

MOTOR DYNAMICS OF EMBODIED COGNITION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sarah Elizabeth Anderson

January 2011

© 2011 Sarah Elizabeth Anderson

MOTOR DYNAMICS OF EMBODIED COGNITION

Sarah Elizabeth Anderson, Ph. D.

Cornell University 2011

Predominant theories of cognition have previously emphasized the modularity of processing, in which individual isolated modules process information free from the influence of other types of information. However, more recent theories suggest that cognition is much more linked to motor and sensory processes than modular theories suggest. In this dissertation, I outline the history of these two approaches in language processing and review recent evidence in favor of the central role of modality specific information in language comprehension. This review suggests that motor representations and actions can influence language processing in predictable ways. Next, I present new data to address some of the outstanding questions raised by the embodied perspective. Chapter 3 investigates the developmental changes of embodied cue use, presenting new data from sentence-processing in five-year-old children. Chapter 4 investigates the interaction of subtle grammatical information, specifically verbal aspect, with motor information. Taken together, the chapters of this dissertation suggest that, while many questions remain unanswered, the embodied cognition approach contributes to our understanding of language processing.

BIOGRAPHICAL SKETCH

Sarah Anderson is from Mayfield, Kentucky. In 2003, she graduated from Randolph-Macon Woman's College in Lynchburg, Virginia Phi Beta Kappa and with honors in Psychology. She completed her PhD in Psychology in 2010, and will join the faculty in Psychology at University of Cincinnati in Spring, 2011. Her research focuses on the way perceptual and motor information interact and ground both language and conceptual processing.

This is dedicated to all of my family, especially Mom, Dad, and Archit Darling.

ACKNOWLEDGMENTS

Thank you to all of my academic family. My advisor, Michael Spivey, has been infinitely patient with me, and I hope to be half as good an advisor as he's been. His assistance is on every page of this dissertation.

Thank you to Thomas Farmer, Michael Goldstein, and Jen Schwade for their work in Chapter 3. Thanks to Teenie Matlock for her work in Chapter 4. Thanks to all of you for being great people and for helping outside of the lab as well.

Thank you to my family. Thanks to Mom and Dad, Peggy and John Anderson, for putting up with me and for always supporting and believing in me.

Finally, thank you to Archit Lal. You are my favorite.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH

iii

DEDICATION

iv

ACKNOWLEDGMENTS

v

TABLE OF CONTENTS

vi

LIST OF FIGURES

viii

LIST OF TABLES

x

CHAPTER 1: INTRODUCTION AND OVERVIEW

1

CHAPTER 2: THE ENACTMENT OF LANGUAGE: DECADE OF INTERACTION BETWEEN LINGUISTIC AND MOTOR PROCESSES

8

CHAPTER 3: INDIVIDUAL DIFFERENCES IN MEASURES OF LINGUISTIC EXPERIENCE ACCOUNT FOR VARIABILITY IN THE SENTENCE PROCESSING SKILL OF FIVE-YEAR-OLDS

36

CHAPTER 4: ON THE PATH TO UNDERSTANDING VERBAL ASPECT

62

CHAPTER 5: CONCLUSIONS AND DISCUSSION

94

REFERENCES

102

LIST OF FIGURES

Figure 1: Example of one- (top) and two- (bottom) referent displays shown to children.	45
Figure 2: Time-normalized ambiguous- and unambiguous-sentence trajectories elicited in the one-referent (top) and two-referent (bottom) contexts.	52
Figure 3: Relationship between vocabulary and number of errors in the two-referent, ambiguous-sentence condition. Children with higher vocabulary scores (richer linguistic experience) made fewer errors only in this condition	56
Figure 4: Visual scene accompanying sound files in Anderson, Fausey, Matlock, & Spivey (2008).	67
Figure 5: Example target visual scene accompanied by target sentences in Experiments 1, 2 and 3.	70
Figure 6: Movement duration differences for the overall trajectory, left panel, and in the region of the visual scene corresponding to the path, right panel, for Experiment 1.	73
Figure 7: Drop locations in response to past perfective (a), past imperfective (b), future perfective sentences (c), and future imperfective sentences (d), for Experiment 2.	77
Figure 8: Average time-normalized trajectories, for Experiment 2.	79
Figure 9: Movement duration differences for the overall trajectory, left panel, and in the region of the visual scene corresponding to the path, right panel, for Experiment 2.	81

Figure 10: Drop locations in response to simple past sentences, left panel, and past progressive sentences, right panel, for Experiment 3.	86
Figure 11: Average time-normalized simple past and past progressive sentence trajectories in rough and smooth terrain contexts, for Experiment 3.	86
Figure 12: Movement duration differences in the region of the visual scene corresponding to the path, for Experiment 3.	88

LIST OF TABLES

Table 1: Error types causing a trial to be excluded from all analyses, per condition (Chapter 3).	49
Table 2: Target stimuli accompanying visual scene for Experiment 2 (Chapter 4).	76
Table 3: Target stimuli accompanying visual scene for Experiment 3 (Chapter 4).	84

CHAPTER 1

INTRODUCTION AND OVERVIEW

Do language and action influence one another? If so, what is the nature of this interaction? Is language processed in isolation, only sharing completely processed information with action areas of the brain? Or do language and action influence one another continuously, such that action does not simply execute a completed command from language processing? The goal of my dissertation is to investigate and begin providing answers to these types of questions.

Research between the 1950's and the 1990's was dominated by modular views of cognition. In this framework, symbols processed within the cognitive system are both amodal and arbitrary (as described in Barsalou, 1999). Similarly, the modular view of cognition implies that information is processed in an isolated and encapsulated fashion. Only when these encapsulated processors have completed their operations do they share their outputs with other encapsulated processors. Specifically, modular theories of cognition hypothesize that language is processed in dedicated language modules, free from the influence of other sources of information. Only when these language modules have completed their processing do they share this information with other areas of the brain. Hence, modular views of cognition call for separate processing of different types of information (Fodor, 1975; Pylyshyn, 1984). According to these theories, language and action may interact, but only after each module has completely finished its own processing.

Despite the predominance and power of these theories, there is a great deal of evidence in support of more interactive approaches. Take, for example, the well-known McGurk effect (McGurk & MacDonald, 1976). In this experiment, participants saw a face repeating the syllable "ga" while simultaneously hearing the syllable "ba" via an auditory stream. Despite the fact that the auditory information

unambiguously supports the perception of “ba,” participants reported perceiving the syllable “da.” Although the auditory input supports “ba,” the visual input of the lips remaining open is incongruous with this perception. The best fit for the auditory and visual stimuli then is “da”. Extending this work by digitally altering faces and sound files along the “bah-dah” continuum, Massaro (1999) provided evidence that the visual perception of the speaker’s mouth had an immediate and graded effect on which of the two phonemes was reported. Hence, visual information significantly alters participant’s perception of auditory information.

Importantly, this interaction is bi-directional. Not only does visual information alter auditory perception, but also auditory stimuli influence the perception of visual information. For example, when a single flash of light is accompanied by two auditory beeps, the visual flash is often perceived as *two* flashes of light (Shams, Kamitani, & Shimojo, 2000). Moreover, auditory information influences the perception of a dynamic motion event. When a leftward-moving disk and a rightward-moving disk are animated on a computer screen so that they pass each other on slightly different depth planes, participants perceive the event as the disks passing by each other. However, if a simple 2.5 millisecond auditory click is delivered at the point of visual coincidence, observers more often perceive this same dynamic visual event as the two disks bouncing off each other and reversing their directions of movement (Sekuler, Sekuler, and Lau, 1997). These results then indicate that auditory stimuli also influence the perception of visual information.

While these examples of interactions are exciting, they only speak to the interactions between so called “lower-levels” of perceptual processing. That is, such research provides evidence for interactions between perceptual modalities, but does not speak to how such low-level information participates in so-called “higher-levels” of cognitive processing. To answer these types of questions, a great deal of recent

empirical evidence suggests that perceptual and motor simulations are automatically and unconsciously engaged while performing higher-level cognitive tasks, like language comprehension. These theories of embodied cognition contend that interactions between an organism and its environment facilitate cognitive abilities generally (Barsalou, 1999; Calvo-Merino, Grezes, Glaser, Passingham, & Haggard, 2006; Thelen & Smith, 1994; Zwaan, R.A., 2004), reminiscent of Gibson's (1979) approach.

Gibson (1979) proposed that a sensory analysis of the world generates not a copy of the world's structure but a pragmatic mapping onto the action opportunities the world's structure makes available. He called these pragmatic mappings affordances and suggested that the process of decision-making underlying voluntary behavior is, at least in part, selecting from these pragmatic representations of motor options. Because one of these will be selected and released into overt execution, this process should be somewhat embedded within the neural systems associated with motor control.

Expanding on Gibson's theory, embodied cognition claims that the cognitive abilities of an organism are grounded in an organism's experiences in the world. Specifically, higher-levels of cognitive processing, like language, should necessarily rely on perceptual and motor groundings in the environment. Hence, embodied cognition researchers suggest that language comprehension is perceptually and motorically grounded (Barsalou, 1999; Langacker, 1987).

Despite many recent advances in our understanding of embodied cognition and in how language and action interact, many questions still remain. In each of the following chapters, I will investigate some of these outstanding questions. Chapter 2 reviews both historical and more recent data in support of interactions between language and action. These results, as a whole, suggest that language not only

influences motor output, but action also influences language processing. Chapter 3 investigates changes in sentence processing across developmental time. These results suggest that children are like adults in that comprehending syntactically ambiguous sentences results in graded spatial differences in motor output. Unlike adults, the results demonstrate that children do not use the same types of visual cues to disambiguate these sentences. Chapter 4 looks at the way subtle temporal differences in language manifest in continuous motor output. Whereas other research has focused on sentence comprehension and word processing interactions with motor output, these data explore the way verbal aspect interacts with visual scene information, impacting continuous motor output. Chapter 5 outlines and discusses outstanding questions that need to be addressed to more fully understand the interactions between language and motor processing.

Chapter 2

Do language and action interact? If so, what is the nature of this interaction? This chapter reviews research regarding these questions. As introduced here, many long-standing predominant theories of cognition have viewed higher levels of processing, such as language and cognition, as free from the influence of lower levels of processing, such as action and perception. However, many recent experiments have found evidence that the delineation between traditional modules is not so clean or precise, with motor output and language comprehension interacting much more fluidly than traditional theories predict. Evidence for this account includes findings of systematic activation of motor cortex while processing action words, as well as functional consequences of language on action and of action on language. It is worth noting that this recent spate of interest in the embodiment of cognition is not without historical precedence. Chapter 2 reviews such evidence coming from previous decades of research on the interaction between language and action, in addition to

exploring the empirical results of more recent experiments and methodologies. It seems that motor representations and actions can influence language processing in predictable ways. Not only is there continuous competition between simultaneously active alternatives in language processing, with multiple sources of information interacting immediately, this competition is apparent in the motor output produced as a response to language.

Chapter 3

Does children's sentence processing influence their motor output in a similar way to what is observed in adult sentence processing performance? Do children rely on the same non-linguistic sources of information as adults rely on when processing sentences? What predicts the transition from child- to more adult-like sentence processing skills?

Prior research suggests that while adults use information from the visual scene to disambiguate garden-path sentences (i.e., Put the apple on the towel in the box), children do not. However, the nature of these differences between children's and adults' sentence processing skills is unclear. One possible mechanism for this transition is linguistic experience: children with greater linguistic experience, and correspondingly higher vocabulary scores, may be better able to utilize scene-based cues in processing structurally complex sentences. Here, the continuous and non-ballistic properties of computer mouse movements are used to investigate how young children incorporate multiple sources of information. Participants heard structurally ambiguous sentences while viewing scenes with properties that did or did not support the less frequent noun phrase modifier interpretation. As previously reported (Snedeker & Trueswell, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999), children have difficulty in utilizing scene-based context cues, especially in the two-referent, ambiguous sentence condition. However, vocabulary scores significantly predicted

performance in this condition. Children with higher vocabulary scores, and presumably greater linguistic experience, were better able to move the appropriate referent to the correct destination in the two-referent, ambiguous sentence context. Using the results from a novel behavioral testing method combined with a measure of linguistic experience, the data suggest that linguistic experience is one mechanism underlying the development of syntactic processing skill.

Chapter 4

How is subtle temporal information about a verb, specifically verbal aspect, processed? Does this influence motor output in systematic ways?

How does grammatical aspect influence the understanding of everyday motion events? Research shows processing differences between the past progressive (e.g., *was walking*) and the simple past (e.g., *walked*), but details about the temporal dynamics remain unexplored. The current work uses computer-mouse tracking (Spivey, Grosjean, & Knoblich, 2005) to explore the processing of aspect. In Experiment 1, participants heard motion descriptions in the past progressive or simple past while viewing a scene with a path that terminated at a destination, and placed a character in the scene to match the descriptions. Overall, the past progressive sentences yielded slower movement durations, suggesting differences in information made prominent by the different aspectual forms. Experiment 2 expanded these results by examining past tense and future tense situations (*will walk*, *will be walking*). Experiment 3 further expanded the results of Experiment 1 by examining the interaction of contextual descriptions and verbal aspect in past tense situations. Together, these experiments provide new insights into the role of aspect in event understanding, demonstrating that perceptual simulation occurs in systematic and predictable ways at the level of grammar.

Chapter 5

What is the overall significance of this program of research? What questions remain to be answered? This chapter will outline the contributions of this research before turning to the questions this research raises. Specifically, I will address previous concerns regarding the mouse-tracking methodology as a reliable index of continuous cognitive decisions leaking into continuous motor output. Similarly, I will suggest future directions for research to investigate the nature of the relationship between motor and language tasks.

CHAPTER 2

THE ENACTMENT OF LANGUAGE: DECADES OF INTERACTION BETWEEN LINGUISTIC AND MOTOR PROCESSES¹

"The world shows up through the enactment of the perceptuomotor regularities."

-Francisco Varela

Many long-standing theories of cognition have assumed that higher levels of cognition are amodal and modular (Dietrich & Markman, 2003; Fodor, 1975; Frazier & Clifton, 1996; Frazier & Fodor, 1978; Pylyshyn, 1984). In such theories, cognitive processes, such as language and action, are handled in two separate systems, each working according to its own principles. Under this view, symbols processed within the cognitive system are both amodal and arbitrary because their content bears no systematic relationship to the original sensory processes that gave rise to them (see Barsalou, 1999b, for review and critique). These theories have dominated cognitive science for decades because they are quite powerful in their ability to account for representations of types and tokens, combine symbols productively, and represent abstract concepts (Fodor, 1975; Pylyshyn, 1984).

While such explanations are quite clean and powerful, evidence against such a modular view of cognition has been accruing for some time. One such example comes from a recent compelling case, worthy of showcasing at the beginning of this review, which illustrates the representational entanglement between motor movement and cognition via a series of fMRI experiments involving dancers. Calvo-Merino and colleagues (2006) exploited the fact that within classical ballet there are movements

¹ Published as: Anderson, S.E. & Spivey, M.J. (2009). The enactment of language: Decades of interactions between linguistic and motor processes. *Language and Cognition*, 1, 87-112.

that are typically only performed by women (pointe work, for example) and movements that are typically only performed by men (cabriole, or jumping into the air and beating the calves, for example). The researchers used these gender-specific steps to investigate whether participants would exhibit different patterns of brain activity when viewing such movements. When dancers passively watched stereotypically male ballet steps, greater premotor, parietal, and cerebral activity, including the mirror neuron system, was observed in the male dancers compared to their female counterparts. Similarly, when the dancers watched stereotypically female ballet steps, greater premotor, parietal, and cerebral activity was observed in the female dancers compared to the male ballet dancers. These results suggest that the brain activity of someone simply watching a movement they have mastered is dramatically different from the brain activity of someone who watches these movements everyday (so they are equally familiar) but who does not perform them herself. This implies that the neural circuits involved in purely motor responses are involved in recognition, over and above visual representations, suggesting that we understand actions not only by visual recognition, but also through activation of brain areas associated with motor activity. Traditional encapsulated theories of cognition, which assume clean separations between different modules, do not predict different patterns of motor cortex activation, based on experience, in such visual recognition tasks.

