53 μV) than the visual no-switch (5.21 μV) condition (Figure 3). Although this difference showed up as a main effect in the analysis, visual inspection suggests it was largest over centro-parietal sites characteristic of the classic N400 effect (Figure 4).

500–800 ms

Overall analysis of ERPs measured 500–800 ms after stimulus onset revealed a significant interaction between the modality and the switch factors [$F(1,19) = 5.27, p < 0.05, MSE = 162.78$], as well as a marginal interaction between modality and electrode site [$F(28,532) = 1.81, p = 0.10, \epsilon = 0.20, MSE = 3.49$]. Follow-up analyses suggested the interaction between modality and switch results from a positive-going switch effect evident only for auditory properties. Separate analysis of the visual modality revealed no effect of the switch factor, either as a main effect ($F < 1$; visual switch = 6.00 μV ; visual no-switch = 6.22 μV), or in interaction with electrode site ($F < 1$). Separate analysis of the auditory modality suggested a trend for switch trials to elicit a slightly larger positivity (switch = 6.70 μV) than did no-switch trials [5.86 μV ; $F(1,19) = 3.02, p = 0.098, MSE = 201.31$; see Figures 5 and 6].

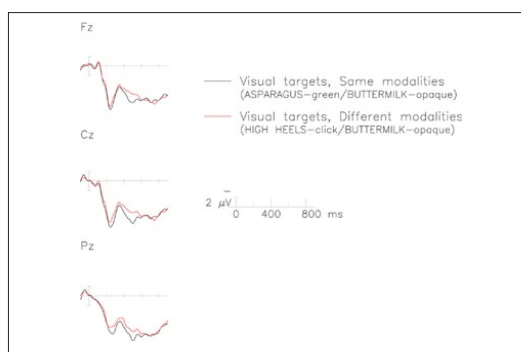


FIGURE 3 | The N400 elicited by visual property verification targets in the switch (red) and no-switch (black) conditions. Each graph represents data recorded from a midline electrode over frontal (top), central (middle), and parietal (bottom) regions of the scalp. Time is plotted on the x-axis against voltage on the y-axis. By convention, negative polarity is plotted upward.

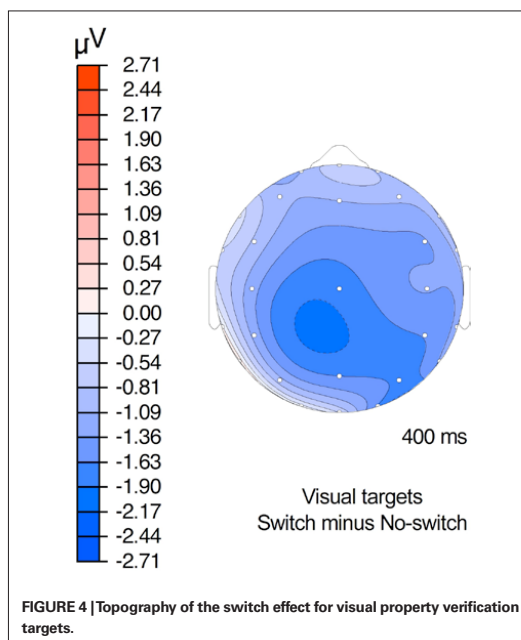
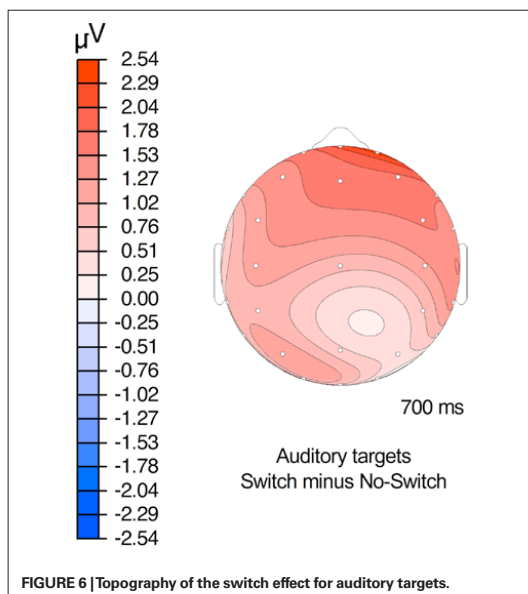
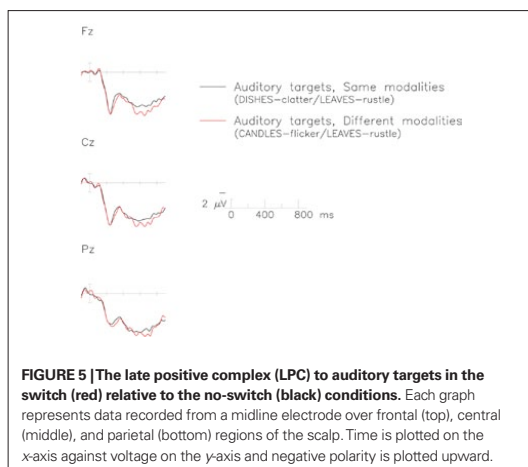
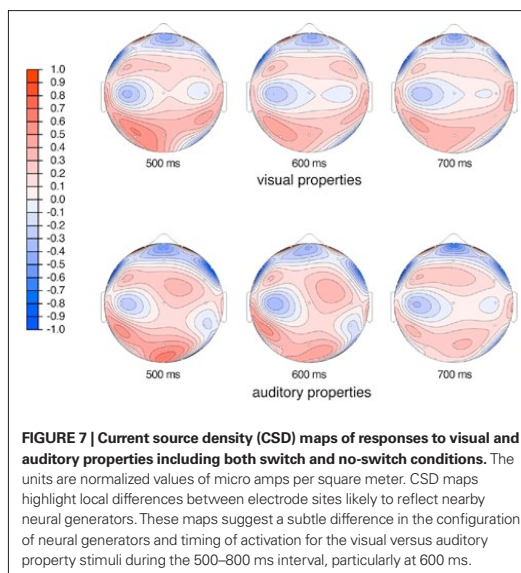


FIGURE 4 | Topography of the switch effect for visual property verification targets.

We also followed up on the marginal interaction between modality and electrode site as the possible topographic differences were of interest to our question of access to underlying perceptual modalities by property words. We tested midline, medial, and lateral sites separately. Our midline test included factors of modality (visual, auditory) and anteriority (seven midline electrodes, see Figure 2). This test revealed a marginal interaction between modality and anteriority [$F(6,114) = 2.67, p = 0.057, \epsilon = 0.45, MSE = 1.35$]. Our test of the medial sites



was similar and also included a factor of hemisphere (right, left). This test also revealed a difference between modalities that interacted with anteriority [$F(6,114) = 3.55, p < 0.05, \epsilon = 0.41$ MSE = 3.70], but no hemispheric differences were significant ($F_s < 1.8$). No differences at the lateral sites were observed ($F_s < 2$). The interaction effects between modality and scalp location can be seen in **Figure 7** with the current source density (CSD) plots. These figures show that the visual and auditory properties result in different patterns of voltage change during this time interval.



DISCUSSION

The present study investigated the electrophysiological correlates of the conceptual modality switch effect, an effect used to argue that conceptual tasks recruit perceptual processing systems. We predicted that the sequencing of property verification trials in same modality versus different modality pairs would be reflected in semantic processing of target properties, and thus would modulate the amplitude of the N400 component of the ERP. While this was indeed the case for the visual properties we tested, it was not the case for the auditory properties. Relative to the no-switch trials, visual properties in the switch condition elicited a larger negativity in the N400 time window; by contrast, auditory properties elicited a larger positivity 500–800 ms after stimulus onset in the switch condition. No early differences emerged for N1–P2 components arguing against the suggestion that the switch effect involves low-level visual processing.

N400 EFFECT

The first effect of interest was the negativity observed 200–500 ms after the onset of visual property terms. As predicted, no-switch trials elicited a smaller negativity than did the switch trials during a time interval typically associated with the semantic processing of words and the elicitation of the N400 component. Experts differ on the exact functional significance of this component, with some arguing it indexes lexical access (Kutas and Federmeier, 2000; Lau et al., 2008), and others contextual integration processes (e.g., Hagoort, 2008). There is widespread agreement, however, that the N400 indexes processing events associated with the construction of meaning, and, further, that its amplitude is related to processing difficulty (see Wu and Coulson, 2005 for a review). In general, contextual factors that facilitate processing lead to reduced amplitude N400; for example, words elicit smaller N400 when preceded by related than unrelated words, and smaller N400 when preceded by supportive than unsupportive sentence and paragraph contexts (see Kutas and Federmeier, 2011 for extensive review).

Results of the present study suggest that the perceptual modality of the property term on a previous trial can comprise a supportive semantic context, and that N400 priming effects can be observed between subsequent decisions disguised to participants as completely independent trials. The smaller negativity observed here for the no-switch trials thus suggests that semantic processing of visual target properties was facilitated by processing a visual prime property relative to an auditory prime property. We attribute this facilitated processing to the use of modality specific sensory simulations to mentally represent objects. While perceptual modalities are recruited automatically during concept processing in general, attention can focus more or less on specific modalities. In the property verification task, the presentation of a modality specific property can direct attention to the relevant modality. If the next trial has a property from a different modality (as in the switch condition) the focus shifts to a simulation in the newly relevant modality in order to represent the property. This shift incurs a processing cost which is evident in the ERP differences observed in the present study and reaction time differences of previous studies (Pecher et al., 2003).

Our results are consistent with those of a recent study by Hald et al. (submitted). Hald et al. (submitted) also used a modality switch paradigm in which they presented visual and tactile properties and obtained N400 differences between switch and no-switch trials. Thus, it seems that the N400 effect for modality switching is robust. The identification of the N400 as an ERP index of the conceptual modality switch effect suggests that the cost of shifting between modalities, in this case driven by visual property words, is reflected in semantic processes. This further implies that the semantic activation indexed by this ERP component includes the activation of perceptual features. Results of the present study are thus consistent with ERP studies that have demonstrated modulation of the N400 based on categorical relations that imply similar visual features (Federmeier and Kutas, 1999), and so-called perceptual priming between items such as *pizza* and *coin* that share a salient visual feature (Kellenbach et al., 2000). In sum, results of the present study are in keeping with an account of concepts as involving sensorimotor simulations (e.g., Barsalou, 1999) and suggest that the access of visual features occurs during meaning processing.

LPC EFFECTS

Two effects of interest were observed in the interval 500–800 ms after the onset of property terms. First, visual versus auditory properties elicited ERPs with subtle topographic differences (modality effects). Second, the switch manipulation modulated the ERPs to auditory but not visual property terms (modality switch effects).

Modality effects

Between 500 and 800 ms ERP patterns differed across midline and medial electrode sites for auditory versus visual property decisions. The positivity elicited by auditory properties was more frontocentrally focused than that elicited by visual properties. **Figure 7** illustrates this relatively subtle difference in the scalp topography particularly visible at 500 and 600 ms after stimulus onset. The CSD maps plot the second spatial derivative of the ERP waveforms, and as such highlight differences in the voltage recorded at adjacent electrode sites. The electrode montage used in the present study was too

sparse to allow localization but the observed scalp topography differences imply differences in the neural generators underlying the brain response to visual versus auditory property terms. These differences observed between visual and auditory processing are compatible with related fMRI studies that show areas of unique brain activity for properties describing different modalities (Goldberg et al., 2006). The timing of observed topographic differences is later than initial semantic activation implicated in the generation of the N400 component. Semantic and pragmatic manipulations have, however, been observed to modulate the amplitude of the ERP in this interval (see e.g., Regel et al., 2010 for a review). Differences in the brain response to visual and auditory properties are consistent with the hypothesis that perceptual networks help subserve the neural representations of concepts, and the corollary that such networks would be different for concepts that predominantly activate one perceptual modality over another.

Modality switch effects

The other effect of interest in the present study was a positive deflection of the LPC for auditory switch trials relative to the auditory no-switch trials between 500 and 800 ms, primarily at anterior electrode sites (see **Figure 6**). This effect is likely related to the P3, a family of ERP components that index memory processing, whose amplitude reflects the allocation of attention, and whose latency is proportional to the task-relevant stimulus evaluation process (see Polich, 2007 for a review). In view of the relatively long reaction times on the property verification task (>900 ms), the timing of the late positivity observed in the present study (500–800 ms after the onset of the auditory property term) is consistent with its interpretation as an index of the property verification decision. In studies of the P3, the same target stimulus has been shown to elicit a larger positivity in the ERP in difficult than in easy discrimination tasks (Comerchero and Polich, 1999). On this interpretation, the larger late positivity on the switch trials suggests the auditory property verification judgments were more difficult when preceded by a visual prime trial than another auditory one. Alternatively, the anterior distribution of the LPC switch effect suggests the predominance of the P3a sub-component associated with attentional orienting to novel stimuli (see Polich, 2007 for review). On this interpretation, the larger late positivity we observed need not imply greater processing difficulty, but rather an appreciation of the switch trials as involving more novelty than the no-switch trials – presumably because the switch trials required participants to activate semantic features from a different modality.

Hald et al. (submitted) also found a positivity for switch items elicited by a conceptual modality switch task but only over posterior electrodes, differing from the distribution described here (**Figure 6**). Their finding of a posterior positivity co-occurred with a larger negativity for switch trials over anterior electrodes in the same time intervals. The timing and scalp distribution of these effects were interpreted as a unified frontal N400 effect similar to that elicited by pictures. The different ERP patterns found by Hald et al. (submitted) at anterior and posterior electrode sites were revealed as a topography difference but this scalp difference cannot be compared to that reported in the current study because the topographic differences reported here were driven by different modalities, a dimension not tested by Hald et al. (submitted).

DIFFERENCES BETWEEN VISUAL AND AUDITORY PROPERTY VERIFICATION

The most surprising result of the present study was the observed difference in the conceptual modality switch effect for visual versus auditory properties. As noted above, visual properties elicited reduced N400 in no-switch relative to switch trials, suggesting our experimental manipulation affected semantic processing of the targets. Auditory properties, however, elicited an enhanced LPC, suggesting the manipulation impacted neural processes occurring later than those indexed by the N400, and were more likely related to making the decision about whether the property was typical of the concept.

Whereas neither finding is surprising alone – that is, a conceptual modality switch might reasonably be predicted to impact either the semantic processing of the stimuli, or the difficulty of decisions regarding property verification, or, indeed, both sets of processes – our finding of semantic effects for visual properties and decision-related effects for auditory properties was unexpected. Prior reports of the conceptual modality switch effect using reaction time measures have found similar sized switch effects for properties from different modalities (Pecher et al., 2009). Similarly, studies of the perceptual modality switch effect also report similar sized switch effects for visual, auditory, and tactile stimuli, with the only difference being a trend for tactile primes to yield longer reaction times for subsequent visual and auditory probes (Spence et al., 2001). However, reaction times measure only the end point of a property verification process, while ERPs provide an index of brain activity from the onset of the stimulus until the generation of the behavioral response on the task. ERP data in the present study suggest the switch manipulation affects different aspects of processing in the verification of visual versus auditory properties.

Our observed differences between auditory and visual switch effects are consistent with a prior ERP study of the perceptual modality shift effect by Gondan et al. (2007) in which stimuli involved either LED flashes (visual targets) or bursts of white noise (auditory targets). They found that visual targets following visual primes compared to when they followed auditory primes elicited ERP effects similar to those found for increased visual attention – namely, an amplified N1 component. In contrast, auditory targets elicited smaller N1 and P2 components when they followed auditory primes than when they followed visual primes. The fact that ERP differences for the switch effect were opposite in the visual and auditory domains was an unexpected asymmetry. The authors explain this asymmetry by suggesting different mechanisms driving the switch effects in the two perceptual domains. They suggest a “neural trace” explanation for the auditory domain in which residual activity from an auditory prime speeds the response and processing for a subsequent auditory stimulus. The result of this priming is a smaller ERP component for the target auditory stimulus. For the visual targets, ERP amplification for the same modality condition is explained through attentional mechanisms because increased attention tends to result in amplified perceptual ERP components. These different patterns suggest that different mechanisms might be driving the modality switch effect in the visual and auditory domains. Likewise, results of the present study suggest that different mechanisms were involved for the conceptual modality switch in the case of visual versus auditory property terms.

One account for why different mechanisms would drive the conceptual switch effects in the present study is that the particular visual and auditory property words we used access the perceptual domains differently. In particular, the visual property words may refer to relatively pure visual experiences, whereas auditory properties may refer to mixed visual and auditory experiences. For example, green (as for asparagus) might refer to a purely visual perception while clicking (as for high heels) might refer to a combined auditory and visual experience. We examined this possibility using the Lynott and Connell (2009) norms. Lynott and Connell (2009) asked participants to what extent each of 423 property words were experienced via each of the five sensory modalities. Of the 48 property words used in each modality category of our study, 37 visual properties and 27 auditory properties were represented in their list. Our subset of visual property words had an average visual ranking of 4.65 (out of 5.0 possible) and the subset of auditory words had an average auditory ranking of 4.60 (two-tailed t -test, $t < 1$), verifying the experimental conditions used in our study.

However, when considering both visual and auditory rankings for each of these sets our auditory properties appear more multimodal than our visual properties as indicated by a smaller difference between their auditory and visual rankings (auditory property difference = 2.44, visual property difference = 4.18; $t(41) = 8.99$, $p < 0.01$). This classification is also consistent with a *modality exclusivity* score available in Lynott and Connell's (2009) norms. For each property word the modality exclusivity score factors the strength of the rating for an individual modality relative to ratings for all five sensory modalities. The visual properties used in our study had a higher modality exclusivity ranking (0.73, of possible values between 0.0 and 1.0) than our auditory properties [0.58; $t(59) = 4.31$, $p < 0.01$].

While it is clear that our auditory properties are characterized as typically experienced via hearing [as indicated by values derived from the Lynott and Connell (2009) norming study], their greater multimodal characteristic might have led to a weaker switch effect than that seen for the visual properties. In perceptual studies of the modality switch effect, a bimodal target stimulus (e.g., simultaneous beep and flash) following a unimodal stimulus (e.g., a flash) produces a smaller switch cost than unimodal targets following unimodal primes (e.g., a beep following a flash; Gondan et al., 2004). The reduction in the switch effect is presumed to result because only one of the two modalities making up the bimodal target stimulus requires a switch from the modality of the previous stimulus; the other, in fact, is primed. The absence of an observed N400 effect in our ERP results for auditory properties could reflect a lack of power to see an attenuated modality switch for these auditory properties that are more multimodal than the visual properties for which we did find an N400 effect. The decision-related LPC effect on the other hand would thus index more effort required to attribute the multimodal (auditory) property to a concept in the context of a visual prime.

Using a combination of published norms and dictionary definitions, we identified four of the visual target properties and eight of the auditory target properties employed in our study as being multimodal, that is, either having a modality exclusivity score (as defined by Lynott and Connell, 2009) of less than 0.51, or a dictionary definition that mentioned more than one modality. We elimi-

nated multimodal items from ERP waveforms and conducted a *post hoc* analysis of ERPs elicited by the remaining unimodal stimuli. In the 200–500 ms window the same pattern of significant effects was obtained as for the complete dataset. Reanalysis restricted to unimodal items thus suggested the N400 switch effect for visual properties was robust and slightly larger than that measured for the full set of experimental items, but still failed to reveal an N400 switch effect for the auditory properties.

A similar reanalysis of the LPC interval failed to reveal either the modality by switch interaction ($F < 1$) or the auditory switch effect ($F < 1$) observed in the original analyses. This raises the possibility that the LPC switch effect observed in the present study primarily reflects the brain response to the multimodal items. Consistent with this, further analysis also suggested a trend for multimodal visual and multimodal auditory properties to elicit slightly more positive ERPs in the 500–800 ms interval than unimodal visual [$1.13 \mu\text{V}$, $F(1,19) = 3.57$, $p = 0.074$, $\text{MSE} = 370.59$] and unimodal auditory [$0.84 \mu\text{V}$, $F(1,19) = 2.86$, $p = 0.107$, $\text{MSE} = 205.51$] properties, respectively. According to this interpretation, the auditory switch effect observed in our original analysis reflects the greater difficulty of responding to multimodal auditory properties following (more likely unimodal) visual than (more likely multimodal) auditory primes. The greater multimodality of the auditory properties also suggests an alternative explanation for the different topography of ERPs elicited by auditory and visual properties measured 500–800 ms post-stimulus onset (discussed in “Modality Effects”). These reanalyses must, however, be interpreted with caution since the comparison of unimodal versus multimodal stimuli involve ERPs derived from different numbers of trials, and the number of visual multimodal trials was particularly low. More definitive conclusions

regarding brain activity elicited by multimodal versus unimodal items in a property verification task would require a stimulus set specifically designed for this purpose.

CONCLUSION

The present study contributes to evidence demonstrating that concepts referring to perceptual properties are recruiting perceptual processing resources. Whereas previous studies have shown similar modality switch effects in conceptual processing, the present study informs us in a more detailed way on the locus of this switch effect. ERP measures showed that the elicitation of perceptual meaning, as typically demonstrated by switching costs, is evident at the semantic level or at later decision-making stages of processing. The switch effect for visual properties was different from the switch effect for auditory properties due to either different underlying mechanisms driving the processes or different modal representations of these properties. Both explanations support a theory of concepts as a reactivation of brain areas important for the perception of the world. Just as seeing candles flicker generates different neural activity from hearing high heels click, we expect the concepts representing these events to differ as well.

ACKNOWLEDGMENTS

This research was supported by grants from the Netherlands Organization for Scientific Research (NWO) to Diane Pecher and from the Erasmus University Trustfonds to Diane Pecher and René Zeelenberg. We would like to thank Lauren Cardoso, Nafes Hamid, and Rubén Moreno for their assistance collecting data. The authors would also like to thank Lea Hald and Frederico Marques for their feedback and reviews that allowed us to improve this paper.

REFERENCES

- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., and Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* 7, 84–91.
- Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain Lang.* 31, 308–327.
- Comerchero, M. D., and Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clin. Neuropsychol.* 110, 24–30.
- Federmeier, K. D., and Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. *J. Mem. Lang.* 41, 469–495.
- Goldberg, R. F., Perfetti, C. A., and Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *J. Neurosci.* 26, 4917–4921.
- Gondan, M., Lange, K., Rösler, F., and Röder, B. (2004). The redundant target effect is affected by modality switch costs. *Psychon. Bull. Rev.* 11, 307–313.
- Gondan, M., Vorberg, D., and Greenlee, M. W. (2007). Modality shift effects mimic multisensory interactions: an event-related potential study. *Exp. Brain Res.* 182, 199–214.
- Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 363, 1055–1069.
- Holcomb, P. (1988). Automatic and attentional processing: an event-related brain potential analysis of semantic priming. *Brain Lang.* 35, 66–85.
- Holcomb, P., and Neville, H. (1990). Semantic priming in visual and auditory lexical processing. *Lang. Cogn. Process.* 5, 281–312.
- Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., and Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cogn. Neuropsychol.* 20, 525–540.
- Kellenbach, M. L., Wijers, A. A., and Mulder, G. (2000). Visual semantic features are activated during the processing of concrete words: event-related potential evidence for perceptual semantic priming. *Cogn. Brain Res.* 10, 67–75.
- Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470.
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647.
- Kutas, M., and Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307, 161–163.
- Lau, E. F., Phillips, C., and Poeppel, C. (2008). A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933.
- Lynott, D., and Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behav. Res. Methods* 41, 558–564.
- Marques, J. F. (2006). Specialization and semantic organization: evidence for multiple semantics linked to sensory modalities. *Mem. Cogn.* 34, 60–67.
- Martin, A. (2007). The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45.
- Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.
- Pecher, D., Van Dantzig, S., and Schifferstein, H. N. J. (2009). Concepts are not represented by imagery. *Psychon. Bull. Rev.* 16, 914–919.
- Pecher, D., Zeelenberg, R., and Barsalou, L. W. (2003). Verifying properties from different modalities for concepts produces switching costs. *Psychol. Sci.* 14, 119–124.
- Pecher, D., Zeelenberg, R., and Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: modality-specific effects of prior activation. *Psychon. Bull. Rev.* 11, 164–167.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148.
- Regel, S., Coulson, S., and Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the

- comprehension of irony. *Brain Res.* 1311, 121–135.
- Rodway, P. (2005). The modality shift effect and the effectiveness of warning signals in different modalities. *Acta Psychol.* 120, 199–226.
- Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., and Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 45, 2802–2810.
- Smith, M. E., and Halgren, E. (1989). Dissociation of recognition memory components following temporal lobe lesions. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 50–60.
- Solomon, K., and Barsalou, L. W. (2004). Perceptual simulation in property verification. *Mem. Cogn.* 32, 244–259.
- Spence, C., Nicholls, M. E. R., and Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Percept. Psychophys.* 63, 330–336.
- Urbach, T. P., and Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology* 39, 791–808.
- van Dantzig, S., Pecher, D., Zeelenberg, R., and Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cogn. Sci.* 32, 579–590.
- Vermeulen, N., Niedenthal, P. M., and Luminet, O. (2007). Switching between sensory and affective systems incurs processing costs. *Cogn. Sci.* 31, 183–192.
- Wu, Y. C., and Coulson, S. (2005). Meaningful gestures: electrophysiological indices of iconic gesture comprehension. *Psychophysiology* 42, 654–667.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 16 July 2010; paper pending published: 29 July 2010; accepted: 09 January 2011; published online: 08 February 2011.
- Citation: Collins J, Pecher D, Zeelenberg R and Coulson S (2011) Modality switching in a property verification task: an ERP study of what happens when candles flicker after high heels click. *Front. Psychology* 2:10. doi: 10.3389/fpsyg.2011.00010
- This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.
- Copyright © 2011 Collins, Pecher, Zeelenberg and Coulson. This is an open-access article subject to an exclusive license agreement between the authors and Frontiers Media SA, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

Acknowledgement

Chapter 2, in full, is a reprint of the material as it appears in *Frontiers in Psychology*; Jennifer Collins, Diane Pecher, René Zeelenberg and Seana Coulson, 2011. The dissertation author was the primary investigator and author of this paper.

Chapter 3

An ERP study of the impact of modality on the conceptual modality switch effect

Abstract

The conceptual modality switch effect (Pecher, Zeelenberg & Barsalou, 2003) supports the proposal that concepts involve the reactivation of perceptual experiences (Barsalou, 1999). In a typical conceptual modality switch experiment, each individual trial is a property verification task requiring participants to determine whether a property (e.g., *rough*) applies to a concept (e.g., *STUCCO*). The critical finding is that the second of two property verification trials suffers a processing cost when it follows a trial from a *different* modality (e.g. sound then touch) compared to when it follows a trial from the *same* modality (e.g. touch then touch). Here we recorded event-related potentials (ERPs) as participants performed a property verification task; the critical trials involved auditory and tactile properties. The modality switch effect elicited a late frontal negativity and posterior positivity. Tactile properties in both switch and no-switch conditions elicited a robust N400 effect reminiscent of a concreteness effect. Results are discussed in terms of their implications for theories of grounded cognition, especially regarding the underlying representation of concepts from different modalities.

Introduction

The human brain supports a powerful conceptual system that allows for a rich, organized understanding of the external world. One way we access these concepts is via words such as, *crunchy green apple*. This particular phrase should evoke the concept of an apple as well as its specific color and texture, the details of which are based on life experience with apples. For example, the color suggested in this phrase is not a prototypical, bold green but rather a lighter green (more like chartreuse) characteristic

of green apples. Grounded cognition is a general theoretical approach to concepts shared by several researchers which emphasizes the critical role of experience in conceptual representations (Barsalou, 1999; Glenberg, 1997; Pulvermüller, 1999). Perceptual Symbol Systems (Barsalou, 1999), for example, posits the construct of *simulators*, schematic simulations of perceptual experience that constitute conceptual representations. Activation of simulators takes place in regions of the brain active during the initial experience with the concept's referent, including systems of perception (e.g. vision, audition, somato-sensation, etc.), motor control, proprioception and introspection (Barsalou, 1999). According to this proposal, the concept for *apple* is built from perceptual experiences with apples such as holding them, looking at them and biting into them.

The far-reaching predictions for perceptual symbols have been tested in numerous ways (see Barsalou, 2008; Fischer & Zwaan, 2008 for reviews). With property generation tasks, for example, the object *half watermelon* leads to more internal descriptions, e.g. *seeds, red, or juicy*, than for *watermelon*, which leads to the description of more external features (Wu & Barsalou, 2009). The finding that a modifier can influence participants to think of typically occluded properties of an object supports the claim that we activate concepts as a simulation of our perceptual (in this case, visual) experiences. Similarly, shape and orientation are visual dimensions activated during normal sentence comprehension even when not explicitly mentioned (Stanfield & Zwaan, 2001; Zwaan, Stanfield & Yaxley, 2002). Hand shapes and arm movements are implicitly active during the comprehension of sentences and are particular to the actions described. The sentence, *John gave the calculator to Mary* primes a reaching

arm movement (Glenberg & Kaschak, 2002), and a hand shape for holding a calculator (Bub & Masson, 2010). The production of a smiling facial expression speeds both a valence and sensibility decision for the positive sentence, *You and your love embrace after a long separation*, over the negative sentence, *The police car rapidly pulls up behind you, siren blaring* (Havas, Glenberg & Rinck, 2007). By demonstrating how readily modal features are accessed during sentence comprehension, these studies provide broad support that the modalities of vision, motor control and emotion are part of the multimodal representations underlying our concepts.

Building from the assumption that concepts are multimodal simulations, we suggest linguistic context can direct attention to particular perspectives of the multimodal concept representation. Considering again the example of *a crunchy green apple*, this phrasing directs attention toward its tactile and visual dimensions and away from its taste, for example. The *conceptual modality switch effect* is an experimental demonstration supporting this interpretation of concept activation. The finding first demonstrated that a difference in the modality referenced by two subsequent conceptual decisions begets a cost relative to two subsequent decisions referring to the same modality (Pecher et al., 2003). The decisions in these experiments are property verifications in which a “property” (e.g. *crunchy*), might or might not refer to a “concept” (e.g. *APPLE*). For example, the property verification, *APPLES-crunchy* (for which participants should respond “true”), draws attention to the tactile dimension of the concept *APPLE*. (Note that the properties in this type of study are designed to evoke a particular perceptual modality.) If the next decision were *SPONGES-soft* (also a tactile

decision) the response would generally be faster than if the next decision were *COINS-jingle* (an auditory decision) because the latter recruits a different modality.

As verification for the conceptual modality switch effect, numerous replications show that the reaction time to the second of two property verification trials is longer when the two trials draw attention to different modalities than when they draw attention to the same modality (e.g., Pecher et al., 2003; Pecher, Zeelenberg & Barsalou 2004; Lynott & Connell, 2009; Marques, 2006). This pattern is found when including the modalities of positive and negative affect (Vermeulen, Niedenthal & Luminet, 2007), it is observed in the behavior of children (Ambrosi, Kalenine, Blaye & Bonthoux, 2011), it can be elicited with novel concepts (e.g. *jingling onion*, Connell & Lynott, 2011), and the influence of modality extends over time (Pecher et al., 2004). The conceptual modality switch effect is not modulated by mental imagery abilities (Pecher, van Dantzig & Schifferstein, 2009) and cannot be explained either by word associations (Pecher et al., 2003) or by the category membership of the test items (Marques, 2006). These negative findings verify the stance that the influence of modality on these conceptual decisions is driven by the perceptual features described by these words, and is not a byproduct of another process. An important finding supporting the claim that concepts are grounded in perceptual systems rather than activating them indirectly shows the switch effect between conceptual and perceptual stimuli (van Dantzig, Pecher, Zeelenberg & Barsalou, 2008). For example, responses are slower for a beep following the tactile verification, *SPONGES-soft* than for a beep following the auditory verification, *COINS-jingle*.

The first conceptual modality switch experiment was motivated by a finding in the perceptual literature analogously demonstrating that the detection of perceptual targets from the visual, auditory or tactile modalities (the flash of an LED, a burst of white noise, or the press of a rod on the forefinger, respectively) were slowed when preceded by a target from a different modality (Spence, Nicholls & Driver, 2001). The attentional system has been regarded as the driving force for these effects, as attention is required to redirect resources to a different perceptual modality (Spence et al., 2001; Rodway, 2005). In fact, an ERP study (Töllner, Gramann, Müller, & Eimer, 2008) compared the switch effect between visual and tactile targets showing that they both elicited early (140-180 ms post-stimulus onset) ERP modulation. The authors claimed that the switch effect was generated by a general purpose attention process that weights attention toward one modality or another and thus requires extra processing when switching between modalities. The consequence of this attentional mechanism is to prepare the relevant perceptual system. The preparation of the different perceptual systems can lead to modality-specific processing after the switch as well (Gondan, Vorberg & Greenlee, 2007). For example, Gondan et al. (2007) found that visual switch conditions elicited attenuated perceptual ERPs whereas auditory switch conditions elicited amplified perceptual ERPs.

The key elements of this mechanism are the attentional mechanism allocating focus between modalities and the perceptual system (or systems) in focus. In regards to the analogous conceptual mechanism, the concept *COINS* can elicit dimensions such as the motor experience of holding and tossing coins in one's hand, their color, shape, or the sound they make when clinking together. According to the perceptual

symbols view, each of these aspects of this concept should have an underlying representation grounded in the particular perceptual modality. The property verification task then serves to focus attention on a particular perceptual modality. The conceptual modality switch effect is the behavioral observation of the underlying perceptual activation in focus for a particular concept derived from the attentional process required to switch to a different perceptual activation.

The suggestion that the property verification task directs processing to relevant perceptual regions is corroborated by cognitive neuroscience work. Property verifications made for different modalities lead to relatively more brain activation in areas related to processing the related perceptual modality (Goldberg, Perfetti & Schneider, 2006; see also Martin, 2007 for a review). Visual identifications result in amplified activation in fusiform gyrus, a brain area important for visual processing (Kan, Barsalou, Solomon, Minor & Thomspen-Schill, 2003). In another study, tactile properties were associated with somatosensory and motor activation, gustatory properties were associated with orbitofrontal activation, auditory properties were associated with superior temporal sulcus activation and visual properties were associated with left ventral temporal activation (Goldberg et al., 2006). The recent inclusion of event-related potentials (ERPs) to studies of conceptual modality switching has provided supporting evidence that visual and auditory switch patterns differ (Collins, Pecher, Zeelenberg & Coulson, 2011) and has also suggested that the conceptual switch is related to semantic and/or decision making processes (Collins et al., 2011; Hald, Marshall, Janssen & Garnham, 2011).

The present study aims to replicate these previous ERP studies, and to clarify discrepancies in their findings. ERPs are used to address questions of cognitive processes based on timing, polarity and scalp distribution differences between experimental task conditions. The observed differences can be compared with functionally-specified ERP components and lead to hypotheses about the processing stage(s) at which the modality switch is manifested. For example, in studies of semantic processing, the N400 is a frequently targeted ERP component reflecting the intersection of current, meaningful stimulus analysis (e.g., words, pictures, sounds) with access to prior context, both recent and long-term (see Kutas & Federmeier, 2011 for a recent review). Its modulation is associated with semantic processing of stimuli in many modalities. One of its striking characteristics is the similarity of ERP patterns evoked by stimuli of different modalities (e.g. visual words, auditory words, pictures, gestures). There are, however, some variations when elicited by stimuli of different modalities with somewhat earlier onset and frontal distributions observed for semantic processing involving pictures (Ganis, Kutas & Sereno, 1996; Kutas & Federmeier, 2011).

In the first ERP study testing switch effects for visual and auditory property verification trials, we found an N400 switch effect elicited by the second property of a pair of trials (Collins et al., 2011). This effect was observed only for visual properties, i.e., auditory-visual trial pairs elicited a larger negativity than visual-visual trial pairs. The N400 finding for the conceptual modality switch effect was interpreted to mean that the perceptual dimensions driving this effect are an inherent part of semantic, long-term memory. Conceptual modality switch conditions also elicited a late positive

complex (LPC) when restricted to auditory properties (Collins et al., 2011). The LPC is a family of ERP effects that index updating mental context such as in attentional allocation and decision making (Polich, 2007). The late positivity was interpreted as indexing greater difficulty in making judgments about the auditory properties. The fact that the LPC and the N400 were modulated by the switch manipulation for different modalities led to the conclusion that auditory and visual concepts have different underlying semantic distributions as demonstrated with fMRI (Goldberg et al, 2006). Also in support of this claim, visual and auditory property decisions in a late window (500-800 ms) showed different scalp topographies suggesting a timing or location difference in the neural assemblies responding to those experimental trials.

In an ERP study of the conceptual modality switch effect using visual and tactile properties, Hald et al. (2011) found a frontal negativity for switch trials in early (160-215 ms; 270-370 ms) and late (500-700 ms) intervals. They compared the timing and distribution of their pattern to an N400 effect to pictures as meaningful completions of sentences (Ganis et al., 1996). Hald et al. (2011) suggested that their early, frontal N400 effect likely indexed a similarity between the underlying representations of property words and visual images. The findings of the two ERP studies, while both implicating semantic processing for the conceptual modality switch, differ in the type of N400 associated with this effect. In one case it was observed with a centroparietal topography associated with the classic N400 (Collins et al., 2011); in the other case it was found with a frontal distribution related to picture or imagery processing (Hald et al., 2011). Further clarifications are needed to explain the distribution differences.

The present study employed the method of Collins et al. (2011) using auditory and tactile properties. This design allowed for a new pairing of conceptual modalities as a way to further assess ERP differences arising from the hypothesized perceptual substrates of concepts. Our primary prediction was that switch trials would elicit more negative ERPs than no-switch trials as in the two prior ERP studies of this phenomenon and consistent with the claim that the perceptual activations underlying the conceptual modality switch effect play a role in semantic processing. Alternatively, auditory switch trials might elicit an LPC, replicating the findings of Collins et al. (2011), suggesting the switch to auditory properties requires broadly different neural patterns than does the switch to visual or tactile properties. A supplementary prediction of the Perceptual Symbol System hypothesis was that the comprehension of auditory and tactile property terms requires the activation of auditory and somato-sensory cortex, respectively, and that this difference might be detectable in the ERPs as subtle differences in the scalp topography of the N400 component elicited by auditory versus tactile properties.

Material and Methods

The protocol for this study was approved by the University of California, San Diego Social and Behavioral Science Institutional Review Board. As such, informed consent was obtained from all participants prior to their enrollment.

Participants

Twenty undergraduates from the UCSD community (10 women) participated as part of a course requirement. Data from five additional participants were not included in the analysis due to the presence of an excessive number of artifacts (greater

than 30% of trials in a critical condition). As reported in a screening questionnaire, all participants had normal vision, and none had any history of neurological or psychiatric disorders within the previous ten years.

Materials

Each trial in the study consisted of a concept-property combination such as *COINS-jingle*. Experimental trials included 48 tactile properties (such as *smooth*), and 48 auditory properties (such as *jingle*). These were repeated once during the experiment with different concepts – all concept-property pairs in the experiment were unique. The only methodological difference from the study by Collins et al. (2011; see Materials and Methods for full details on the organization of experimental trials and list arrangement) was that here we substituted tactile concept-property trials for visual concept-property trials in experimental conditions. The 192 experimental trials were organized into 96 trial pairs in which the modality of the target property was crossed with the modality switch dimension (24 pairs each: auditory prime/auditory target; tactile prime/auditory target; tactile prime/tactile target; auditory prime/tactile target).

The experiment also included 384 filler trials. Tactile and auditory modalities were each represented by 96 filler items eliciting “false” responses (e.g. *PEANUTS-thorny*). These were included so that the experimental items (all “true” trials) would be balanced with “false” items from the same modality. Visual trials were used as filler items, half of which were valid for their concept (e.g. *CROWS-ebony*), and half of which were not (e.g. *ELEPHANTS-glossy*). Associated trials were also used as filler items used to dissuade participants from using word association strategies (e.g. *BALLERINAS-dance*, “true”; *STRAWBERRIES-cream*, “false”). Overall there were four

types of items in the experiment: 192 auditory, 192 tactile, 96 visual and 96 associated, with half of each of these categories requiring “true” responses on the property verification task.

Procedure, EEG recording and analysis

Participants were seated in front of a 17 in computer screen in a dimly lit room insulated to limit external sounds and electrical noise. After receiving instructions and practice, participants watched 10 blocks of experimental trials taking 3-4 min each. Trials began with the presentation of a fixation cross for 250 ms to prepare participants thereby limiting the number of eye movements and blinks in critical trials. At the offset of the fixation cross a randomly selected inter-stimulus interval (ISI) between 200 and 400 ms preceded presentation of the concept which stayed on the screen for 150 ms. After a 250 ms ISI the property was presented for 200 ms, at which point participants could respond “true” or “false”. Participants then had 2600 ms to make their response and get ready for the next trial. “True” responses were indicated by a button press with the right hand and meant that the property was highly typical of the concept. “False” responses were made with the left hand and meant the property was not typical of the concept. Figure 3.1 summarizes the experimental timing.

Participants’ electroencephalogram (EEG) was recorded with tin electrodes mounted in an electrode cap with 29 scalp sites (for procedural details and EEG processing techniques see Materials and Methods-EEG Recording and Analysis of Collins et al., 2011). Our standard analysis was a 2x2x29 repeated measures ANOVA with the factors switch (switch/no-switch), target property modality (visual/auditory),

and electrode site (29 levels) for the time intervals of interest: 100-200 ms, 200-500 ms and 500-800 ms.

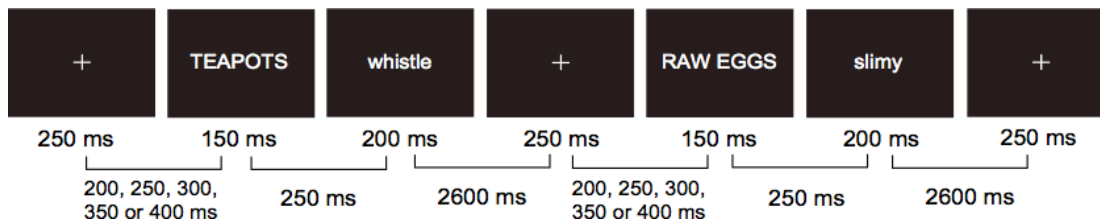


Figure 3.1. Trial arrangement for the conceptual modality switch. Participants were presented with concept-property pairs that were either auditory or tactile, as the two examples here illustrate. For each trial participants responded “true” or “false” depending on whether the property was typical of the concept. In both trials illustrated here the correct response was “true”. The pair of trials presented here would fall into the “switch” condition because the properties evoke different modalities (auditory and tactile). ERPs were time-locked to the second property in a pair of trials.

Results

Reaction times

Analysis of reaction times revealed a significant effect of modality with tactile properties eliciting longer reaction times (switch = 989 ms, SD = 172 ms; no-switch = 965 ms, SD = 134 ms) than auditory properties (switch = 925 ms, SD = 150 ms; no-switch = 908 ms, SD = 134; $F(1,19) = 23.40$, $p < 0.0001$). The pattern of reaction times was in the expected direction for the modality switch effect but this difference was not statistically significant (switch v. no-switch, $F(1,19) = 2.54$, $p = 0.13$). Nor was there a significant interaction between switch and modality ($F < 1$).

Accuracy rates did not show any significant differences (all F s < 1), although numerically the switch conditions yielded slightly lower accuracy scores (auditory switch = 0.87, SD = 0.08; tactile switch = 0.88, SD = 0.09) than the no-switch conditions (auditory no-switch = 0.89, SD = 0.07; tactile no-switch = 0.89, SD = 0.09).

ERP Results

Target properties elicited ERPs typical of visually presented words, an N1-P2 complex followed by the N400 and a late positive complex (LPC). The switch manipulation affected responses 500-800 ms post stimulus onset but was limited to particular scalp locations. In the 200-500 ms window, the typical N400 window, we found a larger negativity for tactile properties than auditory properties but no modulation by switch condition.

100-200 ms: Analysis of ERPs measured 100-200 ms after stimulus onset did not reveal significant effects of modality or switch. There was, however, a trend for a difference between switch and no-switch trials ($F = 3.76$, $p = 0.067$, $MSe = 170.50$) driven by a more positive waveform for switch trials (1.46 microvolts) than for the no-switch condition (0.91 microvolts), (all other F s < 2.5).

200-500 ms: Overall analysis of ERPs measured 200-500 ms after stimulus onset revealed a significant effect of modality ($F(1,19) = 8.63$, $p < 0.01$, $MSe = 435.72$) reflecting less positive (more negative) ERPs elicited by tactile (2.97 microvolts) than auditory (3.84 microvolts) targets. We also found an interaction between modality and electrode ($F(28,532) = 2.87$, $p < 0.05$, $\epsilon = 0.16$, $MSe = 3.52$, see Figure 3.2).

Follow-up analyses on the modality by electrode interaction were done with tests of midline, medial and lateral sites separately to identify the source of the interaction. Midline and medial tests both revealed only a main effect of modality reflecting the larger negativity for tactile properties but no further interactions between modality

and electrode to reveal scalp differences ($F_s < 2$). Tests at the lateral sites revealed, again, a main effect of modality ($F(1,19) = 5.02, p < 0.05, MSe = 22.20$) as well as a marginal interaction of modality and hemisphere ($F(1,19) = 3.15, p = 0.09, MSe = 4.89$) due to a larger negative going response to tactile properties on the right side of the head. This rightward distribution of the negativity for tactile items can be seen in Figure 3.3.

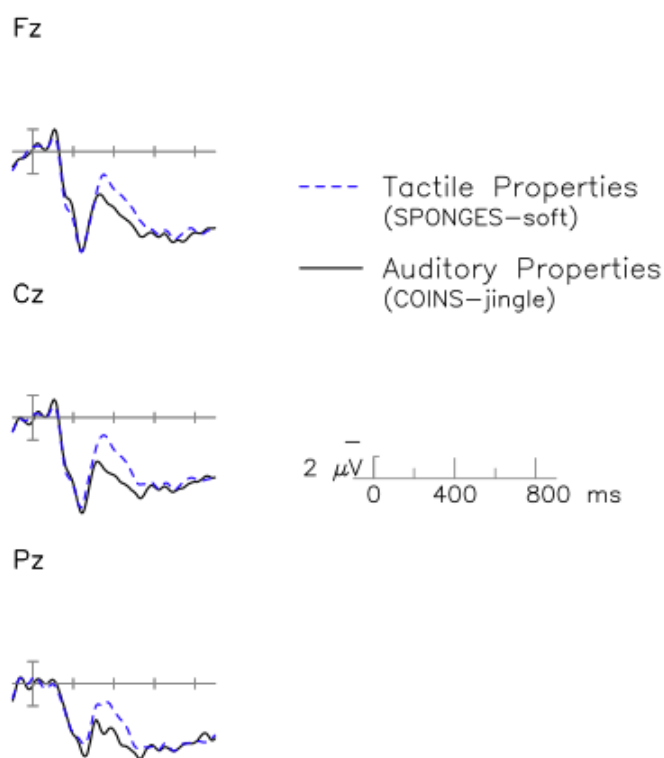


Figure 3.2. Tactile N400 effect. The N400 effect time-locked to target properties in the tactile (blue, dashed) and auditory (black, solid) conditions at three representative midline electrodes. The conditions switch and no-switch are combined. The tactile targets elicit a larger negativity relative to the auditory target properties. Time is plotted on the x-axis against voltage on the y-axis. By convention, negative polarity is plotted upwards.

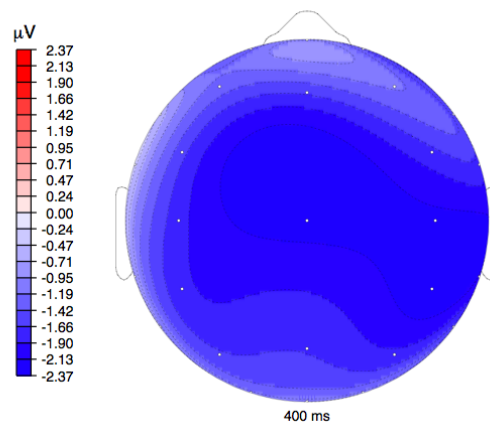


Figure 3.3. Topography of the tactile N400 effect. This modality difference showed a greater negativity for tactile properties than auditory properties with a broad central scalp distribution and a somewhat anterior distribution from the typical centroparietal distribution and a rightward skew.

500-800 ms: Overall analysis of ERPs measured 500-800 ms after stimulus onset revealed a significant interaction between the switch condition and electrode site ($F(28, 532) = 2.56, p < 0.05, \epsilon = 0.24, MSe = 2.49$). No other effects were observed (all $F_s < 1.5$).

To investigate the interaction between modality switch and electrode we did analyses at midline, medial and lateral electrodes in separate tests. No differences for the switch condition were observed at midline ($F_s < 2$) or lateral sites ($F_s < 1.5$). At the medial sites the interaction between the switch effect and anteriority was statistically significant ($F(6,114) = 3.41, p < 0.05, \epsilon = 0.49, MSe = 2.64$). At 6 anterior sites, three on either side of the head (FP1, F3, FC3, FP2, F4, FC4), there was a marginal switch by electrode interaction ($F(2, 38) = 2.55, p = 0.08, \epsilon = 0.59, MSe = 2.55$). At 6 posterior sites, three on either side of the head (CP3, P3, O1, CP4, P4, O2), there was a significant three-way interaction of switch, hemisphere and electrode ($F(2,38) = 6.03, p < 0.01, \epsilon = 0.97, MSe = 0.44$). These effects of anteriority can be seen in Figure 3.4.

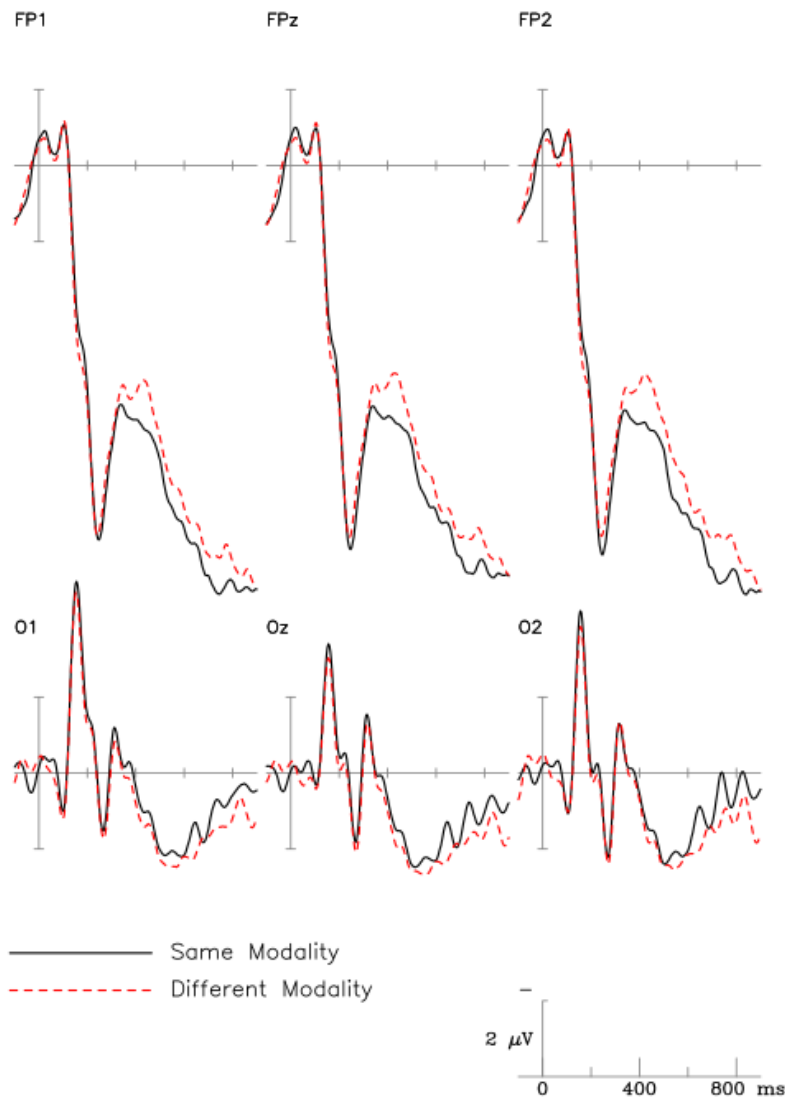


Figure 3.4. Conceptual modality switch effect. This shows an anterior negativity and posterior positivity to switch targets (red, dashed) relative to no-switch targets (black, solid). Both auditory and tactile trials are included. Time is plotted on the x-axis against voltage on the y-axis and negative polarity is plotted upwards.

Discussion

This experiment was designed to test for ERP indices of the conceptual modality switch effect. Properties were chosen to describe auditory and tactile experiences and were presented in property verification trials. As with prior conceptual modality switch studies, pairs of trials were manipulated to refer either to the same or to two

different perceptual modalities. Previous ERP experiments using similar paradigms have tested visual and auditory properties (Collins et al., 2011) and visual and tactile properties (Hald et al., 2011). Here we employed tactile and auditory properties to further assess ERP effects for the conceptual modality switch. This new pairing of property types also allowed us to compare how their hypothesized perceptual bases elicited different ERP patterns.

The largest effect in this study was not the modality switch effect, but rather a main effect of modality: a larger amplitude N400 for tactile than auditory properties. This effect was not modulated by the switch manipulation suggesting both switch and no-switch trials for tactile properties evoked a larger negativity than did auditory properties. We did, however, observe a subtle effect of the modality switch manipulation. In the late window (500-800 ms), the switch trials were slightly more negative than no-switch trials over the front of the head, and were more positive over posterior electrode sites.

The perceptual modality effect

The negativity observed for tactile decisions relative to auditory decisions in the 200-500 ms window is characteristic of an N400 effect. The N400 generally indexes processing of meaningful stimuli such as words, pictures and faces, but it shows characteristic differences, particularly in topography, for different types of semantic processing as well (Kutas & Federmeier, 2011). We propose that the N400 effect observed in the current study is driven by the different perceptually-based representations in semantic memory for tactile versus auditory properties and is represented by an N400 similar to a concreteness effect. A common strategy for indexing semantic

differences is to use psycholinguistic variables such as word frequency and cloze probability; the less likely item typically elicits a larger N400. Our stimuli, however, are not different in ways that would drive such an effect. Based on word frequency, the tactile properties used in this study were actually more common (35.7 per million) than our auditory properties (6.8 per million, from the Kucera-Francis measures in the MRC Psycholinguistic Database; Wilson, 1988). Similarly, values representing the relationship between the terms of each concept-property pair (using latent semantic analysis; Landaur & Dumais, 1997) were not significantly different between the two perceptual domains (auditory: 0.15; tactile: 0.18; $t(173) = 1.26, p = 0.21$).

The particular instantiation of the N400 observed in the current study, rather than being driven by psycholinguistic differences between our stimulus conditions, was likely related to the concreteness effect. The ERP literature on concreteness demonstrates an N400 effect for concrete words (e.g. *rose*) relative to abstract words (e.g. *instance*) with an anterior distribution and a more temporally extended time window relative to a typical N400 effect (Holcomb, Kounios, Anderson & West, 1999). The difference has been used to argue that concrete and abstract words both activate a verbal semantic representation while concrete words activate additional image-based semantic representations (Holcomb et al., 1999; Swaab, Baynes & Knight, 2002). In comparison with the concreteness effect, the topographical distribution for our tactile N400 was also more anterior than the typical centroparietal distribution of the N400 (see Figure 3.3). Explicit norming would be necessary to make the claim that the tactile properties were more concrete than the auditory properties. However, as different representational forms are proposed to drive the concreteness N400, perhaps different

perceptually-based representational forms (specifically parietal somatosensory activation for tactile properties and temporal cortex activation for auditory properties) underlie the N400 difference we observed for the two modalities.

This finding that tactile properties generally evoke different responses than auditory properties supports the suggestion that conceptual representations rely on perceptual systems. As reviewed in the introduction, fMRI studies have demonstrated that a property verification referring to different perceptual modalities can elicit activity in different regions of the brain related to the organization of perceptual processing systems (Goldberg et al., 2006). As such, we should expect to find differing distributions of semantic processes when attention draws focus to meaning supported by one perceptual system over another. The N400 observed for tactile properties relative to auditory properties suggests a different process underlying the interaction of the tactile properties and their context (e.g. concepts, previous trial, baseline word representations) from that of auditory properties. The evidence derived from this experiment suggests the cost for tactile property decisions observed in this time window is not modulated by whether an individual tactile decision was preceded by another tactile one or an auditory one.

The claim that the tactile decisions are unique is supported by previous behavioral findings as well (Connell & Lynott, 2010). Participants were asked to make an explicit decision whether a briefly flashed word (presented for 17-100 ms followed by a visual mask) could be perceived via certain modalities. Every block of trials required participants to focus on a different modality, so in the “tactile” block, *jagged* elicited a “yes” response while *loud* elicited a “no” response. Tactile decisions, at all presenta-

tion times, resulted in the worst accuracy relative to the four other modalities – a “tactile disadvantage” (Connell & Lynott, 2010). Connell and Lynott cited analogous findings in the perceptual literature showing larger costs for perceptual shifts to the tactile modality (Spence et al., 2001; see also Eimer, 2001). The present study revealed a similar conceptual result to the “tactile disadvantage” in that responses to tactile decisions took longer overall than their auditory counterparts. Together, these findings suggest that tactile sensations are processed in a different manner than visual and auditory sensations, and that this difference also emerges for conceptual items.

The modality switch effect

The late effect (500-800 ms) was an anterior negativity and a posterior positivity for switch trials relative to no-switch trials. ERP modulation based on the modality switch manipulation supports the claim that meaning is grounded in perceptual systems by demonstrating that the perceptual modality implicit in a property verification task provides sufficient neural and cognitive activation to impact subsequent decisions. The topographic distribution and timing of the effect are consistent with the findings of Hald et al. (2011) who found a frontal negativity accompanied by a posterior positivity in both late intervals as we found here, as well as the same pattern in an earlier time interval. They argued that their effect was similar to an N400-like effect observed for pictures (Ganis et al., 1996), and used this relationship to support the similarity of linguistic and perceptual representations of meaning (in line with Barsalou, 1999).

However, the late frontal negativity observed in the present study is more similar to the N700 in timing and in scalp distribution than it is to the N400. The N700 is a

late negativity seen exclusively over frontal electrodes and evoked with explicit task demands to engage in mental imagery (West & Holcomb, 2000). It can also be evoked by words intrinsically eliciting more mental imagery in certain conditions such as concrete words (Holcomb et al., 1999). On this interpretation, participants might have recruited more mental imagery on switch than no-switch trials, thus resulting in our observed late negativity. This is unlikely in view of a behavioral study that demonstrated mental imagery abilities do not correlate with performance on the conceptual modality switch task (Pecher et al., 2009). Nonetheless, it seems possible that participants might strategically employ mental imagery for task performance, even if it did not speed their reaction times. Thus, the present study leaves open the question of whether the conceptual modality switch arises from fundamental perceptual elements that are inherent in meaning interpretation (as indicated by an N400 effect), or reflects more strategic processing related to active mental imagery (as indicated by an N700 effect; West & Holcomb, 2000).

The next question in comparing the findings of the conceptual modality switch ERP studies is why earlier switch effects were not observed with the current experimental setup. Hald et al. (2011) suggest that the early frontal negativity they observed could have been driven by a similarity between their stimuli and visual images which tend to elicit earlier-onset ERPs compared to those elicited by words. “The proposed similarity with pictorial stimuli makes it likely (but not necessary) that the modality mismatch effects are stronger for the visual than for the tactile dimension. Qualitative inspection of the frontal waveforms broadly supports this view...” (Hald et al., 2011, pg. 11). Similarly, the findings by Collins et al. (2011) show an N400 in a more typi-

cal time window (200-500 ms) and not in a later window, but only for the visual switch condition. These differences suggest that the particular modalities tested in the conceptual modality switch paradigm are a key factor in determining the ERP pattern of switch effect, particularly, that the visual properties might be driving early switch effects.

This contrasts with behavioral findings that have not found differences between the switch effects for different modalities (Pecher et al., 2009). If the visual trials do indeed elicit the early effect one possible reason could be because the experiment is conducted with visually-presented words. Using stimuli presented in the auditory modality would be an interesting way to test how modality of presentation could influence this task. This would allow for another test of embodiment claims to see how perceptual modality can impact conceptual processing (as in van Dantzig et al., 2008).

Discrepancies between the various ERP studies of the conceptual modality switch effect might also be related to differences in the extent to which words used in the property verification task are exclusively unimodal. We used the Lynott and Connell (2009) norms to identify ratings for 27 of our auditory properties, which had an average auditory ranking of 4.60 out of a possible 5.0. The 34 tactile properties used in our study and also present in the norms received an average rating of 4.29. These high values verify the classification of our items. We also looked at the *modality exclusivity* score designated by Lynott and Connell (2009) to indicate how unimodal the property representations were. Auditory properties had a significantly higher modality exclusivity score (0.58; as defined by Lynott and Connell on a scale from 0 to 1) than tactile properties (0.35; $t(45) = 7.81, p < 0.001$). These modality exclusivity scores were both

lower (i.e. more multimodal) than the visual properties used by Collins et al. (2011; visual modality exclusivity: 0.73). This comparison suggests the current study uses two sets of somewhat multimodal properties whereas the previous study used strongly unimodal visual properties and multimodal auditory properties. One implication is that multimodal properties have more similar underlying representations and thus elicit a weaker switch effect and/or require mental imagery that would result in an N700 modulation. Consistent with this consideration, in a test of switch effects for novel concepts Connell and Lynott found no effect for multimodal to multimodal transitions but a significant effect for the same stimuli in unimodal to multimodal transitions (Connell & Lynott, 2011).

Considerations of multimodality are relevant to the potential underlying process of modality switching. When a concept broadly activates related perceptual modalities as part of its neurally distributed, grounded meaning, the access of one or more modalities should make a difference. A property verification task focuses the representation on the perceptual modality made relevant by the property. Thus, the concept of a *LEMON* can be restricted primarily to visual conceptual processing if described as *yellow* or to gustatory conceptual processing if described as *sour*. The modality switch is the process of changing the dominant modality accessed by the next concept-property pair, engendering a cost as the underlying meaning distribution changes. When the modal features of a property are salient (e.g. unimodal and/or visual) this cost will be observed as a semantic processing cost as a negativity in the N400 window. For the cases when properties are multimodal the presentation of a subsequent trial should change the distribution of the perceptual modality strengths slightly

but the overall multimodal characteristic is similar therefore not eliciting a change in the underlying semantic representation. However, in either case the property verification task could be implicitly requiring mental imagery resources, more so for switch conditions. This would explain the observation of only the late frontal effect in the switch condition for the current study that employed auditory and tactile properties with a somewhat multimodal quality.

Conclusion

The findings of the current study are consistent with theories of grounded cognition suggesting that perceptual systems are critical for conceptual processing. As evidence for this view, our findings show that the conceptual processing of modality matters. Even though the task requirements and degree of relatedness are the same for the decisions *COINS-jingle* and *SPONGES-soft*, there are extra costs involved for verifying the tactile property. The hypothesis put forth here suggests that this is a factor of underlying differences in the tactile perceptual system from other perceptual systems (see Eimer 2001; and Connell & Lynott, 2011 for related perspectives). The ERP modulation we observed from 500-800 ms also suggests that the perceptual modality relevant for a conceptual decision can impact subsequent decision processing. The timing and distribution of this modulation suggests that mental imagery processes might be engaged in this decision, but the nature of the stimuli required to potentially engage mental imagery require more investigation. One possibility is that properties with multimodal characteristics might be impacting conceptual modality switch results. Two other studies have also recently reported that post hoc comparisons of the unimodal and multimodal nature of the tested properties can influence the modality

switch results (Collins et al., 2011; Connell & Lynott, 2011). The dimension of multimodality will be an important direction to study if the full nature of the conceptual modality switch mechanism is to be understood.

References

- Ambrosi, S., Kalenine, S., Blaye, A., & Bonthoux, F. (2011). Modality switching cost during property verification by 7 years of age. *International Journal of Behavioral Development, 35* (1), 78-83.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-660.
- Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.
- Bub, D.N. & Masson, M.E.J. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition, 116*, 394-408.
- Collins, J., Pecher, D., Zeelenberg, R. & Coulson, S. (2011). Modality switching in a property verification task: an ERP study of what happens when candles flicker after high heels click. *Frontiers in Psychology, 2* (10), 1-10.
- Connell, L. & Lynott, D. (2010). Look but don't touch: Tactile disadvantage in processing modality-specific words. *Cognition, 115*, 1-9.
- Connell, L., & Lynott, D. (2011, in press). Modality switching costs emerge in concept creation as well as retrieval. *Cognitive Science*.
- Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related potentials. *Neuropsychologia, 39*, 1292-1303.
- Fischer, M.H. & Zwaan, R.A. (2008). Embodied language: a review of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61* (6) 825-850.
- Ganis, G., Kutas, M. & Sereno, M.I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8* (2), 89-106.
- Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 1-55.
- Glenberg, A.M. & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin and Review, 9* (3), 558-565.
- Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *The Journal of Neuroscience, 26*, 4917-4921.