Simple motor output has been shown to influence recall of memories as well. First, it was demonstrated that having participants smile and stand in an upright position facilitates the retrieval of pleasant memories, while having participants slump and make downcast expressions facilitates the retrieval of unpleasant memories (Riskind, 1983). These results suggest that both facial expression and posture facilitate recall information associated with autobiographical memories that have emotional valence. These results have recently been expanded to demonstrate that

posture in general can facilitate the retrieval of non-valenced memories as well. Dijkstra, Kaschak, and Zwaan (2006) asked young and older adult participants to assume postures congruent with or incongruent with childhood events. Therefore, participants who experienced a congruent condition were asked to lie down while recalling their first visit to a dentist, while participants experiencing the incongruent condition were asked to lie down while they recalled opening the door for a visitor. Participants in the congruent condition recalled these autobiographical events with significantly shorter response times than did participants in the incongruent condition. Additionally, free recall of these events after two weeks was also better for participants in a congruent posture than in an incongruent posture. These results provide compelling evidence that posture and motor processes influence the retrieval of memories when the memory is retrieved in a laboratory setting and after a lag of two weeks. Again, traditional theories of encapsulated cognition would not predict an influence of motor activity on memory retrieval, instead suggesting that memory is modular and free from the influences of motor activity created simply by postural differences.

In light of such evidence against traditional views of cognition, more recent theories of cognition suggest a continuous interaction between so called higher and lower levels of cognition. These theories of embodied cognition contend that interactions between an organism and its environment facilitate cognitive abilities generally (Barsalou, 1999b; Calvo-Merino, Grezes, Glaser, Passingham, & Haggard, 2006; Thelen & Smith, 1994; Zwaan, 2004). In other words, embodied cognition claims that the cognitive abilities of an organism are grounded in an organism's sensorimotor experiences in the world. Demonstrating fluid interactions between these processes, which were once viewed as removed from one another, blurs the delineation imposed by modularity of traditional cognitive theories. Also, embodied

theories suggest a role for low-level processes, like action, on high-level processes, like language.

While a great deal of attention has recently been paid to theories that incorporate motor movements and action into their account of language and cognition, it is certainly not a new idea. In the history of psychology, there is a great deal of empirical support for this interaction. Despite being somewhat downplayed during the cognitive revolution and its ascendancy, evidence from earlier decades steadily set the stage for the interaction of language and action that recent research has turned into a cottage industry. Here, we will review just some of these early findings for the interaction of language and action that laid this foundation, before turning our attention to some more recent research in the field.

A history of support

Although a great deal of research between the 1950's and the 1990's was dominated by modular views of cognition, interest in theories of embodied cognition have increased rapidly over the course of the last decade. However, these embodied theories are by no means without precedent. Approaches in philosophy predating psychology were dominated by this interaction of mind and body, even from as far back Aristotle and Epicurus (for review, see Barsalou, 1999b). These classical approaches in philosophy viewed the representations underlying thought as naturally and necessarily tied to perceptual processes. However, with the cognitive revolution of the mid-twentieth century, this view was swept aside by theories proposing more encapsulated views of cognition, in areas from visual imagery to sentence processing. Even as the zeitgeist of the cognitive revolution rolled into many areas of cognitive research, evidence for a more interactive approach was almost secretly being obtained.

Evidence during the height of the Cognitive Revolution

Despite the prevailing trend at the time for treating cognition as something

independent of perception and action (e.g., Neisser, 1967; Newell, Shaw, & Simon, 1958), several notable experiments from that time provide evidence for a more interactive account. Predominant modular accounts of cognition predict that so called higher levels of processing, like language, should be encapsulated from perception and action, such as the spatial reference implied by a word. Hence, perceived location of a visual stimulus should not influence language processing. On the other hand, more interactive approaches would predict a functionally important interface between processing a word and its spatial reference in the world. Under this view, the perceived location of a visual stimulus may be influenced by comprehending a word, with the spatial reference necessary to comprehend the word interfering with the perception of the visual stimulus's location. In order to explore this, Kaden, Wapner, and Werner (1955) found that subjective eye level is influenced by the spatial components of words. While subjects sat in a dark room, they saw luminescent words at their objective eye level, and were asked to then move these words up or down, until they were at their subjective eye level. Words with an upward connotation ('climbing', 'rising') were placed slightly lower than objective eye level to be perceived as being at eye level; on the other hand, words with a downward component ('falling', 'plunging') were placed slightly above the objective eye level. Hence, it seemed that words were not as free from the influence of spatial perceptual reference frames as would be expected from the predominant symbol-processing theories of the time.

Almost ten years later, evidence for embodied sensorimotor priming in problem solving emerged from an experiment investigating individual differences in functional fixedness. Glucksberg (1964) studied people's actions while attempting to solve Duncker's (1945) candle-mounting problem. In this task, participants were given a box of tacks, a candle, and a matchbook, and then asked to mount the candle on the wall using only the items available on the table. While many people struggled with

the problem for some time, those who did arrive at the correct answer (using the tacks to mount the box itself onto the wall as a sconce for the candle) reported a moment at which the answer seemed to “pop” into their head. When watching a participant’s actions, Glucksberg found that participants who successfully solved the problem happened to accidentally touch the box a greater number of times than participants who failed to solve the problem. This suggests that before the moment the answer seems to pop into the participant’s head, she is paying a little extra attention to the item on the table that can afford solving the problem. This kind of perceptual-motor priming suggests that even incidental touching of the full box of tacks may facilitate the insight of using the box itself as the wall mount.

In the 1970’s, while amodal theories continued to dominate the field, some researchers were exploring more interactive approaches to cognition. Notable exceptions to the amodal trend include Greenwald's (1970) ideomotor theory, which proposed that the performance of a voluntary action is mediated by an image of the action's sensory feedback, and Neisser's (1976) proposal that cognition is actually the result of Perception-Action Cycles. Theories such as these sought to link “lower levels” of perception and action to “higher levels” of cognition and thinking more seamlessly. Another notable exception to the amodal theories of the time is the work of Gibson (1979), who proposed that a sensory analysis of the world generates not a copy of the world’s structure but a pragmatic mapping onto the action opportunities the world’s structure makes available. He called these pragmatic mappings affordances and suggested that the process of decision-making underlying voluntary behavior is, at least in part, selecting from these pragmatic alternatives of motor actions made available by the dynamic kinematic relationship between organism and environment.

Affordances in vision and language

More recently, the role of affordances in object identification has been explored. It seems that physical properties of an object that determine how that object can be manually interacted with, or its micro- affordances, may influence object identification (Ellis & Tucker, 2000; Tucker & Ellis, 1998). In these experiments, participants were asked to make simple perceptual judgments regarding whether an object was right-side-up or upside-down. The reaction times for this judgment were influenced by whether the handle of the object was on the left or right side and which hand was used for responding. The response times were slightly faster when the hand responding was the same hand that would interact with the handle of the object, as indicated by which side of the object the handle was on – despite the fact that these were two-dimensional images of objects on a computer screen, and thus never availed actual affordances. Therefore, it seems that part of recognizing an object, without being asked to interact or even having the opportunity to interact with it, involves at least partial activation of the motor actions needed to interact with it. Hence, to some degree, object identification is tied to the actions required for interacting with that object.

Affordances also influence the processing of temporarily ambiguous sentences. When a participant hears a sentence such as “Pour the egg in the bowl over the flour,” he or she will experience the garden path effect (Chambers, Tanenhaus, & Magnuson, 2004). This sentence is temporarily ambiguous because the prepositional phrase “in the bowl” could either attach to the noun phrase, describing which egg to pour (less preferred), or to the verb phrase, describing where to pour the egg (more preferred). Since the less preferred option is the ultimately correct one, readers should experience the garden path effect, reflected by either inflated reading times or an increased proportion of looks to a second (irrelevant) bowl, when they encounter the

disambiguating information “in the bowl.”

Chambers and colleagues (2004) varied the affordances of the objects in the visual scene. Hence, the visual workspace corresponding to the example sentence “Pour the egg in the bowl on the flour” would either contain one or two eggs in liquid, or pour-able, form. If the visual scene only contained one candidate referent that could accomplish the task (one egg in pour-able form and the other still in the shell), then participants initially misinterpreted the first prepositional phrase, “in the bowl”, as the goal location for the pouring action, since there was only one egg that afforded pouring and thus it needed no further specification. Therefore, they incorrectly attached the prepositional phrase to the verb. However, if the visual scene contained two action-affording candidate referents (two eggs in pour-able form, one in a bowl and the other in a transparent cup), then participants did not initially attach the first prepositional phrase to the verb. Instead, they correctly used the first prepositional phrase to distinguish between the two plausible referents for the pouring action and did not experience the garden-path effect. These results provide evidence both for interactive and continuous processing in sentences and for the important role of affordances in this aspect of sentence comprehension, as would naturally fall out of the earlier theories of Neisser (1976) and Gibson (1979).

The beginnings of a cottage industry

In the early 1990’s, evidence for interactions between action and language began to accumulate rapidly. As findings emerged to suggest that language processing was more continuous and interactive than assumed by amodal symbol-processing theories, the putative barriers between modular areas in the brain created through carving up cognition began to break down. Not only were multiple sources of linguistic information contributing immediately to language processing at the level of phoneme (Elman & McClelland, 1988), word (Allopenna, Magnuson, & Tanenhaus, 1998) and

sentence (Tanenhaus & Trueswell, 1995; Trueswell, Tanenhaus, & Kello, 1993), but non-linguistic cues were also able to influence early processing. Cues from the visual scene were shown to influence early syntactic decisions (Snedeker & Trueswell, 2004; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell, Sekerina, Hill, & Logrip, 1999), thus further blurring the line between subsystems that had been thought of as isolated and independent from one another. As a result, an interaction between action and language comprehension no longer seemed so far fetched.

During this time, several researchers provided evidence for sensorimotor components underlying cognitive representations. Klatzky and colleagues (Klatzky, Pellegrino, McCloskey, & Lederman, 1993) showed that people have explicit knowledge about how to manipulate and grasp objects. In one experiment, participants were told to rate different actions on six dimensions: portion of limb moved, distance moved, forcefulness, effectors involved, size of the contact surface, and resemblance to grasp. Participants' ratings were systematic, mainly relying on factors related to limb movement and effector (usually the hand) configuration. In a second experiment, the researchers found results from sorting tasks indicating the six motor-haptic dimensions used in the first experiment contributed to similarity judgments. The results support the existence of cognitively accessible, but still relatively specific, representations of functional actions. These results provide a relatively early demonstration of how people's cognitive representations contain sensorimotor components.

Also, before the early 90's, much of the research investigating human cognition viewed action as trivial, suggesting that the motor system is simply a slave of higher-level cognition, such as language processing and problem solving, and that the motor system merely executes a discrete and fully completed command that is

delivered from these “higher levels” of processing. However, in investigating the differences in online processing, evidence suggests that factors influencing latency to respond also influence later aspects of response dynamics. In a lexical decision task, Abrams and Balota (1991) had participants make rapid limb movements in opposite directions in order to indicate whether a string of letters was a word or not. As expected, they found that the frequency of the word influenced the speed of reaction time, with high frequency words being responded to more quickly than low frequency words. In addition, they showed that word frequency also influenced response kinematics after the response was initiated, specifically finding that responses to high frequency words were more forceful than responses to low frequency words. These findings suggest that word frequency not only influenced the time required to recognize a word, but also influenced the subsequent response dynamics, implying that the response system is *not* slavishly executing a complete command regarding the status of the word. In a later experiment, they found that high-frequency words are also responded to with greater acceleration than low frequency words (Balota & Abrams, 1995). They found a consistent pattern of results in a memory scanning paradigm, where both onset latency and response dynamics after the response was initiated were influenced by the memory set size and by the presence of the probe word (Abrams & Balota, 1991). These results make a compelling case for looking at both reaction time and other response variables from motor output, including acceleration and force of movement, after a movement to respond is initiated.

In the second half of that decade, the number of experiments investigating interactions between language and motor action exploded. With the work of researchers like Barsalou (1999a, 1999b), Pulvermuller (1999), and Glenberg (1997), it seemed that the ideas of earlier researchers like Neisser (1976) and Gibson (1979) were finally coming to fruition. The discovery of mirror neurons in 1996 (Rizzolatti,

Fadiga, Gallese, & Fogassi, 1996) played no small part in fueling this increase. Mirror neurons are neural structures that are active when a monkey either observes or performs certain motor activities (Rizzolatti, et al., 1996), and these findings have also been extended to humans via neuroimaging (e.g., Buccino, et al., 2001; Stevens, Fonlupt, Shiffrar, & Decety, 2000). Prompted by the discovery of mirror neurons, the momentum of the resurgence of embodied theories of cognition in mainstream cognitive psychology was increased yet again.

Barsalou (1999b) provided a contemporary theoretical framework within which to view the interaction of perception and action with cognition. For example, rather than viewing a concept as a list of features that have been extracted away from the original perceptual experience of that object, Barsalou described how perceptual simulations can capture the object from perception. Specifically, he postulated that when an object is visually perceived, neurons in the visual system fire in response to features of that object. Neurons in nearby associative areas then conjoin the active visual features' firing responses in memory so that later, when the object is not physically present, the conjunctive neurons partially reactivate the original set of feature detectors that were active when the object was perceived. This approach allows that these simulations are not complete and may be biased, but also suggests that these perceptually grounded simulations are enough to support memory, language, and thought, bypassing the need for redundant conceptual representations in the brain (Barsalou, Simmons, Barbey, & Wilson, 2003).

Further evidence for the grounding of language in lower levels of processing came from research comparing performance of participants who were expressly told to use imagery in order to list features of a concept verbally to that of participants who were simply asked to list features (Barsalou, Solomon, & Wu, 1999). Amodal symbol-processing theories assume that differences should arise between the two

groups, with participants who are told to use imagery to create the list engaging modal simulations while the non-imagery participants rely on the default amodal simulations. On the other hand, modal embodied accounts of cognition hypothesize that no differences should arise as a result of the different instructions: modal imagery *is* the default method for creating such lists. Indeed, the data supported the second hypothesis. Both groups produced comparable distributions of complex features, and similar regression equations were able to account for reaction times when members of each group were asked to verify properties of a concept. These results taken together make a strong case for the role of lower levels of processing on cognition.

Also during this time, the role of action was not only applied to adult cognition but also to that of young infants (Smith, Thelen, Titzer, & McLin, 1999; Thelen, Schoener, Scheier, & Smith, 2001). In one such study, an experimenter hid a toy under one container while a nine-month old participant watched, and then the infant was given an opportunity to retrieve the toy (Thelen, et al., 2001). The infant was able to do this correctly and would continue to successfully retrieve the toy on successive trials as long as the toy continued to be hidden under the same container. After multiple trials, though, if the experimenter then hid the toy under a second container, the child would persist in reaching to the container that the toy had been hidden under in previous trials, even though he or she had seen that the toy was no longer there.