- Gondan, M., Vorberg, D., & Greenlee, M.W. (2007). Modality shift effects mimic multisensory interactions: an event-related potential study. *Experimental Brain Research, 182*, 199-214.
- Hald, L.A., Marshall, J., Janssen, D.P. & Garnham, A. (2011). Switching modalities in a sentence verification task: ERP evidence for embodied language processing. *Frontiers in Psychology, 2* (45), 1-15.
- Havas, D.H., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14* (3), 436-441.
- Holcomb, P.J., Kounios, J., Anderson, J.E. & West, W.C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory & Cognition, 25*, 721-742.
- Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology, 20*, 525-540.
- Kutas, M., and Federmeier, K.D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.
- Landauer T.K. & Dumais S.T. (1997). A solution to Plato's problem: the latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-40.
- Lynott, D. & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods, 41* (2), 558-564.
- Marques, J. F. (2006). Specialization and semantic organization: Evidence for multiple semantics linked to sensory modalities. *Memory & Cognition, 34*, 60-67.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25-45.
- Pecher, D., van Dantzig, S., and Schifferstein, H. N. J. (2009). Concepts are not represented by imagery. *Psychonomic Bulletin & Review, 16*, 914-919.
- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying properties from different modalities for concepts produces switching costs. *Psychological Science, 14* (2), 119-124.
- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review, 11*, 164-167.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118* (10), 2128-2148.

- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22, 253-336.
- Rodway, P. (2005). The modality shift effect and the effectiveness of warning signals in different modalities. *Acta Psychologica*, 120, 199–226.
- Spence, C., Nicholls, M.E.R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, 63, 330-336.
- Stanfield, R. A. & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12 (2), 153-156.
- Swaab, T.Y., Baynes, K. & Knight, R.T. (2002). Separable effects of priming and imageability on word processing: an ERP study. *Cognitive Brain Research*, 15, 99-103.
- Töllner, T., Gramann, K., Müller, H.J. & Eimer, M. (2008). The anterior N1 component as an index of modality shifting. *Journal of Cognitive Neuroscience*, 21 (9), 1653-1669.
- van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, 32, 579-590.
- Vermeulen, N., Niedenthal, P. M., & Luminet, O. (2007). Switching between sensory and affective systems incurs processing costs. *Cognitive Science*, 31, 183-192.
- West, W.C. & Holcomb, P.J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience*, 12 (6), 1024-1037.
- Wilson, M.D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers*, 20, 6-11.
- Wu, L.L. & Barsalou, L.W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, 132, 173-189.
- Zwaan, R.A., Stanfield, R.A., & Yaxley, R.H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13 (2), 168-171.

Acknowledgment

Chapter 3, will be submitted for publication: Jennifer Collins, Diane Pecher, René Zeelenberg, Seana Coulson, “An ERP study of the impact of modality on the conceptual modality switch effect”. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Processing specifications for the influence of vertical language on visual motion perception

Abstract

Prior research has shown that verbs of motion influence tasks of motion perception (Meteyard et al., 2007) and activate motion processing areas MT+ (Saygin et al., 2010). The three experiments presented in this chapter address whether these effects reflect access to low-level or high-level visual processing systems. Participants were presented with verbs describing either upward, downward or horizontal motion, such as *rise*, *plummet* or *glide* followed by random dot kinematograms (RDKs) that elicit low-level motion perception. Several effects showed that the direction of motion conveyed by the verbs was facilitative for detecting RDKs moving in the same direction. Between Experiments 1 and 2 the findings also showed that reaction times dissociate from d' , C , and error rate in that direction-specific effects on the latter require blocked design, whereas direction-specific effects in the former measure emerged with randomly inter-mixed trials. Experiment 3 simultaneously recorded ERPs to look at electrophysiological indices of low-level and high-level effects. Participants who showed high sensitivity for detecting coherent RDKs showed amplification of the low-level component N1, similar to attentional gating effects (Luck et al., 2000). Low-sensitivity participants showed an amplified P300 for downward motion in the context of downward motion verbs suggesting a high-level, decision making effect driving the interaction of language and perception. Together these findings suggest the directionality of motion language impacts perception but in a task-dependent way modulated by individual differences.

Introduction

Grounded approaches to meaning suggest sensorimotor brain areas participate in semantic aspects of language comprehension (Barsalou, 2008). Comprehension of the word *tumble*, for example, by itself and in sentence contexts, should elicit activation in perceptual and motor brain regions that have been relevant for previous experiences of tumbling. For *tumble* in particular, visual brain areas used for perceiving moving objects should be activated, similarly with words such as *prance*, *fall*, and *stagger*. Much evidence suggests that there are indeed strong connections between language meaning and perceptuo-motor systems giving support to the idea that multi-modal activations form the basis of conceptual representations (see Barsalou, 2008 for a review). The goal of this chapter is to understand the nature of these connections with a particular focus on whether motion language accesses motion perception via low-level, perceptual systems or high-level ones involved in attention and decision-making.

Motion is an important domain in which questions of language-perception integration can be explored. On the language side of this interaction there are many different ways we can describe motion events. In English, we can talk about how things move (*flutter*, *writhe*, *limp*, *bounce*), the direction they move (*exit*, *plunge*), their spatial configuration when moving (*coil*, *dangle*), and what they move with (*pour*, *kick*, *hammer*), among other variations (Levin, 1993). These dimensions allow us to parameterize language within experiments and focus on certain features. The emphasis here will be on the direction of motion. Another reason that motion is a key domain to investigate how language and perception interact is that there is extensive work on the

complex physiology of motion perception using various methodological approaches (Born & Bradley, 2005).

The importance of perceptual activation for language comprehension is supported by a series of behavioral studies that showed an influence of motion perception on language comprehension (Meteyard et al., 2008) and an influence of language on motion perception (Meteyard et al., 2007). In particular, verbs describing upward motion impaired the detection of motion in a downward direction, and verbs describing downward motion impaired the detection of motion in an upward direction (Meteyard et al., 2007). This work built on previous studies that demonstrated motion toward and away a speaker like, *the kids tossed the beach ball over the sand toward you*, can influence visual detections (Zwaan et al., 2004) as well as the inverse showing that continuous visual presentation of moving displays can impact sensibility judgments made about sentences (Kaschak et al., 2005). In sum, the behavioral work on the interaction of language and perception, particularly focused on directionality, strongly supports the claim that these two cognitive abilities interact during real-time processing. There are varied approaches to this question, however, showing that the interaction can be observed in different paradigms that implicate different perceptual processing mechanisms such as spatial attention (Bergen et al., 2007; Richardson et al., 2003), motion inference (Chapter 5; Zwaan et al., 2004), eye movement control and imagery (Spivey & Geng, 2001) and low-level motion processing mechanisms (Dils & Boroditsky, 2010; Meteyard et al., 2007).

The application of neuroimaging techniques (e.g. fMRI, PET, MEG) has provided insight into the brain areas that are activated during the comprehension of mo-

tion language. Numerous studies have reported findings of topographically-related motor cortex activation when comparing motion words that describe actions implemented with either the hand, foot or mouth (Aziz-Zadeh et al., 2006; Buccino et al., 2001; Hauk & Pulvermüller, 2004; Tettamanti et al., 2005). Work on the perceptual aspects of motion language has shown that sentences with verbs of motion activate brain area MT+ (Revill et al., 2008; Saygin et al., 2010), a region important for visual perception of motion. Whether there is direct activation of MT by motion language is, however, a contested issue as several other research groups have not elicited MT activation with motion verbs. Rather, they have found activity in posterolateral temporal cortex (PTLC), an area anterior and superior to MT (Kable et al., 2002; Kable et al., 2005; Chatterjee, 2010). One suggestion is that a gradient from posterior to anterior regions represents a gradient of concreteness for concepts (Chatterjee, 2010). Under this framework, meaningful language about motion is processed in neighboring regions anterior to those used to process meaningful pictures about motion. The details will be critical in eventually understanding the nature of conceptual representations and how they are built from perceptual experiences, but the general neuroimaging findings do demonstrate that motion language and motion perception are intimately related.

The combination of functional imaging and behavioral results is important because each methodology leads to different implications about language-perception interactions. The former reveals brain areas that are active for processing motion language relative to non-motion language. The behavioral work demonstrates that the cognitive processes engaged for perception and language have sufficient overlap dur-

ing real-time processing to show interference in motion-specific conditions. Together, the findings from these separate methods suggest that the loci of the processes inferred from the behavioral findings are likely the systems activated in the neuroimaging studies. However, the behavioral studies are frequently dual-task experiments and the same experiments have not typically been run in neuroimaging studies. Furthermore, various processes could lead to activation in motion perception brain areas. For example, motion imagery (Goebel et al., 1998), motion inference (Kourtzi & Kanwisher, 2000) and low-level motion viewing (Tootell et al., 1995) are three examples of tasks leading to MT+ activity. Thus, just because we know activation in MT+ can be driven by motion language, we cannot infer whether a high-level or low-level process is driving an interference effect between language and perception.

Some behavioral studies have made steps in the direction of determining the level(s) at which language and perception interact. Meteyard, Vigliocco and colleagues, in particular, have addressed this question in their series of two studies using random-dot kinematograms (Meteyard et al., 2007; 2008). Random-dot kinematograms (RDKs) in these studies present a field of dots with a percentage moving either upward or downward while the rest move randomly. Higher percentages of coherently-moving dots result in stronger percepts of directed motion (Gros et al., 1998). These stimuli have been used extensively in work on visual perception both with humans and primates and are known to activate low-level systems of visual motion processing as well as brain area MT (Born & Bradley, 2005). The conclusion made by Meteyard and colleagues, as mentioned above, was that language describing either upward or downward motion interfered with detection of motion for RDKs moving in the opposite di-

rection. The low-level characteristics of these visual stimuli suggest that the impact of language on perception could be a low-level activation process that proceeds automatically in parts of the visual system.

Besides the characteristics of the visual motion stimuli, Meteyard and colleagues also referred to the dependent measures they used from signal detection theory (Wickens, 2002) in support of the position that language impacts perception at a low level of processing. Their two reported measures were d' and criterion (C), representing perceptual sensitivity and decision bias, respectively. The values are derived from error rates for a task with two decisions, e.g., participants detect coherently moving RDKs as a “target” and reject randomly moving RDKs. d' is a value transformed from error rates as a measure of perceptual sensitivity independent of response bias and is considered a low-level measure (Wickens, 2002). On the other hand, C is a value representing the bias to make one response or another and is thus considered a high-level measure. Both of these measures are important because a participant could show perfect accuracy for coherent motion by responding “coherent” for all trials of an experiment. But in this hypothetical case the participant also has an extreme response bias and low sensitivity for detecting the target while rejecting noise. d' is supposed to account for this; high d' values are observed for behavior showing both good target detection and good noise rejection. Meteyard et al. (2007) found worse perceptual sensitivity (lower d' values) when the direction of motion mismatched the direction of the concurrent verbs (e.g. an upward RDK with verbs like *plunge* and *fall*). This interference effect was claimed to represent a low-level visual phenomenon

rather than a high-level attention process because perceptual sensitivity (d') modulation was found in the absence of reaction time effects.

One challenge to the stance that the language-perception interference reflected low-level visual activation is that Meteyard et al. (2007) found criterion modulation as well as d' modulation but primarily emphasized the latter. Matched trials (same directionality for verbs and RDKs) showed a lower decision bias relative to mismatched trials. This suggests that their d' effects could have been the result of a high-level modulation of attention or decision making with an accompanying effect of perceptual sensitivity. Another reason to consider the results of this study as being due to attention or decision-making is that the experimental design allowed for an easy decision-making strategy. In day one of testing, participants were required to detect only upward (or only downward) motion, while on day two, the coherent motion was oriented in the other direction. Furthermore, within each block, participants listened to only one class of verbs presented in a continuous stream throughout the block. The observed effects, particularly the decreased C values for matched language-perception trials, might reflect strategic factors engendered by the blocked design.

It is clear that the perception of motion can be influenced by concurrent language. This effect can be observed for simple, low-level visual stimuli such as RDKs. Furthermore, there is evidence that low-level perceptual sensitivity can be impacted by the interaction. However, in order to better understand the processes underlying the language-perception relationship, other interpretations of this effect still need to be explored such as whether impaired perceptual sensitivity can be partially explained by procedural specifications. Building this understanding is the goal addressed in the fol-

lowing three experiments. First, Experiments 1 and 2 will determine the extent to which procedural differences affect performance when vertical motion language is paired with RDKs. Second, Experiment 3 will use event-related potential (ERP) methodology to assess the timing and underlying processes involved when language and perception are engaged simultaneously.

ERPs are processed electric potentials recorded from scalp electrodes and indexing online electrical activity of the brain evoked by a particular stimulus. This method can provide time-sensitive details of the different electrical patterns elicited by different experimental conditions, e.g., congruent (upward verb + upward RDK) versus incongruent (upward verb + downward RDK). Unlike behavioral measures that primarily provide information about the ultimate response, ERPs can provide insight into the different constituent processes that could be affecting this response during processing. With their good temporal resolution, ERPs will be an appropriate measure for investigating the current question of the processes underlying the detection of a coherently (and vertically) moving RDK in the context of vertical motion language.

In particular, we can divide the ERP signal into early and late components. The early components include the P1 and N1, positive and negative components, respectively, generally evoked by the visual onset of a stimulus and are completed within the first 200 ms of processing. These components are thought to index activity in extrastriate regions of visual cortex, and are amplified with increased attention (Hopfinger et al., 2004). The N2 is a negative-going waveform that peaks around 200 milliseconds post-stimulus onset with a broad posterior distribution. This ERP component responds to the onset of motion stimuli and can be modulated by the strength of mo-

tion that is being detected; it is only slightly modulated by attention (Niedeggen & Wist, 1999; Hirai et al., 2003; Wang et al., 1999). The late component likely to be elicited by the RDKs with potential modulation by language is the P300, a positive-going waveform that peaks at central-parietal sites. This component has been described as indexing context updating or working memory (Rugg & Coles, 1995). The amplitude of this component can be modulated by stimulus frequency while latency differences can be seen for different degrees of difficulty of the task (Kutas et al., 1977). The modulation of the P300 by language-perception condition would argue against a purely perceptual locus for this effect, suggesting the involvement of decision-level processes.

Current Experiments

All three experiments used RDKs and vertical motion verbs as modeled after Meteyard and colleagues (2007). Experiments 1 and 2 used event-coupled word presentation for each RDK trial. Verbs were classified into three categories: upward (e.g., *rise, surge, elevate, climb*), downward (e.g., *collapse, sink, dump, fall*), and horizontal or neutral (e.g., *depart, rust, exchange, tremble*). In all experiments the three verb types were randomly intermixed in blocks so that verbs were completely uninformative about the presence of coherent motion. In Experiment 1, we also randomly intermixed upwards, downwards, and random RDK trials, while in Experiment 2, trials were blocked according to their direction of motion.

We expected to find worse sensitivity for trials with mismatched directionality on the language-perception setup. We also expected to find lower C values for trials with matched language-perception directionality. This pattern would demonstrate a

replication of previous findings (Meteyard et al., 2007). We also predicted that the blocked organization would more closely replicate the previous findings. In particular, we predicted blocked organization would show C differences, while randomized ordering would not. This would suggest that the extended perceptual context (e.g. only seeing downward motion) can influence the way that perception and language interact. Following the reasoning of Meteyard et al. (2007), if our effects occur at the levels of error rate, reaction time or decision bias, this would suggest that there are high-level effects of language on perception. If they are only observed for d' then we would conclude that language can affect low-level perceptual mechanisms used for distinguishing a target from noise.

Experiment 3 was designed to replicate Experiment 2 while recording electroencephalogram (EEG). EEG to critical stimuli were later processed and averaged into ERPs representing brain activity for particular cognitive events. If we were to see modulation of the ERP components involved with visual detection such as the P1, N1 and N2 then this will indeed support the suggestions that language can affect low-level perceptual detection, similar to the claims made for observed d' modulation. If later ERP effects such as modulation of the P3 were to arise, then this would indicate that language influences the decision process involved with detecting the motion. This observation would suggest that language interacts with perception at a high-level during decision-making processes. Our predictions for this experiment were based on the results of Experiment 2. If we were to find modulation of only d' behavioral measures then we would expect to analogously only find modulation of early ERP components. On the other hand, if we were to find both modulation of d' measures as well as C

and/or reaction times, then we would expect to find both early and later modulation of the ERPs.

Experiment 1

Experiment 1 presented upward, downward and neutral motion verbs with target trials of both upward and downward RDKs. Filler trials were those for which RDK motion was completely random. All trials were randomized so participants could have no expectation for the words they would hear or the motion they would see. The question addressed was, in the absence of any expectations (i.e. with a randomized design), whether upward verbs would facilitate upward RDK detection and downward verbs would facilitate downward RDK detection.

Methods

Participants

Twenty undergraduates from the UC San Diego community participated for course credit.

Materials

Participants were presented with random dot kinematograms (RDKs) composed of 1000 white dots on a black background subtending a square with dimensions of approximately 7.1cm. Following Meteyard et al. (2007; 2008) we created the RDK stimuli with the Cogent Graphics package for Matlab developed by John Romaya at the LON, Wellcome Department of Imaging Neuroscience. The procedure was run with Presentation software so the individual RDK frames were created in advance at various coherence percentages from 3% to 15%. Each RDK was nine frames long and

lasted 144 ms total, 16 ms apiece. For each coherence percentage we created ten different RDKs which were sampled from randomly.

Half of the RDK stimuli presented throughout the course of the experiment showed coherent motion. Of the coherent motion trials half showed upward motion and half showed downward motion. These were both classified as “coherent” motion for the purposes of the experimental task. Thus, directionality was not a dimension ever explicitly distinguished by participants. A coherent motion stimulus is represented schematically in Figure 4.1. The other half of the RDK stimuli were “random” motion trials created with 0% coherence.

Verbs described either upward, downward or horizontal motion, such as *rise*, *plummet* or *glide*, respectively. Each verb set was comprised of 30 verbs and was categorized according to prior norms (Meteyard & Vigliocco, 2009; Meteyard et al., 2007). All but one of these verbs were normalized to be 549 ms long using Adobe Audition software. The only exception for the length normalization was *deteriorate* which sounded natural at a minimum of 700 ms long. It was the longest word of the set with five syllables while all other verbs had 1-3 syllables (average = 1.4, standard deviation = 0.56).

Each of the 90 verbs was presented four times throughout the experimental phase – once in a congruent visual condition, once in an incongruent visual condition and twice with randomly moving dots. Verb congruency depended on whether the motion of the verb and RDK were in the same direction (e.g. upward-moving RDK with an upward verb), or in different directions. Horizontal verbs served as control stimuli. The 360 total trials were randomized and divided into twelve blocks of 30 trials each.

Each verb was presented with two different voices – one male and one female. Two experimental lists allowed for verbs spoken by each speaker to occur in every experimental condition.

Timing between trials varied from 2000 ms to 2500 ms with 100 ms intervals. During this time participants made their response and prepared for the next trial. Every trial began with the verb presented for 350 ms prior to the RDK onset. The RDK lasted for 144 ms leaving 55 ms of each sound file at the offset of the RDK (see Figure 4.1).

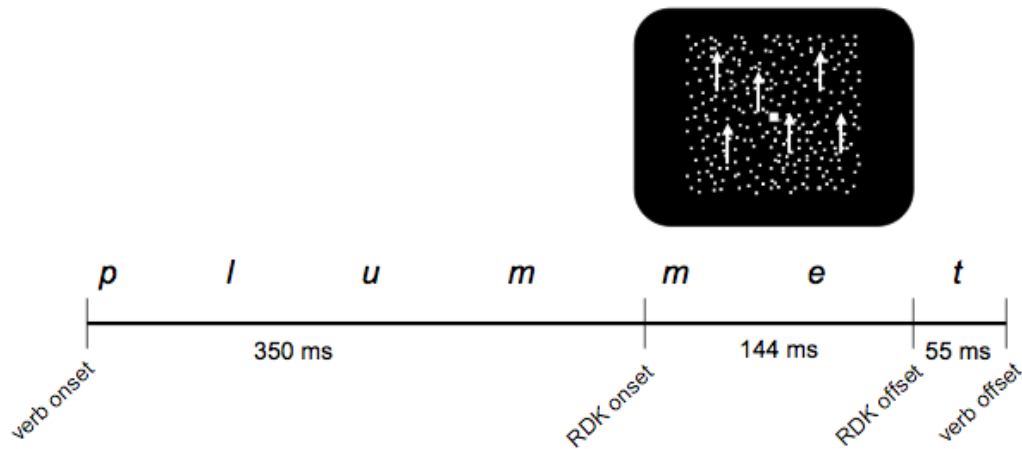


Figure 4.1. The timing of a single trial from all three experiments. Verbs were presented auditorily for 549 ms during which time the RDK was presented. Each coherently-moving RDK presented a predetermined proportion of dots moving upward or downward, as illustrated here schematically with white arrows. The trial represented here is an “incongruous” trial because the direction of motion conveyed by the verb (down) is opposite the direction displayed in the RDK (up).

Procedure

Calibration Phase: Calibration occurred in a dark room lit only by the light from a small nightlight. Participants sat with an experimenter who read a standard set of instructions out loud and adjusted the coherence levels for each participant individually. The task was to determine if each trial showed coherent or random motion.

Participants pressed the “x” key on the keyboard for coherent motion and the “.” key on the keyboard for random motion.

The calibration phase began with a block of 16 RDK trials, half of which were presented with 50% coherence, the other half were completely random (0% coherence). At a 50% level of coherence the motion is quite evident and participants typically made few errors. If any errors were made the experimenter had them repeat that practice block. The experimenter then presented participants with blocks using 25% and 20% coherence levels and continued to adjust the coherence for each block until the participant reached a threshold of 75-81% accuracy on two subsequent blocks of test trials. That coherence percentage was used for the remainder of the experiment (Experiment 1 average threshold: 7.7%; Experiment 2 average threshold: 8.6%).

Experimental Phase: During the experimental phase participants responded to each RDK trial with an accompanying verb presented over headphones (see Figure 4.1 for details). The task was the same as during calibration – to determine whether the dot motion was random or coherent. The participants were left alone in the darkened testing room for the experimental phase. They were self-paced in advancing between blocks but were required to take breaks after every eight blocks.

Data analysis

We analyzed the behavioral measures error rate and reaction time, as well as signal detection measures d-prime (d') and criterion (C). D' is a measure of discrimination computed from participants' performance on both targets (i.e. correct responses for coherently moving RDK) and distractors (i.e. false alarms to randomly moving RDKs). The d' values reported here represent participants' ability to discriminate be-

tween coherent and random motion, as they index both the ability to correctly detect coherent motion and the ability to not perceive random motion as coherent. C , on the other hand, is a measure of decision bias with 0 reflecting no decision bias, negative values reflecting a bias toward coherent responses and positive values reflecting a bias toward random responses. For reaction time measures, values over two standard deviations from an individual's mean were filtered out of the data. Our analyses for all experiments used ANOVA repeated-measures design with factors of Dot Direction (upward/downward) and Verb (upward/downward).

Results

All four dependent variables indicated that participants were better at detecting downward than upward motion. Participants made 9.3% errors for downward motion and 18.8% errors when responding to upward motion ($F(1,19) = 24.20, p < 0.001$). Reaction times showed only a marginally significant difference ($F(1,19) = 3.77, p = 0.067$) with faster responses to downward trials (684 ms) than to upward trials (708 ms). Perceptual sensitivity was better for downward trials with higher d' values ($d' = 2.34$) than for upward trials ($d' = 1.91; F(1,19) = 23.11, p < 0.001$). Finally, the decision criterion was lower for downward ($C = -0.23$) than upward motion ($C = -0.02$), indicating a greater tendency for “coherent motion” responses when seeing downward motion ($F(1,19) = 23.11, p < 0.001$).

The concurrently presented verbal stimuli did not affect error rates, d' or C (all $F_s < 1$). However, verbal stimuli did affect participants' reaction times, as trials after upwards verbs were faster than downwards verbs (up verbs: 686.9 ms; down verbs: 706.0 ms; Verb: $F(1,19) = 4.62, p < .05$). The main effect of Verb type was also quali-

fied by an interaction with Dot Direction (Verb x Dot Direction: $F(1,19) = 5.183$, $p < .05$), due to differences in the pattern of results for upward and downward motion trials (see Figure 4.2b). Response times in downward motion trials were similar for both types of verbs (up verbs: 683.1 ms, SD = 78.0; down verbs: 685.1 ms, SD = 85.0). By contrast, for upward motion trials, faster responses were observed after congruent (690.7ms; SD=86.2) than incongruent verbs (727ms; SD=85.1).

In summary, all measures showed a main effect demonstrating that upward motion was a harder experimental condition than downward motion. Only reaction time showed a statistically significant interaction suggesting that the verbs describing motion in a direction congruent to the RDK direction would speed reaction times. In follow-up contrasts this was only demonstrated for upward motion.

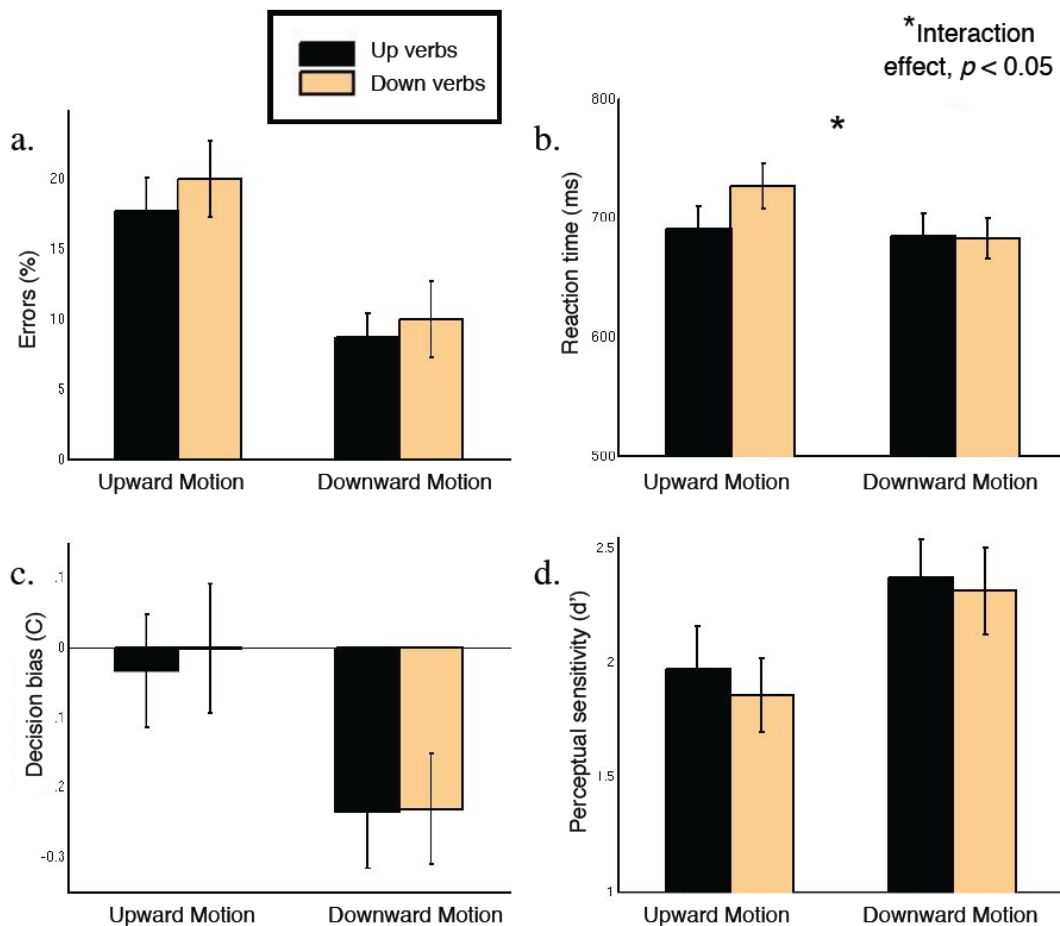


Figure 4.2. Behavioral results for Experiment 1. Measures of (a) error rate, (b) reaction time, (c) decision bias: C , and (d) perceptual sensitivity: d' are displayed. All measures showed a main effect of Motion Direction demonstrating that upward motion was more difficult to detect than downward motion. Verb type did not modulate participants' error rates (a), decision bias (c), or d' (d). Analysis of reaction times shown in (b) revealed a statistically significant interaction between Motion Direction and Verb. Reaction times for upwards motion were faster after up verbs than down verbs; no significant Verb effect was observed for downwards motion trials.

Discussion

The findings for this study suggest two conclusions. First, the language manipulation affected the responses to dot motion in the upward motion condition. This effect was observed as an interaction in the reaction time measure with faster responses for upward motion following upward verbs than downward verbs. The effect is another demonstration that language directionality affects directional motion percep-

tion. The particular interaction pattern observed does seem similar to one reaction time difference observed by Meteyard (2007) who found an increase in reaction times for upward motion with incongruent language. They found the opposite, inexplicable effect for downward motion (a decrease in response time with incongruent language) and did not elaborate on the possible significance of this finding. Perhaps in our case, the symmetric effect for downward motion trials was not observed because of the differences between upward and downward motion perception.

The observation of the reaction time effect with no concurrent observations of error rate, d' or C modulation suggests that the interaction of language and perception is not necessarily a low-level process. Reaction times can be considered a relatively late index of cognitive processing that can index many aspects of the processing of the task at hand such as perceptual identification and/or decision-making. With randomized trial organization it seems that a low-level process might not be behind the interaction of language and perception.

The second finding revealed that performance for detecting downward motion was superior to detecting upward motion on all levels of analysis. This difference was surprising because the original study showed no such differences. One difference between our procedure and that used by Meteyard et al. (2007) was that we incorporated both upward and downward motion in the same blocks of trials whereas they had participants come into the lab on separate days to perform the upward and downward motion tasks. This procedural difference was necessary as one goal of our experiment was to eliminate the potential confound that could be contributed by doing only one type of motion perception at a time. We calibrated people on upward and downward

motion within the same block of trials during the calibration phase, and our findings suggest that with the same percentage coherence, participants' perception for upward and downward motion differs. While many studies show no differences between upward and downward RDK motion detection (Gros et al., 1998) when the stimuli are presented in different hemifields there is better detection for motion toward the center (Giaschi et al., 2007). The consistent effect demonstrated here for better detection of downward motion simply suggests that the position of our testing monitor was positioned in participants' upper visual field. The differences in upward and downward motion perception elicited in the current study could have influenced the asymmetric way language modulated perception as indicated by the reaction time effect described above.

In summary, this study shows that the blocked design used by Meteyard et al. (2007) to investigate the differences between language and perception was not necessary to elicit effects demonstrating that these two cognitive abilities interact. However, as our findings did not replicate their findings of d' and C differences, the blocked design could have been driving those particular effects. Experiment 2 will group trials in blocks according to the direction of motion presented. The task for every trial will always be to detect coherent motion but for some blocks this coherent motion will be directed upwards and for other blocks coherent motion will be directed downwards. All other conditions, such as randomized verb types, will be the same in Experiment 2 as in Experiment 1.

Experiment 2

Methods

Twenty-two undergraduates from the UC San Diego community participated; none had participated in Experiment 1.

In Experiment 2 all materials were exactly the same as in Experiment 1, save the trial organization. RDK motion was separated so that 6 of the 12 blocks presented “coherent” trials with upward motion, and 6 presented “coherent” trials with downward motion. Blocks were randomly ordered, and participants were not informed that blocks would include only one direction of coherent motion. Within each block participants heard 10 upward, 10 downward and 10 horizontal verbs, half with coherent RDK motion, half with random RDK motion. The timing of the verbs relative to each RDK was the same as in Experiment 1, viz., each verb preceded RDK onset by 350 ms.

Results

As with Experiment 1, we found the downward moving condition resulted in lower error rates (10.2%; $F(1,21) = 16.69, p < 0.005$) and faster reaction times (705.6 ms; $F(1,21) = 8.22, p < 0.01$) than for upward moving dots (error rate: 23.8%; RT: 728.2 ms). We found better sensitivity for downward ($d' = 2.39$) than upward motion ($d' = 1.87; F(1,21) = 16.31, p < 0.005$). The trials with downward-moving dots showed a stronger decision bias for coherent motion responses ($C = -0.16$), whereas responses to upward-moving dots resulted in C values that were positive, suggesting a random motion response bias ($C = 0.10; F(1,21) = 16.31, p < 0.005$).