The results of the A not B task were originally interpreted as a discontinuity of cognitive development in children, with children not being able to inhibit the response that had become salient in the previous trials, even though they have seen that this option was no longer the correct one. Piaget suggested that progress from one stage of development to another is precipitated by something disrupting the stability of the current pattern of behaviors associated with that stage (Piaget, 1952). If the shift from a current stable set of behaviors to another is caused by disrupting the stability of the

current behaviors, then what creates the disruption? In the context of the A not B task, the inability to inhibit an inappropriate response was originally thought to be due to differences in maturational rates in possible dual systems of perception and action (Ahmed & Ruffman, 1998; Bertenthal, 1996). However, Thelen and colleagues (2001) found that making a slight adjustment to the child's reaching motion eliminate the erroneous action. By placing a weight on the baby's arm right before he or she moved to retrieve the toy from the new hiding place, the child was able to reach to the correct container and retrieve the toy successfully, thus acting like older children. Changing the biomechanics of the situation was sufficient to disrupt the reaching pattern that had become stable during subsequent trials and allow the child to retrieve the toy from the correct hiding place. These results suggest that the child's experience and motor movements through the world had become tied to retrieving the toy in this experimental situation. The change was not simply one of a maturation of cognitive control influencing the ability to inhibit a salient response, and developing in a discontinuous step-like manner. Changing the immediate biomechanics of the child by attaching a weight to the arm prior to reaching to the new hiding location allowed the young children to successfully complete the task.

Grant and Spivey (2003) provided evidence that motor movements of the eye made to relevant aspects of a visual scene can influence problem solving. One such example comes from a recent extension of Duncker's (1945) radiation problem. In this problem, participants were asked how to cure a person with an inoperable stomach tumor using lasers, which destroy the surrounding tissue if applied at sufficient intensity to kill the tumor. In addition to the word problem, participants are provided a diagram depicting the tumor, the healthy tissue, and the skin separating the stomach from the outside. The answer is that the lasers need to be fired from multiple locations outside the healthy tissue so that they converge on the tumor, thereby applying

sufficient intensity to the tumor but leaving the healthy surrounding tissue unharmed. Grant and Spivey (2003), using eye-tracking, found that just before participants solved the problem successfully, they made a number of fixations on the stomach lining. This finding suggested that attention to this particular visual feature influenced successful answering of the problem. Further supporting this inference, in a second experiment, they found that by highlighting these relevant visual areas, participants were more likely to arrive at the correct answer. They conclude from this that making relevant aspects of the visual display more visually salient drew visual attention to these regions and made solving the problem more likely, with eye movements helping to coordinate spatial locations in the visual field with internally represented elements of the mental model (e.g., Spivey & Geng, 2001). The movement of the eyes to the relevant aspects of the scene assisted in building the appropriate representation necessary to solve the problem (see also Thomas & Lleras, 2007).

Recent behavioral evidence for the interaction of language and action

While the 90's saw a surge in evidence to support the interaction of language and action, the past eight years have continued this trend. The body of literature in support of an interaction between language and action has continued to grow as more encapsulated views of language have been called into question. The next point to explore, beyond whether or not there is an interaction, is how this interaction occurs. However, the exact nature of this interaction has been largely unexplored (for review, see Fischer & Zwaan, 2008; Zwaan & Taylor, 2006). If language and motor actions do interact fluidly, then there should be evidence of facilitation (or in some cases inhibition) when processing language and executing motor movements that rely on the same underlying neural architecture.

Evidence for the motor underpinnings of language, using both neuroimaging techniques and behavioral experiments, have made great strides in beginning to

answer these questions. In relation to action words, similar patterns of activation are found in motor and pre-motor cortex in perceiving an action word and in performing that action (Hauk & Pulvermuller, 2004; Pulvermuller, 1999; Pulvermuller, 2001; Tettamanti, Buccino, Saccuman, Gallese, Danna, Scifo, Fazio, Rizzolatti, Cappa & Perani, 2005). One such experiment investigated how activity in the left temporal lobe differed when the participant heard an action word or performed the action. While in the fMRI scanner, participants passively read sentences containing action words associated with different parts of the body or performed actions associated with those same parts of the body. When performing these actions, pre-motor and motor cortex show a topographical arrangement of activation, with activity involving the face located in one area, and with activity involving the feet located in another area, etc. Interestingly, and most importantly here, when participants were passively reading words about actions involving the face (kiss), arm (pick), and leg (kick), the brain activity overlapped with parts of motor cortex associated with actually performing the actions (Hauk, Johnsrude, & Pulvermuller, 2004). In a follow-up experiment, these results were expanded to demonstrate that the activity in motor cortex is available early in processing (Hauk & Pulvermuller, 2005). By exploring the time course of the spread of activity in motor cortex during word processing, it seems that this firing is not simply a by-product of the comprehension process that happens further downstream from processing the word. Instead, this information is immediately available to influence word comprehension. These results taken together indicate not merely a “spreading activation” from linguistic representations to motor representations, but also suggest a functional role for neural feedback from motor areas to language comprehension. Further, these results suggest that the brain basis for action words is grounded in transcortical cell assemblies (Hebb, 1968) that are distributed across language areas and motor areas.

Influence of language on action

To investigate the behavioral consequences of these polymodal neuronal ensembles, Boulenger and colleagues (Boulenger, Roy, Paulignan, Deprez, Jeannerod, & Nazir, 2006) explored the effect of processing action words on the response dynamics of a reaching movement. In one experiment, participants moved their dominant hand from a central location when a fixation cross appeared. Upon moving the hand from this home pad, either a word or a pseudoword replaced the fixation cross. If the letter string was a word, participants were instructed to continue the hand movement and grasp a cylinder located away from the home-pad, but if the letter string was a pseudoword, they were required to return to the home-pad. The procedure of the second experiment was identical, except that the letter string was presented in place of the fixation cross before initiation of hand movement, thus serving as the go-signal. In this way, the experimenters were able to investigate the relative impact of the movement both during and after word comprehension. The results showed that when the words appeared after the onset of the movement, the comprehension process seemed to *interfere* with the reaching movement. Specifically, the latency to reach to the cylinder was longer, and the amplitude of the wrist acceleration was smaller when the word that appeared after the movement had been initiated was a verb than when it was a noun. However, when the word appeared as the go-signal itself, the verbs, but not the nouns or the pseudowords, seemed to *facilitate* the response, with the peak wrist acceleration appearing earlier. These results suggest that processing action words recruits the cortical regions also involved in programming and executing motion. Hence, the action words, but not concrete nouns, interfered with overt motor behavior when the two tasks were performed simultaneously, but facilitated the same overt motor behavior when the lexical decision task came first.

In a follow up experiment, these findings were extended to show that even

after the movement had been initiated, the reaching motion is disturbed by the visual presentation of a verb (Nazir, Boulenger, Roy, Silber, Jeannerod, & Paulignagn, 2007). When an action verb was visually presented at either 50ms or 200ms after the movement to reach the cylinder was initiated, the motor output was disrupted, resulting in greater deceleration rates when the presented word was a verb than when it was a noun. In the first experiment, peak wrist acceleration (which occurred between 160 and 177 ms after the word was presented) showed interference from the action word when the word was presented as movement was initiated (Boulenger, et al., 2006). Here, when the word was presented either 50 or 200 ms after the movement was initiated, the interference of the action word manifested itself in deceleration peaks of the reaching movement, presumably because peak wrist acceleration happens too early to reflect motor perturbations initiated at these longer intervals (Nazir, et al., 2007). The results of these two experiments taken together suggest that the cross-talk between language processing and movement execution does not simply occur after a constant interval; such cross-talk appears to manifest itself differentially over the course of motor programming and the movement itself.

There is also evidence to suggest that comprehension of a whole sentence implying movement interferes with motor output (Glenberg & Kashak, 2002). In a series of experiments, participants were asked to judge whether or not sentences made sense by making a movement away or towards themselves. For example, after reading a sentence such as “Jay rolled the marble to you,” participants were asked, in one condition, to move their hand from a central fixed location towards themselves to indicate that the sentence made sense; but if a participant read a sentence such as “Jay swept the beach bum now,” they were asked to move their hand away from themselves and the central location to indicate that the sentence did not make sense. The results of these experiments showed that when participants read sentences that implied

direction, reaction times to respond that the sentence made sense were significantly shorter when the movement they were required to make matched the direction of motion implied by the sentence. Therefore, a participant would be quicker to respond “yes” if they moved toward themselves (matched direction) to indicate the sentence “Jay rolled the marble to you” made sense than if they were required to move away from themselves (mismatched direction). The pattern of results across this series of experiments provides evidence that the action implied by an entire sentence facilitated movements that were compatible with the verbal description.

Similar effects showing differences in motor output as modulated by differences in language input have been demonstrated in the investigation of eye movements and fictive motion sentences. Fictive motion sentences are sentences that contain an action verb but no actual movement or action takes place. For example, “The road ran through the valley” contains the action verb *ran*, but the road is not actually performing the act of running, or any action at all. Participants heard a context sentence describing the terrain, and then a target sentence containing fictive motion. When participants first heard context sentences describing the terrain as difficult, inspection times and eye-movement scanning along the path were increased (as though a hint of actual motion was being perceptually simulated), compared to when participants heard the terrain described as easy (Richardson & Matlock, 2007). These results support earlier reaction time results, in which participants read narratives describing a terrain, and then made decisions about whether a fictive motion sentence was related to the preceding context (Matlock, 2004). Participants were slower to respond to fictive motion sentences when they had read context sentences describing slow travel, long distances, or difficult terrain, than when they had read context sentences describing fast travel, short distances, or easy terrain. However, no such differences in context descriptions were found when the target sentence did not

contain fictive motion. Taken together, these results provide evidence that fictive motion descriptions affect both reaction times and eye movements by evoking mental representations of motion, and that this is then influenced by descriptions of that motion. The eye-movement data allows for a closer look at the way the constructed mental model and the linguistic description are coordinated.

Influence of motor movement on language

The next question is one of bi-directionality: does behavioral action influence language processing similar to the way language influences action? It may be possible to reconcile findings supporting language's influence on action by claiming that language is central, but activation in language areas spread to supplementary areas, such as motor areas, giving rise to these results. In other words, it could be the case that the language system does process information separately from other systems, but that this information then spreads into motor output. Therefore, parts of the brain associated with language are central to language processing, and activity in motor areas is simply peripheral and redundant. Following this line of argumentation, despite evidence from neuroimaging studies and behavioral studies suggesting language's influence on action, these effects may not be indicative of a meaningful interaction between the two levels of processing. Determining whether action is central to language processing then requires bi-directional influence between action and language processing.

One way of exploring this issue is to create temporary and reversible changes in an otherwise healthy brain through transcranial magnetic stimulation (TMS), and then observe the differences in language processing. If activation of motor areas is peripheral to language processing, then there should be no observed differences between processing action words associated with different areas when these motor areas are disrupted by TMS. On the other hand, if TMS does influence results of

language processing, then this would provide evidence that these motor areas are not peripheral or epiphenomenal to language processing, but are indeed functionally integrated with it. Pulvermüller and colleagues (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005) explored this hypothesis by applying TMS to motor areas while subjects made lexical decisions on action related words. TMS of hand and leg areas influenced the reaction time to arm and leg words differentially. Action words related to the arms (“fold,” “grasp,” and “write,” for example) were responded to more quickly than leg action words when arm motor areas were subjected to TMS. Similarly, action words related to the legs (“run,” “walk,” and “step,” for example) were responded to more quickly than action words related to the arms when the leg motor areas were subjected to TMS. By demonstrating that motor activity influences language processing in a category specific manner, this evidence suggests a functional and non-peripheral role for motor activity in the processing of language related to action words.

Motor output also influences language production in surprising ways. The fact that people gesture as they speak can be observed anywhere two people are having a conversation, although the degree to which individuals gesture may depend on a number of different factors. In order to more formally investigate the communicative value of gestures, Casasanto and Lozano (2007) devised a series of experiments. In the first experiment, participants were videotaped as they told a confederate a story using either literal spatial language, which described movement of physical objects either upward, downward, right or left (the rocket went higher), metaphorical spatial language, which described non-spatial events that are often expressed using spatial descriptions (my grades went higher), and non-spatial language stories, which contained no spatial language (my grades got better). They found that in all cases, the overwhelming majority of gestures were in the same direction as the implied

metaphor, critically including the non-spatial descriptions (“My grades got better,” accompanied by an upward gesture). This suggests that the participants were using spatio-motor representations to convey information, even when that information was not spatial in nature.

However, it was unclear whether these gestures were for the benefit of the listener (in order to help convey the story) or for the benefit of the speaker (using spatio-motor representations in order to bolster concrete and abstract spatial language). By having participants hold down a button while they told their stories, thus preventing gesturing, a significantly greater number of verbal disfluencies were found compared to the condition in which participants were able to freely gesture. Because these findings could be the result of performing two tasks, that of holding down a button and speaking, Casasanto and Lozano (2007) conducted a third experiment. In this experiment, participants were asked to move balls from one container to another: either from left to right, right to left, up to down, or down to up. As they moved the balls, they were instructed to tell a story that included all four types of movement. Participants created many fewer verbal disfluencies when performing the task of moving the balls in a direction that was congruent with that part of the story than when they were moving the balls in a direction that was incongruent with that part of the story. These results suggest that the gesture, even when it was an arbitrary movement the participant was required to make throughout the session, facilitated verbal fluency in telling the parts of the story congruent with that movement. Thus, in conjunction with Pulvermüller and colleagues’ (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005) evidence that neural activity in motor areas can influence lexical decision tasks for verbs that involve the corresponding limbs, Casasanto and Lozano’s findings support the conclusion that motor actions are an active part of cognitive functions of both literal and metaphorical language use and support the hypothesis that motor

movements do influence linguistic processing.

Toward a continuity among language and action

The fluid bi-directional interaction between motor processes and language processes is particularly evident when looking at the millisecond-by-millisecond timing of information integration during language comprehension. One area in which the continuous competition between multiple, partially activated representations can map onto continuous motor output comes from research in lexical decision tasks. Just as sentences may have temporary ambiguities because of their incremental nature, the same is true for words. As a word unfolds over time, it is ambiguous with other words that share similar sounding onsets (Marlsen-Wilson, 1987; Zwitserlood, 1989). For a brief period after the onset of a word, all words beginning with the same acoustic-phonetic input show partial activation. As more acousti-phonetic input is received, the target word is further identified, and some words drop out of the cohort (the set of words with similar sounding onsets). For example, the acoustic-phonetic input for the words *candy* and *candor* is more overlapping than the acoustic-phonetic input for *candy* and *cable*. Hence, *candy* would be differentiated from *candor* after a longer period of time than it would from *cable*, since *candy* and *candor* share more overlap and become unique further into each word. Likewise, when presented with a target word *candy*, *cable* would drop from the cohort faster than *candor*. Therefore, it seems that lexical decisions are based on a competition between multiple, partially active candidate words. Recent eye-tracking and reach-tracking work has investigated how these multiple competing candidates appear to continuously map onto multiple competing motor response options.

As a spoken word unfolds over time and exhibits temporary ambiguity, a second potential referent in the visual scene illustrates the dynamic competition between the two possibilities. Tanenhaus and colleagues used such a visual scene to

investigate the processing of phonetically similar words (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). Participants heard auditory commands such as “Pick up the candy.” When the visual scene only contained a single possible referent (for example a piece of candy, a fork, a pincushion, and a pair of scissors), the eyes moved straight to the correct referent, the candy. In fact, the speed with which the eyes arrived at the appropriate location meant that the saccadic motor program had been initiated before the word was finished. However, when the pincushion was replaced with a phonologically similar item, such as a candle, a different pattern of results emerged. The participants looked to something else in the visual scene on about a third of the trials, and on about two-thirds of these trials, that something was the phonological competitor, the candle. The competition between the two possible referents, due to their overlapping phonological similarity, was played out in the eye movements (see Gold & Shadlen, 2000, for related evidence of multiple competing oculomotor signals).