Verbal stimuli did not affect RT as in Experiment 1 ($F_s < 1.5$), but did affect all other dependent measures (see Figure 4.3). In general these results showed interaction effects because of facilitation for the congruent language-perception conditions. For example, analysis of error rates revealed a significant interaction of Verb and Dot Direction ($F(1,21) = 14.18, p < 0.005$), with better performance (fewer errors) in the congruent conditions than their respective incongruent ones (dots up, verb up: 20.0% vs. dots up, verb down: 27.4%; dots down, verb down: 8.9% vs. dots down, verb up: 11.4%).

Analysis of d' revealed no main effect of Verb ($F < 1$), but rather an interaction between Verb and Dot Direction ($F(1,21) = 9.61, p < 0.01$). This interaction reflects better perceptual sensitivity for RDKs preceded by verbs conveying congruent directionality (dots up, verb up = 1.92; dots down, verb down = 2.50) than for those preceded by incongruent verbs (dots up, verb down = 1.81; down dots, verb up = 2.28).

Analysis of Criterion showed a main effect of Verb ($F(1,21) = 5.70, p < 0.05$), qualified by an interaction with Dot Direction (Dot Direction x Verb ($F(1,21) = 9.61, p < 0.01$)). The interaction was driven by a large difference between the verb conditions when making decisions about upward motion (up verb = 0.01; down verb = 0.19), compared to a small difference in the downward motion conditions (down verb = -0.16; up verb = -0.17).

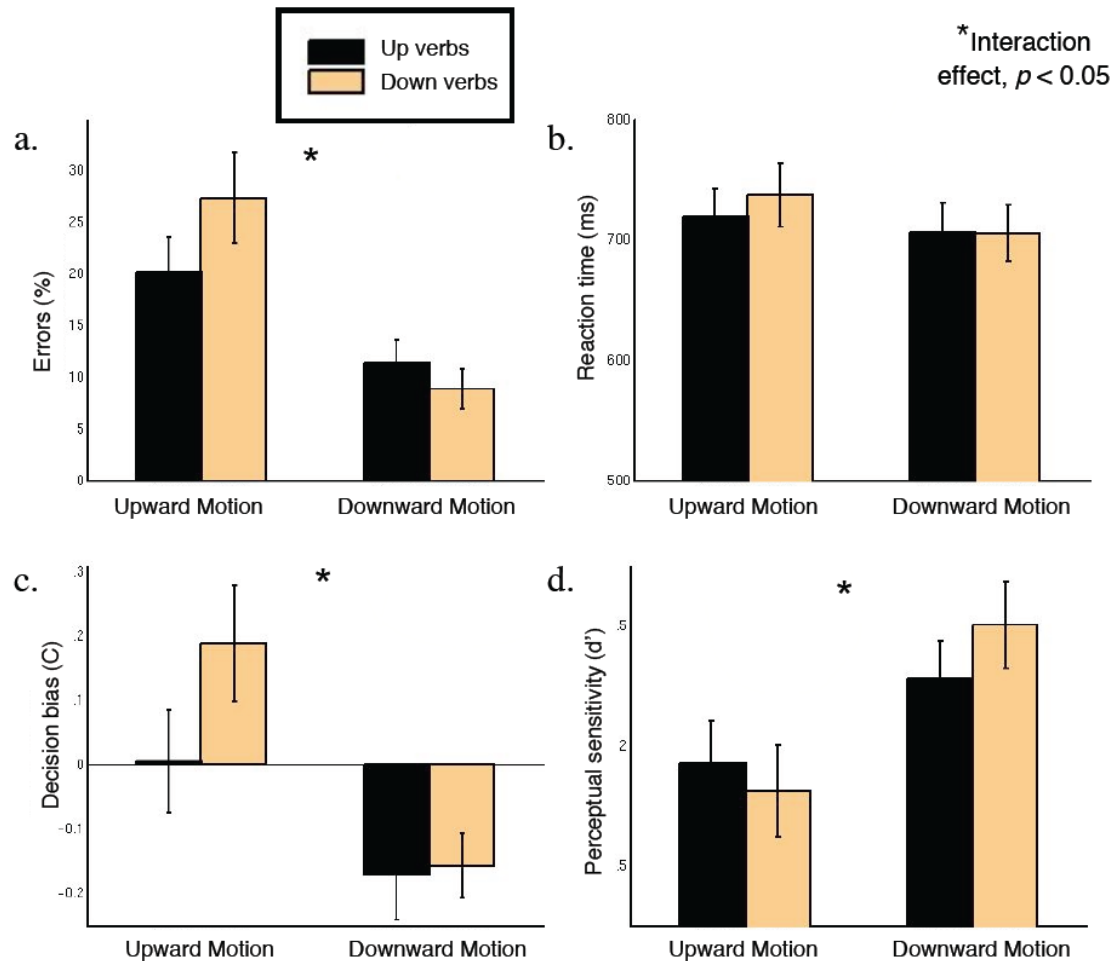


Figure 4.3. Behavioral results for Experiment 2. Measures show (a) error rate, (b) reaction time, (c) decision bias: C , and (d) perceptual sensitivity: d' . As with the previous experiment, all measures demonstrated that responses to upward motion were harder than responses to downward motion. Error rate, C , and d' all resulted in significant interaction effects. (a) Error rates showed better performance for conditions when the RDK motion and the direction described by the verb matched. (b) Reaction times did not show any significant effects based on verb type. (c) C showed a tendency for less bias when the direction of motion and the direction conveyed by the verb matched. This effect was negligible for downward motion. (d) d' , too, showed an interaction effect demonstrating that congruent motion for the words and RDKs resulted in better sensitivity.

Discussion

The findings of Experiment 2 generally showed facilitation of motion perception in the context of verbs conveying congruent directionality. These findings replicate those found by Meteyard et al. (2007) with a lower d' for incongruent conditions

and a lower C for congruent conditions. Generally, the findings show that at several levels of analysis the direction described in language can facilitate perception.

In contrast with Experiment 1, the findings here make it clear that the experimental setup can make a big difference in the way language impacts perception. The only difference between the methods of these two experiments was that in Experiment 1, within an experimental block, trials that elicited “coherent” responses could be moving upward or downward but this dimension was never explicitly mentioned. In Experiment 2, within a block of trials those that elicited “coherent” responses were consistent within a block, but this organization was also never explicitly mentioned nor was it task relevant. The effects found in Experiment 2 and the parallel lack of findings in Experiment 1 suggest that the perceptual context surrounding an individual’s task can influence the way in which language modulates perception. Furthermore, this modulation is seen both with high-level measures (error rate, C) and low-level measures (d'). The consistency between the d' and C measures as well as the additional observation of error rate differences suggest that the d' differences, rather than indexing a purely low-level process that drives the interaction of language and perception, might be a low-level effect that arises from other high-level effects. For example, within a block of upward motion trials, perhaps participants attend to upward language and this leads to improved detection.

Experiment 2 showed the same directional differences that were observed in Experiment 1. Upward motion was harder to detect than downward motion. The calibration phase or the particular angle of the computer screen relative to the participants’ eyes are likely to be factors driving this difference. The particular circumstances are

not important though because Experiment 2 replicated previous findings for both upward and downward motion.

Experiment 3

Experiment 3 used a different methodology to further explore the possibilities of the processing level at which language and perception interact. We used an experimental setup designed similarly to Experiment 2 while recording EEG. Previous studies have used this method to provide evidence for how visual motion processing takes place leading to a more detailed account of how fast cognitive processes are occurring beyond what information is available with behavioral measures. Event-related potentials (ERPs) measure the online electrical activity recorded from a number of electrodes connected to the scalp. ERPs can provide information about the time course of the processing before the reaction is made and can also help experimenters distinguish between qualitatively different processes that might be happening with similar time courses. The particular ERP components of interest in Experiment 3 are early perceptual components associated with motion processing (e.g. N2, N1), and later decision-related components (e.g. P3b).

Methods

Participants

Twenty-four undergraduates (13 female, average age = 21) from the UC San Diego community participated in this experiment for course credit. In the analyses presented below 16 participants are included (8 each in upward and downward conditions). Six of the others were excluded from the analysis due to excessive eye move-

ments (artifacts exceeding 27% of critical trials), and two were excluded because of experimenter error.

Materials

Random dot kinematograms (RDKs) were images of 640x480 pixels composed of 480 white dots on a black background subtending a rectangle with dimension of 320x240 pixels. Cogent Graphics package for Matlab (developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience) was used to create frames for the random dot displays. Each RDK was composed of nine frames, presented 17 ms each, for a total length of 153 ms. The coherent motion trials took a random selection of dots equal to an individual's coherence percentage (e.g. 48 dots for 10% coherence) from the first frame and in each subsequent frame those dots were moved up or down by 6 pixels. The remainder of the 480 dots were randomly refreshed to different locations in each subsequent frame.

The verbs were exactly the same as those used in Experiments 1 and 2, classified into the three direction categories: upward, downward, horizontal, with 30 verbs each.

In the experimental phase participants heard all 90 verbs spoken by both a male and female speaker (180 tokens) with coherent and random motion for a total of 360 trials with four repetitions of each verb. In order to increase the number of trials per critical condition for ERP averaging we used a between-subjects design with participants observing either upward or downward motion as coherent motion trials throughout the experiment. As in Experiments 1 and 2, experimental conditions were defined by the direction of motion of the dots relative to the direction of motion de-

scribed by the verbs. For each individual there were 60 trials in the congruent, incongruent, and control conditions. For all participants the control condition was defined by the horizontal verbs, regardless of whether they observed upward or downward motion. A corresponding 60 trials presented the exact same verbs for each test condition with random dot motion.

Procedure

Calibration Phase: Participants were seated in a small, dimly lit experimental testing room in front of a 14 in computer monitor. An experimenter read a standard set of instructions describing the task. Participants were given a demonstration block with 16 trials, half of which showed RDKs with 50% coherence, the other half were completely random (0% coherence). The task was to decide for each trial whether the RDK had coherent motion (upward or downward) or completely random motion. Participants completed this block with the goal of 100% accuracy which was typically reached on the first or second try. They then practiced with a block of trials at 25% coherence. These practice trials were included to allow participants to gain an understanding of the difference between coherent and random motion and to practice the task as would be presented to them later during the testing phase.

The second part of the calibration phase was a forced choice task modeled after the calibration used in (Meteyard et al., 2007). Participants were presented with 12 blocks of 20 trials. For each trial they saw two subsequent RDKs. Their task was to determine which one presented coherent motion. Coherence percentages ranged from 0.1% to 20.1% with steps of 2.2%; each of which was presented twice per block. Participants pressed “1” on a standard QWERTY keyboard if the first of the two trials

was coherent motion, and “2” if the second of the two trials was coherent motion. The presentation time between trials was between 500 and 1000 ms. The presentation did not advance until participants made their decision and the onset of the next trial was between 200 and 500 ms after participants’ response.

As implemented by Meteyard et al., 2007, each individual’s threshold was determined by fitting the Weibull cumulative distribution function (CDF) to their accuracy rates at the 10 coherence values on which they were tested. We defined the threshold as the coherence percentage where the function predicted participants’ accuracy to be 81%. This threshold was rounded to its nearest half-percent and was used for that participant’s experimental phase. The Weibull CDF is defined as $p = 1 - \exp(-(x/a)^b)$, where x represents the coherence values, p the accuracy rates, and a and b are the fitting parameters, determined by a least squares fit. Once the parameters were determined, the CDF was used to find the value of x for which p was 81%. The parameters were found with Matlab version 2010a.

EEG Recording and Experimental Phase: Before the experimental phase participants were outfitted with an electrode cap with tin electrodes mounted at 29 scalp sites. Electroencephalogram (EEG) was recorded from these electrodes as well as those placed at right and left mastoids, right and left canthi and under the right eye. EEG from scalp electrodes were referenced on-line to the left mastoid, and subsequently re-referenced to the average of the left and right mastoid electrodes off-line. The signal was amplified with an SA Instruments isolated bioelectric amplifier at a bandpass of 0.1 and 100 Hz, digitized online at 250 Hz, and stored on a hard drive for subsequent averaging. Artifacts from blinks were monitored with the electrode below

the right eye and horizontal eye movements with electrodes placed over the outer canthi. Trials with eye movement artifacts were rejected manually off-line.

Participants were seated in a dimly lit, sound attenuating chamber approximately 30 inches from a 17-inch computer monitor. Verbs were presented over speakers placed at either side of the monitor. Trials began with 350 ms of the verb presentation followed by RDK onset as in Experiments 1 and 2. Participants made the decision whether the motion was coherent or random with a button box by pressing a button with their right hand for random motion and with their left hand for coherent motion. Timing between trials varied from 1945 to 2445 ms during which participants made their response and prepared for the next trial. Between blocks they were given time to rest, as needed.

Behavioral Results

We began by testing the behavioral results in a mixed 2x2 design with factors of Dot Direction (between subjects: upward/downward) and Verb (within subjects: upward/downward). This was the same as Experiments 1 and 2, except that here Dot Direction was a between-subjects variable.

Error Rate

The difference in error rates between groups who saw upward and downward motion was not significant (Dot Direction $F(1,14) = 2.52$, $p = 0.135$). There was, however, a significant main effect of Verb with more errors in the context of downward (29.7%) than upward verbs (23.8%; Verb $F(1,14) = 11.55$, $p < 0.005$). We also found a marginal interaction of Verb and Dot Direction ($F(1,14) = 4.25$, $p = 0.058$), due to a larger modulation of error rate for upward motion perception (up verbs: 26.4%, down

verbs: 35.9%) than for downward motion perception (up verbs: 21.1%, down verbs: 23.4%). Only the case of upward motion showed the expected effect of Verb, with fewer errors for the congruent (up verbs) than the incongruent (down verbs) condition.

Reaction Times

Although neither main effect was significant ($F_s < 1$), analysis of reaction times revealed the predicted crossover interaction between Dot Direction and Verb ($F(1,14) = 6.06, p < 0.05$), due to faster reaction times in the context of congruent directional language. Those who saw upward motion showed faster responses with upward verbs (up verbs: 826.7 ms, down verbs: 856.5 ms). Those who saw downward motion showed faster responses with downward verbs (up verbs: 781.8 ms, down verbs: 754.7 ms).

Signal Detection Measures

D' analyses did not show either a main effect of Verb, or an interaction with Dot Direction ($F_s < 2$). The between-subjects effect of Dot Direction also failed to reach significance ($F(1,14) = 2.61, p = 0.128$).

Analysis of criterion revealed only a main effect of Verb ($F(1,14) = 6.02, p < 0.05$), due to a bias to respond “coherent” in the context of up verbs ($C = -0.09$) and “random” in the context of down verbs ($C = 0.04$). Dot Direction was not significant either as a main effect ($F < 1$), or in interaction with Verb ($F(1,14) = 2.94, p = 0.11$).

Interim Behavioral Results Discussion

Experiment 3 was designed to replicate Experiment 2 while using a between-subjects design with participants either discriminating random motion from coherent motion upward or random motion from coherent motion downward. The between-

subjects design was intended to increase the number of trials per condition for ERP averaging. However, just as the procedural manipulation between Experiments 1 and 2 changed the observed interactions between language and perception, the results in Experiment 3 differed from those in Experiment 2. In Experiment 2, facilitation for the congruent conditions was evident in d' scores, but not RTs. In Experiment 3, it was evident in RTs, but not d' scores. In general, sensitivity scores in Experiment 3 were lower (down group $d' = 1.57$, $SD = 0.63$; up group $d' = 1.06$, $SD = 0.62$) than for our behavioral participants in Experiments 1 (average $d' = 2.08$, $SD = 0.70$) and 2 (average $d' = 2.04$, $SD = 0.69$). In fact, the findings of Experiment 3 better resemble those of Experiment 1, with reaction time differences in the absence of perceptual sensitivity differences, than they resemble the findings of Experiment 2.

A few methodological differences between Experiments 2 and 3 might account for the observed differences. First, the calibration used in Experiment 3 was calculated by a fitting function, whereas previously it was determined by gradual, manual adjustment on a block-by-block basis. This change was implemented to better follow the methods of Meteyard et al. (2007) but instead this difference appears to have resulted in lower average coherence values in Experiment 3 (5.7, $SD = 3.98$) than for Experiment 1 (7.7, $SD = 2.22$) or Experiment 2 (8.6, $SD = 2.30$). These lower coherence values are likely to be related to the lower sensitivity scores observed for Experiment 3. In other words, in Experiment 3 participants were calibrated to lower thresholds subsequently making the task more challenging. In addition, Experiment 3 involved the ERP capping procedure requiring a different testing environment, different hardware, and a delay between calibration and experimental testing.

Post-hoc Analysis

All the factors presented above suggest that the task presented for Experiment 3 was qualitatively different from Experiment 2. To best account for the difference we included another behavioral dimension in the ERP analyses – Sensitivity. This split participants into groups based on how sensitive they were in performing the task overall (high sensitivity mean $d' = 1.80$, low sensitivity mean $d' = 0.84$) with the high sensitivity group better resembling the participants from Experiment 2.

To test this comparison, we re-analyzed the reaction time results with this additional factor. We found interaction effects of Verb with Direction (as reported above, $F(1,12) = 9.76$, $p < 0.01$) as well as Verb with Direction and Sensitivity ($F(1,12) = 8.63$, $p < 0.05$; other $F_s < 2$).

Low sensitivity participants showed a facilitative interaction effect of language on perception, namely, verb context speeds responses to congruently moving dots ($F(1,6) = 14.65$, $p < 0.01$; see Figure 4.4b: down motion, down verb: 808.5 ms; down motion, up verb: 849.8 ms; up motion, up verb: 832.5 ms; up motion, down verb: 901.4 ms;). (When tested as pairwise comparisons with two-tailed t-tests, only the group who saw downward motion showed a significant difference: $t(7) = 2.25$, $p < .05$.) The modulation of reaction time for the low sensitivity group resembles the reaction time differences observed in Experiment 1.

High sensitivity participants do not show an interaction between verb type and dot direction, or a main effect ($F_s < 1.5$; see Figure 4.4a). The null effect for reaction time difference for the high sensitivity group resembles the null findings of Experiment 2.

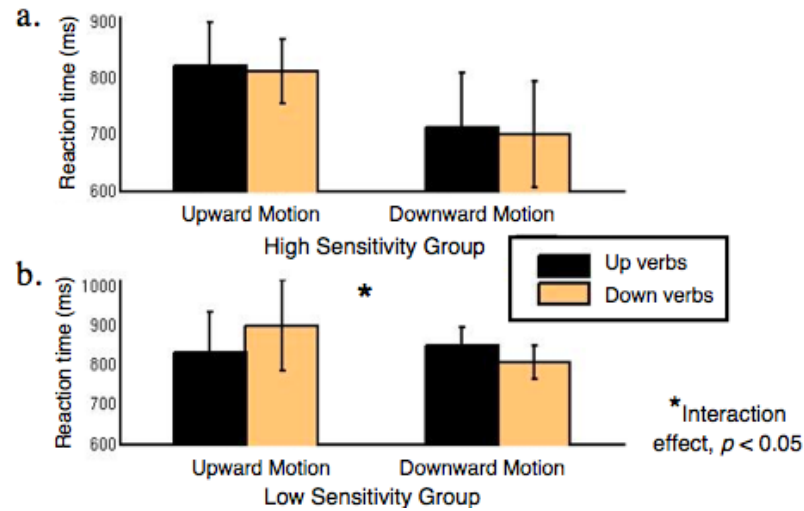


Figure 4.4. Reaction time results split by sensitivity. The groupings along the horizontal axis represent the direction of motion viewed by participants. The top figure shows reaction times for participants with high sensitivity. The bottom figures represents participants with low sensitivity who show an interaction effect due to a facilitation in reaction times for motion perception when it was consistent with the direction of motion in the concurrently presented verb.

ERP Results

ERPs were time-locked to the onset of the RDKs. For each time interval of interest we performed a repeated measures ANOVA with between-subjects factors of Dot Direction (up/down) and Sensitivity (high/low) and within-subjects factors of Coherence (coherent/random), Verb (up/down) and Electrode (various levels depending on the analysis). The dependent measure was the mean amplitude within the time intervals of interest. The Hunyh-Feldt correction was applied where relevant. In all cases the analyses were done on trials for which participants accurately identified the motion of the RDKs.

100-175ms, Electrodes O1, Oz, O2

In this early interval our Electrode factor included the three occipital channels measured with our electrode array: O1, Oz and O2. This analysis was targeted at early

components indexing visual perception. The significant effects were: Sensitivity by Coherence ($F(1,12) = 11.96, p < 0.005$), Sensitivity by Coherence by Electrode ($F(2,24) = 14.91, p < 0.05$) and Sensitivity by Coherence by Verb by Electrode ($F(2,24) = 6.04, p < 0.01$).

We divided participants by sensitivity to further explore these effects. Participants with low sensitivity scores showed a significant effect for Coherence ($F(1,6) = 7.61, p < 0.05$) as well as an interaction of Coherence and Electrode ($F(2,12) = 5.94, p < 0.01$). Coherent motion elicited a larger negative response ($-6.01 \mu\text{V}$) than the random motion ($-5.66 \mu\text{V}$). This difference was larger on the left side of the head (O1: -0.87 ; Oz: -0.69 ; O2: -0.50 microvolt differences). There was no effect of language or dot direction in this group.

Participants with high sensitivity scores showed an interaction of Dot Direction, Coherence, Verb and Electrode ($F(2,12) = 4.93, p < 0.05$) as well as Coherence with Verb ($F(1,6) = 9.09, p < 0.05$) and Coherence, Verb and Electrode ($F(2,12) = 7.85, p < 0.01$). These interactions were investigated by dividing the participants based on whether they observed upward versus downward motion. For those who saw only downward motion we found a main effect of Verb, with downward verbs eliciting a larger negativity ($-4.45 \mu\text{V}$) than upward verbs ($-3.75 \mu\text{V}$; $F(1,3) = 51.22, p < 0.01$). In other words, we found an amplification of an early visual component elicited by RDKs when the preceding verb described motion congruent with that which the participants were trying to detect. There was no interaction between Verb and Coherence, suggesting this amplification occurred both for coherent (downward) trials as well as for random motion trials (see Figure 4.5a).

For those high sensitivity participants who saw upward motion we found a 3-way interaction of the variables Coherence, Verb and Electrode ($F(2,6) = 11.57, p < 0.01$). However, splitting by coherence revealed no significant effects of verb type for coherently-moving dots (all $F_s < 2.5$) or for randomly-moving dots (interaction of verb by electrode: $F(2,6) = 3.77, p = 0.12$, other $F_s < 2.5$). Numerically, and as can be seen visually in Figure 4.5b, the verbs describing information congruent to the expected dot direction results in a larger negativity for both coherent motion (upward verbs: $-1.77 \mu\text{V}$; downward verbs: $-1.18 \mu\text{V}$) and random motion (upward verbs: $-2.42 \mu\text{V}$; downward verbs: $-2.05 \mu\text{V}$).

Summary – early effects

Participants with high sensitivity scores (and watching downward motion) showed an amplification of the N1 component for motion detection in the context of language in a consistent direction. Participants with low sensitivity scores did not show any change in this response as a result of language. They did, however, show larger amplitude N1s for coherent motion compared to random motion.

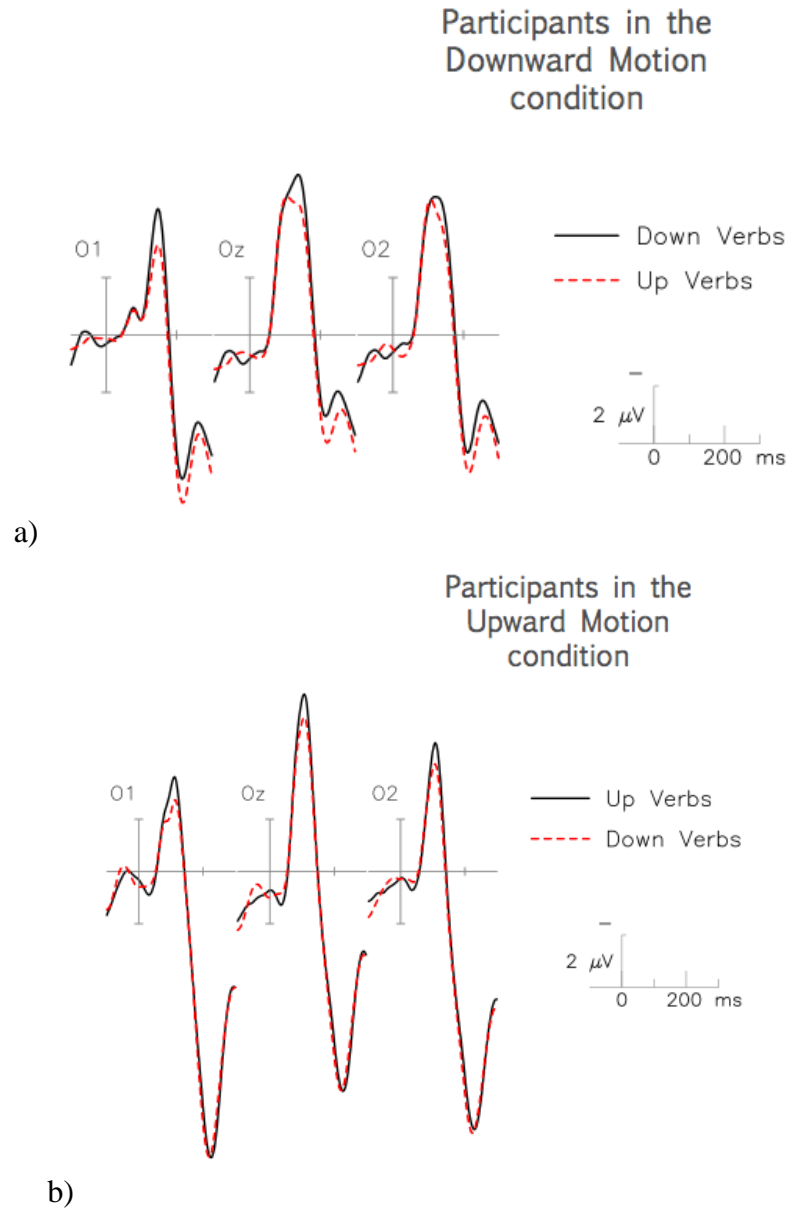


Figure 4.5. N1 plots for high sensitivity participants. Figure a plots ERPs for participants who saw downward motion and Figure b plots ERPs for participants who saw upward motion. These occipital ERPs evoked by the onset of RDK were amplified when the direction conveyed by the preceding verb was consistent with the direction of motion the participant was looking for throughout the experiment. In both cases the coherent motion as well as the random motion trials are displayed. For those watching downward motion (Figure a) the downward verbs amplified these early perceptual ERPs for both coherent and random motion. For those watching upward motion (Figure b) statistically significant results were not found for either coherent or random motion.

400-600, 16 Posterior electrodes

At 16 posterior electrodes we tested for later, decision-related effects such as the P300. Our targeted time interval was 400-600 ms. We began with the same analysis as described in the previous section: between-subjects factors, Dot Direction (up/down) and Sensitivity (high/low) and within-subjects factors, Coherence (coherent/random) and Verb (up/down). This analysis also included the Electrode factor with 16 levels for the posterior electrodes sites (O1, Oz, O2, T5, P3, Pz, P4, T6, TP7, CP3, CPz, CP4, TP8, C3, Cz, C4). The significant interactions included: Sensitivity by Coherence ($F(1,12) = 10.69, p < 0.01$), Coherence by Electrode ($F(15,180) = 3.33, p < 0.05$), Sensitivity by Coherence by Electrode ($F(15,180) = 6.38, p < 0.001$), Dot Direction by Sensitivity by Coherence by Electrode ($F(15,180) = 3.37, p < 0.01$), Verb by Electrode ($F(15,180) = 3.04, p < 0.01$) and marginally Dot Direction by Coherence by Verb by Electrode ($F(15,180) = 1.82, p = 0.094$).

As with our analysis of the early component, and based on the observed interaction effects, we tested participants with high and low sensitivity separately.

Follow-up: High Sensitivity Group: Those with high sensitivity showed an interaction effect based on the type of verb presented for individual trials (Verb x Electrode: $F(15,90) = 3.24, p < 0.05$). We explored this interaction by looking at central/posterior as an additional factor in the analysis. We found a marginal interaction of Verb and Location ($F(1,6) = 4.79, p = 0.071$) which suggested that up verbs show a positivity at posterior sites (up: 6.97 μ V, down: 6.43 μ V) with no differences at central sites (up: 3.64 μ V, down: 3.70 μ V). In this analysis there was also an interaction between verb and electrode ($F(7,42) = 2.88, p < 0.05$) suggesting that the distribution of

these effects likely have a lateral component too (see Figure 4.6). In a targeted post hoc analysis over six posterior electrodes (P3, Pz, P4, O1, Oz, O2) verb type again showed a significant interaction with electrode ($F(5,35) = 3.32, p < 0.05$).

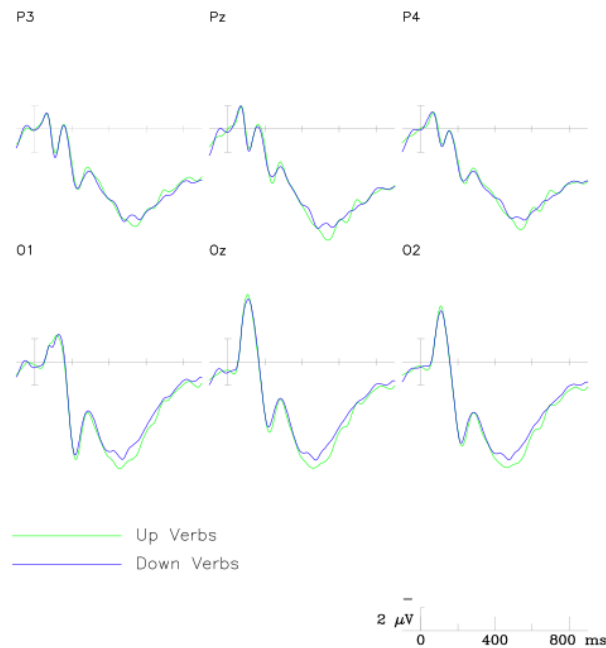


Figure 4.6. Participants with high sensitivity scores showed a positive ERP deflection for upward verbs at a subset of 6 posterior electrodes in the 400-600 ms interval.

The participants with high sensitivity also showed three effects centered around the Coherence factor (Coherence: $F(1,6) = 16.85, p < 0.01$; Coherence x Electrode: $F(15,90) = 12.11, p < 0.0001$; Direction x Coherence x Electrode: $F(15, 90) = 3.29, p < 0.01$). The coherence difference observed for these participants with high sensitivity showed a greater positivity for coherent motion ($5.81 \mu\text{V}$) than the random motion ($4.57 \mu\text{V}$).

The positivity for coherent motion was qualified by differences depending on whether upward or downward motion was observed. For the group watching downward motion we found a main effect of coherence ($F(1,3) = 31.55, p < 0.05$), and an

interaction of coherence and electrode ($F(15,45) = 7.83, p < 0.01$; see Figure 4.7a). We included a post hoc analysis testing the effect of differences of anteriority (2 levels: 8 posterior sites versus 8 central sites) for this effect, and found interactions of verb with anteriority ($F(1,3) = 17.04, p < 0.05$), and verb with both anteriority and lateral electrode site ($F(7,21) = 3.92, p < 0.05$). There was a slightly larger coherence effect over posterior electrodes (coherent = 7.67 μV , random = 5.72 μV) than over central electrodes (coherent = 5.15 μV , random = 3.60 μV); see Figure 4.7a for nine representative electrodes.

The downward motion group also showed a verb by electrode interaction ($F(15,45) = 7.02, p < 0.01$) with more positive ERPs for upward verbs than downward verbs. In a follow-up analysis with anteriority as a factor (2 levels: 8 posterior sites versus 8 central sites) we found that the interaction was due to a difference in anteriority ($F(1,2) = 26.61, p < 0.05$) and anteriority by lateral electrode ($F(7,21) = 3.52, p < 0.05$) with the effect observed only at posterior electrode sites (verb up, posterior: 7.09; verb down, posterior: 6.14; verb up, central: 4.29; verb down, central: 4.29). This was consistent with the effect described above for both the downward and upward groups combined (Figure 4.6).