Further evidence for this continuous accrual of lexical evidence cascading directly into a continuous accrual of motor response activations comes from reaching tasks. For example, Spivey and colleagues used computer-mouse movements to track continuous responses in investigating this task (Spivey, Grosjean, & Knoblich, 2005). Participants were presented with an auditory instruction such as “Click on the candy” and asked to use a computer mouse to click on the corresponding picture. In the visual scene on the computer screen, participants either saw pictures of two items that were phonologically similar, such as a piece of candy and a candle, or they saw pictures of two items that were not phonologically similar, such as a piece of candy and a ladle. During the movement of the mouse to click on the corresponding picture, the continuous streaming x, y coordinates of the mouse as it moved with the goal-directed hand motion to click on the appropriate object were recorded, and the competition

between the partially activated lexical alternatives was revealed in the shape and curvature of the hand-movement trajectories. When the pictures showed items that were not phonologically similar, the average trajectories of the mouse-movement trajectories moved in a relatively straight path to the appropriate object. However, when the pictured objects had phonologically similar names, each individual trajectory, as well as the overall average trajectory, showed some graded spatial attraction to the competing object with the phonologically similar name. As the partially active lexical representations competed over time, that continuous competition was reflected in the smooth curve of the motor output.

Indeed, evidence from cognitive neuroscience supports the necessary claim of competing underlying motor commands to produce this smooth motor output. One experiment investigated this question using multi-cell recordings. By presenting monkeys with two potential reaching actions and recording from dorsal premotor cortex, Cisek and Kalaska (2005) provide evidence that two populations of neurons - each corresponding to a reach toward one of the two targets - respond simultaneously in preparation for a reach to either targets.

The task began with the monkey moving a cursor to a central green circle, which was then followed by red and blue cue circles that appeared at two of eight possible target locations surrounding the start location, usually opposite from one another. Next, the color cues disappeared, and the central circle changed color to red or blue, indicating which of the two memorized color coded spatial cue locations was the target. After this central color cue disappeared, green circles appeared at all eight peripheral locations. The trial ended when the monkey moved the cursor to the correct location. This task was compared to performance in a one-target task, which was similar to the two-target task except that only one cue circle appeared and only one movement direction was specified.

By making the color cue centrally located, the task did not encourage or require that two movements be simultaneous specified prior to the onset of the actual reaching motion. In other words, the monkeys could have waited until they were given the central cue indicating which target to move toward to begin motor planning, selecting only one target. However, the experimenters found evidence that when the monkeys faced two potential reaching targets, two separate directional signals reflecting both options arose in dorsal premotor cortex. When new information was given to select one of the targets, then the corresponding directional signal increased while the other decreased. These results support the hypothesis that at least two motor commands can be prepared until sufficient information is provided to allow one of them win the competition. When one of the two becomes more salient, it wins the right to be executed and ramps up, while the other command gradually dies off.

Tipper, Howard, and Houghton (2000) developed an account based on neural ensembles for how irrelevant visual stimuli can influence overt motor output in reaching motions, describing the effect of this competition process on the motor program that is launched into execution. In their description, competing and overlapping population codes, representing different actions, are activated in response to visual stimuli. Because the codes overlap, inhibition of the irrelevant population code must occur, but this does not mean that the overlap of population codes does not influence the motor program that is executed. Due to the overlap, the movement that is ultimately executed is influenced by competition from the population code readied to respond to the irrelevant stimulus. Therefore, trajectories of the hand movement to the correct stimulus deviate toward the incorrect stimulus. Also, the degree of the deviation toward the irrelevant stimulus is influenced by how salient the distracter is. Hence, if a distracter is particularly salient, the degree of deviation toward the distracter will be greater than if the distracter is less salient.

Therefore, arm movements seem to be a competitive, dynamic process, with the result of this process being curves in trajectories toward distracters. When these stimuli are close together in the environment, the influence of the distracter is greater than if they are farther away, resulting in greater curvature of the hand toward the distracter (Houghton & Tipper, 1996). When one of the two potential motor commands wins, the other is suppressed via lateral inhibition (Meegan & Tipper, 1998). This inhibition of nearby locations successfully suppresses the competing action, but shifts the population distribution in such a way to cause trajectories of hand movements to curve toward the distracter item (Houghton and Tipper, 1996). This evidence converges to suggest that the competition between two underlying motor commands produces a smooth, curving arm movement.

Conclusion

This paper has attempted to review relevant data for the interaction of action and language processing through the years. During the cognitive revolution, most empirical language research focused on the encapsulation of language from lower levels of processing. However, even during this time, evidence for an interaction between language and action was occasionally being reported. As the prevalence of these modular theories decreased, the corresponding number of investigations into this interaction increased.

More recent research suggests that cortical areas devoted to motor planning and execution play a role in language processing. Not only are some of these neural architectures shared, language seems to influence execution of motor commands, and execution of a motor command seems to influence language comprehension. However, this does not imply that the basis of lexical access to a word is based only in the motor cortical areas of the brain. The evidence provided here suggests that motor planning is involved with language processing, but of course many other brain areas

are also involved. While motor planning is implicated in the process, it would be ridiculous to suggest that it is solely responsible. After all, people who cannot scuba dive or ride a horse can still talk about these events.

In addition to such unanswered questions about the exact functional role of motor cortex in language comprehension, there are unanswered questions about the embodied approach more broadly. For example, there is evidence to suggest that effects like those reported here can be detached from the physical body, casting doubt onto how situated these systems really are. For example, Chen and Bargh (1999) demonstrated that participants are quicker to pull a lever to respond to a word if that word is positively valenced than if it is negatively valenced. On the other hand, people are quicker to respond to a negative word than to a positive word if they have to push a lever to do so. This is in keeping with theories of embodied cognition, suggesting that pulling movements toward the body are more associated with positive objects, and pushing actions are more associated with pushing negative objects away from the body (Caciopoo, Priester, & Bernston, 1993; Chen & Bargh, 1999). However, recent evidence shows that these preferences can be altered by representing the “self” outside of the body, for instance by labeling a box with a participant’s name on a computer screen (Markman & Brendl, 2005). Here, when participants were shown a box labeled with their name in a hallway on the computer screen, participants were faster to push in response to positive words if their labeled box was far away on the screen. Therefore, participants were more likely to perform an action that was typically associated with negative words if that motion resulted in their virtual-self pulling. Theories of embodied cognition are not yet developed enough to explicitly account for such results, where the self can be represented so freely outside of the actual physical body.

Despite such unanswered questions, by providing evidence that action and

perception influence language processing, we gain a richer understanding of how dynamic the process of language comprehension really is. Instead of treating language as a process of logical rules operating on amodal symbols, and transitioning from one static state to another, the majority of experimental findings support a view of language wherein the linguistic representations are coextensive with sensorimotor representations and change continuously in time. Thus, rather than focusing overmuch on the rare and impressive feats of cognition that only humans can accomplish (such as mental arithmetic and formulating chess strategies), a great deal of progress in understanding the mind can be achieved by focusing on the ubiquitous ordinary feats of perception-action cycles (such as engaging in interactive conversations that integrate visual and linguistic cues), behavior that philosophers are now calling “everyday coping”.

CHAPTER 3
INDIVIDUAL DIFFERENCES IN MEASURES OF LINGUISTIC EXPERIENCE
ACCOUNT FOR VARIABILITY IN THE SENTENCE PROCESSING SKILL OF
FIVE-YEAR-OLDS²

Over the past fifteen to twenty years, a great deal of evidence has accrued in support of the notion that when a sentence is heard or read, the adult language comprehension system rapidly accesses many different sources of linguistic and/or non-linguistic information in pursuit of extracting structure and meaning from the signal (e. g. Altmann & Steedman, 1988; MacDonald, Pearlmutter, & Seidenberg, 1994; Snedeker & Trueswell, 2004; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell, Tanenhaus, & Garnsey, 1994). Until recently, however, a lack of child-appropriate behavioral techniques has hindered the study of young children's on-line syntactic processing and development. Here, we consider whether non-linguistic cues facilitate the comprehension of syntactically ambiguous sentences in children, present data from a new technique to support previous findings from the children's sentence processing literature, and explore the role of individual differences in linguistic experience in facilitating the transition to more adult-like sentence processing.

In order to study on-line language comprehension, researchers often present adult subjects with syntactically ambiguous sentences and then examine the sources of information that influence the manner in which these sentences are initially interpreted. Take, for example, the two following sentences.

- 1a) Put the apple on the towel in the box.
- 1b) Put the apple that's on the towel in the box.

² In press: Anderson, S.E., Farmer, T.A., Goldstein, M., Schwade, J., & Spivey, M. (in press). Individual differences in measures of linguistic experience account for variability in the sentence processing skill of five-year-olds. In I. Arnon & E.V. Clark (Eds.), *How children make linguistic generalizations: Experience and variation in first language acquisition*. Trends in Language Acquisition Research: John Benjamins Publishing Company.

In example (1a), the prepositional phrase (PP) *on the towel* creates a temporary syntactic ambiguity in that it can be initially interpreted as a destination (or Goal) for the referring expression *the apple*, thus attaching to the verb phrase (VP-Attachment). Alternatively, the PP *on the towel* could be interpreted as a modifier of the noun phrase (NP), such as *Put [the apple on the towel] in the box* (NP-attachment).

Referential context is a non-linguistic cue that has been shown to influence how adults initially attach an ambiguous PP. When participants hear ambiguous sentences like (1a) in the presence of visual scenes where only one referent is present (an apple already on a towel), along with an incorrect destination (an empty towel), and a correct destination (a box), adults often look to the incorrect destination until the second disambiguating PP is heard, at which time eye-movements tend to be re-directed to the correct destination (Tanenhaus et al., 1995; Trueswell, Sekerina, Hill, & Logrip, 1999). Looks to the incorrect destination indicate “garden-pathing,” or, initially incorrectly attaching the PP to the verb phrase. Looks to the incorrect destination do not occur when the instruction is unambiguous as in (1b).

The looking patterns are, however, markedly different when the visual context contains two possible referents (e.g., an apple on a towel and another apple on a napkin). When hearing an ambiguous sentence like (1a) in a “two-referent” visual context, adults tend to look at the correct referent (the apple on the towel) and move it to the correct destination with few looks to the incorrect destination. In accordance with various instantiations of referential theory (Altmann & Steedman, 1988; Spivey & Tanenhaus, 1998), when two possible referents are present, an expectation is created that the two similar entities will be discriminated, thereby forcing a modifier interpretation of the initial PP (NP-attachment).

While adults use visual context to resolve temporary ambiguities with less

common but correct interpretations of sentences, children do not immediately use the visual context to disambiguate temporarily ambiguous sentences in the same way that adults do. Regardless of the number of referents in the visual scene, children continue to garden-path when hearing an ambiguous sentence (Trueswell, et al., 1999). However, children behave similarly to adults in regard to another cue. Lexical biases, or the frequency with which a verb (like *Put*) takes a prepositional object (such as *on the towel*) in naturally occurring language, influence the degree to which adults and children experience a garden-path effect. When the lexical bias of a verb strongly supports VP-attachment, both adults and children prefer the VP-attached interpretation of the PP (Britt, 1994; Snedeker & Trueswell, 2004). For example, when hearing a globally ambiguous sentence with a VP-attachment biased verb like “tickle,” e.g., “Tickle the pig with the fan,” both adult and child participants will pick up a fan provided in the visual display and use it to tickle a pig (VP-attachment) more often than tickling a second pig in the visual display holding a fan (NP-attachment), demonstrating a preference for resolving the ambiguity with VP-attachment. If, however, a verb does not typically take a prepositional object, such as *choose*, which favors NP-attachment (Snedeker & Trueswell, 2004), children and adults prefer the NP-attached interpretation of the PP. Therefore, in an example like, “Choose the cow with the stick,” participants attach the PP to the NP and pick up the cow holding the stick in the visual display (rather than using another stick present in the visual display to pick a second cow). Unlike adults, however, the scene-based referential context does not interact with the lexical-bias of the verb in determining which interpretation children initially entertain. That is, although children show remarkable sensitivity to the biases of the verbs, context does not further facilitate children’s initial attachment preference (Snedeker & Trueswell, 2004).

However, under some experimental circumstances, children do seem to use referential context in an adult-like way. For instance, when a child does not see a visual scene containing two referents until hearing the entire ambiguous sentence, she performs more like an adult in the two-referent, ambiguous-sentence condition (Crain & Meroni, 2002). Thus, hearing an ambiguous sentence before seeing the visual scene may simplify the task, allowing the child to deal with the inputs one at a time and ultimately to incorporate the visual scene information into her comprehension of the sentence. These results may imply that processing efficiency is linked to not only language comprehension (Fernald, Perfors, & Marchman, 2006; Hurtado, Marchman & Fernald, 2008; Marchman & Fernald, 2008), but also to the ability to use multiple cues during language comprehension.

Children are not adult-like in their sentence processing ability, but it is still unclear what mechanisms underlie such a transition from child- to adult-like sentence processing. Snedeker and Trueswell (2004) suggest that the sources of information that are most reliable in the input children receive, specifically the lexical bias of the ambiguity producing verb, are relied on more heavily than less reliable information, such as the referential context. Such an explanation implies a role for linguistic experience: a child with greater linguistic experience may be more likely to have come across either more diverse examples of syntactic constructions or more examples of visual context as a reliable disambiguating cue. As a result, a child with more linguistic experience may be more likely to behave in a more adult-like way in the two-referent, ambiguous condition that is typically most difficult for children (Snedeker & Trueswell, 2004; Trueswell, et. al, 1999).

Previous research supports this prediction, demonstrating that individual differences in linguistic input predict children's mastery of more complex sentences (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). Specifically, individual

differences between children's comprehension of multi-clause (more complex) sentences is related to the proportion of such multi-clause sentences they hear in parental speech. Children whose parents use multi-clause sentences a greater proportion of the time have greater mastery of such sentences. These data suggest that exposure to more complex syntactic input results in the child's increased mastery of such syntactically complex sentences. The present experiment aims to establish a more explicit link between individual differences in linguistic experience and more adult-like use of visual cue information in processing complex sentences. If exposure to the input is what drives the comprehension system toward adult-like behavior by differentially weighting the reliance on one or more sources of information as the databases for each cue grows via experience with language, children with more linguistic experience should have a more adult-like use of context than children with less linguistic experience.

In the adult literature, researchers use vocabulary scores to index linguistic experience (see MacDonald & Christiansen, 2002 for a discussion). Similarly, child research suggests that vocabulary growth is related to the extent and diversity of linguistic experience, at least until the age of two-years-old (Hoff, 2003). These data suggest then that vocabulary scores are a reasonable approximation of linguistic experience, in that children with richer linguistic experiences also have correspondingly higher vocabulary scores. More recent research demonstrates a link not only between early maternal input and later vocabulary size (Hurtado, et al., 2008), but also between vocabulary size and language processing efficiency (Fernald, et al., 2006; Hurtado, et al., 2008; Marchman & Fernald, 2008). Eighteen-month-old children of parents who spoke relatively more often and produced relatively more complex constructions had correspondingly larger vocabularies at twenty-five months (Hurtado, et al., 2008; Fernald, et al., 2006) and at eight years of age (Marchman &

Fernald, 2008). Also, children who heard more early input had greater processing speed in language comprehension tasks at a later age (Hurtado, et al., 2008). Although there are likely many factors that drive the development of the system over time, previous research suggests that linguistic experience is a legitimate starting place for looking at these individual differences in sentence processing ability within a single age group.

A new methodology for use with children

Although the eye-tracking paradigm has provided invaluable child-performance data, the technique can be expensive, can involve slow hand-coding of data, and can sometimes incur parental objections. As a supplement to eye-tracking, research has demonstrated that monitoring the continuous nonlinear trajectories recorded from the streaming x, y coordinates of computer-mouse movements can serve as an indicator of the underlying cognitive processes in spoken word recognition (Spivey, Grosjean, & Knoblich, 2005), categorization (Dale, Kehoe, & Spivey, 2007), adult syntactic processing (Farmer, Anderson, & Spivey, 2007), and various higher-level social phenomena (Freeman, Ambady, Rule, & Johnson, 2008; Wojnowicz, Ferguson, Dale & Spivey, 2009). Unlike saccadic eye-movements, mouse movements are generally smooth, continuous, and can curve substantially mid-flight. This dense sampling methodology then allows graded spatial attractions to emerge within single trials. For example, although eye-movement data afford approximately 2-4 data points (saccades) per second, mouse-tracking yields somewhere between 30-60 data points per second, depending on the sampling rate of the software used. By recording the x, y coordinates of the mouse as it moved with the goal-directed hand motion to click on the appropriate object, smooth competition between the partially active underlying representations are revealed in the shape and curvature of the hand-movement trajectories. These properties of tracked mouse-movements provide crucial benefits in

that they allow a fine-grained and graded response pattern to emerge within an individual trial.

Although children can use a computer mouse at 3.5 years, on average, and the onset of autonomous computer use is approximately 3.7 years (Calvert, Rideout, Woolard, Barr, and Strouse, 2005), the degree to which this cheap, portable, and accessible mouse-tracking methodology can be used to study complex cognitive phenomena in young children remains to be seen. Here, we exploit these mouse-movement trajectories, in the visual-world paradigm (Tanenhaus et al., 1995), to determine its efficacy in detecting children's processing differences in complex cognitive tasks.

Purpose of the current experiment

Here, we employed the mouse-tracking methodology to monitor the motor output of children as they moved objects around a natural scene in response to ambiguous (1a) and unambiguous (1b) spoken instructions. To hold the lexical-bias variable constant, sentences contained only the verb *Put*, which is strongly biased toward VP-attachment (for example, attaching the first PP “on the towel” to the verb “put” resulting in the garden-path effect). The computer mouse-movements made towards either target or distractor objects in the visual scene provide an index of what the child is attending to, moment to moment, as they interpret the utterance. Specifically, trajectories that veer towards the incorrect goal location while the correct object is picked up and moved to the ultimately correct location in the display indicate a garden-path effect, signaling some consideration of the VP-attachment.

This is the first experiment to use the mouse-tracking methodology with children, and we predicted results similar to those of previous eye-tracking studies. In the one-referent context, we predicted significant curvature toward the ultimately incorrect destination for ambiguous sentences (1a), relative to the unambiguous

control condition (1b), signaling some consideration of VP-attachment. In the two-referent context, we predicted that if mouse-movement data provide results similar to the eye-movement data elicited from children in previous studies (Trueswell et al., 1999), then we would expect to detect significant ambiguous-sentence trajectory curvature toward the incorrect destination, relative to the unambiguous control condition.

In addition to verifying that mouse-tracking is a suitable methodology for examining children's comprehension of complex sentences, we also examined the role of linguistic experience on children's performance in this task. While previous research suggests that linguistic experience plays a crucial role in the transition from child- to adult-like sentence-comprehension (Snedeker & Trueswell, 2004), this is the first experiment to explicitly test such predictions within a single age group using individual differences in the visual world paradigm. To examine the role of linguistic experience on children's performance, children also completed a receptive vocabulary test. As described above, we use vocabulary scores as a proxy for linguistic experience, predicting that children with higher vocabulary scores would have more adult-like use of context in the two-referent, ambiguous-sentence condition.

Method

Participants

Forty-three participants, nineteen females and twenty-four males, between the ages of 56 and 70 months old ($M=63.38$ months) participated in this experiment. We recruited participants from the developmental lab's birth announcement database and through the local school systems. The children came from a mostly middle to upper middle SES background. Each child received a small toy for participating. One additional child participated, but was excluded from the analysis due to refusal to participate in the mouse-tracking task. Twelve additional children were excluded from

the trajectory analysis because they produced too many incorrect mouse trajectories (see *Data screening and coding* for details).

Materials and Procedure

We adapted and digitally recorded sixteen experimental items from Spivey, Tanenhaus, Eberhard, & Sedivy (2002). We made each item “child-friendly” by substituting potentially unfamiliar objects with objects included in the MacArthur-Bates Communicative Development Inventory, a widely used parental-report of productive vocabulary for children up to thirty months (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994). We recorded ambiguous (1a) and unambiguous instructions (1b) that corresponded to each of sixteen experimental scenes (see Spivey, et al., 2002 for details). We varied each of the visual scenes corresponding to the sixteen experimental items to produce a one-referent condition and a two-referent condition. The one-referent visual context (Figure 1, top) contained a target referent (e.g., an apple on a towel), an incorrect destination (e.g., a second towel), the correct destination (e.g., a box), and a distracter object (e.g., a flower). In the two-referent context (Figure 1, bottom), all items were the same except that we replaced the distracter with a second possible referent (such as an apple on a napkin). We also created sixteen distracter scenes, designed to accompany the filler sentences, and these scenes used different combinations of the objects from the experimental trials and from a set of new, easily recognized objects.

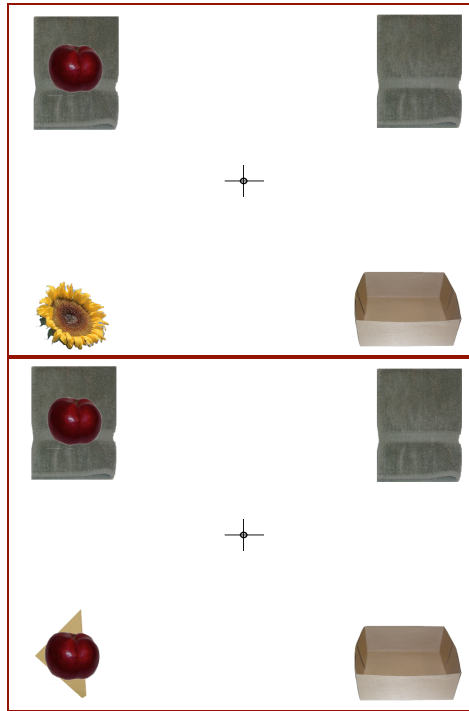


Figure 1. Example of one- (top) and two- (bottom) referent displays shown to children.

A female adult, using age-appropriate child-directed speech, recorded the spoken instructions using the speech synthesizer program Audacity. Children sat at a child-sized table in front of a computer monitor. Macromedia Director MX presented the visual context and sound files on an Apple G4 computer. We used a small, portable mouse to collect mouse movements. A visual display containing all items associated with each auditorily-presented instruction appeared on the monitor. See Figure 1. For each scene, at the beginning of the sound-file participants first heard “Place the arrow at the center of the cross.” Once the child moved the cursor to the center of the cross, an experimenter repositioned the mouse to the center of a small sticker on the table. This repositioning ensured that the mouse cursor always started in the same place for each trial and minimized the possibility that the participants would run out room to move the mouse on the table during each trial. Sound-files accompanying experimental scenes always played the experimental sentence first, followed by two additional filler instructions. Thus, for experimental items, participants viewed the

appropriate scene while hearing, for example:

- 1) Place the cursor at the center of the cross.
- 2) Put the apple on the towel in the box (experimental trial).
- 3) Now put the apple beside the flower (filler sentence).
- 4) Now put the flower in the box (filler sentence).

We also created sixteen additional filler scenes, and participants heard three scene-appropriate unambiguous filler instructions accompany these. In all cases, six seconds separated the offset of one sentence from the onset of the next within each trial.

Between trials, children saw a large yellow star centered on the screen and heard the enthusiastically spoken instruction “Click on the star to go on!” This step provided a natural break in the experiment and helped keep the child motivated.

In both the one- and two-referent conditions, the target referent always appeared in the top left corner of the screen, the incorrect destination always appeared in the top right corner, and the correct destination was always located at the bottom right corner (as in Figure 1). The bottom left corner of the screen showed either the distracter object in the one-referent trials or the second referent in the two-referent trials. Filler sentences prevented participants from detecting the regularity created by the object placements in the experimental trials. That is, in addition to the movement used in the experimental instructions, eleven distinct movements were possible in the visual scene across trials (i.e., bottom right-hand corner to top right-hand corner), and an approximately equal number of filler sentences were assigned to each of these. Therefore, for each child, ten sentences required an object in the upper left-hand corner to be moved to the upper right corner of the display, eight sentences required an object in the upper left-hand corner to be moved to the bottom left-hand corner, and so on.

In each scene, participants saw four to six color images, depending on the instructions. We used pictures of real objects, taken by a digital camera and edited in

Adobe Photoshop, to construct the images in the visual scenes. Visual stimuli subtended an average of 5.96 X 4.35 degrees of visual angle, and were 14.38 degrees of visual angle diagonally from the central cross. Mouse movements were recorded at an average sampling rate of 40 Hz.

We counterbalanced the experimental items across four presentation lists. Each list contained four instances of each possible condition, but only one version of each sentence frame and corresponding visual context. Participants were randomly assigned to one of the four presentation lists, and the presentation order of the items within each list was randomized for each participant. Each participant saw three practice items at the beginning of the experiment.

Each child also completed the Peabody Picture Vocabulary Test—Third Edition, PPVT-III (Dunn & Dunn, 1997), a widely-used and reliable test of receptive vocabulary. Participants heard a spoken word and pointed to the correct referent from a display of four pictures. Half of the subjects received the PPVT before the computer portion of the experiment, and the other half of the subjects received it after. Parents completed a form providing information about the child's computer use and their demographic information. The entire session lasted approximately 30 minutes.

Results

Data Screening and Coding

Mouse movements were recorded during the grab-click, transferral, and drop-click of the referent object in the experimental trials. As a result of the large number of possible trajectory shapes, the x,y coordinates for each trajectory from each experimental trial were plotted in order to detect the presence of errors or otherwise aberrant movements. A trajectory was considered valid and submitted to further analyses if it was initiated at the top left quadrant of the display (location of the correct referent) and terminated in the bottom right quadrant (location of the correct

destination), signaling that the correct referent had been picked-up and placed at the correct destination. Twelve children were excluded from all trajectory analyses because they either produced more than six errors on the sixteen experimental trials or committed errors on each of the four trials in one condition. The error types of all forty-three children, along with their frequency per condition, are included in Table 1 (with the numbers outside of the parentheses indicating the errors of children who were included in the trajectory analyses, and the number in the parenthesis indicating the error frequencies for all the children). No significant differences existed between the included versus the excluded children in age, vocabulary score, gender, or number of hours using the computer at home or at school (computer familiarity), all $ps > .15$. Thus, age, vocabulary, and computer familiarity were not likely causes of refusal to participate or exclusion from the on-line analysis.

Table 1: Error types causing a trial to be excluded from all analyses, per condition. The number in parentheses is the number of error trials from children excluded from the final on-line analysis; the number outside the parentheses is the number of error trials the thirty-one children contributing to the online analyses.

Error Type	1 Referent, Ambiguous	1 Referent, Unambiguous	2 Referent, Ambiguous	2 Referent, Unambiguous
Target Referent Moved to Incorrect Destination	19 (24)	3 (3)	6 (14)	5 (6)
Incorrect Referent Moved to Incorrect Destination	1 (6)	1 (3)	23 (37)	4 (11)
Incorrect Referent Moved to Correct Destination	0 (0)	1 (1)	2 (3)	4 (6)
Picture Representing a Destination was Moved	1 (1)	0 (1)	1 (2)	1 (3)
Erratic Movement Yielding an Uninterruptible Trajectory	5 (13)	3 (10)	10 (15)	5 (11)

In assessing overall error rates in each condition, children made significantly more errors, $F(1,42) = 11.96, p < .01$, in the two-referent condition ($M = 1.26, SD = 1.19$) than in the one-referent condition ($M = .72, SD = 1.26$). Similarly, children made significantly more errors, $F(1,42) = 20.65, p < .01$, in the ambiguous-sentence condition ($M = 1.34, SD = 1.13$) than in the unambiguous sentence condition ($M = .64, SD = .95$). However, the interaction was not significant, $p > .37$. A similar trend emerges when looking at only the thirty-one children included in the on-line analyses. These children also made significantly more errors, $F(1,30) = 9.105, p < .01$, in the two-referent conditions ($M = .98, SD = 1.03$) than in the one-referent conditions ($M = .55, SD = .76$). Similarly, these children made significantly more errors, $F(1,30) =$

17.17, $p < .01$, in the ambiguous-sentence conditions ($M = 1.01$, $SD = .99$) than in the unambiguous-sentence conditions ($M = .44$, $SD = .74$). However, the interaction was not significant, $p > .48$. These data are similar to the error rates found in previous research using eye-tracking (Snedeker & Trueswell, 2004; Trueswell, et al., 1999). Ambiguous sentences elicit more off-line errors than the unambiguous sentences due to the increased linguistic complexity of the ambiguous sentences. Similarly, the two-referent conditions elicit more off-line errors than the one-referent conditions, because, in part, two identical possible referents in the visual scene increase the complexity of the referential context.

All analyzable trajectories were time-normalized to 101 time-steps by a procedure originally described in Spivey et al. (2005). All trajectories were spatially aligned so that the first observation point corresponded to the x,y coordinates of (0, 0). Then, across 101 normalized time-steps, the corresponding x and y coordinates were computed using linear interpolation.

The Context and Garden-path Effects

The mean ambiguous- and unambiguous-sentence trajectories at each of the 101 time-steps in the top panel of Figure 2 demonstrate that in the one-referent context, like adults, the average ambiguous-sentence trajectory was more curved toward the incorrect destination than the average trajectory elicited by the unambiguous sentences. Unlike adults, however, in the two-referent condition (Figure 2, bottom), there still appears to be noticeable attraction toward the incorrect destination for the ambiguous-sentence trajectories. Both of these trends support the notion that the children were garden-pathed by the syntactic ambiguity manipulation, regardless of context.

In order to determine whether or not the divergences observed across the ambiguous- and unambiguous-sentence trajectories in the one-referent and two-

referent contexts were statistically reliable, we conducted a series of t-tests. Due to the horizontally elongated shape of the overall display, differences in x-coordinates of the mouse movements are more indicative of velocity differences, and differences in the y-coordinates are more indicative of genuine spatial attraction toward the incorrect referent. As such, analyses were conducted separately on the x- and the y-coordinates at each of the 101 time-steps. In order to avoid the increased probability of a Type-1 error associated with multiple t-tests, and in keeping with bootstrap simulations of such multiple t-tests on mouse trajectories (Dale, et al., 2007), an observed divergence was not considered significant unless the coordinates between the ambiguous- and unambiguous-sentence trajectories elicited p -values $< .05$ for at least eight consecutive time-steps.

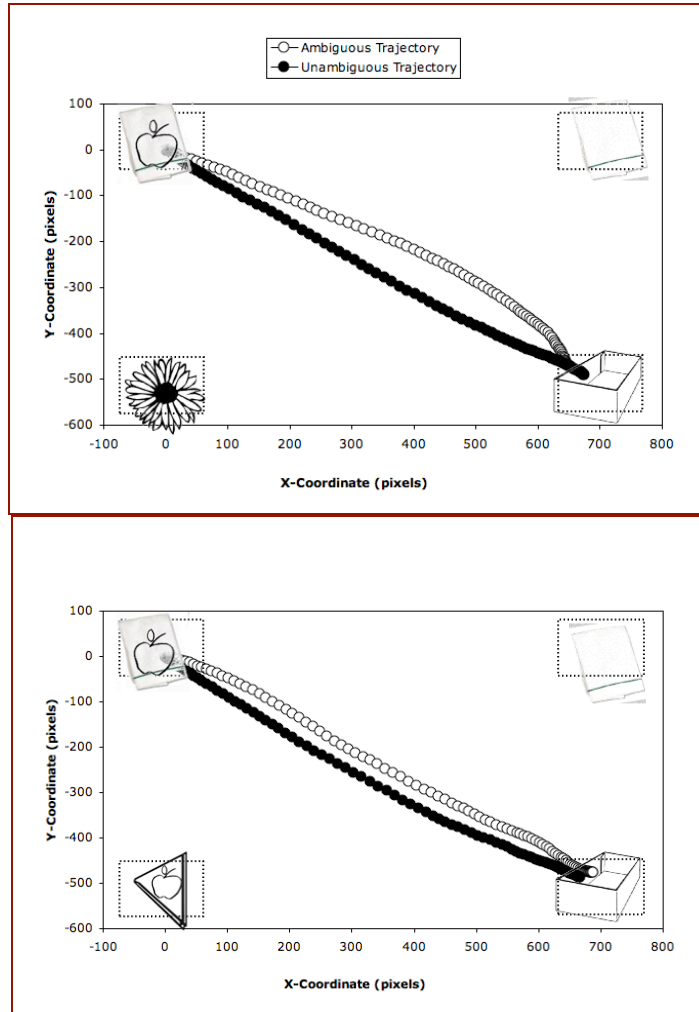


Figure 2. Time-normalized ambiguous- and unambiguous-sentence trajectories elicited in the one-referent (top) and two-referent (bottom) contexts.