For the group who watched upward motion we found a coherence by electrode interaction ($F(15,45) = 7.63, p < 0.001$). Again, with a post hoc analysis testing the effect of differences of anteriority for this effect we found an interaction of verb with anteriority ($F(1,3) = 18.98, p < 0.05$) and verb with lateral electrode ($F(7,21) = 3.92, p < 0.05$). In this case (see Figure 4.7b), we found a positivity for coherent trials over

posterior electrodes (coherent = 7.52 μV , random = 6.38 μV) and essentially no effect over central electrodes (coherent = 2.91 μV , random = 3.24 μV).

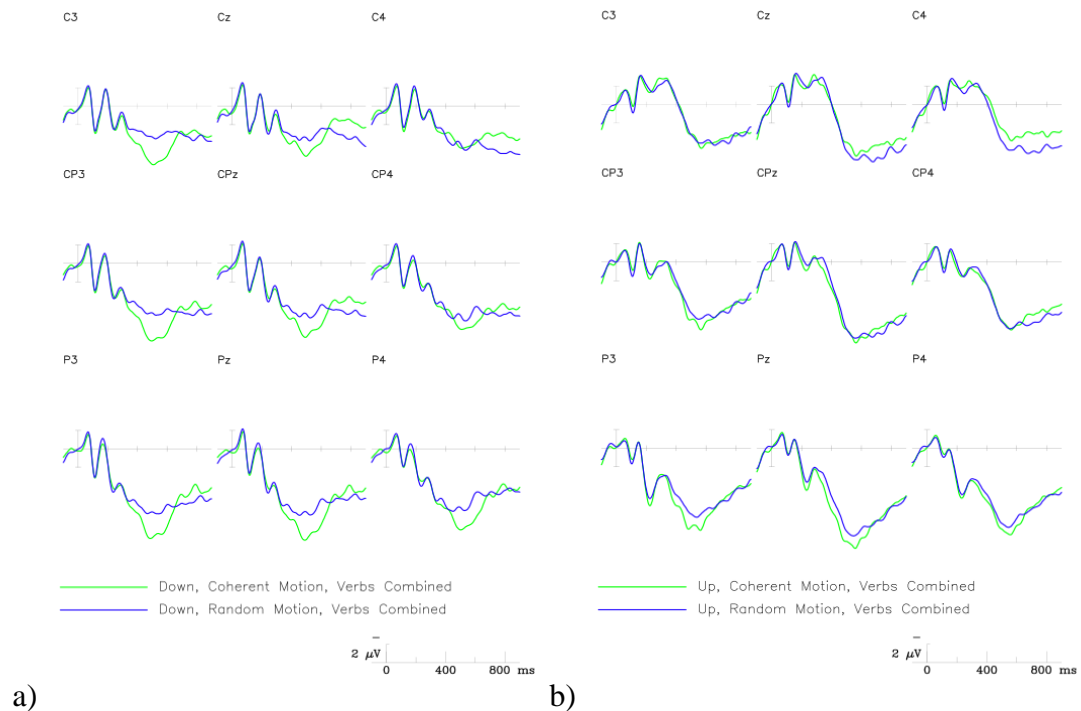


Figure 4.7. The coherence effect for participants with high sensitivity. Figure a shows the group watching downward motion who showed a main effect of coherence as well as a coherence by electrode interaction. Figure b shows the group watching upward motion who showed a coherence by electrode interaction. At posterior electrodes in the 400-600 ms window these groups showed a more positive P3 for coherently moving RDKs relative to random RDKs. This dimension (coherent versus random) was the task-relevant dimension on which participants were making their decisions.

Follow-up: Low Sensitivity Group: Participants with low sensitivity showed a 4-way interaction effect (Dot Direction x Coherence x Verb x Electrode: $F(15,90) = 2.84, p < 0.05$). Focusing the analysis on the coherent motion we found an interaction effect of Dot Direction, Verb and Electrode (see Fig 2a and 2b; $F(15,90) = 2.36, p < 0.05$; other $F_s < 2$). Only the group watching downward motion showed a marginal interaction of verb and electrode (see Figure 4.8a; $F(15,45) = 2.32, p = 0.076$) because of more positive ERPs to downward verb conditions (2.69 μV ; upward verbs:

2.03 μV). Those who saw upward motion did not show such an effect ($F_s < 1.5$). This pattern of effects mirrors the pattern of results found in the reaction time data.

For the test of random motion we found only a marginal interaction between Dot Direction and Electrode (see Figure 4.8b; $F(15,90) = 3.18$, $p = 0.052$; all other $F_s < 1.5$).

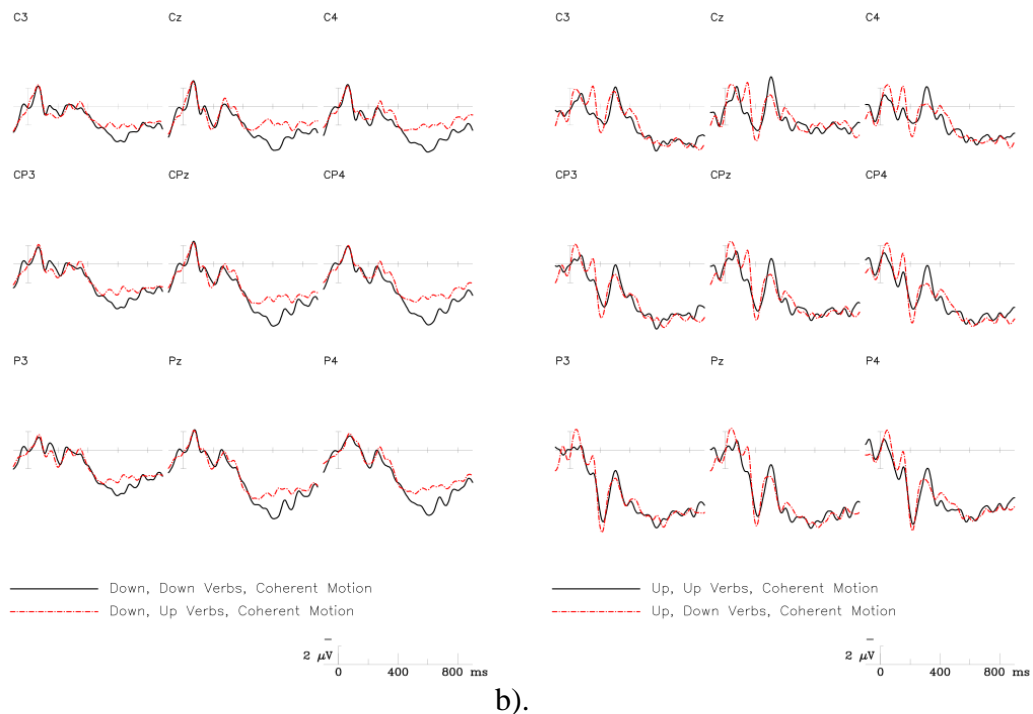


Figure 4.8. ERPs from the low-sensitivity group seeing coherently-moving dots. In the 400-600 ms window there is an interaction of dot direction and verb direction. The interaction was driven by the effect for the participants who saw downward motion (a). For this group congruent verbs elicited a greater positivity than incongruent verbs did, but in both conditions the stimuli eliciting the ERPs were downward RDKs. Those who saw upward motion did not see this effect (b)

Summary – late effects

The high sensitivity group showed a strong P300 positivity for detecting coherent versus random motion. The difference between these conditions is task-relevant (i.e. participants are actively trying to notice coherent versus random trials) which is a typical case for eliciting a P300. Those detecting upward and downward motion both

showed this positivity though it showed a broader scalp distribution for those watching downward motion and was restricted to posterior regions for those watching upward motion. In this population there was also an effect of verb type with the upward verbs eliciting a slightly stronger positivity over select posterior sites.

The low sensitivity group did not show the same task-dependent positivity as seen in the high-sensitivity group. They did show an interaction effect demonstrating a differential modulation of the P300 based on the interaction between verb direction and dot motion. In particular, downward motion language resulted in a positive deflection of the P300 for downward motion trials.

Summary – Experiment 3

Error rates showed a main effect of verb type with more errors in the context of downward verbs (29.7%) than in the context of upward verbs. Similarly, C showed more bias to respond to coherent motion in the context of upward verbs. No d' differences were found. Reaction time showed the expected interaction effect, but later analysis showed that this was particularly significant for those participants with low sensitivity. This interaction demonstrated language with congruent directionality speeded reaction times. The low sensitivity group also showed an ERP effect related to their reaction time difference with an amplified P300 for downward motion detection in the context of downward verbs. In the early interval they only showed a task-related difference of amplified N1 component for coherent RDKs relative to random RDKs. The high sensitivity group showed no effect of reaction time, but they did show an amplification of the N1 component influenced by congruent verb type, again

for the downward motion group. They also showed a decision-related P300 effect likely related to their high performance and a posterior positivity for upward verbs.

Discussion

Summary of findings

In Experiments 1, 2, and 3 (E1, E2 and E3) we tested the effect of vertical language on motion perception as indexed by vertical RDK identification. The experimental variations looked at randomized trial presentation (E1), blocked motion organization (E2 and E3) and ERP indices of the motion responses (E3). The results of these three experiments show that the influence of language on perception is affected by the task demands but can affect low-level access of perceptual resources as well as high-level decision-making processes.

The first two experiments showed a dissociation between reaction time effects and other dependent measures (d' , C, and error rate). The dissociation was driven by the organization of experimental trials. When trials were blocked by motion direction, we saw both low-level (d') and high-level (C, error rate) interactions between the direction of motion described in the language and perceived direction of motion. This finding was consistent with previous findings that incongruent directionality in language can impair detection of directed motion (e.g. Meteyard et al., 2007). When trials were randomized (E1), we only found a high-level modulation of reaction time. In other words, when every trial could display upward, downward or random motion participants' responses were influenced by congruent language, but this influence did not affect perceptual detection, nor was there a decision bias for responses.

The third experiment was an extension of the previous two experiments using ERP measures as a window into the low- and high-level characteristics of the cognitive processes involved in the interaction of language and perception. The experimental design was meant to replicate E2 but we did not find a d' or C effect as expected. This was probably due to several factors such as the lower number of participants, the between-subjects design, and the different calibration procedures discussed in the methods and behavioral results of E3. As a way to accommodate for the fact that we did not replicate the effect emphasized in previous work (i.e. Meteyard et al., 2007) we divided participants in low and high sensitivity groups.

With this division we found several interesting differences – the high sensitivity group showed amplification of the N1 component, an early perceptual component, in the context of congruent language stimuli. They also showed a larger P300 positivity for detecting coherent (vertical) RDKs relative to random motion. The low sensitivity group showed an amplified N1 only for coherent versus random RDKs and a larger P300 positivity for the detection of target RDKs in the context of congruent language stimuli, an effect that mirrored their reaction time differences. In both groups the observed effects were more pronounced for the group of participants who watched downward motion, in some cases this was the only group to show significant effects. The different patterns for high- and low-sensitivity participants also relate to the differences observed in E1 and E2. The following sections will consider the early and late ERP effects in turn and then compare the interpretations with findings of the behavioral experiments (E1 and E2).

Early ERP Effect (N1 amplification)

There have been many studies looking at the role of attention in modulating early sensory processes (see Luck, Woodman & Vogel, 2000 for a review). The typical visual components evoked at the onset of a visual stimulus are the C1, P1 and N1 components. The P1 and N1 are amplified when attention is directed to a certain location or stimulus before the visual onset or with methods such as spatial cueing (Luck et al., 2000). The P1 is the earlier component peaking between 80 and 120 ms and reflecting an early process of increasing sensory gain employed to suppress noise (Hopfinger et al., 2004). The N1 peaks between 100 and 200 ms and is thought to be involved in making visual discriminations, independent of response requirements, arousal and difficulty of the task (Vogel & Luck, 2000).

The negativity we observed showed the timing and posterior distribution of the posterior N1 (see Hopfinger et al., 2004 for a description of the anterior N1). Our task demands required participants to determine whether the array of dots was moving randomly or coherently which is a typical discrimination task for eliciting an N1. What we found was an early modulation of this N1 component in the context of downward motion verbs for high sensitivity participants responding to downward motion (Figure 4.5a). The modulation of this component suggests that the language stimuli describing a congruent directionality with the expected visual direction served as an attentional stimulus to impact early perceptual processing. The fact that this finding is demonstrated for both random and coherent motion is support that this finding is an attentional effect. Regardless of what kind of stimulus is presented, the preceding language has already enabled an increase in the visual signal. As demonstrated with the beha-

vioral experiments, perceptual sensitivity effects are observed only when direction of motion is blocked, so this increase in sensory gain with preceding verbs is probably particularly possible because participants were only responding to one type of motion throughout the experiment.

The timing and distribution of this effect, as well as previous studies showing that the N1 is generated in extrastriate areas (Di Russo et al., 2001) suggest that language can affect processing in relatively early visual areas, a key issue in theories of grounded cognition. These findings suggest that the direction of motion conveyed by a verb is accessed while the verb is being spoken and can affect extrastriate regions within 175 ms of visual processing. Furthermore, this effect occurs on a trial-by-trial basis because every trial presents a new verb, none of which are statistically valid cues because all language types occur equally often and are randomized. Verbal cues have been found to be more effective at enhancing difficult visual discrimination tasks than visual cues even when the cue is not necessarily informative (Lupyan & Spivey, 2010). The current finding adds to this finding by demonstrating that the direction of motion described in a verb can be used as an implicit cue for visual discrimination.

Another ERP component that could be related to our observed N1 is the N2 elicited by motion perception (Heinrich, 2007). The N2 peaks around 200 ms post stimulus onset, is observed at occipital and posterior temporal sites, often shows a rightward skew and likely reflects MT activity (Heinrich, 2007). The amplitude of the N2 increases with larger levels of coherence in response to RDK motion onset (Niedeggen & Wist, 1999) and is not correlated with attention to visual motion (Niedeggen et al., 2002). Relative to the negativity we observed, the N2 is generally later and more

temporally distributed suggesting our effect is instead the posterior N1 as discussed above (Hopfinger et al., 2004). However, there are several early negative components that can be variations on the N2 (see Di Russo et al., 2001 for a description of four different N2 components) with slightly different timing and topographies. Perhaps the negativity we observed reflects both a visual discrimination process and motion perception. This would explain the larger amplitude for the coherent trials (without any effect of verb type) seen for the low sensitivity group, similar to the N2 modulation observed by Niedeggen and Wist (1999). The influence of motion language on the N1 and the N2 as separable components will be an interesting issue to address in the future with a denser occipital and temporal electrode array.

Late ERP Effect (P3 modulation)

The P300 has been described as indexing context updating (Polich, 2007; Rugg & Coles, 1995). This well-studied, centroparietal ERP component has a peak latency that varies from 300 ms to 900 ms, depending on the difficulty of the task (Comerchero & Polich, 1999; Kutas et al., 1977). It is elicited when participants must detect a target stimulus as in the study here – determining if an RDK has coherent motion (the target) as opposed to random motion. When the target decreases in frequency the amplitude of the P300 increases. If there are more cognitive demands (i.e. fewer attentional resources available for target detection), then the P300 amplitude decreases (Polich, 2007).

The P300 can be subdivided into components with more specialized processing specifications. The P3b has a parietal maximum, is task-relevant, and is highly correlated with reaction time. This effect is likely that observed for the high-sensitivity par-

ticipants detecting coherent motion (we will refer to this as the P300 for the rest of the discussion even though it probably reflects the P3b subcomponent). Coherent motion in both directions elicited a larger posterior P300 relative to random motion demonstrating that they were paying attention to the task-dependent stimulus differences and making decision accordingly (Figure 4.7). There were no differences based on verb type. This pattern was a further reflection that the high-sensitivity group was good at performing the task of detecting coherent RDKs.

Both high sensitivity participants who saw upward and downward motion showed the P300 for coherent motion, but the downward group showed a broader scalp distribution. In comparing the size of these effects at the posterior sites we see that the downward motion group (coherent = 7.67 μV , random = 5.72 μV) relative to the upward motion group (coherent = 7.52 μV , random = 6.38 μV) shows a larger positive deflection for random motion. Perhaps this means the decision was more difficult for the upward motion group leading to less of a difference and a more limited distribution of the effect.

The high-sensitivity group also showed a positive P300 deflection with upward verbs, restricted to only a few posterior electrode sites (Figure 4.6). This effect was not related to a particular direction of motion though it was observed specifically for the downward motion group when tested separately. The behavioral results for all participants showed a stronger decision bias (negative C values, on average) in the context of upward verbs. This suggests that people had a bias to respond that they perceived coherent motion in the context of upward verbs. (This bias was not observed in E1 or

E2.) The positivity for upward verbs could be an electrophysiological index of the process by which upward verbs facilitated a coherent motion decision.

The P300 effect was also observed for the low-sensitivity group of participants with a language-specific late ERP modulation. They showed a positivity for downward motion detection in the context of downward motion verbs relative to upward motion verbs. The same group showed an interaction of reaction times with a facilitation for downward motion in the context of downward motion verbs. The ERP difference is particularly interesting because it does not relate to task demands – in both conditions being compared (upward and downward language) participants detected coherent motion. What this P300 amplification suggests then is that more attention was allocated to the decision process after hearing a downward verb. After an upward verb, the attenuated P300 leads to the suggestion that there was less attention allocated or higher cognitive demands for the incongruent condition. Furthermore, this pattern of effect was only observed for participants who were less sensitive at detecting the coherent motion overall. Perhaps the observation of this pattern suggests that those who are not good at the task take a high-level attention-based approach to making coherence decisions. If the P300 modulation does indicate a strategic effect (whether conscious or not), this process could be driving the observed reaction time differences as well.

If a strategic approach were driving the P300 and reaction time effects in E3 then we could infer that the reaction time effect observed in E1 might also have been due to a strategy. Randomized trials could encourage participants to use the language information (though a non-predictive cue for the actual target response) in speeding

their decision to coherent motion detection. The reaction time effects contrast with the findings of E2 that did not show such differences when the motion trials were blocked; in this case blocking motion direction apparently precludes the use of a similar strategy. With blocked trials (E2) clearly the motion language still influenced how participants performed the task (as evidenced by the d' , C and error rate effects) but the mechanisms by which this interaction occurred were likely different.

Grounded Cognition and Motion Language

This investigation began from questions generated by grounded theories of meaning that state language meaning recruits perceptual systems. With respect to the present study, these theories suggest that motion language is subserved by visual motion processing systems. These claims have been demonstrated by showing that language and perception can interfere with each other's processing (Meteyard et al., 2007; Bergen et al., 2007; Dils & Boroditsky, 2010). The implication is that language and perception must employ similar mechanisms at some point in processing. With existing support for this general claim in the literature the next issue to address is at what processing stage(s) this interaction takes place. The broad possibilities that we can contrast are early perceptual processes (low-levels) or decision-making and attention processes (high-levels). Another consideration is the procedural manipulations that can influence these effects.

One important finding presented here is that language can influence perception at low levels of visual discrimination, as seen by N1 amplification for visual motion perception in the context of congruent language. This finding supports previous claims that low-level perceptual sensitivity, as indexed by d' modulation for example, is a

point at which language and perception can interact (Meteyard et al., 2007). Other studies that have investigated this question have primarily dealt with motor representations in language. For example, action verbs, performed with the hand, arm, leg, foot or mouth, can change the pattern of the first 200 ms of reaching movements (Boulinger et al., 2006). Sentence comprehension of transfer events such as, *Anna gives the sandwich to you*, have been found to implicitly involve motor areas generating evoked potentials measured at distal hand muscles, thus suggesting that low-level motor regions are activated in language comprehension (Glenberg et al., 2008). These studies, like ours, demonstrate that low-level processing areas can be accessed during language comprehension.

Other studies have demonstrated that low-level effects can show variation on different task demands or depending on individual differences. For instance, verbal cuing for visual discrimination can be correlated with an individual's mental imagery ability when the location is predictable, but imagery abilities do not correlate with visual discrimination when the location of the target is not predictable (Lupyan & Spivey, 2010). In the motor domain, activity in motor planning regions (prefrontal regions) is involved in comprehending sentences about hockey activities, particularly for participants with more expert hockey knowledge. Lower-level activity in primary motor cortex is activated for comprehension of these same actions for novices. Furthermore, only the prefrontal activation was correlated with hockey sentence comprehension (Beilock et al., 2008).

Consistent with the idea of multiple methods of impact between language and perceptuo-motor systems, we also found high-level decision-making processes for

visual identification was affected by language. The language-specific P300 modulation for the low-sensitivity group demonstrates that language can modulate decision-making. This could simply represent variety in approaches - participants who are worse at detecting coherent motion can attempt to use language cues to aid their decision-making strategy. The low sensitivity group who showed this P300 modulation did not show N1 modulation tempering an argument that low-level perceptual activation is automatically recruited for directional motion language. Individuals can show different types of perceptual and motor representations (e.g. based on experience, Beilock et al., 2008) in comprehension and low-level activation is one of these options. The combination of the findings presented in the current study suggests that the interaction of language with perception can be low-level but can also vary with individual differences.

Language-perception interactions – especially data showing that language affects low-level perceptual or motor processing – are typically interpreted as support for the suggestion that language (perhaps automatically) cues perceptual reactivations. These simulations of perceptual experience are posited to be constitutive of meaning. The data presented here do not support the claim that meaning for motion verbs automatically recruits visual motion processes as part of a unified perceptual meaning system. For one, the experimental design modulates language-perception interactions. If perceptual systems are evoked consistently and automatically then the impact of the verb *plummet*, for example, on a downward-moving RDK, should not differ based on the 29 other experimental trials surrounding that decision. The different patterns of results found in E1 and E2 show that the experimental context does matter and thus

the perceptual evocation is not consistent and possibly not automatic. The findings in E3 suggest that word comprehension can influence attention allocated for a motion-detection task or the decision-making approach to the task. Furthermore, the role of language can differ on an individual basis. Often the findings that language and perception interact are interpreted to mean that language is grounded in perceptual systems but the findings presented here suggest that language can be used in the service of perception. Future claims about language-perception interactions should be moderated and include considerations of language as an attentional and/or decision making cue used for interacting in the world.

Acknowledgment

Chapter 4, will be submitted for publication: Jennifer Collins, Seana Coulson, “Processing specifications for the influence of vertical language on visual motion perception”. The dissertation author was the primary investigator and author of this paper.

References

- Aziz-Zadeh, L., Wilson, S., Rizzolatti, G., Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*, 1818-1823.
- Barsalou, LW. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.
- Beilock, S.L., Lyons, I.M., Mattarella-Micke, A., Nusbaum, H.C. & Small, S.L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences*, *105* (36), 13269-13273.
- Bergen, B.K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*, 733-764.

- Born, R.T., & Bradley, D.C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, 28, 157-189.
- Boulenger, V., Roy, A.C., Paulingnan, Y., Deprez, V., Jeannerod, M. & Nazir, T.A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of Cognitive Neuroscience*, 18, 1607-1615.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., & Freund, H.-J (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13, 400-404.
- Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition*, 2 (1), 79-116.
- Comerchero, M.D. and Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neuropsychology*, 110, 24-30.
- Di Russo, F., Martínez, A., Sereno, M.I., Pitzalis, S. & Hillyard, S.A. (2001). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, 15, 95-111.
- Dils, A.T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences*, 107 (37), 16396-16400,
- Giaschi, D., Zwicker, A., Young, S.A. & Bjornson, B. (2007). The role of cortical area V5/MT+ in speed-tuned directional anisotropies in global motion perception. *Vision Research*, 47, 887-898.
- Glenberg, A.M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D. & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experiment Psychology*, 61 (6), 903-919.
- Goebel, R., Khorrám-Sefat, D., Muckli, L., Hacker, H. & Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, 10, 1563-1573.
- Gros, B.L., Blake, R. & Hiris, E. (1998). Anisotropies in visual motion perception: a fresh look. *Journal of the Optical Society of America A*, 15 (8), 2003-2011.
- Hauk, O. & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the front-central cortex. *Human Brain Mapping*, 21, 191-201.
- Heinrich, S.P. (2007). A primer on motion visual evoked potentials. *Documenta Ophthalmologica*, 114, 83-105.
- Hirai, M., Fukushima, H. & Hirakia, K. (2003). An event-related potentials study of biological motion perception in humans. *Neuroscience Letters*, 344, 41-44.

- Hopfinger, J.B., Luck, S.J. & Hillyard, S.A. (2004). Selective attention: Electrophysiological and neuromagnetic studies. In M.S. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed.). Cambridge: MIT Press.
- Kable, J.W., Kan, I., Wilson, A., Thompson-Schill, S.L., & Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *Journal of Cognitive Neuroscience*, *17* (12), 1855-1870.
- Kable, J.W., Lease-Spellmeyer, J. & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, *14*, 5,795-805.
- 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*, B79-B89.
- Kourtzi, Z. & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*,1, 48-55.
- Kutas, M, McCarthy, G, & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science* *197*: 792-795.
- Levin, B. (1993). *English verb classes and alternations: a preliminary investigation*. University of Chicago press, Chicago.
- Luck, S.J., Woodman, G.F., Vogel, E.K. (2000). Event-related potential studies of attention. *Trends in Cognitive Science*, *4* (11), 432-440.
- Lupyan, G. & Spivey, M.J. (2010). Making the invisible visible: Verbal but not visual cues enhance visual detection. *PLoS ONE*, *5* (7), e11452.
- Meteyard, L. & Vigliocco, G. (2009). Verbs in space: Axis and direction of motion norms for 299 English verbs. *Behavior Research Methods*, *41*, 565-574.
- Meteyard, L., Bahrami, B., Vigliocco, G. (2007). Motion detection and motion verbs: language affects low-level visual perception. *Psychological Science*, *18* (11), 1007-1013.
- Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, *18*, R732-733.
- Niedeggen, M. & Wist, E.R. (1999). Characteristics of visual evoked potentials generated by motion coherence onset. *Cognitive Brain Research*, *8*, 95-105.
- Niedeggen, M., Sahrie, A., Hesselmann, G., Maarten, M., Blakemore, C. (2002). Is experimental motion blindness due to sensory suppression? An ERP approach. *Cognitive Brain Research*, *13*, 241-247.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118* (10), 2128-2148.
- Revill, K.P., Aslin, R.N., Tanenhaus, M.K., & Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences*, *105* (35), 13111-13115.

- Richardson, D.C., Spivey, M.J., Barsalou, L.W. & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, 27, 767-780.
- Rugg, M.D. & Coles, M. (1995). Event-related brain potentials: An introduction. In Rugg, Michael D. & Coles, Michael (Eds.), *Electrophysiology of Mind: Event-Related Potentials and Cognition*. Oxford: Oxford University Press.
- Saygin, A.P., McCullough, S., Alac, M., & Emmorey K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *Journal of Cognitive Neuroscience*, 22, 2480-2490.
- Spivey, M. & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235-241.
- Tettamanti, M., Buccino, G., Saccuman, M.C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. & Perani, D. (2005). *Journal of Cognitive Neuroscience*, 17 (2), 273-281.
- Tootell, R.B.H., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R. & Belliveau, J.W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15:4, 3215-3230.
- Vogel, E.K. & Luck, S.J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, 37, 190-203.
- Wang, J., Jin, Y., Xiao, F., Fan, S. & Chen, L. (1999). Attention-sensitive visual event-related potentials elicited by kinetic forms. *Clinical Neurophysiology*, 110, 329-341.
- Wickens, T.D. (2002). *Elementary Signal Detection Theory*. Oxford University Press, New York.
- Zwaan, R.A., Madden, C.J., Yaxley, R.H., & Aveyard, M.E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science*, 28, 611-619.

Chapter 5

The impact of pictorial and verbal motion content on representational momentum

Abstract

Theories of embodied cognition suggest that language comprehension involves re-activation of relevant perceptual and motor experiences and thus predict that the comprehension of motion verbs recruits cognitive and neural resources associated with motion perception. We tested the hypothesis that the particular visual motion processing systems recruited for motion language comprehension are high-level mechanisms also used for motion inferences. The primary task in a series of three experiments was a representational momentum task, for which high-level motion processing systems are recruited to extrapolate an implied motion trajectory. Perceptual sensitivity on this task was modulated by accompanying photographs of moving objects, sentences describing horizontal motion, and sentences describing vertical motion. Together this set of experiments enables us to make new claims about the types of cognitive and neural processes involved in the comprehension of motion language, namely that the interface between motion language and motion perception is supported by high-level motion inference processes. The data also suggest that the comprehension of motion language is similar to motion inferences derived from static photographs.

Introduction

A topic given much attention in recent studies of language and mental representation is whether perception and action brain systems support language meaning (see Barsalou, 2008 and Fischer & Zwaan, 2008 for recent reviews). The thesis of this approach, often referred to as “grounded” or “embodied cognition”, is that meaningful concepts consist of simulations that involve the reactivation of brain systems that

would be used to perceive that concept's referent (Barsalou, 1999; Zwaan, 2004). Cognitive linguists have also long suggested that meaningful language comprehension is based in conceptualization processes and is not an autonomous cognitive ability. They suggest instead that the representations and processes critical for language meaning are tied to processes used for interaction with the world such as spatial abilities, perspective taking, and attention (Croft & Cruse, 2004; Talmy, 2000). Consequently, much research on grounded cognition has addressed spatial language, and focused on how understanding spatial language impacts visual and motor processes.

Findings that language modulates spatial behavior have been used to indicate that spatial characteristics, as processed by visual and motor systems, are an intrinsic dimension of language meaning. For example, sentences like, *you handed Courtney the notebook* and *Andy delivered the pizza to you*, were presented to participants who had to make a sensibility judgment, a standard type of response in a psycholinguistics laboratory (Glenberg & Kaschak, 2002). But participants were required to indicate the "sensible/not sensible" response with a non-standard button arrangement so their actions were either toward or away from the body. The findings showed that sensibility judgments were faster when sentence direction and arm movement matched (Glenberg & Kaschak, 2002). In the visual domain the orientation of an object implied in a verbal description such as, *the nail was pounded into the wall* (suggesting a nail with a horizontal orientation), affects immediate recognition (Stanfield & Zwaan, 2001) as well as long-term recognition (Pecher, Zanolie & Zeelenberg, 2007) of a line drawing of the relevant object. After hearing such a sentence, participants were asked to do a simple object recognition task in which any picture of a nail should receive a positive re-

sponse. What they found, however, was that pictures with an orientation consistent with that described in the sentence (viz. a horizontally oriented nail) were recognized faster than those without a consistent orientation (viz. a vertically oriented nail) – even though the orientation was not relevant to the task (Stanfield & Zwaan, 2001). These and other studies implicate visual and motoric processes in the comprehension of meaningful spatial language.

The literature on grounded language also contains studies of motion verbs that together suggest the comprehension of motion language both results in the activation of perceptual brain areas (e.g. Kemmerer, Castillo, Talavage, Patterson & Wiley, 2008), and affects people’s performance on perceptual tasks (e.g. Meteyard, Bahrami & Vigliocco, 2007; Bergen, Lindsay, Matlock & Narayanan, 2007; Zwaan, Madden, Yaxley & Aveyard, 2004). Missing, however, is a unified account of the processing system underlying the interaction of motion language and motion perception. Below we propose that motion language interacts with the visual system via a high-level mechanism, and describe three experiments that demonstrate the relationship between motion language, motion images, and high-level motion inference.

Motion language

One early behavioral study on motion language aurally presented sentences such as, *for the first pitch of the softball game, the pitcher hurled the ball towards you*. Participants were required to judge whether two pictures were the same or different objects while listening to the sentences (Zwaan et al., 2004). In the critical trials, both the sentence and picture judgment were about ball movement. When the second ball was larger than the first, the implied movement was toward the listener (“zooming

in”), and when the second ball was smaller than the first, the implied movement was away from the listener (“zooming out”). Zwaan and colleagues (2004) found that reaction times were faster when the implied direction of ball movement was consistent with the direction described in the accompanying sentence. They concluded that when participants hear sentences such as, ...*the pitcher hurled the ball towards you*, they activate a perceptual simulation that allows for priming of ball movement in a particular direction (Zwaan et al., 2004).

In a similarly motivated study, Meteyard and colleagues (2007) found that vertical motion verbs (e.g., *raise*, *plunge*) relative to control motion verbs (e.g., *glide*) impaired sensitivity to motion in a direction-specific pattern. Participants were asked to detect coherent motion in a random dot kinematogram (RDK), a display of hundreds of randomly moving dots, a proportion of which coherently move (in their study, upward or downward). In this paradigm, task difficulty for detecting the coherent movement increases as the proportion of coherently moving dots decreases. Meteyard and colleagues found that perceptual sensitivity to motion in the display was impaired in the context of vertical motion verbs conveying motion in the direction opposite from the RDK. Participants’ decision threshold for responding to the RDKs was also greater when the perceived motion was opposite that conveyed by the verbs. The direction-specific interference on this low-level visual task was interpreted as support for the proposal that the interaction of language and perception occurs at an early perceptual neuro-cognitive processing stage (Meteyard et al., 2007).