In the one-referent context, no significant divergence occurred between the x-coordinates of the ambiguous- and unambiguous-sentence trajectories indicating that, across time, trajectories progressed toward the right side of the screen at approximately the same speed in both sentence conditions. For the y-coordinates, however, the ambiguous- and unambiguous-sentence trajectories diverged significantly from time-steps 43-78, all t 's > 2.08 , all p 's $< .05$, with higher y-coordinates (closer to zero, thus closer to the top of the screen) in the ambiguous than in the unambiguous sentence condition. See Figure 2. In the one-referent context, the divergence of the average y-coordinates of the ambiguous-sentence trajectory away

from the unambiguous-sentence trajectory and towards in the incorrect destination is indicative of the garden-path effect observed also in eye-tracking research (Snedeker & Trueswell, 2004; Trueswell, et al., 1999).

In the two-referent context, significant x-coordinate divergences between the ambiguous- and unambiguous-sentence trajectories occurred from time-steps 9-50, all t 's > 2.07, all p 's < .05, with ambiguous-sentence trajectories traveling *more quickly* toward the correct destination. However, in the two-referent context, there was no statistically reliable y-coordinate divergence at any of the 101 time-steps. This null effect suggests that the two-referent context did not induce the expected garden-path effect. This was surprising given the data from previous eye-tracking studies, demonstrating that children look often to the incorrect destination, temporarily considering it as the goal location of the putting action, even when two referents are present.

Thus, the t-test analyses provide mixed support for the expectation that the data obtained by tracking streaming x,y coordinates in mouse-tracking would align with the saccadic eye-movements of children in the visual world paradigm (Trueswell et al., 1999). In the one-referent context, children do consider, at least temporarily, the destination interpretation of the ambiguous PP. In mouse-tracking the curvature towards the incorrect destination is commensurate with the large number of looks to the incorrect destination in this condition when examining eye-movements. In the two-referent condition, however, there appears to be no statistically significant attraction toward the incorrect destination in the presence of a syntactic ambiguity. This result is surprising given previous research and given that in Figure 2 (bottom), there appears to be a divergence between ambiguous- and unambiguous-sentence trajectories.

One explanation for the incongruence of our results with previous eye-tracking results in the two-referent condition is that the amount of variability in the y-

coordinates of our trajectory data exceeds the power to detect divergence, should it be present. In order to reduce the variability surrounding each participant's mean y-coordinate movement in each condition, to avoid concerns associated with multiple comparisons in the t-tests above, and to assess directly the statistical reliability of the crucial Context X Ambiguity interaction, we averaged the y-coordinates recorded from time-steps 34-67 (the middle portion of the movement where, on average, the ambiguous-unambiguous divergences appear most extreme) and used the average y-coordinate response as the dependent measure in a 2 (Context) X 2 (Ambiguity) repeated-measures ANOVA. Average "middle-segment" y-coordinates were closer to the incorrect destination in the one-referent context, suggesting greater uncertainty in the one-referent than in the two-referent context, $F(1, 30)=7.58, p=.01$. The y-coordinates were also closer to the incorrect destination for the ambiguous over the unambiguous condition, $F(1, 30)=8.26, p=.007$. However, the Context X Ambiguity interaction was not significant, $F(1, 30)=1.83, n.s.$, suggesting that context does not modulate the magnitude of the garden-path effect. These data support earlier eye-tracking data and suggest that children still experience the garden-path effect in the two-referent ambiguous context.

Taken together, these results reveal spatial attraction toward the ultimately incorrect destination (i. e. garden-pathing) in response to ambiguous sentences when compared to the unambiguous condition, regardless of context. They also suggest that, like adults, children experience the garden-path effect in the one-referent, ambiguous-sentence condition, and that, unlike adults, children do not readily utilize context to disambiguate the sentence in the two-referent ambiguous-sentence condition.

The Influence of Individual Differences in Linguistic Experience on the Use of Referential Context

As the trajectory data demonstrate, children do not readily incorporate visual context into their comprehension of an ambiguous sentence like adults do. Previous research (Snedeker & Trueswell, 2004) suggests a role for linguistic experience in accounting for the transition to more adult-like behavior in this task. It may be the case that the most reliable sources of information in the input children receive are relied on more heavily than other information sources of information, such as referential context.

This explanation implies that children with more linguistic experience may have come across either more diverse syntactic constructions or more instances of visual context providing a reliable disambiguating cue. Therefore, children with more linguistic experience should behave in a more adult-like way, specifically in the two-referent, ambiguous-sentence condition that is typically most difficult for children. Here, we aim to establish a more explicit and empirical link between individual differences in linguistic experience and more adult-like processing performance. Vocabulary scores served as a proxy for linguistic experience, with higher vocabulary scores indicating greater linguistic experience (e.g. Hoff, 2003).

In accessing the vocabulary scores, we explored the degree to which those scores could account for difficulty associated with the most difficult condition for this age group, the two-referent context (Snedeker & Trueswell, 2004; Trueswell, et al., 1999). First, we examined the error rates and raw (non-transformed) vocabulary scores of all the participants. The raw vocabulary scores significantly predicted the number of error trials only in the two-referent ambiguous-sentence condition, $t(38)=-2.77$, $p=.009$, $\beta=-.42$, $R^2=.17$. See Figure 3. The relationship is negative, such that children with higher vocabularies have fewer error trials in the two-referent ambiguous condition than do children with lower vocabularies. This significant relationship

implicates vocabulary, and thus linguistic experience, as a strong predictor of who could move the correct referent to the correct destination in the critical two-referent, ambiguous-sentence condition.

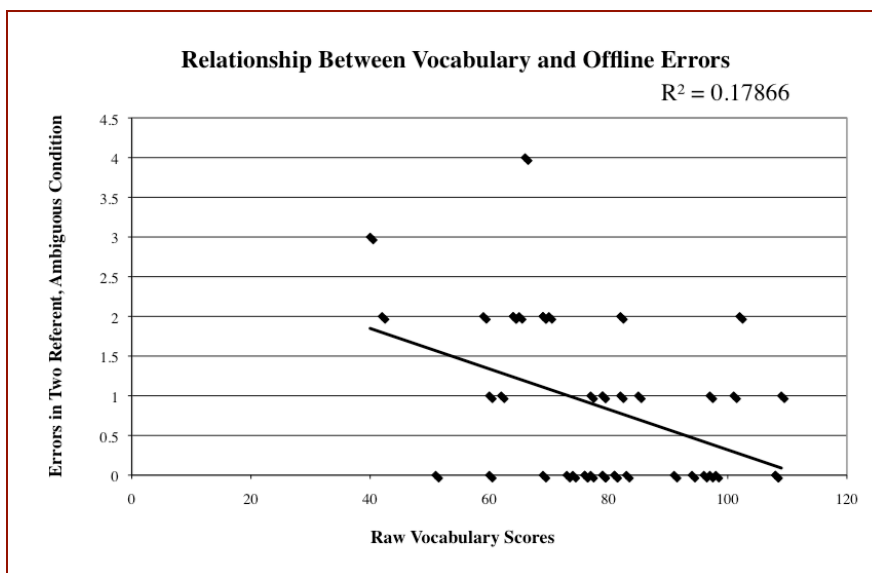


Figure 3. Relationship between vocabulary and number of errors in the two-referent, ambiguous-sentence condition. Children with higher vocabulary scores (richer linguistic experience) made fewer errors only in this condition.

We also examined the relationship between the raw vocabulary scores and an on-line measure of trajectory divergence. Of the children who contributed to the on-line analyses, raw vocabulary scores did not predict the degree of trajectory curvature toward the incorrect destination (indexed by calculating the average maximum deviation value for ambiguous-sentence trials, per subject) in either the one- or two-referent ambiguous-sentence conditions.

Discussion

These data present an emerging picture whereby the average five-year-old child does not employ referential context to override the VP-attachment bias associated with this manipulation. This observation, along with the large number of errors made in the two-referent ambiguous-sentence condition (Table 1), is consistent with the eye-

movement data reported in Trueswell et al. (1999) and Snedeker and Trueswell (2004). The fact that the trajectory data are commensurate with previously-reported eye-tracking data lends support to the notion that the mouse-tracking methodology is a feasible and reliable methodology for use in understanding the underlying cognitive processes of young children. It is important to note, however, that we do not advocate, or foresee, the usurping of eye-tracking methods in lieu of the advantages of mouse-tracking enumerated here. Instead, we believe that the two techniques can be used in a complementary (even simultaneous) fashion in order to more fully unlock the nature of the complex interactions associated with high-level cognitive processes, even in young children.

Further, the results reported here provide insight into the nature of the developmental mechanisms responsible for the transition from child-like to more adult-like sentence processing skill. Linguistic experience significantly predicts off-line errors in the condition that is most difficult for children, the two referent, ambiguous-sentence condition. Children with more linguistic experience produced fewer errors in this condition than did children with less linguistic experience, but not in any of the other conditions. To the degree that variability in receptive vocabulary serves as a suitable proxy for variability in linguistic experience, these data suggest that as children are exposed to richer linguistic experiences, they become more adult-like in their use of referential context.

More broadly, our data dovetails with research into the relationship between early input and later language processing efficiency (Hurtado, et al., 2008). In one longitudinal study, Hurtado and colleagues (2008) found a link between maternal speech and children's vocabularies. Spanish learning eighteen-month-old children whose mothers spoke relatively more to them had larger vocabularies at twenty-four months than children whose mothers spoke relatively less to them. Also, children with

larger vocabularies at twenty-four months were faster to identify familiar words in speech. These data further support our assertion that children with richer linguistic experience have correspondingly larger vocabularies, and that these early experiences influence later language comprehension. Additionally, Hurtado and colleagues (2008) found a link between language input at eighteen months and processing efficiency at twenty-four months. Both quality and quantity of the mother's speech to the child at eighteen months predicted children's language processing at twenty-four months. Children who received relatively more maternal input at the first time point were faster in online comprehension tasks at the second time point. Our research compliments this work, extending the link between early input and later processing efficiency to five-year-old children. In our research, larger vocabularies, resulting in part from relatively more linguistic experience, predicted off-line error rates in the difficult two-referent, ambiguous-sentence condition.

While these data are an important step towards understanding the mechanisms underlying the transition from child to adult-like sentence processing abilities, several questions remain. For instance, how does linguistic experience mechanistically facilitate this change (Snedeker & Trueswell, 2004)? There are at least three possible explanations: linguistic experience increases reliance on visual scene information, linguistic experience decreases reliance on lexical bias information, or linguistic experience does both. First, linguistic experience may increase reliance on visual scene information. As children are exposed to a greater variety of sentence constructions in an increasing number of communicative contexts, their sensitivity to and reliance on cues other than verb-based lexical biases, such as visual context, may increase. The statistical relationships that exist between language and properties of a co-occurring visual scene are likely to be substantially sparser, and thus more difficult to learn, than the statistical information associated with verb-biases in child-directed

speech. As a result, efficiently utilizing scene-based referential information during language comprehension may simply require more experience with contextualized linguistic processing before visual-scene based information can act as a viable cue to accurate interpretation. A second possibility is that linguistic experience decreases reliance on lexical bias information. As children are exposed to richer linguistic experiences, their reliance on the verb's lexical bias decreases as a function of encountering a greater number of instances where lexical bias is not the only reliable cue. Third, linguistic experience may do both. As linguistic experience increases, reliance on other cues, like visual scene information, is increased *and* reliance on lexical bias cues is decreased. Hence, if children have had more linguistic experience, they may have more varied experiences with syntactic constructions, specifically with the verb "put" being followed by a PP as both VP-attachment and as NP-attachment, and they may also have had more experience with the relationship between visual context and co-occurring language. Therefore, this third possible explanation suggests that a child with greater linguistic experience is more likely to have experienced a) more diverse examples of syntactic constructions (decreasing reliance on verb bias) and b) more examples of visual context as a reliable disambiguating cue (increasing reliance on visual context cues).

The current results do not suggest which of these three possibilities is most accurate, and computational modeling is likely needed to more explicitly examine the impact of changing weights according to each of the three alternative explanations. By adapting the model used in earlier adult sentence processing research (Farmer, Anderson, & Spivey, 2007), the influence of changing individual weights within the model can be examined and compared to the behavioral data reported here. Hence, work is underway to examine the impact of changing individual weights in the adult sentence processing model in order to more thoroughly explore the mechanistic role of

linguistic experience.

Another unanswered question is what other sources of individual-based variability may contribute to this transition? For example, another factor that may contribute to the development of sentence processing skill is the ability to inhibit a salient response. In other words, cognitive control, which develops into late adolescence, may contribute to a child becoming more adult-like in their sentence comprehension. As cognitive control increases, the child may be better at inhibiting an appealing, or more frequent, but ultimately incorrect response in order to select the correct referent and move it to the correct destination without experiencing the garden-path effect. Individual differences in cognitive control have been shown to reliably predict adult performance on ambiguous sentences (Novick, Trueswell, & Thompson-Schill, 2005). For example, adults scoring higher on cognitive control tasks exhibit less difficulty recovering from misinterpretation of an ambiguity.

Similarly, recent cross-linguistic research suggests that cognitive control may be even more important in accounting for differences between children's and adult's sentence processing skills than lexical bias information (Choi & Trueswell, 2010). In this research, the experimenters used eye-tracking to compare the performance of Korean speaking adults and children. The data demonstrate that Korean five-year olds do not use late-arriving verb bias information to override early-arriving syntactic information, producing results similar to those of earlier eye-tracking research looking at English-speaking five-year-olds. These results are striking because English-speaking children produce similar looking data, but for different reasons; English-speaking children do not use late arriving syntactic information to override early arriving verb bias information. Even when the lexical bias information is presented late, children still persist in garden-pathing in the two-referent, ambiguous-sentence condition. Hence, other sources of individual differences, namely cognitive control,

may be more important than verb bias in accounting for children's use of referential context and more adult-like performance.

Cognitive control may also influence how quickly children can move their attention around the visual scene. As discussed earlier, previous research suggests that the ability to quickly shift visual attention may play an important role in children's ability to use referential context effectively. For instance, when a child is prevented from viewing the visual scene until she has heard the entire sentence, that child is able to perform in a more adult like fashion in the two-referent, ambiguous-sentence condition (Crain & Meroni, 2002). Hence, inhibiting visual attention to a salient object, thus increased cognitive control, may be crucial in decreasing the complexity of the two-referent ambiguous-sentence condition.

While there is still much work to be done, the current results provide further evidence, converging with previous eye-tracking data, that five-year-olds do not use visual cues to disambiguate an ambiguous sentence. These data also suggest that individual differences in linguistic experience may be an important aspect of why this is so. Vocabulary scores predict the degree to which children are able to move the appropriate item to the appropriate destination in the traditionally difficult two-referent, ambiguous-sentence condition, implicating linguistic experience as one contributor to the transition from child-like to more adult-like processing.

CHAPTER 4

ON THE PATH TO UNDERSTANDING VERBAL ASPECT³

People often talk about past events. While they sometimes highlight the ongoing nature of the event, other times they emphasize the beginning or end state. The goal of this paper is to examine the role of grammar in this process. Specifically, how does grammatical aspect influence processing in real time? How might the use of a verb such as *walked* differ from a verb such as *was walking*? To what extent do different aspectual forms engage perceptual-motor simulations? Here, we explore the influence of verbal aspect on the conceptualization of motion events. The main question is how grammatical aspect differentially influences motor output as people are understanding language.