Similarly, Dils and Borditsky (2010) found that the motion aftereffect, another low-level visual effect, can be caused by stories about vertical motion. The motion af-

tereffect is a phenomenon that occurs after extended exposure to motion in a particular direction, whereby the perception of motion in that direction is impaired. Interestingly, Dils and Boroditsky (2010) found that the motion aftereffect was not immediately elicited by sentences in their stories, even though early sentences in the paragraphs contained vertical motion language. Motion aftereffects were only observed for sentences presented toward the end of a full story context. Further, the effect was strongest in participants who also showed that mental imagery evoked a motion aftereffect. These data were interpreted as suggesting that the efficiency of visual system recruitment for language might vary across individuals, and the interaction of these systems can occur at the level of the narrative rather than being driven by individual words (Dils & Boroditsky, 2010).

Another set of experiments showed that short motion sentences with upward verbs (*rising*) and downward verbs (*sinking*) impaired a perceptual decision in the related region of space (Bergen, Lindsay, Matlock & Narayanan, 2007, modeled after Richardson, Spivey, Barsalou & McRae, 2003). The particular perceptual decision required participants to decide the shape of an object. When the object was in the region of space cued by the preceding motion verb, (i.e. upper quadrant for *rising*, or the lower quadrant for *sinking*), decisions were slowed. The series of experiments was used to argue that during comprehension people automatically and unconsciously simulate experiences represented by the language. The authors claimed the observed spatial effect was not driven solely by the lexical items because the same results were not found using the verbs in metaphorical contexts (*rising prices*, or *sinking market*). By using

an object detection task in different quadrants of space, these data suggest a spatial attention mechanism supports the visuo-spatial simulation of vertical motion language.

The research reviewed above suggests that the recruitment of perceptual information is part of normal language comprehension, but it leaves open the question of neural mechanisms involved in the interaction of motion language and perception. Language might be integrating with low-level or high-level motion processing systems within the complex, multiple neural pathways supporting visual motion processing. Some possibilities include low-level motion perception (Meteyard et al., 2007), spatial attention (Bergen et al., 2007), motion imagery (Dils & Boroditsky, 2010), and motion inference (Zwaan et al., 2004). By low-level motion processing, we mean motion starting with luminance changes in small receptive fields detected by neurons in V1 and V3 in the occipital cortex and earlier in the visual processing stream (Lu & Sperling, 2001). This low-level type of motion processing is the kind of motion perception evoked by RDKs. These areas subsequently send many projections to MT, an area commonly thought to be essential for motion processing in humans and monkeys (Born & Bradley, 2005). MT is also implicated in high-level motion processing in the absence of low-level luminance differences. For example, certain motion stimuli can be perceived even in the face of lesions to areas V1 and V3 (Vaina, Cowey & Kennedy, 1999; Lu & Sperling, 2001). MT has even been implicated functionally when participants look at static images that only imply motion (e.g. an athlete in mid-stride), but that cannot be activating low-level visual processing mechanisms (Kourtzi & Kanwisher, 2000; Senior et al., 2000).

The embodied language literature has suggested that comprehension of motion language both activates cortical motion processing areas and impacts perception behaviorally. The proposal presented here is that the particular processing mechanism by which language taps into the perceptual system does not activate low-level processing systems, but rather, recruits top-down processing mechanisms used to infer motion. We can think of this mechanism as similar to that underlying motion inference from still images. Static photos of motion events evoke processing in motion-specific areas of visual cortex (MT) relative to equally complex non-motion images (Kourtzi & Kanwisher, 2000). Static photos of motion in a particular direction have also been shown to produce the motion aftereffect, an impairment of subsequent motion judgments in the same direction during the refractory (viz. recovery) period for neurons that code for the relevant direction (Winawer, Huk & Boroditsky, 2008; see also Lor-teije et al., 2007). Although motion images lack the low-level cues that accompany actual motion, they apparently contain enough high-level cues to activate cortical areas that subserve motion processing via interaction between form and motion perception areas (Kourtzi et al., 2008). Motion language likely does the same.

Representational momentum

In order to test the proposal that a high-level motion processing mechanism is recruited for language processing, the following set of studies used a representational momentum (RM) task that results in a well-established perceptual phenomenon of motion perception and expectation (Freyd & Finke, 1984; Freyd, 1987; Hubbard, 2005). RM induces motion processing as indexed by the overestimation of a trajectory's implied endpoint (see Hubbard, 2005 for a review). The RM effect was discovered by

Freyd and Finke (1984) who presented participants with a series of three static rectangles (the “inducing stimuli”) each rotated approximately 17 degrees in a consistent direction from the position of the previous. This series could be interpreted as a single rotating rectangle (e.g., rotating clockwise). The participants’ task was to determine whether a fourth rectangle (“the probe”) was in the same or different orientation as that in the third frame of the sequence. The probe rectangle orientation was manipulated to be in one of three orientations - at the same orientation, rotated 8 degrees further in the implied direction of rotation, or 8 degrees backwards. Participants made more discrimination errors and took longer to respond when the rectangle was rotated in the direction consistent with the implied motion (i.e. “forward”). This effect has been subsequently replicated in many other paradigms with implied motion made by various objects and shapes translating in vertical and horizontal directions, rotations, or non-linear paths (see Hubbard, 2005 for a review).

The RM phenomenon was named to suggest that cognitive processing could reflect, and intrinsically represent, the dynamic processes of the physical world (Freyd, 1987; Freyd & Finke, 1984; Freyd & Finke, 1985; Hubbard, 2005). However, subsequent research has shown that some RM effects are inconsistent with real physical principles (Kerzel, 2003) and performance can be modulated by explicit expectations (Vinson & Reed, 2002). Hubbard (2005) suggests RM is the result of a predictive mechanism that extrapolates perceptual information (e.g. of expected motion) and uses this prediction for action in a dynamic world. Neuroimaging studies of RM show prefrontal areas and temporal areas in the ventral visual stream support this high-level

motion inference process (Amorim et al., 2000; Senior et al., 2000; Senior, Ward & David, 2002; Rao et al., 2004).

RM tasks activate motion processing mechanisms in the absence of low-level visual activation because the timing between the inducing stimuli is too long to evoke the perception of apparent motion (Beck, Elsner & Sliverstein, 1977). Thus, this task is an ideal way to test the proposal that motion language activates visual motion processing at a high level and does not directly impact low-level perceptual systems. Static photos of motion have also been used to elicit effects showing the overestimation of a figure's suggested trajectory (Freyd, 1983; Senior et al., 2000; Senior et al., 2002). The underlying mechanism for motion photographs and RM have been claimed to be motivated by the same process at a computational level, namely perceptual extrapolation in the service of action planning (Hubbard, 2005). RM is also an appropriate paradigm for assessing the hypothesis that motion language and still images activate a similar high-level mechanism of motion extrapolation.

The present study

If language comprehension recruits high-level motion inference processing, such as that used to process static images implying motion, motion language should modulate a perceptual task of motion inference – representational momentum. The following set of studies capitalized on the flexibility of RM and its status as a high-level motion processing task to test if mental simulations of motion language also recruit high-level perceptual systems. In the following experiments, we presented both motion photographs as well as short sentences to assess the impact of meaningful motion stimuli on this task. This set of studies is intended to help clarify how brain sys-

tems underlying motion perception are recruited for relevant language comprehension and potential factors modulating this recruitment.

In Experiment 1 we used a translating triangle to induce the RM effect paired with preceding still images conveying motion events. As in previous work demonstrating that still images can elicit motion-like processing (Winawer et al., 2008; Lorteije et al., 2007; Kourtzi & Kanwisher, 2000; Freyd, 1983), Experiment 1 established our RM paradigm as one that can be affected by motion content. In addition, because static images are known to involve the top-down activation of motion processing areas (e.g., MT), Experiment 1 served as a good baseline for subsequent experiments intended to test our proposal that language interacts with perceptual systems via high-level motion processing systems. Experiment 2 followed as a test of whether horizontal motion verbs, such as *wandering* or *fleeing*, can modulate high-level motion perception. In Experiment 2, we presented short sentences describing horizontal and vertical motion while participants made perceptual discriminations of triangles implying horizontal motion. Based on findings in other studies of vertical (Bergen et al., 2007; Meteyard et al., 2007d) and sagittal movement (Zwaan et al., 2004), we predicted that language describing horizontal motion would modulate motion perception for this task, while vertical motion would not. In Experiment 3 we tested whether vertical verbs (upward/downward) with salient directionality, as tested with other methods, would impact this task of RM.

The first hypothesis presented here is that language comprehension involving motion verbs and the comprehension of static snapshots of motion are subserved by similar high-level motion processing mechanisms. Thus we would expect motion pho-

tographs and motion sentences to similarly influence the processes indexed by the RM phenomenon. The second hypothesis presented here is that motion language recruits the same mechanism as that used for RM and accordingly, motion sentences will result in an amplification of the RM effect, i.e. a larger final position overestimation. The top-down mechanism hypothesized to support language processing presumably works via MT which is direction-sensitive. Thus, the amplification of the RM effect should be observed for conditions in which the directionality of the language and RM stimulus are aligned (see also Bergen et al., 2007 and Meteyard et al., 2007). If, on the other hand, we see consistent effects of all language types on the RM task then we would infer that task impairments reflect the attentional demands of the task, and that the neural processes underlying language semantics are not sufficiently similar to the RM process to evoke selective interference.

Experiment 1

Experiment 1 was designed to examine the impact of still motion images on the RM paradigm. The use of still photos in Experiment 1 allowed us, first, to establish the details of the experimental paradigm eventually used with linguistic stimuli (Experiments 2 and 3); and, second, to compare results associated with different types of motion semantics – pictorial and verbal. Photographs that conveyed either motion or non-motion events were paired with a RM task that induced motion processing with a translating triangle. Participants' task was to indicate whether or not a probe triangle was in the same position as the final triangle in the inducing stimulus. Because the comprehension of motion images presumably recruits the motion inference processes

operative in the RM paradigm, we expected the motion images to impair task performance more than the non-motion images.

Methods

Participants

40 UCSD undergraduates participated in this experiment for course credit.

Materials

Materials included 8 different photographs depicting a moving object, and 8 depicting an object at rest. Each photograph was followed by stimuli from the RM paradigm. The latter involved a rightward-pointing triangle displayed in a series of three static stimulus frames, “the inducing stimuli”, and then a probe frame. Each inducing stimulus was displayed with a 2 cm rightward displacement from the previous figure, evoking the impression of a rightward moving triangle. The probe frame involved the presentation of the triangle in either the same position as the third inducing frame, or at a nearby position.

Photographs were chosen from a larger set of 152 images that were evaluated in a separate norming study. The initial set of photographs was selected to comprise a diverse representation of both motion and non-motion activities. All images were converted to grayscale using Adobe Photoshop and were cropped to 864x720 pixels.

The 30 participants in the norming study were presented with the images projected in front of a room. For each image they responded on a 5-point scale to the questions, “How much motion is in this image?” where 1 indicated “very little”, and 5 indicated “very much”, and “How much speed is evoked by this image?” where 1 indicated “very slow”, and 5 indicated “very fast”. Participants were also asked a third

question, “What direction is evoked by this image?” and given three options: “diagonal”, “horizontal” and “vertical” for which there were corresponding bi-directional arrows for illustration. Although the option of “not applicable” was available for each of the three questions, participants were encouraged to make responses for as many questions as possible.

Eight “motion” images were chosen from this set, with average ratings of 4.2 for motion content and 4.0 for speed. Eight “non-motion” images were chosen, with average ratings of 1.2 for motion, and 1.2 for speed. Motion and non-motion images thus differed significantly in average ratings of motion ($t = 28.72$, $p < 0.0001$) and speed ($t = 13.76$, $p < 0.0001$), but not for the direction evoked (motion set = 26.3° , non-motion set = 28.5° , $t < 1$). See Figures 5.1a and 5.1b for sample motion images and Figures 5.1c and 5.1d for sample non-motion images.

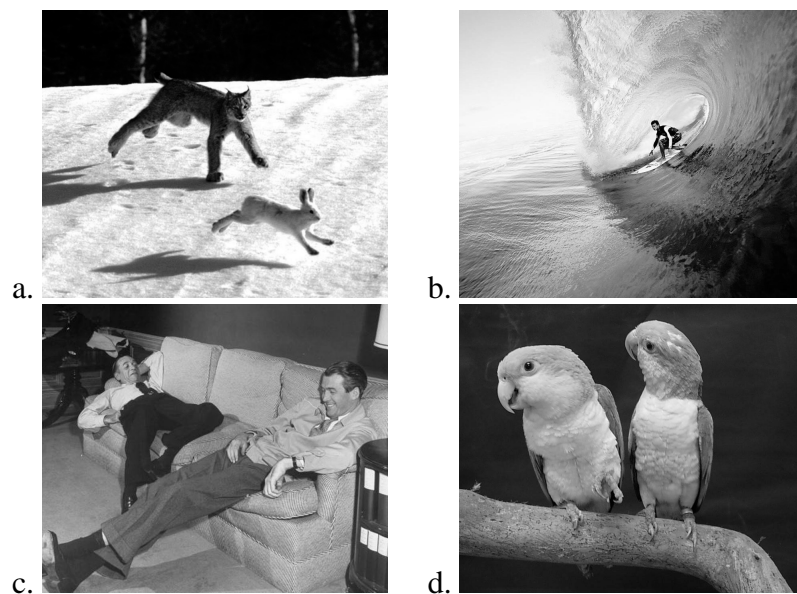


Figure 5.1. Sample images. Panels a and b show two sample images used as primes in the motion condition preceding RM stimuli. Panels c and d show two sample images used as primes in the non-motion condition.

Each trial began with the presentation of a photographic image for 1350 ms, followed by the presentation of the first inducing stimulus, a rightward-pointing triangle. Each of the three inducing stimuli was presented for 250 ms followed by 250 ms of blank screen. Each inducing stimulus was displayed with a 2 cm rightward displacement from the previous figure, evoking the impression of a rightward moving triangle. A fourth triangle, the probe, remained on the screen until participants made their response (see Figure 5.2). We chose to use only rightward moving triangles in this study because rightward movement has been suggested to elicit larger displacement effects, whereas leftward RM has shown contentious results (Halpern & Kelly, 1993).

The task, as with other RM studies, required participants to identify whether the probe triangle was in the same location or a different location on the screen as the third inducing triangle (e.g., Freyd & Finke, 1984 or Vinson & Reed, 2002). Participants pressed the “x” key on a standard qwerty keyboard to respond “same,” and pressed the period key (“.”) for “different” responses. The probe triangle was in one of seven positions relative to the position of the last stimulus triangle: +/-0.9 cm, +/-0.6 cm, +/-0.3 cm, and 0 cm. RM effects are less likely at large displacements (e.g. +0.9 cm) but we chose several displacement values for Experiment 1 to test the paradigm. We doubled the number of 0 cm shifts so the proportion of “same” trials to “different” trials was 1:3 to make the response distribution more even.

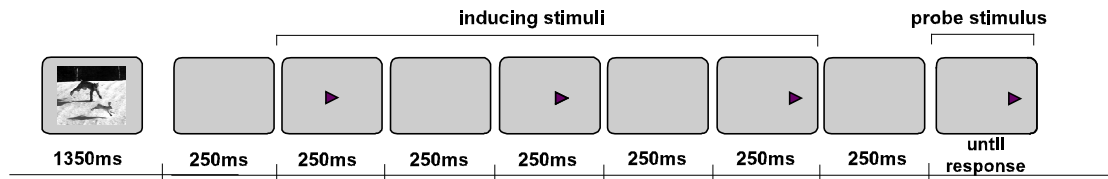


Figure 5.2. The timing of stimulus presentation in Experiment 1. Three triangle stimuli were presented in succession to induce the appearance of a rightward trajectory. Each was presented 2 cm to the right of the previous. The probe stimulus displayed here represents the “same” condition because it is located at the same position on the screen as the final inducing stimulus. The other possible positions for the probe were forward or backward relative to the final inducing stimulus. In the representational momentum paradigm more errors are expected for decisions about forward-shifted probes

The sixteen image primes occurred twenty-four times each, three times with each of the shifted probe positions and six times with the 0 cm final probe position. The starting position of the first inducing stimulus was varied at one of three positions so participants could not make their final judgment relative to a fixed location on the monitor and were instead required to track the triangles to make their response. The 384 total trials were divided into four equal blocks presented between participants in a Latin Squares design. They were presented in pseudo-randomized orders so that trials with the same image prime were always at least three trials apart. The triangle for each RM display was red, orange, purple, blue, green or brown. The color was varied randomly between trials to maintain participants’ interest throughout the experiment.

Procedure

Participants were seated in a small testing room in front of a 14 in computer monitor. They were read a standard set of instructions and tested with a practice set of trials with automated feedback as many times as they needed to feel comfortable with the procedure. During each block they were left alone in the testing room and alerted the experimenter when they were ready to advance to the next block.

To ensure participants were paying attention to the images, after 12.5% of the trials in every block they were given the memory question: “Did the image you just saw have....?”. Some of the memory questions referred to actions portrayed in the image and others referred to a prominent object in the image.

Analyses

The dependent measure was d-prime (d'), as defined by signal detection theory (Wickens, 2002). D' is a transformation of participants' error rates and represents sensitivity to detect changes in the position of the triangle. This value accounts for correct judgments about the probe triangle being in a “different” position independent of participants' overall bias to respond “different”. The use of an unbiased response measure is particularly important for this paradigm, because the correct answer in Experiment 1 was “different” three times as often as it was “same”.

Results

Comprehension Questions

All participants performed between 70.1 and 100% correct (mean = 89.9%, S.D. = 0.07%) on the questions probing for their memory of the images. This suggests participants were paying attention to the content of the images, and remembered what was in the images after they responded to the RM task.

Representational Momentum Task

Our initial analysis involved a 6-way ANOVA with the factor of shift position at the levels, +/-0.9 cm, +/-0.6 cm, +/-0.3 cm, all of which were conditions eliciting a “different” response. To verify a standard RM effect we would expect worse perceptual sensitivity (lower d') for forward shifts (+0.3, +0.6, +0.9 cm) than for the back-

ward shifts. This analysis revealed a robust effect of probe position on d-prime (d') scores ($F(5,195) = 116.18, p < 0.001, \epsilon = 0.49$). This effect was highly expected and represents a replication of the RM effects similar to those described by Freyd and Finke (1984) and Vinson and Reed (2002).

D-prime (d') values at all six probe positions were compared using Bonferroni correction for multiple comparisons. The small forward shift (+0.3 cm), as expected, showed the worst sensitivity ($d' = 0.05$) which was significantly lower than d' for any of the other positions (all $ps < 0.001$). This was followed by the +0.6 cm shift that received an average d' of 1.31, also differing from sensitivity values at all other positions (all $ps < 0.001$). The small backward shift (-0.3 cm, $d' = 1.76$) and large forward shift position (+0.9 cm, $d' = 2.00$) were not statistically different from each other ($p = 0.817$) due to large variability. The medium backward shift showed the next best d' (2.34), differing from all other shift positions (all $ps < 0.01$). The large backward shift (-0.9 cm) showed the best sensitivity ($d' = 2.49$), which also differed from d' at all other positions (all $ps < 0.01$). Mean d' scores at each shift position are shown in Figure 5.3.

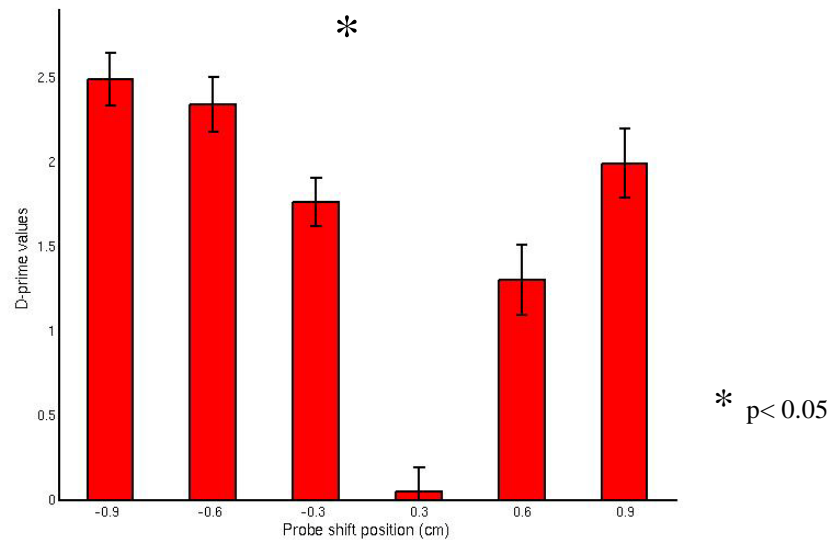


Figure 5.3. The effect of probe position on sensitivity. D' values show a standard RM effect with worse sensitivity seen at forward shift positions relative to backward shift positions of the same absolute size.

Analysis also included planned 2x2 ANOVAs at each shift magnitude with the factors of motion (motion image prime/non-motion image prime) and shift direction (forward/backward). These analyses were designed to explore potential modulation of the RM effect by the image prime. The results of these analyses are presented in Figure 5.4.

Analysis at the large shift positions revealed only a main effect of direction ($F(1,39) = 29.73, p < 0.0001$), due to worse sensitivity (smaller d' scores) at the forward than the backward shift position. Neither the motion effect nor the interaction between motion and shift were significant ($F_s < 1$), indicating that the size of the RM effect was similar after motion and non-motion events.

Analysis of the medium shift positions again revealed a reliable effect of direction ($F(1,39) = 60.42, p < 0.0001$), indicative of RM, as well as a marginal effect of the motion prime ($F(1,39) = 2.90, p = 0.097$) with motion images eliciting worse sen-

sitivity ($d' = 1.77$) than non-motion images ($d' = 1.88$, Figure 5.4). The interaction of motion and direction was not significant ($F < 2$).

Analysis of the small shift position revealed a reliable effect of direction ($F(1,39) = 145.11, p < 0.0001$) reflecting lower d' scores at the forward shift position, and a marginal effect of motion prime ($F(1,39) = 3.16, p = 0.084$) due to worse sensitivity with motion image primes ($d' = 0.85$) than with non-motion image primes ($d' = 0.97$). No interaction of motion and direction was observed ($F < 1$).

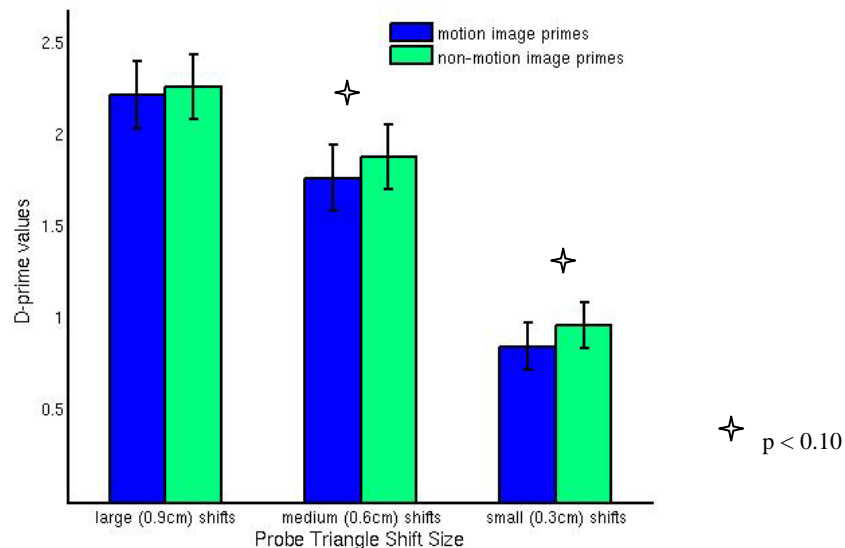


Figure 5.4. Effect of image prime on sensitivity to the probe. Motion image primes resulted in marginally worse sensitivity (d') for triangles at forward and backward positions for small (± 0.3 cm) and medium (± 0.6 cm) shifts, but not for large (± 0.9 cm) shifts.

Discussion

In Experiment 1, participants viewed still photographs of both motion and non-motion events, followed by an inducing stimulus from the RM paradigm (see Figure 5.2). Two effects were observed. First, results replicate the standard RM effect. Namely, when the probe was shifted forward from the final inducing stimulus, participants

were less able to detect that it was in a different position from the final inducing stimulus. In other words, the three stimulus triangles, each translated slightly to the right induced a motion illusion causing participants to be worse at identifying the location of triangles in the suggested direction of motion. More importantly, we also observed a trend for worse task performance following images depicting motion than non-motion events. This marginal effect was observed at the two positions where the RM effect was the strongest (at the small and the medium shifts).

These data indicate that motion picture primes impaired performance on the perceptual discrimination in the RM paradigm. These findings are consistent with research demonstrating that motion images impact subsequent motion processing (Winawer et al., 2008), and those that show motion images activate motion-processing brain areas (Kourtzi & Kanwisher, 2000).

Results, however, did diverge somewhat from our prediction that motion stimuli would amplify the RM effect (worse performance for probe shifts in a direction consistent with the inducing stimuli). The RM effect is characterized, as was found here, by worse performance for probe shifts in a direction consistent with the inducing stimuli. Our prediction was that motion content would amplify this effect resulting in a larger difference between forward and backward shifts of the same magnitude. Instead, we found that motion content impaired performance similarly at the forward and backward shift positions. The RM task has two components to it, the inducing stimuli and the probe presentation. Worse performance at the forward probe positions relative to backward positions indexes motion extrapolation after the presentation of the inducing stimuli. The more general perceptual errors observed here (i.e. equally lower sensi-

tivity at forward and backward shifts) caused by motion semantics suggest that the interaction of motion meaning and perception took place during the induction process of the motion inference rather than as a consequence of the motion inference. In other words, it seems the interference in RM performance caused by the motion images degraded the perception for the inducing stimuli enough to create equal performance decrements at both forward and backward probe positions.

In sum, Experiment 1 showed that images of motion events impaired sensitivity on the perceptual discrimination task in the RM paradigm more so than did images of non-motion events. These findings confirm our hypothesis that the visual inference processes required to extract motion information from static images overlap with the top-down motion inference process indexed in the RM paradigm, and suggest this interference paradigm can be used to assess whether motion language also recruits top-down motion inference processes. Accordingly, Experiments 2 and 3 addressed whether the effect of motion language on task performance in the RM paradigm was similar to that observed in Experiment 1 for the still motion images.

Experiment 2

Experiment 2 was intended to test whether motion language recruits aspects of the same processing resources recruited in the RM paradigm. Consequently, we presented sentences describing motion along horizontal and vertical axes of space as participants performed a RM task with triangles moving in rightward and leftward trajectories. The sentence, *the hiker is striding*, is an example from the horizontal motion condition whereas, *the hiker is climbing*, is an example from the vertical motion condition. The procedure in Experiment 2 was similar to that used in Experiment 1, except

that the presentation of spoken sentences describing motion was substituted for the picture primes. We also eliminated both large shift positions of the probe triangle, in order to focus on the positions at which Experiment 1 suggested motion content was most likely to modulate performance (the small and medium shift positions). In anticipation of Experiment 3 in which both upward and downward motion would be tested we also looked at both rightward (like Experiment 1) and leftward triangle motion.

If the comprehension of sentences describing horizontal motion recruits top-down motion inference processes, like those underlying the horizontal RM effect, then participants should experience more difficulty making judgments when paired with sentences describing horizontal than vertical motion. Based on results from Experiment 1, we expected interference to be manifested as impaired performance at all final shift positions of the probe triangle (both forward and backward shifts). Alternatively, if motion language recruits cognitive and neural resources distinct from those invoked by the RM task, task performance on trials accompanied by sentences describing horizontal motion would be expected to be similar to that for trials accompanied by sentences describing vertical motion. As horizontal language (e.g. *striding*, *creeping*, *leaving*) does not explicitly represent leftward/rightward directionality we expected this set of language stimuli to affect motion perception in these directions to the same degree.

Methods

Participants

Forty-one undergraduates from UCSD initially participated in this experiment. Data from one of these participants were removed because of experimenter error. We report data from forty participants balanced across two lists of stimuli.

Materials

The language materials included 30 sentences constructed by combining each of ten agents (*the tiger, submarine, hiker, man, fire, girl, plane, kite, missile, and raft*) with an upward moving verb, a downward moving verb, or a horizontal verb (see Table 5.1).

Table 5.1. Sentences used in Experiments 2 and 3. All verbs were normed in a separate study so the categories differ only on their implied direction of motion and are similar on measures of motion quantity, speed, concreteness, familiarity and frequency.

	<u>Upward</u>	<u>Downward</u>	<u>Horizontal</u>
The tiger is ...	jumping	squatting	Wandering
The submarine is ...	rising	descending	Fleeing
The hiker is ...	climbing	dangling	Striding
The man is ...	ascending	groveling	Driving
The fire is ...	increasing	diminishing	Creeping
The girl is ...	improving	ducking	Leaving
The plane is ...	flying	plunging	Gliding
The kite is ...	soaring	falling	Trailing
The missile is ...	rocketing	dropping	Coasting
The raft is ...	floating	sinking	Departing

The verbs were part of a larger set normed by a separate group of 61 UCSD undergraduates on several indices: amount of motion, concreteness, familiarity, speed and directionality. For the first four classifications participants responded on a 6-point rating scale. One option was always “none” or “not applicable” and the other five op-

tions (1-5) ranged from “very little” to “very much”. The verbs in the current study were chosen so that the categories did not differ significantly on measures of concreteness (mean = 4.06, SD = 0.56) or familiarity (mean = 4.74, SD = 0.41; both F s < 1). The categories did not differ by frequency (mean = 1528, SD = 2747, F < 1) based on the COBUILD corpus of 18 million words as represented in the Celex database (Baayen et al., 1995). Furthermore, the categories were not significantly different on measures of amount of motion (mean = 2.79, SD = 0.91, F < 1.5) or speed (mean = 2.59, SD = 1.11, F < 1). The only measure for which the three classes of verbs differed was on the measure of directionality ($F(2,27) = 531.60$, $p < 0.001$). In our norming study participants were given the option of “not applicable” as well as five unidirectional arrows pointing to 90°, 45°, 0°, -45° and -90° allowing for quantifiable averages for the upward (mean = +63.4°, SD = 13.1°), downward (mean = -76.9°, SD = 6.2°) and horizontal (mean = -1.0°, SD = 8.2°) verb categories.

Sentence stimuli were recorded by both a male and female speaker in a sound-attenuating chamber using a Beyer Dynamic Soundstar Mk II unidirectional dynamic microphone and Adobe Audition 3.0. All sound files were cropped to the exact sentence length then normalized to be 1375 ms. They were edited and screened to ensure that no distortions occurred during the normalization process. Sentences were presented over headphones while participants watched and responded to the visual display of triangles designed to induce a RM effect.

A picture of a rightward- or leftward-pointed triangle was displayed in a series of three static frames. Each of these three “inducing stimuli” was displayed 2 cm to the right or left of the previous figure, always consistent with the direction the triangle

was pointing. Each inducing stimulus was presented for 250 ms with a 250 ms inter-stimulus interval (ISI), see Figure 5.5 for an illustration of the trial timing. A probe triangle followed the inducing stimuli with a 250 ms ISI in one of five horizontal displacements relative to the position of the previous triangle: ± 0.6 cm, ± 0.3 cm, and 0 cm. These positions were chosen based on the results of Experiment 1. Participants' task was to determine whether or not the probe triangle was in the same position as the immediately preceding triangle. The probe triangle occurred twice as often at the "same" position (0 cm) than it did at each of the shifted positions for a ratio of 1:2 "same" and "different" target responses. The initial position of the triangle was randomly varied around the center of the screen so the final location of the triangle could not be predicted relative to an absolute location on the monitor. The triangle color varied randomly between red, orange, purple, blue, green or brown to increase participants' interest.

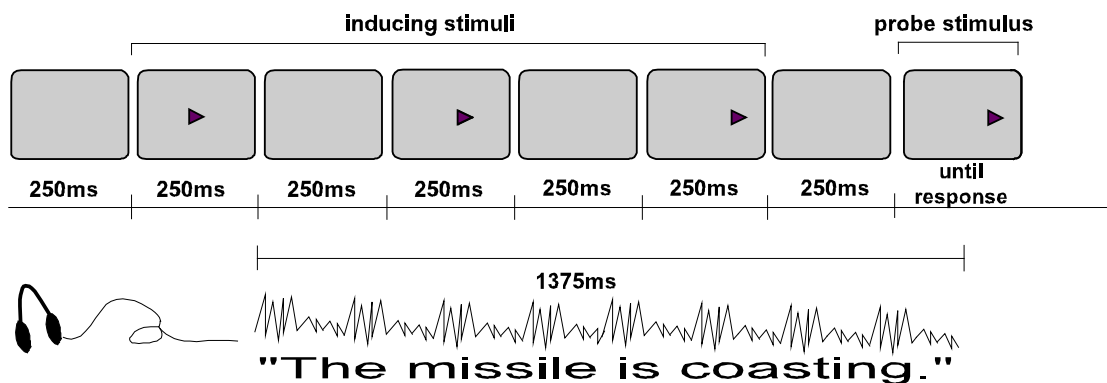


Figure 5.5. The timing of stimulus presentation in Experiments 2 and 3. In Experiment 2 the orientation and direction of displacement was always horizontal, either rightward, as in this illustration, or leftward. Experiment 3 employed the same method except that vertically oriented triangles were displaced either upward or downward.

Half of the test trials (90) presented rightward-moving triangles, 30 of which occurred with horizontal sentences, 30 with upward sentences, and 30 with downward

sentences. Similarly, the 90 leftward-moving triangles occurred with 30 of each type of sentence. The onset of the sentences always followed the offset of the first inducing stimulus (see Figure 5.5). The 180 test trials were completely randomized and separated in two blocks, presented in a Latin Square design between subjects. Thus, every horizontal verb was presented six times total, three in a rightward visual context, and three in a leftward visual context. These test trials were compared with the upward and downward sentences also presented three times each with rightward and leftward movement. Two lists were used so that when collapsed across both lists, each verb occurred with trials ending in all six probe positions of both rightward and leftward moving triangles.

Control trials (without simultaneous sentences) were presented between the two experimental blocks in two control blocks of 60 trials each. Triangles were presented in the six final positions ten times for each direction of motion (right/left).

Procedure

Participants were seated in a small testing room in front of a 14 in computer monitor. Participants pressed the “x” key on a standard QWERTY keyboard to respond “same”, and pressed the period key (“.”) for “different” responses.

Participants were read a standard set of instructions and tested with a set of practice trials with automated feedback. Practice trials were repeated until participants felt comfortable with the procedure. The first test block was always a control block with no sentences. Before the second block, participants donned headphones for the sentence presentation and were given another set of practice trials allowing them to get

accustomed to doing the task while listening to sentences. The third and fourth blocks were control and experimental blocks, respectively.

To ensure that participants were paying attention to the sentences and processing their meaning, they were asked a memory question about the agent or action of the previous sentence on approximately 30% of randomly-selected trials in each experimental block.

Analyses

D-prime (d') values were calculated from the accuracy rate at each shift position, as in Experiment 1, representing perceptual sensitivity. D' scores were subjected to repeated measures ANOVA with the following three factors: triangle direction (right, left), verb type (horizontal, upwards, downwards, no verb) and probe position (+/- 0.6cm, +/- 0.3cm). Degrees of freedom were adjusted with the Greenhouse-Geisser correction where appropriate. As our primary test was to compare horizontal and vertical sentences in regards to their impact on the horizontal motion of the triangles, we used a customized measure for follow-up simple effects that combined upward and downward conditions as a “vertical” factor. This allowed for horizontal and vertical language to be directly compared while compensating for different numbers of trials.

Results

Comprehension Questions

All participants performed at or above 94.2% accuracy (mean = 98.9%, S.D. = 0.02) on the memory trials suggesting they were attending to and comprehending the sentences.

Representational Momentum Task

The omnibus repeated-measures ANOVA on d' values included three factors: triangle direction (right, left), verb type (upward, downward, horizontal, no verb), and probe position (+/-0.6 cm, +/-0.3 cm). Analysis revealed a main effect of probe position, replicating the RM finding ($F(3,117) = 68.83$, $p < 0.001$, epsilon = 0.64). The small forward position (+0.3 cm) of the probe triangle showed the worst sensitivity ($d' = 0.43$) followed by the medium forward shift (+0.6 cm, $d' = 1.16$), and the small backward shift (-0.3 cm, $d' = 1.36$). The large backward position (-0.6cm) showed better sensitivity ($d' = 1.62$) than the other positions.

The omnibus analysis also revealed a main effect of verb type ($F(3,117) = 8.25$, $p < 0.001$, epsilon = 0.85), qualified by an interaction with probe position ($F(9,351) = 5.39$, $p < 0.001$, epsilon = 0.74). The main effect of verb type was due to a greater sensitivity for the control condition without accompanying verbs ($d' = 1.31$) than all the other verb types (all $ps < 0.05$, adjusted for multiple comparisons). Similarly, the interaction between verb type and probe position occurred because the better sensitivity for the control condition relative to the sentence conditions was attenuated for the small forward shift.

Finally, the omnibus analysis revealed a marginal interaction between triangle direction and verb type ($F(3,117) = 2.52$, $p = 0.065$, epsilon = 0.94) suggesting sentence context differentially affected the discriminability of rightward and leftward motion. Rightward and leftward moving trials were subsequently analyzed separately.

Using a repeated-measures ANOVA with the customized, within-subjects contrast of vertical (upward and downward) versus horizontal verb types, we found signif-

ificantly lower d' values for rightward moving triangles following horizontal verbs ($d' = 0.93$) than vertical verbs combined ($F(1,39) = 4.80, p < 0.05$; upward $d' = 1.10$, downward $d' = 1.09$). Analysis of the leftward moving triangles revealed no effects of verb type with this same analysis ($F < 1$). Data for each of the comparisons of vertical versus horizontal sentence types is shown in Figure 5.6.

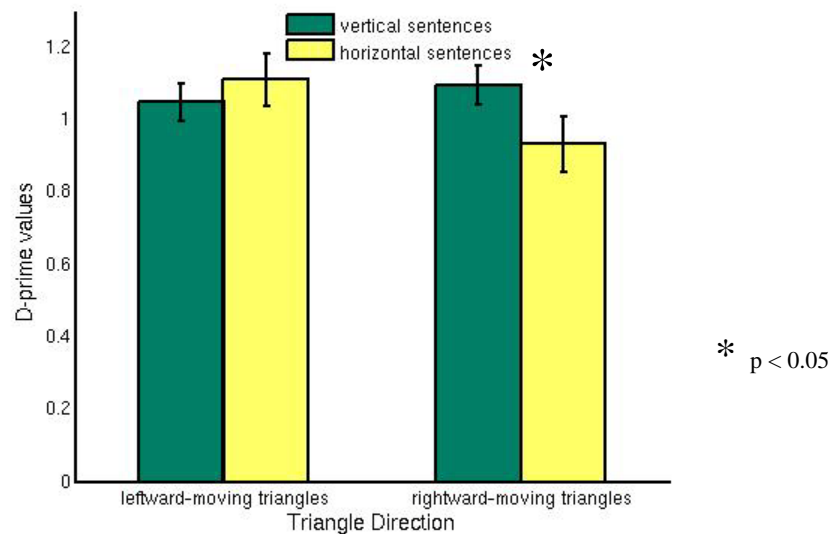


Figure 5.6. The interaction of language context and triangle direction. Sensitivity to triangle shifts for rightward moving triangles was impaired in the context of horizontal sentences relative to vertical sentences. The perception of leftward-moving triangles was not affected by language type.

Summary

These results demonstrate first, a robust RM effect evidenced by lower d' scores for triangles in the forwards than backwards shift positions. Second, the results show better performance on this task for the control condition when no language is presented than for trials accompanied by sentences. Critically, however, direct comparison of the language conditions suggested that perceptual sensitivity was worse for rightward trials accompanied by sentences describing horizontal than vertical motion. No differences were observed for leftward trials.

Discussion

Experiment 2 showed that sentences describing horizontal motion interfere with task performance in the RM paradigm, particularly for rightward moving stimuli. This impairment is relative to conditions with concurrent vertical sentences and is consistent with our prediction that descriptions of motion can impact a task involving motion inference. Although all three motion language categories described the same amount of motion, performance on the perceptual task was worst with the horizontal sentences which described motion along the same axis as the inducing stimulus. Moreover, the pattern of results was very comparable to that in Experiment 1 in which decreased discriminability was evident for probes at both forward and backward shift positions. Taken together, Experiments 1 and 2 suggest that motion content, presented in the form of verbal language or still images, results in worse performance in the RM paradigm. Overall, Experiment 2 suggests horizontal language processing accesses the visual system to a sufficient degree to affect simultaneous, high-level motion processing.

While the findings of Experiment 2 showed the expected difference between horizontal and vertical motion language, this effect was only evident when the inducing triangles moved from left to right. The impact of language on this task could be revealed in an asymmetric way (i.e. stronger in the rightward direction) because of asymmetries in the perceptual process itself. In keeping with this possibility, other studies have shown larger RM effects in the rightward direction (Halpern & Kelly, 1993). There was, however, no evidence for this perceptual asymmetry in the present study as we observed similar RM effects for rightward and leftward movement.

Results of the present study are perhaps more consistent with previous reports of a rightward bias in the schematic representation of motion verbs. Studies in which participants were asked to create schematic representations of motion sentences have found a left-to-right bias (e.g., a propensity to put the agent on the left and patient on the right or drawn an arrow toward the right indicating motion). This bias has been observed particularly for speakers of languages whose writing system also goes left to right (Chatterjee, Southwood & Basilico, 1999; Maass & Russo, 2003; Dobel, Diesendruck & Bólte, 2007). The current results add to these conclusions by suggesting that the representation underlying our intransitive horizontal sentences may be a left-to-right dynamic mental simulation.

Experiment 3

Experiment 2 showed that sentences describing horizontal motion impaired performance on the RM task more so than vertical sentences. Although we attributed the directional sensitivity of the effect to left-to-right simulations of the horizontal motion described in the sentences, an alternative explanation is possible. Horizontal sentences might, for example, be more difficult to process, and thus be more disruptive than the vertical sentences on any concurrent task. We find this explanation unlikely, in view of the similarity between the vocabulary and the structure of the sentences in the three conditions (see Table 5.1). This possibility is addressed in Experiment 3, by using the same sentence conditions paired with a RM task in which the inducing stimulus is translated in the vertical direction.

If task performance impairments observed in Experiment 2 were due to greater complexity of the horizontal sentences, we should expect to replicate our finding of

worse performance with horizontal than vertical sentences. However, if the results of Experiment 2 were due to the hypothesized simulation process, we should expect to observe the reverse pattern, with worse performance for trials accompanied by sentences describing vertical than horizontal motion. Further, as in other studies targeting high-salience directional language (Bergen et al., 2007; Meteyard et al., 2007; Richardson et al., 2003; Zwaan et al., 2004), we predict that it will be more difficult to accurately perform the vertical RM task when it is paired with verbs that describe the same direction of motion, (i.e., upward language for upward motion and downward language for downward motion).

Methods

Participants

Forty-four undergraduates from UCSD initially participated in this experiment. Data from six of these participants were removed because of experimenter error or computer malfunction. We report data from thirty-eight participants balanced across two lists of stimuli.

Materials & Procedure

The language materials used in Experiment 3 included the same set of 30 sentences used in Experiment 2 (see Table 5.1 and Experiment 2 Methods for details).

Participants were presented with a series of three triangle pictures making up the “inducing stimuli”. The triangles were the same as those presented in Experiment 2 except these pointed either upward or downward. Each of the three inducing stimuli was displayed either 2 cm above or below the previous figure, suggesting movement

upward and downward, respectively. Other methodological manipulations were the same as in Experiment 2 including the timing, apparatus and trial organization. Participants' task was also the same: decide whether the fourth frame, viz. the probe, was in the same or different position as the previous triangle.

The final shift manipulation of the probe triangle was analogous to that in Experiment 2. The probe triangle was displaced vertically ± 0.6 cm, ± 0.3 cm, and 0 cm relative to the previous inducing stimulus. All other manipulations including the timing (see Figure 5.5), variation in initial position of the first triangle, the color of the triangles, the distribution of trials into blocks, list differences, and memory questions were the same as in Experiment 2. Because we were interested in whether the language would affect the particular direction of motion, the classifications of interest were those in which the language and triangle direction were congruent versus those in which they were incongruent. Every upward and downward verb was presented six times total, three in a congruent visual context, (i.e. upward sentences paired with upward motion, and vice versa for downward sentences), and three in an incongruent visual context (i.e. upward sentences paired with downward motion, and vice versa for downward sentences). Similarly, horizontal verbs were presented three times with upward and three times with downward movement.

Statistical analysis was similar to that in Experiment 2 – we used an omnibus repeated measures ANOVA with the following three factors: triangle direction (up, down), verb type (upwards, downwards, horizontal, no verb) and probe position (± 0.6 cm, ± 0.3 cm). Unlike Experiment 2, however, we did not include an adjustment to combine the upward and downward verb conditions into a single factor because the

horizontal verbs in this experiment served as a control for the upward and downward language conditions. Instead, we did a planned comparison of the upward and downward motion sentences, motivated by the asymmetric results found in Experiment 2, and because our question of interest was whether the perceptual task would be affected by directionality of the verb.

Results

Comprehension Questions

All participants performed at or above 98.0% accuracy (mean = 99.4%, S.D. = 0.01) on the memory trials suggesting they were attending to and comprehending the sentences.

Representational Momentum Task

In the omnibus analysis, we observed a main effect of probe position ($F(3,111) = 68.72, p < 0.001, \text{epsilon} = 0.54$) indexing the RM effect. The small forward position (+0.3 cm) of the probe triangle showed the worst sensitivity ($d' = 0.57$) followed by the small backward shift (-0.3 cm, $d' = 1.32$) and the medium forward shift (+0.6 cm, $d' = 1.36$). The large backward position (-0.6cm) showed better sensitivity ($d' = 1.58$) than the other positions.

Analysis also revealed a main effect of triangle direction ($F(1,37) = 27.03, p < 0.001$), qualified by an interaction with probe position ($F(3,111) = 9.47, p < 0.001, \text{epsilon} = 0.71$). Similarly, we found a main effect of verb condition ($F(3,111) = 10.24, p < 0.001, \text{epsilon} = 0.96$) that also interacted with probe position ($F(9,333) = 5.02, p < 0.001, \text{epsilon} = 0.72$). The triangle direction effect resulted because participants'

performance was better for upward-moving ($d' = 1.35$) than downward-moving ($d' = 1.07$) triangles, especially at the forward shift positions (see Figure 5.7).

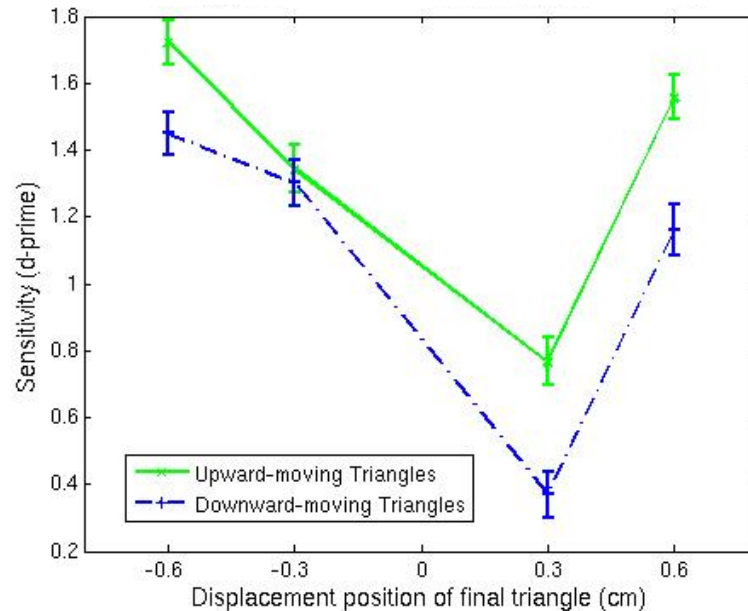


Figure 5.7. D' values for the direction by shift interaction. Better sensitivity (higher d' values) was observed for upward-moving triangles than downward-moving triangles. This difference is largest at the forward shift positions and smallest at the small backward position. The main effect of shift position is also clearly observed in this figure as the small forward shift frequently evokes missed responses (the classic RM effect), which are evident here in lower sensitivity values.

The main effect of verb type reflects a better sensitivity for the control condition with no sentence presentation relative to all conditions including the verbs (all p s < 0.01 , corrected for multiple comparisons; upward condition, $d' = 1.08$; downward condition, $d' = 1.15$; horizontal condition, $d' = 1.17$; no verb control, $d' = 1.43$). The interaction of verb type with shift position indicates that this relationship holds at all probe positions except at the small forward shift where the sensitivity values are much more similar.

Planned Comparisons

Analysis of the upward-moving triangles revealed a main effect of verb type ($F(2,74) = 3.27, p < 0.05, \epsilon = 0.98$) driven by a lower sensitivity for detecting upward triangle movement in the context of upward sentences ($d' = 1.16$) relative to either downward ($d' = 1.35$) or horizontal sentences ($d' = 1.34, ps < 0.05$) as shown in Figure 5.8. The sensitivity for upward moving triangles in the context of downward and horizontal sentences did not differ from each other ($p = 0.97$).

No verb type effect was found in the complementary analysis of downward moving triangles ($F < 1$; Figure 5.8).

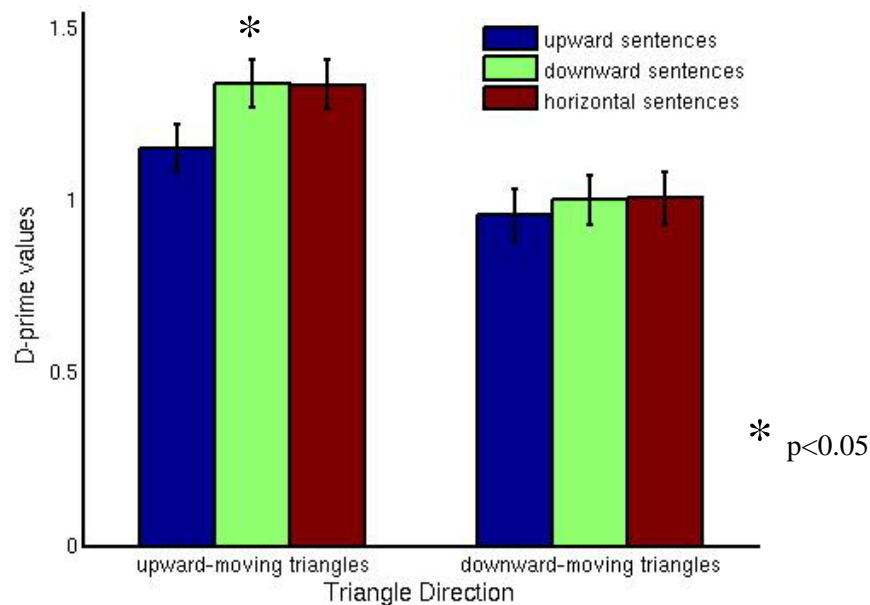


Figure 5.8. Interaction of vertical language and vertical motion. Concurrent upward sentences resulted in worse sensitivity (lower d' values) than did downward or horizontal sentences for the upward direction of motion. No effects of sentence direction were observed on the perception of downward motion.

Discussion

Experiment 3 showed that sentences describing upward motion interfered more with performance on an upward-moving RM task than sentences describing either horizontal or downward motion. Thus language affects high-level motion inference

processes in a direction-specific way. Further, task performance was equivalent for trials accompanied by horizontal sentences and those accompanied by incongruent vertical sentences, arguing against the suggestion that our observation in Experiment 2 of greater interference for horizontal than vertical sentences on the horizontally displaced RM task resulted because the horizontal sentences are intrinsically more difficult to understand. Finally, the interference effect observed here was the same as that seen in Experiments 1 and 2 in that the congruent motion language impaired perceptual judgments similarly at all probe positions. These results are consistent with the hypothesis that comprehension of motion language is related to the high-level motion inference process generating RM effects.

One unexpected result in Experiment 3 was the absence of verb type effects in the downward-moving RM task. Contrary to our prediction that the worst performance on these trials would be for the sentences describing downwards motion, performance levels were similar for all three verb classes. Although unexpected, the observed asymmetry in perceptual interference attributable to verbs describing upwards and downwards motion is consistent with one previously reported by Bergen and colleagues (2007). These researchers attributed the asymmetry to baseline differences in processing difficulty between upwards and downwards motion verbs. This possibility is unlikely in the current study for several reasons. First, verbs were carefully matched for psycholinguistic variables such as word frequency and concreteness, and all three classes of verbs used the same ten agents. Second, in Experiment 2, in which the exact same sentences accompanied horizontal RM displays, we saw no evidence that the upward motion language had particularly high processing requirements relative to the

other two classes of sentences. Post hoc comparisons of the data from Experiment 2, between the upward and downward language conditions, showed no differences in task performance ($F < 1$).

Alternatively, our failure to find the predicted interference effect for sentences describing downwards motion may simply be a function of near-floor performance on the RM task with downwards moving displays. D' scores in this condition were less than 1 for all three verb classes, suggesting near-chance performance. Greater overall performance on upwards than downwards moving displays observed in Experiment 3 is consistent with previous reports, by Vinson and Reed (2002), that downward motion causes larger RM effects than upward motion. Vinson and Reed argue that the asymmetry arises because the underlying motion inference processes tacitly incorporate our knowledge of the way gravity affects downwards motion. Thus the downward RM task in Experiment 3 may have been too difficult, i.e. did not afford enough variation in task performance, for the impact of language to be observable in our paradigm.

Another explanation of the greater impact of verbs describing upwards than downwards motion is that because events described by the upward moving verbs are less common, their simulations require more visuo-spatial processing resources, and thus show stronger interference effects (Bergen, et al., 2007). Perhaps differences in the way we experience upward and downward motion lead to different perceptual tendencies (like those described between upward and downward RM, for example). Theories of grounded cognition hypothesize that language meanings are built from our perceptual experiences, thus upward and downward verbs might well evoke qualitatively different mental simulations.

General Discussion

The present study revealed reduced perceptual sensitivity on RM tasks paired with motion images (Experiment 1), with sentences describing horizontal motion (Experiment 2), and with sentences describing upward motion (Experiment 3). In Experiment 1, images high in motion content impaired sensitivity relative to images low in motion content. Experiment 2 expanded these findings to the domain of horizontal motion language by showing perceptual sensitivity to rightward movement was impaired by auditory presentation of sentences describing horizontal relative to vertical motion. Experiment 3 demonstrated that sensitivity to upward motion was impaired by auditory presentation of sentences describing upward relative to downward vertical motion. The first hypothesis presented in the introduction questioned the similarity of the semantic processing behind motion language and motion images. Our results supported this comparison showing language and images affect performance on RM, both worsening perceptual sensitivity for final location judgments at several positions. The second hypothesis addressed the potential high-level visual motion processing behind motion language comprehension, related to predictions from embodied cognition. The findings presented here showed worse sensitivity when sentences described motion similar to that in the visual task. Because the conditions when the interference happens are only those in which the direction of language and perceived motion are consistent, we conclude that meaning comprehension recruits neural mechanisms used to infer direction-selective motion.

The question of how language meaning is represented neurally is an interesting and important field of inquiry in cognitive science. In line with theories of grounded

cognition we have shown that the process of constructing a meaning representation from language input involves recruiting brain regions used for related experiences. In this case, understanding the meaning of the sentences *the man is driving* or *the tiger is jumping* uses brain regions also used when watching a man drive or a tiger jump. Our results are consistent with prior investigations of motion language and perception suggesting that directional semantic information can interfere with perception and might be represented in a perceptual form (Bergen et al., 2007; Dils & Boroditsky, 2010; Meteyard et al., 2007; Zwaan et al., 2004). By including motion images as a test stimulus we allowed for the comparison between language and images concluding that both of these forms of motion meaning activate similar motion simulations.

An unexpected finding in all three experiments was that the concurrent motion stimuli (images and sentences) impacted performance on the RM task at all probe shift positions. The motion inference in this task is typically indexed by more errors at forward probe shift positions relative to backward shifts. All experiments showed this finding very clearly so we have confidence that visual motion inference occurred. While motion inference is indexed by responses to the probe, it is induced by the three triangles preceding the probe. Perhaps the motion sentences that showed particular interference with the visual task (i.e. those with directionality relating to the visual conditions) caused interference during the inducing phase of the task. The mechanism, as previously proposed, is that the motion language and motion perception both recruit the same motion processing resource. If perception of the inducing motion is degraded due to competition with motion semantics, then perhaps there is more noise regarding the position of the final inducing stimulus, leading to worse sensitivity to detect a

probe at any position. Observed results suggest language content modulates perception of the entire motion stimulus rather than only at the discrete point when a response is elicited.

Processing Considerations

The findings presented in this paper offer a new method for use with motion language to test for the interaction of language and spatial processing. This motion stimulus, with its high-level motion processing characteristics, allows for inferences about the cognitive and neural substrates activated by motion language. In particular, the activation of low-level processing mechanisms is not necessary for the interaction of language and motion perception. RM does not activate low-level perceptual processing because of the relatively long lag between the inducing stimuli. It does, however, activate area MT (Senior et al., 2000; Senior et al., 2002). The neural mechanism inferred from our results – that supporting both language and visual processing – must then be related to MT activity in the absence of low-level motion processing.

The experiments presented here used both static images of motion and motion language to convey motion content. Both of these stimuli types elicited the same pattern of effects on RM. The comparison of these types of motion information provides support to the claim that simulations of motion (by motion language) do not overlap with early visual processing mechanisms. Motion images elicit activation in brain area MT (Kourtzi & Kanwisher, 2000), an effect proposed to be based on the interaction of brain areas involved in human form perception (in superior temporal sulcus), object recognition (in ventral temporal lobe), and their connections with MT (Kourtzi et al.,

2008). In studies with high temporal resolution using EEG and MEG methodologies processing static motion images activates MT later than real motion stimuli by approximately 200 ms (Lorteije et al., 2007; Amorim et al., 2000). This delay is suggested to be the result of feedback connections from visual form to motion processing areas allowing for motion inference and impact on MT by static images (Lorteije et al., 2007; Kourtzi et al., 2008).

Motion language has also been found to activate brain area MT or nearby brain areas (see Chatterjee, 2010 for a review). The activation of this brain region by motion language is likely part of the mechanism resulting in the impact of directed motion on RM. MT, for example, is an important brain region for perceiving the direction of motion (Born & Bradley, 2005). Linguistically evoked motion simulation associated with motion verbs might better resemble the sorts of cognitive and neural processes invoked to infer motion from a static photograph than those used to detect low-level motion elicited by luminance changes in small receptive fields. The inclusion of time sensitive methods to investigate the impact of language on visual motion processing will be an important step in elaborating on the specific processes behind this interaction.

Qualifications for Theories of Grounded Cognition

While an extensive body of work is being built supporting sensory grounding of meaning, there have also been recent appeals for more rigor and specificity regarding the processing mechanisms behind this process (e.g., Chatterjee, 2010). The asymmetries we observed here lead to such qualifications. Horizontal sentences were found to have underlying rightward, dynamic mental simulations. The rightward schematization has previously been argued to be culturally-specific (Dobel et al.,

2007; Maass & Russo, 2003) but the dynamicity presented in this paper is presumably due to visual experiences with motion. Therefore, the simulations supporting language meaning cannot simply be characterized as re-representations of experience with language referents, but rather, arise from a complex processing system continuously integrating a number of our experiences. These studies reinforce the extent that language cues cognitive construals rather than objective states of affairs (Croft & Cruse, 2004), and in view of the highly fictive nature of many semantic phenomena (see Talmy, 2000 for examples) language researchers should not necessarily expect the neurocognitive representation of meaning to mirror objective reality. Many of these experiences will relate to our perception of objects and events but others will be unrelated to the language meaning per se and will be more about experiences with our language system. Presumably the extensive experience that English-speaking, literate adults have with reading (i.e. sweeping our attention over a series of word forms in a left-to-right pattern) impacts the way language simulations run, and also, possibly, the mechanisms underlying the representational momentum effect (Halpern & Kelly, 1993). Similarly, human experiences with upward and downward motion differ, as do the mental simulations of upward and downward motion sentences in terms of their access to visual motion. While continuing to probe the perceptual dimensions of language meaning we must keep an open mind and an open eye for surprises in how experiences affect meaning representation.

Acknowledgment

Chapter 5, will be submitted for publication: Jennifer Collins and Seana Coulson, “The impact of pictorial and verbal motion content on representational momentum”. The dissertation author was the primary investigator and author of this paper.

References

- Amorim, M.A., Lang, W., Lindinger, G., Mayer, D., Deecke, L. & Berthoz, A. (2000). Modulation of spatial orientation processing by mental imagery instructions: a MEG study of representational momentum. *Journal of Cognitive Neuroscience*, 12 (4), 569-582.
- Baayen, R. H., Piepenbrock, R. & Gulikers, L. (1995). The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Beck, J., Elsner, A. & Silverstein, C. (1977). Position uncertainty and the perceptino of apparent movement. *Perception & Psychophysics*, 21 (1), 33-38.
- Bergen, B.K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, 31, 733-764.
- Born, R.T., & Bradley, D.C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, 28, 157-189.
- Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition*, 2(1), 79-116.
- Chatterjee, A., Southwood, M.H., and Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, 37, 395-402.
- Croft, W. & Cruse, D.A., (2004). *Cognitive Linguistics*. Cambridge University Press, New York.
- Dils, A.T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences*, 107(37), 16396-16400,
- Dobel, C., Diesendruck, G., & Bölte, J. (2007). How writing system and age influence spatial representations of actions: a developmental, cross-linguistic study. *Psychological Science* 18(6), 487-491.
- 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.

- Freyd, J.J. & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10* (1), 126-132.
- Freyd, J.J. & Finke, R.A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society*, *23* (6), 443-446.
- Freyd, J.J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, *33* (6), 575-581.
- Freyd, J.J. (1987). Dynamic mental representations. *Psychological Review*, *94*, 427-438.
- Glenberg, A.M. & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9* (3), 558-565.
- Halpern, A.R. & Kelly, M.H. (1993). Memory biases in left versus right implied motion. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19* (2), 471-484.
- Hubbard, T.L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12* (5), 822-851.
- Kemmerer, D., Castillo, G.J., Talavage, T., Patterson, S. & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Language*, *107*, 16-43.
- Kerzel, D. (2003). Centripetal force draws the eyes, not memory of the target toward the center. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29* (3), 458-466.
- Kourtzi, Z. & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12* (1), 48-55.
- Kourtzi, Z., Krekelberg, B. & van Wezel, R.J.A. (2008). Linking form and motion in the primate brain. *Trends in Cognitive Science*, *12* (6), 230-236.
- Lorteije, J.A.M., Kenemans, J.L., Jellema, T., van der Lubbe, R.H.J., Lommers, M.W., & van Wezel, R.J.A. (2007). Adaptation to real motion reveals direction-selective interaction between real and implied motion processing. *Journal of Cognitive Neuroscience*, *19* (8), 1231-1240.
- Lu, Z.-L. & Sperling, G. (2001). Three-systems theory of human visual motion perception: review and update. *Journal of the Optical Society of America A*, *18* (9), 2331-2370.
- Maass, A. & Russo, A. (2003). Directional bias in the mental representation of spatial events: Nature or culture? *Psychological Science*, *14*, 296-301.
- Meteyard, L., Bahrami, B., Vigliocco, G. (2007). Motion detection and motion verbs: language affects low-level visual perception. *Psychological Science*, *18* (11), 1007-1013.