Grammatical aspect provides information about how events unfold in time. It provides information about the completion, duration, or repetition of actions or situations expressed by the verb (Comrie, 1976; Frawley, 1992). Take, for example, the following sentences: “David ran to the university,” and “David was running to the university.” Both convey information about a past event but they use different aspectual forms. The first sentence uses the perfective form, specifically simple past, of the verb “ran” to emphasize the completion of the action. The second uses the imperfective form, specifically past progressive, to emphasize the ongoing nature of that past event. Here after, we will refer to these as simple past and past progressive sentences, respectively. Even though aspect is known to shape the temporal “coloring” of a verb’s information, little is known about the dynamics of cognitive processing during comprehension of different aspectual forms.

³ Submitted: Anderson, S.E., Matlock, T., & Spivey, M. (submitted). On the path to understanding verbal aspect.

Recently, we examined spatial outcome differences in response to simple past or past progressive sentences in several offline studies (Matlock, Fausey, Cargill, & Spivey, 2007). Participants read a sentence like “This morning David walked to the university” (simple past) or “This morning David was walking to the university” (past progressive), and saw a schematic drawing that showed a path leading to the destination described in the sentence and ten unevenly spaced identical silhouette characters on the path (e.g., pedestrian with leg extended forward and arms bent as if in motion). Participants were instructed to “circle the man that the sentence is most naturally referring to.” They were more likely to circle a character in the middle region of the path with sentences containing past progressive verbs (e.g., *was walking*), and more likely to circle a character in the latter region of the path in response to sentences containing a simple past verb (e.g., *walked*). A similar pattern emerged in a subsequent experiment where participants were asked to indicate where along the path an object had been dropped after reading simple past or past progressive sentences. These results show that when participants read simple past sentences, they focus on the end of the path, or the location of the completed action in the scene. In contrast, when participants read past progressive sentences, they focus on the middle section of the path, where the ongoing action would have taken place. These data indicate that different aspectual forms have consequences for thinking about motion events, but questions about processing remain.

Madden and Zwaan (2003) addressed the on-line processing of verbal aspect, showing that simple past and past progressive sentences create reaction time differences in narrative reading. In one experiment, participants were quicker to respond to pictures showing a completed action after they had read a simple past sentence (e.g., The car sped through the intersection) versus a past progressive sentence (e.g., The car was speeding through the intersection). However, no such

latency differences arose when participants read sentences containing past progressive verbs and saw pictures of intermediate action. The authors suggest that the effect was not significant in the past progressive condition because readers represented the ongoing action at different stages of completion. In other words, past progressive sentences could potentially correspond to any of a number of intermediate actions, and these diffuse possibilities were not captured by static visual stimuli used in the picture verification and reaction time task. These results suggest that different aspectual forms lead to processing differences in real time. (For other work on aspect and spatial representation, see Ferretti, Kutas, & McRae, 2007; Magliano & Schleich, 2000; Morrow, 1985). Such reaction time data have revealed valuable insights into the processing of aspect. However, as suggested by the work of Madden and Zwaan (2003), they are somewhat limited when investigating diffuse alternatives made available by past progressive sentences.

In addition to such offline and reaction time experiments, there is a great deal of information about real-time cognitive processing that can be extracted from the motor dynamics of the manual response itself. For example, factors influencing latency to respond also influence later aspects of response dynamics, implying that the temporal dynamics of the motor movement that executes a response contain volumes of virtually untapped data. As a simple example, Abrams and Balota (1991) asked participants to perform a lexical decision task by rapidly sliding a handle leftward or rightward to indicate whether a string of letters was a word or not. As expected, the frequency of the word influenced reaction time, with high frequency words eliciting faster onset of response than low frequency words. Also, they found that word frequency influenced response kinematics after the response was initiated. Responses to high frequency words were executed with greater force than responses to low frequency words. These findings suggest that word frequency not only influences the

time required to recognize a word, but also influences response dynamics, suggesting that the response system is not slavishly executing a stereotyped command in response to the categorical status of the word. This makes a compelling case for looking not only at reaction time differences, but also at variables of the motor movements themselves initiated in response to a stimulus.

A similar case of ongoing motor movements revealing cognitive processing was produced by Richardson and Matlock (2007), when they analyzed eye movements during comprehension of fictive motion sentences. In fictive motion sentences, motion verbs are used to depict the extent of an object or action without there actually being any motion in the situation being described, as in “The fence ran from the house to the garage.” Previous work by Matlock (2004) had shown that response times for reading fictive motion sentences were longer than responses to equivalent sentences that did not use fictive motion. However, Richardson and Matlock were able to uncover the real-time temporal dynamics of processing by showing that listeners made more eye movements scanning along *the specific region of the screen where the mentally simulated motion takes place*. Their results suggested that a fictive-motion use of verbs generates a perceptual simulation that has some form of implied motion in it.

For the present study, aimed at improving our understanding of the perceptual simulations associated with different aspectual forms of verbs, we have employed a methodology that was inspired by the eye-tracking paradigm: *computer-mouse tracking*. Monitoring the streaming x- and y-coordinates of goal-directed mouse movements in response to spoken language is a useful indicator of underlying cognitive processes. In contrast to ballistic saccades, arm movements allow for a continuous, smooth motor output within a single trial to complement eye-tracking research. Spivey, Grosjean, and Knoblich (2005) demonstrated that these mouse

movements can be used to index the continuous activation of lexical alternatives. By recording the x,y coordinates of the mouse as it moved with the goal-directed hand motion to click on the appropriate object, competition between the partially active lexical alternatives was revealed in the shape and curvature of the hand-movement trajectories. Computer-mouse tracking has also been informative for sentence processing studies (Farmer, Anderson, & Spivey, 2007), children's sentences processing (Anderson, Farmer, Schwade, Goldstein, & Spivey, to appear), semantic categorization (Dale, Kehoe, & Spivey, 2007), color categorization (Huette & McMurray, 2010), decision making (McKinstry, Dale, & Spivey, 2008), and social judgments (Freeman, Ambady, Rule, & Johnson, 2008; Wojnowicz, Ferguson, Dale, & Spivey, 2009).

Further, our own earlier data indicates that mouse-tracking is useful and informative for exploring research questions about the on-line processing of grammatical aspect (Anderson, Matlock, Fausey, & Spivey, 2008). In one experiment, participants listened to sentences like, "Tom jogged to the woods and then stretched when he got there," or "Tom was jogging to the woods and then stretched when he got there." While participants heard these sentences, they saw scenes consisting of a path curving upwards from left to right, and terminating at the destination described in the sentence. A character was located to the right of the beginning of the path and under the destination, separated from the scene by a black box framing the destination and path. See Figure 4. The two aspectual forms elicited significantly different movement durations: participants spent a longer period of time moving the character into the scene with past progressive sentences than when they heard sentences containing simple past verbs. These data converge with and further inform earlier research, supporting the idea that past progressive aspect focuses attention on the on-going nature of the action while simple past aspect focuses attention on the end state of that

action, even during real time processing.

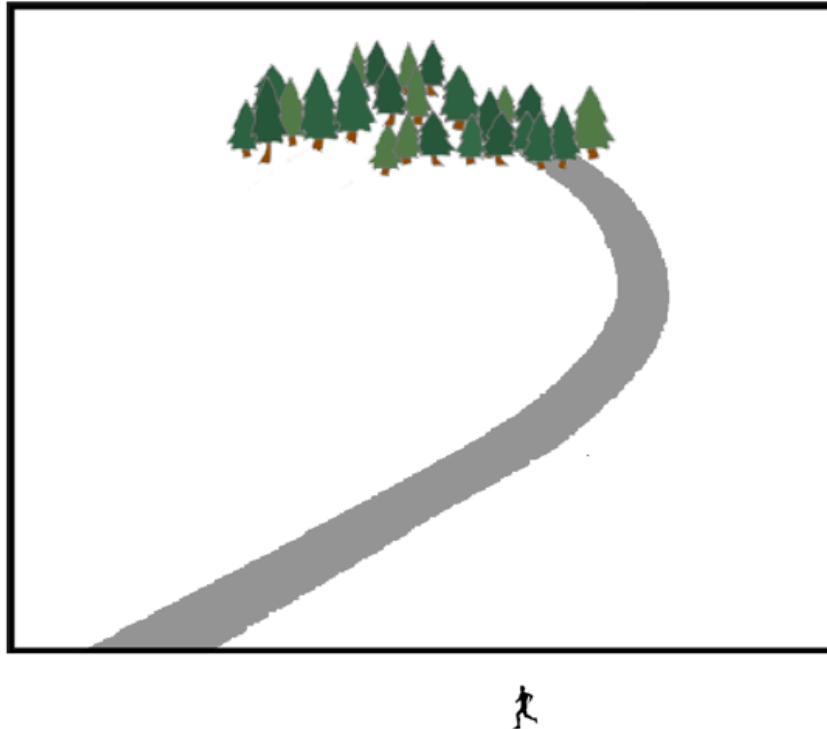


Figure 4. Visual scene accompanying sound files in Anderson, Fausey, Matlock, & Spivey (2008).

In the current experiments, we employed mouse-tracking to further investigate the online processing of different verbal aspectual forms. In Experiment 1, we replicated the earlier results of Anderson et al. (2008), extending them to a different visual scene in which we can test for effects in *the specific region of the screen where the mentally simulated motion takes place*. Experiments 2 and 3 extend these results by examining future tenses (Experiment 2) and the role of contextual descriptions (Experiment 3).

Experiment 1

In the current experiment, we used the mouse-tracking methodology to explore differences in comprehending the two past tense forms (simple past and past progressive) that have been empirically studied before. Prior research suggests that simple past sentences (e.g., *walked*) focus on a completed action (Madden & Zwaan,

2003) and to the location of that completed action (Anderson, et al., 2008; Matlock, et al., 2007). On the other hand, past progressive sentences (e.g., *was walking*) focus on the ongoing nature of the verb and the location of that ongoing action (Anderson, et al., 2008, Matlock, et al., 2007). Here, we extended these findings to investigate the interaction of verbal aspect and visual scene information, by adapting the visual scene used in earlier work to display a shorter, non-curving path. This change allowed us to explore not only overall movement durations but also movement durations within the area of the visual scene corresponding to the path, the specific region of the ongoing action described by the sentence.

Method

Participants

Sixty-four undergraduates at Cornell University participated in the experiment for extra credit in psychology courses. All were right handed and native speakers of American English.

Materials

Twelve sentences were created by adapting the stimuli used in the offline studies of Matlock et al. (2007). As we hoped to elicit movements across the entire scene, from which we could extract differences in motor dynamics between the two conditions, a final clause that described an event at the destination was added, encouraging movement all the way to the destination. Two versions of each of the 12 experimental items were created, as shown below: (1a) simple past sentence, (1b) past progressive sentence.

1a) David walked to the university where he sat in class.

1b) David was walking to the university where he sat in class.

Sentences were recorded using Mac-based sound software. Each of the experimental items was spliced to produce both a past progressive and a simple past version,

ensuring that the prosody for both of the targets was otherwise identical. The experimental items were counterbalanced across two presentation lists. Each list contained six instances of each condition, so that all participants heard all 12 target sentences, but only one version of each target frame.

Corresponding visual scenes were created for each target sentence pair. Each target visual scene consisted of a diagonal path starting halfway up and on the extreme left side of the screen. The path slanted to the right, terminating in the middle top of the screen. A character was located to the right of the beginning of the path and under the destination, separated from the scene by a black box framing the destination and path. See Figure 5. The depicted items in the scene were created by hand or taken from clipart and edited in Adobe Photoshop. The only moveable item in the scene was the character, which subtended an average of 1.53 degrees of visual angle in width by 2.05 degrees in height. The destinations were an average of 11.22 degrees in width by 4.09 degrees in height, and the path itself occupied a square of 8.42 degrees in width by 6.11 degrees in height. The character was located 14.25 degrees from the destination. The stimuli were presented using Macromedia Director MX, and mouse movements were recorded at an average sampling rate of 40 Hz. The display resolution was set to 1024 x 768.

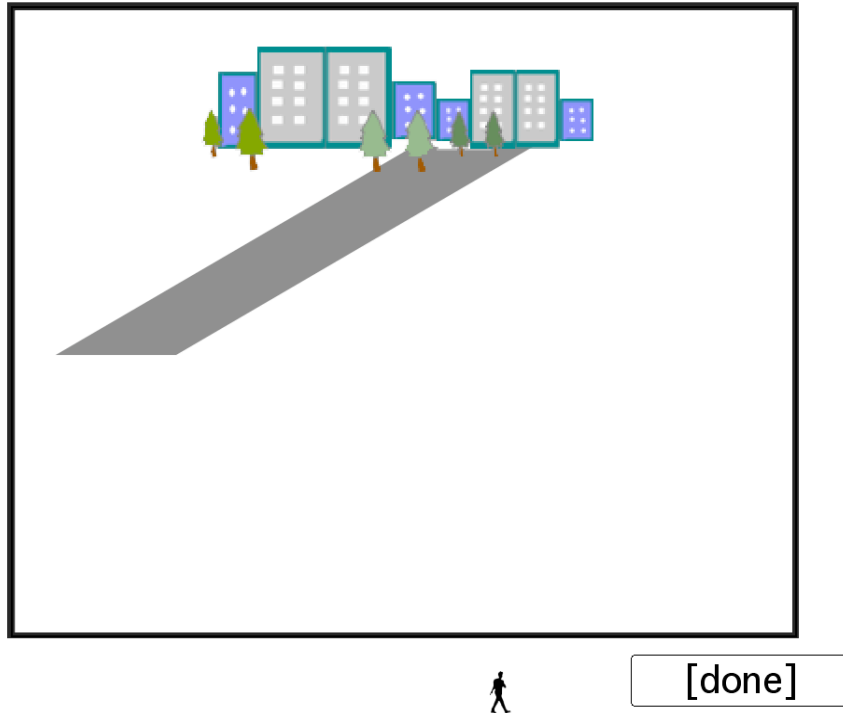


Figure 5. Example target visual scene accompanied by target sentences in Experiments 1, 2 and 3.

To keep participants from developing strategies specific to the experimental sentences, 12 filler items were created. Like the target items, the fillers contained either a past progressive or simple past verb. Unlike, the target items, the filler trials described no movement along the path (e.g., “Janet swam in the pool and then dried in the sun,”). These filler items were accompanied by 12 filler scenes, created using a short path beginning on the right side of the screen and slanting to the top, center of the screen. Other than path direction, each filler scene was similar to the target scenes.

Procedure

Participants were asked to make themselves comfortable in front of the computer and allowed to adjust the mouse and mouse-pad to a location that suited them. First, participants read instructions to place the character into the scene to make the scene match the sentence they heard. After indicating that they understood the task, participants were next presented with two practice trials (similar to the filler trials),

followed by the experimental task. At the onset of each trial, participants were presented with the entire visual scene. After a 500 ms preview, the sound file began. After the participant had moved the character (though not to any particular location), a “Done” button appeared in the bottom left corner of the screen. Participants clicked this button to move to the next trial. A blank screen with a button in the center labeled “Click here to go on” separated the trials. The entire experiment lasted approximately 20 minutes.

Results

Mouse movements were recorded during the grab-click, transferal, and drop-click of the character in the experimental trials. Prior to the analyses, the data were screened to remove extremely long trials (trials longer than 20 seconds). Using this criterion, no trials were removed. However, one subject was removed from the analyses due to computer malfunction.