- Pecher, D., Zanolie, K. & Zeelenberg, R. (2007). Verifying visual properties in sentence verification facilitates picture recognition memory. *Journal of Experimental Psychology*, 54 (3) 173-179.
- Rao, H., Han, S., Jiang, Y., Xue, Y., Gu, H., Cui, Y. & Gao, D. (2004). Engagement of the prefrontal cortex in representational momentum: an fMRI study. *NeuroImage*, 23, 98-103.
- Richardson, D.C., Spivey, M.J., Barsalou, L.W. & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, 27, 767-780.
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E.T., Brammer, M. & David, A.S. (2000). The functional neuroanatomy of implicit-motion perception or 'representational momentum.' *Current Biology*, 10, 16-22.
- Senior, C., Ward, J., & David, A.S. (2002). Representational momentum and the brain: an investigation into the functional necessity of V5/MT. *Visual Cognition*, 9 (1/2), 81-92.
- Stanfield, R. A. & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12 (2), 153-156.
- Talmy, L. (2000). *Toward a Cognitive Semantics, Vol. I: Concept Structuring Systems*. The MIT Press, Cambridge, MA.
- Vaina, L.M., Cowey, A. & Kennedy, D. (1999). Perception of first- and second-order motion: separable neurological mechanisms? *Human Brain Mapping*, 7, 67-77.
- Vinson, N.G. & Reed, C.L. (2002). Sources of object-specific effects in representational momentum. *Visual Cognition*, 9 (1/2), 41-65.
- Wickens, T.D. (2002). *Elementary Signal Detection Theory*. Oxford University Press, New York.
- Winawer, J., Huk, A.C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological Science*, 19 (3), 276-283.
- Zwaan, R.A., Madden, C.J., Yaxley, R.H., & Aveyard, M.E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science*, 28, 611-619.
- Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645.

Chapter 6

Conclusion

The introduction began with two theories that structure the research on language grounded in perception and action systems. The Perceptual Symbol Systems theory (Barsalou, 1999; Barsalou, 2008a; Simmons & Barsalou, 2003) emphasizes that language comprehension and all other conceptual processes will always activate perception and action brain areas. The activation of these areas will represent concepts without perceptual stimulation, but do so in patterns similar to previous perceptual experiences. In the Immersed Experiencer Framework, Zwaan makes predictions about the perceptuo-motor representations of meaning in linguistic contexts (e.g. sentences; Zwaan 2004; Zwaan & Madden, 2004). For example, he predicts that shape, orientation and directionality are part of phrase construal. He and Barsalou both predict modal information (e.g. smells, textures, colors, etc.) is active during language comprehension but Zwaan would probably add that this broad activation would primarily happen early in processing before the narrowing of meaning occurred.

As more evidence is collected supporting particular claims of these theories, requests for more rigor in defining the relevant processes also mounts (e.g., Barsalou, 2008b; Zwaan 2009; Fielder, 2009; Chatterjee, 2010; Mahon & Carramazza, 2008; Taylor & Zwaan, 2009). These papers have an overarching positive theme that the existing information supports embodiment theories inasmuch as perceptual and motor information is accessed during language comprehension. They also claim, to different degrees, that the direction of empirical work needs to change from studies of demonstration to studies that provide processing constraints on such theories. The goal of this dissertation was to focus on questions of when and how language accesses perceptual representations. The three methodologies used here were able to address these ques-

tions to different degrees, as discussed next. The combined conclusions, integrated with other work, make headway in some of the challenging issues of this field while also illuminating avenues for future research.

Review of findings

Modality switching (Chapters 2 and 3)

Chapters 2 and 3 utilized the paradigm of the conceptual modality switch which has been important for demonstrating that perceptual representations underlie meaningful language and concepts (Pecher et al., 2003; see Barsalou, 2008a for a review of these findings as support for Perceptual Symbol Systems). In testing this well-studied method while recording event related potentials (ERPs) we were interested in describing underlying processes and modality differences driving this effect that might be masked when using behavioral measures such as reaction time. We formed pairs of property verification trials in which the property of each trial was characterized as referring to a particular perceptual modality. For example, *STUCCO-rough* refers to a tactile property of an object whereas *LEMON-yellow* refers to a visual property of an object. We measured ERPs to the second property in a pair of property verification trials, enabling us to look at the difference in neural processing patterns for pairs of trials referring to the same modality (“no-switch”) versus pairs referring to different modalities (“switch”). Typical behavioral performance shows a reaction time processing cost for a switch condition relative to a no-switch condition. The perceptual modality evoked by the properties is the only variable that has been shown to drive switch effects between pairs of property verification trials (other variables that have been tested are reviewed in the introductions of Chapters 2 and 3). Because an atten-

tional cost is observed for switching between stimuli of different modalities (e.g. beeps and flashes) and analogous reaction time costs are observed for switching between concepts of different modalities, the conclusion is that the concepts are grounded in perceptual systems.

In Chapter 2, we investigated the conceptual modality switch effect in visual and auditory property processing. We found that the very same visual property terms elicited larger amplitude N400 when preceded by auditory property terms than when they were preceded by other visual property terms. Our finding of an amplified N400 for a modality switch supports the claim that the perceptual information manipulated in the experimental design is part of semantic memory. Curiously, the switch from a visual decision to an auditory decision did not result in an N400. Rather, it elicited a larger P300 (or late positive complex, LPC), important for decision making. In Chapter 3 we investigated the conceptual modality switch effect with auditory and tactile properties and found a later, frontal negativity for the switch condition without differences between the auditory and tactile properties. The morphology of this negative component suggested that perhaps mental imagery is also involved in processing the conceptual modality switch (see also, Hald et al., 2011). It was surprising that across modality types we found such variability in ERP components evoked by the conceptual modality switch effect. Behavioral studies have not shown differences based on conceptual modality type when explicitly tested (e.g. Pecher et al., 2009). Previously, reaction time delays have been used to indicate that switching occurs, driven by the perceptual modality, within a single perceptuo-motor system underlying concepts (Barsalou, 2008b). The inconsistencies in our ERP findings challenge this unified

view by suggesting the processes driving the single reaction time effect might be more complicated and involve different systems (e.g. semantic processing, decision making, and possibly mental imagery) under different circumstances.

We also found main effects of modality. In other words, ERP responses differed depending on the modality of property verification (visual, auditory, tactile) regardless of the switch manipulation. In Chapter 2 we found different scalp distributions of ERP patterns in a late time interval (500-800 ms post property onset) for visual versus auditory decisions. In Chapter 3 the tactile properties evoked an increased amplitude N400 relative to auditory properties for both switch and no-switch conditions. The hypothesis that different perceptual brain networks underlie the decisions *LEMON-yellow* versus *LEMON-sour* leads to the prediction that the electrical patterns observed at the scalp would differ because they are driven by distinct distributions of neural activity. This is what we found and the timing of these effects shows that the differential activation occurs both in the period typically associated with meaning processing and afterwards. These findings support theories of grounded cognition. Also consistent with the hypothesis that perceptual networks support modal concepts, in both Chapters 2 and 3, we raised the possibility that some of our conceptual differences could be driven by somewhat multimodal representations (e.g. *bumpy* can be perceived both visually and tactually). The question of unimodal versus multimodal properties will be an important perspective to address in developing the specifications of PSS.

Vertical verbs and dot motion (Chapter 4)

Chapters 2 and 3 (summarized above) demonstrated that perceptual modality is a relevant dimension in the semantic representation of property words. The experiments in Chapters 4 and 5 tested the specificity of perceptual activation via language by directly comparing motion language with motion perception tasks. In Chapter 4 we paired vertical motion verbs with random-dot kinematograms (RDKs) in several experimental designs in order to understand the circumstances under which language interacts with low-level visual processing. This question is relevant in the goal of developing constraints for embodiment theories under the assumption that the low-level activation of perceptual systems by language reflects more direct connections between language and perception than would be indicated with high level perceptual activation (Meteyard et al., 2007).

We replicated previous findings showing that upward verbs facilitate the perception of upward movement in RDKs, while downward verbs facilitate the perception of downward movement in RDKs (Meteyard et al., 2007). This effect was only found, however, when RDKs within a test block consistently moved in the same direction (e.g. all the coherent trials within a block showed upward motion). Rather than a simple demonstration that language affects perception, this procedural manipulation demonstrates that the effects of language on perception are not observed equally in all situations. In this particular situation it seems that the broad perceptual, perhaps attentional, context created by having systematic directionality within a block of trials is important for driving low-level effects. Furthermore, in our study, as in the original study by Meteyard and colleagues (2007), multiple dependent measures were mod-

ulated by the interaction between language and perception. Of particular interest were the measures d' and C which originate in signal detection theory (see Wickens, 2002) and represent low-level perceptual sensitivity and decision bias, respectively. For blocked trials we found both improved sensitivity and more decision bias when the direction of motion of the verb and RDK matched. In the case when the direction of motion of all trials was randomized, verbs influenced the detection of target RDKs only indexed by reaction time (RT) differences. Together, these varied results lead to the conclusion that regardless of the fact that RDKs activate low-level visual cortices, various processes must be involved in language-perception interactions for us to observed both high level (context effects, C and RT modulation) and low level (d' modulation) effects.

The nature of language-perception interaction was further addressed in Chapter 4 with an ERP study of the RDK motion detection task described above. In this study we used RDK trials, again preceded by upward, downward and horizontal verbs on individual trials. Participants were divided into high- and low-sensitivity groups based on their overall d' scores for the analysis. In the high-sensitivity group we found that low-level perceptual systems activated by the onset of an RDK can be influenced by the presentation of a preceding verb within the first 175 ms of visual processing (an N1 component). Specifically, for those watching downward motion, the N1 evoked by an RDK was larger after a downward verb such as *plummet* than after an upward verb such as *sink*. This is a remarkable finding showing that the effect of language on perception can happen very early in visual processing. The finding is consistent with previous studies in demonstrating that language interacts with perception in a direction-

specific way because the amplification is driven by a certain class verbs (i.e. those with congruent directionality). The increased amplitude was not constrained to just the coherent RDKs, however. The congruent verb direction increased both the N1 amplitude for coherent RDKs and random RDKs perception suggesting that the process was similar to attentional effects that have been shown to amplify early visual perception processes (Luck, Woodman & Vogel, 2000).

A later ERP component, the P300, also showed differences in this task. The P300 indexes attention allocation and working memory updating in decision processes with a larger positivity typically evoked by a target, which in this case would be the coherent motion RDKs. The high-sensitivity participants showed a larger P300 for the task-related condition of detecting coherent motion versus random motion, indexing their ability to do the task well. They also showed a small modulation limited to a few posterior electrodes for coherent motion decisions preceded by upward verbs. This effect suggested a bias to make a coherent motion response in the context of upward verbs. The low-sensitivity participants showed a robust effect of verb type on their P300 as well. Downward verbs amplified the P300 elicited by downward RDKs relative to upward verbs. In both conditions of this analysis participants were responding to coherent (downward) RDKs. A modulation of verb type for the same task decision demonstrates the impact of verb direction on a high level, perhaps strategic, process in which language direction is used when making a task-based decision. Together the findings on the N1 and P300 components suggest that for the low-sensitivity group, the process driving their language-perception interaction was restricted to decision-

making processes (P300) whereas the high-sensitivity group showed low-level attentional modulation (N1).

Representational momentum and motion sentences (Chapter 5)

Low-level versus high-level visual processes that interact with language are considered to be an indicator of the nature of the links between language and perception. Chapter 5 shifted the focus of inquiry from motion words to motion sentences and presents a high-level mechanism that is suggested to better represent underlying linguistic forms of motion semantics than low-level visual systems. As will be discussed later, perhaps several systems at different levels are involved in the process. We used the representational momentum paradigm which begins with a series of discrete stimuli arranged so as to induce motion perception. The method presented here used triangles flashed in a sequence implying translational motion either vertically or horizontally. The fourth and final triangle was the probe. Participants were asked whether the probe's location was the same or different from the location of the previous triangle. People consistently made errors when the probe triangle was displaced in a forward direction from the final inducing stimulus by responding that the probe was in the same location as the previous triangle. The interpretation of this effect is no longer thought to be based on an internal representation of physical momentum as initially suggested (Freyd & Finke, 1985), but it is understood to be part of a prediction mechanism that uses perception in the service of action (Hubbard, 2005).

The experiments presented in Chapter 5 paired a translational representational momentum task with different types of stimuli conveying motion meaning (static images implying motion, horizontal motion sentences and vertical motion sentences).

The goals of this study were to understand how motion language relates to the high-level predictive motion task engendered by representational momentum, and to understand how motion language might be related to the processing of static motion images. We found that all three of the meaningful stimuli representing motion – static motion images, vertical motion language and horizontal motion language – interacted with the representational momentum task in the same pattern. Specifically, motion content impaired performance on representational momentum at several final probe positions. The experiments with motion language demonstrated that the directionality of the language mattered. Upward motion sentences impaired performance for upward representational momentum trials and horizontal sentences impaired performance for rightward-moving trials. The null findings for downward and leftward motion were unexpected asymmetries that we attributed to the complex nature of experience. Neurocognitive representations are more likely to represent cognitive interpretations of experience (e.g. different perception of upward versus downward motion) rather than objective states of the world (e.g. symmetry of upward and downward motion).

These findings demonstrated that language can impact a high-level motion inference process (representational momentum) that probably does not recruit low-level visual motion detectors because of the characteristically long lag between the inducing triangle stimuli. In addressing the question of the nature of meaning representations for motion language, this set of experiments suggests that the visual system accessed by meaning is a high-level system. Secondly, the language stimuli were similar to images of motion events in their patterns of modulation of representational momentum performance. This is an interesting comparison because the tasks involved in keeping

an image of a motion event in mind and keeping the meaning of a sentence in mind are superficially quite different tasks relying on different perceptual systems (vision versus audition) and subsequently different forms of working memory. We can use the characterization of the neural systems involved in inferring meaning from static images (Kourtzi et al., 2008) to better understand the processes involved when building meaning representations for motion language.

Perception is part of the meaning comprehension system

In demonstrating all the different ways in which meaning can interact with perception, the studies presented here and those they were modeled after show that perceptual dimensions of language are key features that are extracted automatically and used in comprehension. Chapters 2 and 3 demonstrated that for the semantic task of property verification perceptual information is an inherent part of the property's meaning. The perceptual information (particularly visual information), though irrelevant superficially for the task, modulated the N400 just as other core meaning manipulations can like sentence context and predictability. Furthermore, the perceptual modality of visual, tactile and auditory properties elicit different ERP patterns for the same general task demonstrating the relevance of their underlying perceptual representations during processing.

Another perceptual dimension that is a fundamental part of the semantic representation of language is motion directionality. Both words (Chapter 4) and sentences (Chapter 5) that convey motion in a particular direction recruit visual motion processing systems. The access to perceptual systems by motion information in language is at a relatively high level of processing as it interferes with a motion inference

task and modulated decision-making ERP components. The modulation of the N1 component by motion language in a direction-specific way shows that perceptual access can also take place at a very low level too.

These findings add to the certainty that part of the meaningful representations accessed by our linguistic symbols (words and sentences) can be and often are based in perceptual and motor brain systems. The grounding problem has a parsimonious solution evident in these results. However, the specifications and constraints for how embodied representations are accessed and used are only slowly becoming available with the state of current research, as has been addressed by many recent critiques (Barsalou, 2008b; Zwaan 2009; Fielder, 2009; Chatterjee, 2010; Mahon & Caramazza, 2008; Taylor & Zwaan, 2009). An alternative to embodiment called “grounding by interaction” has recently been put forth by Mahon & Caramazza (2008). In this perspective concepts are represented in an abstract, all-purpose mode that allows for consistency of conceptual representations across linguistic and conceptual forms. Some instantiations of these concepts are related to sensory and motor processing leading to the activity observed in empirical work on embodied cognition. Mahon and Caramazza’s primary point of differentiation from Barsalou is in arguing against the idea that sensory and motor activity is constitutive of concepts. If, as they suggest, the conceptual system is divided between abstract and perceptuo-motor representations then we need to further consider the extent to which our findings are an observation of language meaning and concepts used in the service of related cognitive processes such as perception.

In Chapters 2 and 3 the modality conveyed in the information of one trial influenced the response to the next trial. In Chapter 4, the direction of motion described by the verbs influenced the visual perception of subsequent visual stimuli as well as the decisions about these stimuli. With these examples, we are demonstrating simultaneously that language is subserved by perception but also that language serves perception. In order to delineate, and/or integrate, the systems relevant for the effects discussed here we need more research clearly defining the separable roles of language and perception. The work presented in this dissertation can speak to topics of *linguistic form, processing level, timing* and *individual differences*, all of which will help to shape the constraints on theories of grounded meaning. In addressing these four issues (described below) and their future directions with a more critical eye we will inevitably find clarifications regarding which perceptual effects reveal underlying language representations versus which represent the application of these representations.

Clarifying and constraining issues

Linguistic form

Zwaan suggests that broad perceptual representations are initially activated in the comprehension process and are narrowed with linguistic focus to arrive at a situation model representation (Zwaan, 2004; Taylor & Zwaan, 2008). This implies that the level of linguistic analysis (e.g. words, sentences, discourse, etc.) will change the type of perceptual representation activated with more specificity perceptual representations given more linguistic context. In support of this claim, Bergen et al. (2007) concluded that comprehending motion sentences describing upward or downward motion language interferes with imagery in related regions of space for sentences but not indi-

vidual lexical items (see also Dils & Boroditsky, 2010). For example, when participants attempted to identify an object in the upper quadrant of the visual field they were worse when presented with an auditory sentence like, *the mule climbed* than, *the pipe dropped*, which in turn, impeded judgments in the lower quadrant. This pattern of effects was not found using the same verbs in metaphorical contexts (*the cost climbed*; *the percentage dropped*) and was consequently attributed to interference from the full sentence meaning rather than the verbs alone.

As just described, however, our studies tested words, sentences and property verifications and showed that for all of these linguistic forms there are underlying perceptual specifications. Studies have differentially demonstrated that language and perception interact at the story level (Dils & Boroditsky, 2010; Spivey & Geng, 2001), the sentence level (Chapter 5; Bergen et al., 2007), the word level (Chapter 4; Meteyard et al., 2007) and even after the first two phonemes of a word (Revill et al., 2008). Chapter 4 showed that individual verbs, based on their directionality, can result in the activation of early extrastriate brain regions similar to the low-level effects seen with attentional modulation (Luck et al., 2000). The fact that individual words can have quite specific effects on perception lends doubt to the idea that with increased linguistic context we get more specific perceptual simulations. One important future direction will be to conduct studies that make direct comparisons between these different linguistic forms (words versus sentences, for example) to further address the details of perceptual specifications and how these specifications change with linguistic context.

Processing Level

Multiple linguistic forms can interact with perceptual systems and furthermore, they can interact with various stages of perceptual processing. We found that vertical verbs can drive perception at a low visual level or, for a different group of people, these same verbs can drive decision-making processes (Chapter 4). Using similar vertical verbs in sentence contexts, as well as using horizontal verbs we found interference with the high-level motion inference process of representational momentum (Chapter 5). Similar language elements (vertical motion verbs) in different linguistic forms (words and sentences) and with different task specifications can lead to different modes of language-perception interactions. Even when using the same motion inference task of representational momentum, results differ for different directions of motion (e.g. left versus right and up versus down). Similarly, in Chapter 2 we found that for the same property verification task the perceptual experiences that are referred to (visual or auditory) differentially lead to modulation of either semantic or decision making ERP components when switching between modalities. The commonalities among these findings are, a) perceptual systems are accessed during language comprehension, and b) the way in which perceptual systems are recruited frequently differ, even under common circumstances.

The research on vertical visual processing in language has not only tested a number of levels of language analysis but has used a number of dependent measures as well. Language comprehension can interface with: visual feedback circuits that generate the motion aftereffect (Dils & Boroditsky, 2010), eye movement planning (Spivey & Geng, 2001), spatial attention (Bergen et al., 2007; Richardson et al., 2003),

motion discrimination mechanisms (Meteyard et al., 2007; Chapter 4), and motion inference (Chapter 5). This work offers mixed results that conclude the interaction of language and perception takes place at any number of different perceptual processing stages. While initially this diversity might appear to demonstrate a lack of mechanistic specificity, our findings of multiple potential mechanisms for similar tasks suggest that this diversity might truly represent the diversity of visual processing mechanisms that are available to the linguistic meaning processing system.

An important approach for comparing the variety of perceptual systems used in language comprehension will be to further employ ERP methods. We can look at the temporal differences across experiments to compare the time course of effects in paradigms implicating different perceptual systems. For example, vertical language interferes with visual detection of objects in the upper and lower quadrants of visual space (Bergen et al., 2007; Richardson et al., 2003). The task in these studies was a simple shape discrimination, in some ways similar to the task of detecting coherent or random motion. Interestingly, their behavioral findings showed that the direction of motion described in a preceding sentence (e.g. *the mule climbed*) impaired visual recognition of an object in a congruent region of space (e.g. the upper quadrant) (Bergen et al., 2007). This is in direct contrast to our findings with random dot motion that showed facilitation for congruent language-motion conditions (Chapter 4). Based on our observed ERP differences, we suggested both low-level and high-level processing mechanisms that could be driving the behavioral facilitation effect. It would be interesting to record ERPs with the former study as well, in which participants discriminated between shapes. Would ERPs for a shape discrimination in the upper quadrant also

show amplification of the N1 and P300 components when preceded by *the mule climbed* as we found with RDKs? Or would these components show reduction of these ERP components because the behavioral effects are opposite in these paradigms? Or would we find similarities with only the early or late effect? A huge benefit of using ERPs to learn about cognitive processing is that their response patterns can dissociate from reaction times. Comparing real-time electrophysiological indices for different perceptual tasks can potentially lead to insight into the question of when facilitation versus inhibition effects are observed in the context of motion language, an issue that has not been resolved with behavioral measures.

Timing

In the framework of embodied cognition, an important research direction will be to characterize the dynamic patterns of activation in the language-perception interface. Studies on the role of motor representations in language have increasingly demonstrated that various types of motor representations are active to different degrees during comprehension. For example, motor activity occurs at points when it is relevant to the focus of the discourse (Taylor & Zwaan, 2008), multiple hand shapes relevant to objects described in a sentence can be activated simultaneously (Bub & Masson, 2010), and sentences can impact motor responses when the motor planning is done early or late during sentence comprehension but not when motor planning is done after (Borreggine & Kaschak, 2006; Kaschak & Borreggine, 2008). In the perceptual domain the work on timing is more limited. Our ERP studies presented several new pieces of information about timing and related processes underlying perceptual activation by language.

Words can drive changes in perception in a direction-specific way before the first 200 ms of visual processing (Chapter 4). A recent study has also demonstrated that MT+ can be engaged by motion language with only a few syllables of lexical information (Revill et al., 2008). These two studies show that perceptual information is available and used very quickly in language comprehension. However, the early attentional-perceptual modulation we found for participants' detection of RDKs based on motion language was only observed for a subset of our participants. It will be key in understanding the dynamic processes of language and perception to know the consistency with which early and late visual processes are affected by language.

Our findings also showed that perceptual information plays a modulating role for conceptual tasks during the typical time course of semantic processes (200-500 ms; Chapters 2, 3). These findings provide strong support for Perceptual Symbol Systems (Barsalou, 1999) because they demonstrate that modality-specific information is part of our semantic long term memory. In order to strengthen the claim that this is the default in our conceptual system more demonstrations that perceptual features drive processing in the N400 time window will be necessary. Other perceptual domains, such as the motion and direction characteristics accessible in words like *rise* and *plummet*, can be integrated into studies similar to the conceptual modality switch to further understand which perceptual dimensions of language are part of the long term memory system available for semantic activity. Several of the studies presented here also showed that perceptual features of language modulate relatively late processes (300+ ms; Chapters 2, 3, 4). The later the effect of language on perception, the more likely it is that this interaction is driven by top-down processes, in which case the in-

herent relationship between perception and language is less likely. It will be a worthwhile challenge to sort out the conditions under which earlier and later processes are engaged.

Another direction the study of timing can take is in understanding longer lasting impacts of perceptual features of language. Pecher and colleagues (2004) had participants first verify a target concept for a particular modality (e.g. visual: *apples can be shiny*). Several trials later, the same target item was tested for the same modality (e.g. visual: *apples can be green*) or a different modality (e.g. gustatory: *apples can be tart*). Responses were faster for the second test item when the modality was the same as that tested initially (e.g. *shiny* and *green* being both visual properties). Similarly, the shape and orientation of an object implicitly described in a sentence can impact picture identification even 50 min after the initial sentence exposure (Pecher et al., 2007). In order to understand perceptual recruitment by language it will also be interesting to investigate how the online dynamics of language impact long-term perceptual memories.

Individual differences

The final important issue for embodiment theories presented here is also a challenging one to address experimentally. We need to understand when perceptual representations might not be activated for language comprehension. The theory currently dominant in this field (Barsalou, 1999; Barsalou, 2008b) suggests that perceptual systems support inherent aspects of meaning, and as such, will always be accessed for meaning processing. Criticisms of this view cite work on patients with brain damage as providing only limited support for the claim that perceptual systems are neces-

sary for meaning comprehension (Mahon & Caramazza, 2008). As described in several critiques of the embodiment literature, studies constraining the working hypotheses, demonstrating where it breaks down, and testing how non-sensorimotor cues like expectations impact embodied comprehension effects will lead to clearer predictions for these theories (Chatterjee, 2010; Fielder, 2009; Mahon & Caramazza, 2008). Another possible path for research on the necessity of perceptual systems for language comprehension is to investigate individual differences.

In one of our studies we divided participants by a performance measure and in doing so found differences in how language influenced perceptual recruitment (Chapter 4). Those who were good at the RDK task (the high sensitivity group) used the directional information of the verbs for increasing the visual signal when detecting the RDK. The low sensitivity group, on the other hand, used this directional information for decision making. (For the meaning of “use” I make no claims about conscious versus unconscious use.) It seems that the verbs could provide extra signal information for those already good at the task, and could play the role of a decision making aid for those worse at the task. This hypothesis requires verification but the question still remains why different patterns were observed and what this means about the perceptual specifications underlying meaning processing systems. Some possibilities include: the different patterns reflect inherently different meaning processing systems, they reflect different strategic implementations of language in the service of perception, or they represent a gradation of an otherwise similar perceptual basis of language.

Barsalou presents a “single-system” hypothesis which suggests there is one unified perceptuo-motor system that underlies meaning of all types and the differences

between cognitive tasks are a function of the access mechanism to the “single system” (Barsalou, 2008b). He also denies perceptual symbols are conscious, veridical, or equivalent to mental imagery (Barsalou, 1999). Thus, he would not predict individual differences based on mental imagery abilities in the types of representations they support. Three experiments have explicitly tested the effects of imagery abilities on the interactions between language and perception (Pecher et al., 2009; Dils & Boroditsky, 2010; Lupyan & Spivey, 2010). The first result did not find that mental imagery abilities measured in a number of different ways correlated with the activation of perceptual features of language, as elicited in the conceptual modality switch task (Pecher et al., 2009). The other two studies did show mental imagery effects. They suggested, in turn, that the influence of language on perception is similar to mental imagery (Dils & Boroditsky, 2010), and that mental imagery ability aids perceptual enhancement driven by a verbal cue only when the location of the perceptual stimulus is predictable (Lupyan & Spivey, 2010). While the mixed results are still inconclusive, these results are presenting more of a challenge for a single system view of perceptual representations for language. They suggest that the activation of perceptual representations by language might not be automatic as initially posited (Barsalou, 1999; Zwaan, 2004) and that multiple perceptual processing systems (possibly playing different roles) need to be considered as possible foundations of meaningful language.

Summary

This dissertation began with the development of theories on language and concepts suggesting that in order for a representational system to be properly grounded it needs to inherently link concepts with experiences. Theories of embodiment do this by

suggesting that concepts are reactivations of perceptual and motor brain patterns as well as other experiential patterns such as introspection and emotion. The common theme of the studies presented here (Chapters 2-5) is that perceptual features are accessible during comprehension of several types of language on several time scales and for various perceptual processes. Experimental evidence is growing at a rapid rate in support of the general premise of embodied language in the last decade (Chatterjee, 2010). With the rapid growth of research a number of criticisms have also converged within the past couple of years emphasizing that this field is ripe for the development of constraints.

This chapter has presented several issues in this field that have the potential for clarifying processing constraints of the systems integrating language and perception. The *linguistic form* issue – whether perceptual representations occur at all levels of language and how they might differ – is one not typically raised in critical discussions of this field. Empirical work tends to use short sentences but language-perception effects are also found with single words, word pairs, and short paragraphs. We have demonstrated the strength of effects elicited by words but must work further to make the relevant comparisons between different linguistic forms. This topic also relates to the topic of *processing level* because along with different linguistic forms, various perceptual processes have been tested. The conclusions of the work presented here suggest that several processing levels are viable options as the integration points of language and perception. Furthermore, the reason that different processes might be recruited could vary based on an individual's needs and circumstances. We have made steps regarding the *timing* of language and perception interactions. The characteriza-

tion of the dynamic processes involved in recruitment of perceptual systems for language will be a powerful product that could emerge with closer scrutiny of these issues. Finally, *individual differences* is a topic that will become more important as we better understand the implicit activation by language of perceptual processes over others. In learning how this activity can differ between individuals we will be able to characterize the core aspects of how perceptual meanings are encoded.

Fielder (2009) describes George Kelly's creative cycle against which to compare the recent boom in embodied cognition research. "Loosening consists in the generation of new ideas, gathering new findings and the invention of new methods and measures. ... In the second stage, tightening, the best exemplars that were generated during the loosening stage must then be filtered out in a strict and thorough selection process" (Fielder, 2009, summarizing Kelly, 1955). I agree with Fielder's message that there is plenty of established work in grounded cognition research that should allow us to now dive into the creative tightening process.

References

- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Barsalou, L. W. (2008a). Grounding symbolic operations in the brain's modal systems. In G. R. Semin, & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 9-42). New York, NY, US: Cambridge University Press.
- Barsalou, L.W. (2008b). Grounded cognition. *Annual Review of Psychology*, 59, 617-645.
- Bergen, B.K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, 31, 733-764.
- Borreggine, K.L. & Kaschak, M.P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science*, 30, 1097-1112.

- Bub, D.N. & Masson, M.E.J. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition*, *116*, 394-408.
- Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition*, *2* (1), 79-116.
- Dils, A.T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences*, *107* (37), 16396-16400.
- Fielder, K. (2009). On embodied cognition and mental simulation: A meta-theoretical comment to Zwaan's treatise. *European Journal of Social Psychology*, *39*, 1156-1159.
- Freyd, J.J. & Finke, R.A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society*, *23* (6), 443-446.
- Hald, L.A., Marshall, J., Janssen, D.P. & Garnham, A. (2011). Switching modalities in a sentence verification task: ERP evidence for embodied language processing. *Frontiers in Psychology*, *2* (45), 1-15.
- Hubbard, T.L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12* (5), 822-851.
- Kaschak, M.P. & Borreggine, K.L. (2008). Temporal dynamics of the action-sentence compatibility effect. *The Quarterly Journal of Experimental Psychology*, *61*, 883-895.
- Kourtzi, Z., Krekelberg, B. & van Wezel, R.J.A. (2008). Linking form and motion in the primate brain. *Trends in Cognitive Science*, *12* (6), 230-236.
- Luck, S.J., Woodman, G.F., Vogel, E.K. (2000). Event-related potential studies of attention. *Trends in Cognitive Science*, *4* (11), 432-440.
- Lupyan, G. & Spivey, M.J. (2010). Making the invisible visible: Verbal but not visual cues enhance visual detection. *PLoS ONE*, *5* (7), e11452.
- 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*, 59-70.
- Meteyard, L., Bahrami, B., Vigliocco, G. (2007). Motion detection and motion verbs: language affects low-level visual perception. *Psychological Science*, *18* (11), 1007-1013.
- Pecher, D., van Dantzig, S., and Schifferstein, H. N. J. (2009). Concepts are not represented by imagery. *Psychonomic Bulletin & Review*, *16*, 914-919.
- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying properties from different modalities for concepts produces switching costs. *Psychological Science*, *14* (2), 119-124.

- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*, 164-167.
- Revill, K.P., Aslin, R.N., Tanenhaus, M.K., & Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences*, *105* (35), 13111-13115.
- Richardson, D.C., Spivey, M.J., Barsalou, L.W. & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*, 767-780.
- Simmons, W.K. & Barsalou, L. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20* (3/4/5/6), 451-486.
- Spivey, M. & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65*, 235-241.
- Taylor, L.J. & Zwaan, R.A. (2008). Motor resonance and linguistic focus. *Quarterly Journal of Experimental Psychology*, *61*, 896-904.
- Taylor, L.J. & Zwaan, R.A. (2009). Action in cognition: The case of language. *Language and Cognition*, *1*, 45-58.
- Wickens, T.D. (2002). *Elementary Signal Detection Theory*. Oxford University Press, New York.
- Zwaan, R.A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In *The Psychology of Learning and Motivation*, Vol. 44, ed. B.H. Ross, New York: Academic Press.
- Zwaan, R.A. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology*, *7*, 1142-1150.
- Zwaan, R.A., & Madden, C.J. (2004). Updating situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 283-288.