Drop Locations

Previous offline results, with sentences that did not contain an extra clause encouraging participants to move the character all the way to the destination, revealed that participants chose a location closer to the middle of the path as the best representative of a sentence containing a past progressive verb, while selecting a location closer to the destination as the best representative of a sentence containing a simple past verb (Matlock, et al., 2007). Here, we plotted the drop point (location along the path where each participant let go of the mouse to “drop” the character) in both conditions. However, the current results do not demonstrate the trend found in earlier offline experiments. There was no effect of aspect when comparing the average drop location x-coordinates ($p > .6$) or y-coordinates ($p > .5$). Thus, due to the extra clause (e.g., “...where he sat in class.”), participants were moving the character across the screen to the destination about equally in the simple past and past progressive

conditions. This actually works well for our purposes because it means that, rather than dropping the character onto the path itself, participants carried the character across the region of the path, thereby providing motor dynamics for us to measure in the specific region of the screen where mentally simulated motion may be taking place.

Spatial Differences

In order to start investigating online aspectual differences, we first looked (as previous mouse-tracking studies have done) at spatial differences between the average trajectories elicited in response to simple past sentences and past progressive sentences. To determine whether the two average trajectories significantly diverged from each other, we time-normalized the trajectories and conducted a series of t-tests at each of the 101 time-steps. These analyses were conducted separately on the x- and the y-coordinates at each of the 101 time-steps to compare spatial differences across participants and across conditions. To avoid the increased probability of a Type-1 error associated with multiple t-tests, and in keeping with Bootstrap simulations of such multiple t-tests on mouse trajectories (Dale, Kehoe, & Spivey, 2007), an observed divergence was not considered significant unless the coordinates between the simple past and past progressive sentence trajectories elicited p-values $< .05$ for at least eight consecutive time-steps.

There were no significant differences in either the x or the y coordinates for the 101-timesteps. That is, the movement trajectories in the past progressive condition did not curve or shift in any direction relative to those in the simple past condition. The spatial extent of the two averaged movement trajectories were essentially overlapping one another, suggesting that participants were not intentionally acting out an explicit adherence to the depicted path in one condition and avoiding the depicted path in the other condition.

Movement Durations

After observing no differences due to aspectual form in the spatial properties of the trajectories, we turned our analysis to the temporal properties of the trajectories. There was a significant difference between the two aspectual forms when comparing overall movement durations (i.e., the length of time from the initial grab of the character to the final drop of the character into the scene), $t(62) = 2.22, p > .05$. See Figure 6 (left panel). This difference was the result of significantly longer movement durations in response to sentences containing past progressive verbs ($M = 4749$ ms, $SD = 2548$) than to sentences containing simple past verbs ($M = 4214$ ms, $SD = 1971$).

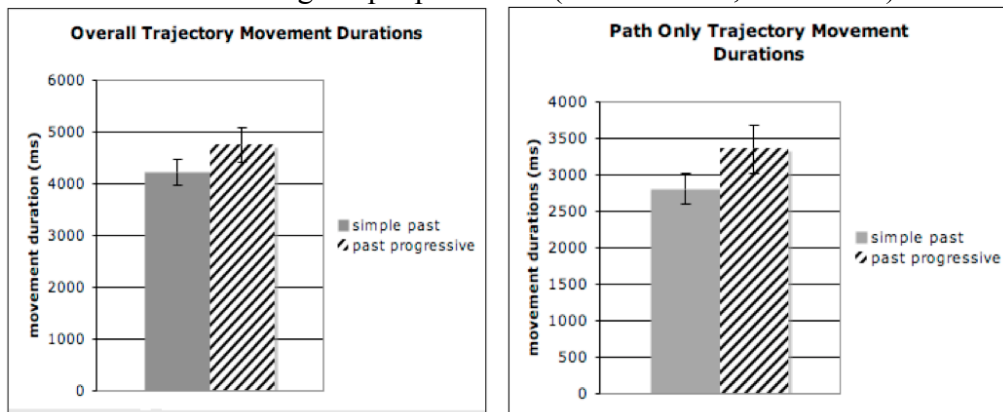


Figure 6. Movement duration differences for the overall trajectory, left panel, and in the region of the visual scene corresponding to the path, right panel, for Experiment 1.

In addition, the shorter path used in the visual scene (Figure 5) enabled us to investigate movement duration differences in the specific region of the screen where we hypothesize the participant is generating a perceptual simulation of motion (as in Richardson & Matlock, 2007). Looking at the movement durations in the region of the computer screen corresponding to the visually presented path, we found significantly different movement durations, $t(62) = 2.07, p < .05$. See Figure 6 (right panel). This difference is due to significantly longer movement durations in the region of the path in response to sentences containing past progressive verbs ($M = 3356$ ms, $SD = 2603$) than to sentences containing simple past verbs ($M = 2805$ ms, $SD = 1692$).

Discussion

Mouse-tracking enables the investigation of many indexes of response/motor dynamics. We looked at several of these indexes, including drop locations, time normalized spatial differences, and movement durations. Drop locations and time-normalized spatial differences did not demonstrate significant differences between processing simple past and past progressive sentences. However, movement duration differences showed that participants move the character across the scene for a longer period of time, both in the overall trajectory and in the trajectory corresponding to the specific region of the visually presented path, in response to past progressive sentences than to simple past sentences. These results complement and extend the results of earlier experiments investigating the processing of verbal aspect (Madden & Zwaan, 2003; Anderson, et al., 2008). Because simple past sentences focus attention on completed action, this makes the location of that completed action, namely the destination in the visual display, most salient. On the other hand, past progressive sentences encourage attention to the ongoing-ness of the action, making the region of that ongoing action (the path in the scene) most salient.

Experiment 2

The results of Experiment 1 allowed us to compare aspectual forms in the past tense descriptions alone, but they say nothing about processing aspect in present or future tense descriptions. For instance, does future tense afford similar motor dynamics to those of past tense? Or do future tenses not afford the same motor differences, since future tense refers to an event that has not yet occurred? While such questions have not yet been addressed empirically, the results of Experiment 1 suggest that mouse-tracking is a viable way to look at potential processing differences created by different verb tenses. Here, we adapted the materials and method of Experiment 1 to investigate how aspect in future tense descriptions influences motor output, and how

future tense might differ from past tense descriptions. To tease apart aspect and tense, we designed the experiment to have separate independent variables: tense (future or past) and aspect (perfective or imperfective).

Method

Participants

Sixty-four undergraduates participated in this experiment: 33 undergraduates from University of California, Merced and 31 undergraduates from Cornell University. University of California, Merced undergraduates participated for extra credit in Cognitive Science courses, and Cornell undergraduates participated for extra credit in Psychology courses. All were right handed and native speakers of American English.

Materials and Procedures

Sixteen sentences were created by adapting the stimuli of Experiment 1. The final clause, added to encourage movement all the way to the destination, was altered here to make the target sentence frames sound natural in both the past and future tenses. Four versions of each of the 16 experimental items were created, as shown in Table 2: 2a) past tense, perfective aspect (simple past), 2b) past tense, imperfective aspect (past progressive), 2c) future tense, perfective aspect (future simple), and 2d) future tense, imperfective aspect (future continuous).

Table 2: Target stimuli accompanying visual scene for Experiment 2.

	Perfective aspect	Imperfective aspect
Past tense	2a) David walked to the university for class.	2b) David was walking to the university for class.
Future tense	2c) David will walk to the university for class.	2d) David will be walking to the university for class.

Sentences were recorded using a Mac-based sound software. Each of the 16 experimental items was spliced to produce each of the four experimental conditions, ensuring that the prosody of the targets was otherwise identical. An additional 20 seconds of silence was added to the end of each target sentence, allowing us to time lock participants' mouse-movements to the raw time stamp of the sound files. The experimental items were counterbalanced across four presentation lists. Each list contained four instances of each condition, so a participant heard all the target sentence frames, but only one version of each.

Corresponding visual scenes were created for each target sentence frame. These visual scenes were identical in their specifications as the visual stimuli of Experiment 1. See Figure 4.

Similarly, 16 filler items were created to prevent participants from developing strategies specific to the experimental items. As in Experiment 1, they were of the same form as the target sentences, containing each of the four verbal aspects, but differed in that they described no motion along the path (i.e., “Janet will swim in the pool for exercise”). The filler items were accompanied by 16 filler scenes, created using a short path beginning on the right side of the screen and slanting to the top,

center of the screen. Besides the direction of the path, each filler scene was quite similar to the target scenes. The procedure was identical to that of Experiment 1.

Results

Mouse movements were recorded during the grab-click, transferal, and drop-click of the character in the experimental trials. Prior to the analyses, the data were screened to remove extremely long trials (trials longer than 20 seconds). Using this criterion, no trials were removed. However, one subject was removed from the analyses due to failure to comply with the instructions; rather than moving the character into the scene, this participant clicked the “done” button when it appeared on the screen for over half of the target trials.

Drop Locations

We began our investigation by looking at the outcome differences between the tenses (future and past) and aspect (perfective or imperfective). We plotted the drop location in each of the four conditions. See Figure 7. We analyzed the x- and y-coordinates for the drop locations separately.

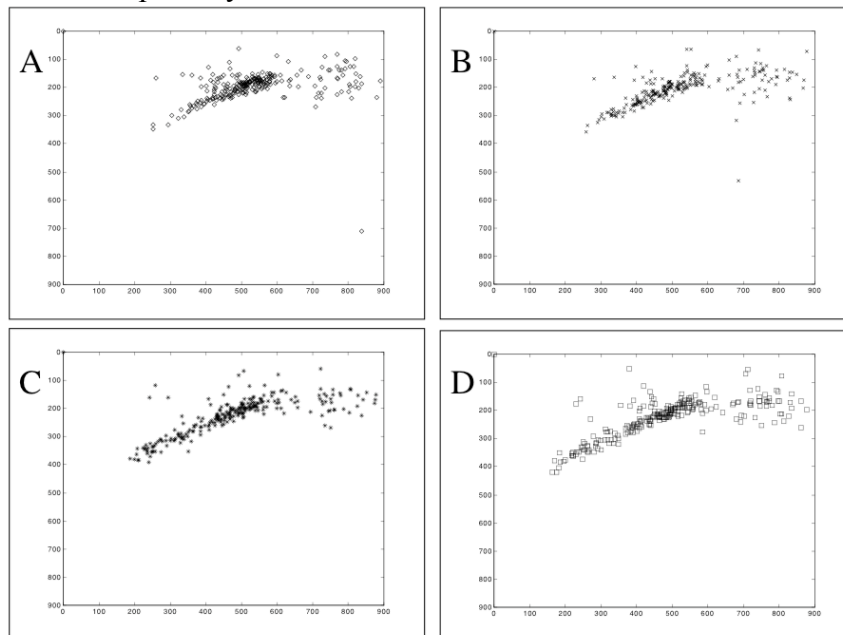


Figure 7. Drop locations in response to past perfective (a), past imperfective (b), future perfective sentences (c), and future imperfective sentences (d), for Experiment 2.

For the x-coordinates, the interaction between tense (future and past) and aspect (perfective and imperfective) was not significant, and there was no main effect of aspect, p 's $>.3$. However, there was a main effect of tense, $F(1,62) = 15.95$, $p <.05$. The x-coordinates of the drop location in response to past tense sentences (both perfective and imperfective), $M = 532$ x-pixel, $SD = 81$, were further to the right (closer to the destination) than the drop locations in response to future sentences, $M = 485$ x-pixel, $SD = 112$.

A similar picture emerges for the y-coordinates. There was no interaction between verb tense and aspect, and there was no significant effect of aspect, p 's $>.1$. However, there is a main effect of tense, $F(1,62) = 18.747$, $p <.05$. The y-coordinates of the drop location in response to past tense sentences, $M = 202$ y-pixel, $SD = 226$, was significantly closer to the destination (higher on the screen) than those produced in response to future tense sentences, $M = 226$ y-pixel, $SD = 55$. These results are commensurate with the drop location analyses of Experiment 1: drop location differences were not significant when comparing past progressive (imperfective) and simple past (perfective) verbs. However, future tense sentences elicit drop locations further away from the landmark, indicating greater uncertainty about this event that has not yet occurred.

Time-Normalized Spatial Differences

Figure 8 shows the average time-normalized trajectories in each of the four conditions. Visual inspection shows that the past tense conditions are closer to the destination than the average trajectories of the future tense conditions. Similar to the drop location differences, this suggests that the future tense may increase the uncertainty about an event that has not yet happened, resulting in average trajectories further from the destination. Similarly, the two average trajectories produced in response to imperfective sentences (past progressive and future continuous) are further from the

destination than the perfective sentences (simple past and future). This supports the hypotheses that imperfective sentences direct attention to the location of the ongoing action.

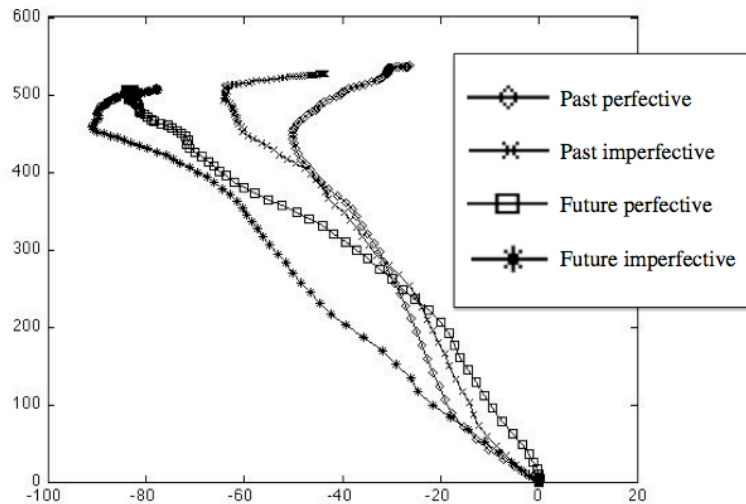


Figure 8. Average time-normalized trajectories, for Experiment 2.

To statistically assess spatial differences, we time-normalized the trajectories. A series of t-tests was conducted on each of 101 time-steps to compare spatial differences across participants and across conditions, as in Experiment 1. In keeping with Bootstrap simulations of such multiple t-tests on mouse trajectories (Dale et al., 2007), an observed divergence was not considered significant unless the divergence elicited p-values $< .05$ for at least eight consecutive time-steps. The x- and the y-coordinates were tested separately.

For the x-coordinates, the interaction and the main effect of aspect were not significant. However, there was a main effect of tense, with a significant divergence of the future tense average time-normalized trajectories away from those of the past tense trajectories and closer to the path between timesteps 49 to 101. The pattern of results is similar when examining the y-coordinates. For the y-coordinates, the interaction of tense and aspect was not significant. However, the main effect of tense

was significant. The future tense trajectories significantly diverged away from those of the past tense trajectories and away from the destination from timesteps 46 to 101. Also, in the y-coordinates there was a main effect of verb from timesteps 4 to 17, and again from timesteps 63 to 78. These y-coordinate differences were a result of the imperfective sentence trajectories diverging from the perfective sentence trajectories and tending towards the lower portion of the screen, or away from the destination.

To examine these spatial differences and preserve a time course reference at the same time, we returned to the raw (non-normalized) timestamps, corresponding to specific elements of the sound files, in the online trajectories. We used these raw timestamps to examine the average x and y coordinates at each of eight time bins. The first time bin corresponded to the time between the offset of the verb and 250 ms after this onset. Each of the following seven time bins corresponded to consecutive incremental 250 ms timebins. Correcting for multiple comparisons, we found no significant effects in either the x- or the y-coordinates, p 's > .05.

Movement Durations

As in Experiment 1, we concluded our investigation of online processing by looking at the temporal dynamics of the movement of the character into the visual scene. In looking at the movement duration for the overall trajectory, there was no significant interaction of tense and aspect, $p > .7$. Similarly, there was no main effect of tense, $p > .5$. However, the main effect of aspect (perfective vs. imperfective) was significant for the overall trajectory, $F(1,62) = 8.129$, $p < .05$. See Figure 9 (left panel). Sentences containing perfective verbs, i.e., *ran* or *will run*, elicited shorter movement durations ($M = 2598$ ms, $SD = 1263$) than sentences containing imperfective verbs, i.e., *was running* or *will be running* ($M = 3025$ ms, $SD = 1832$).

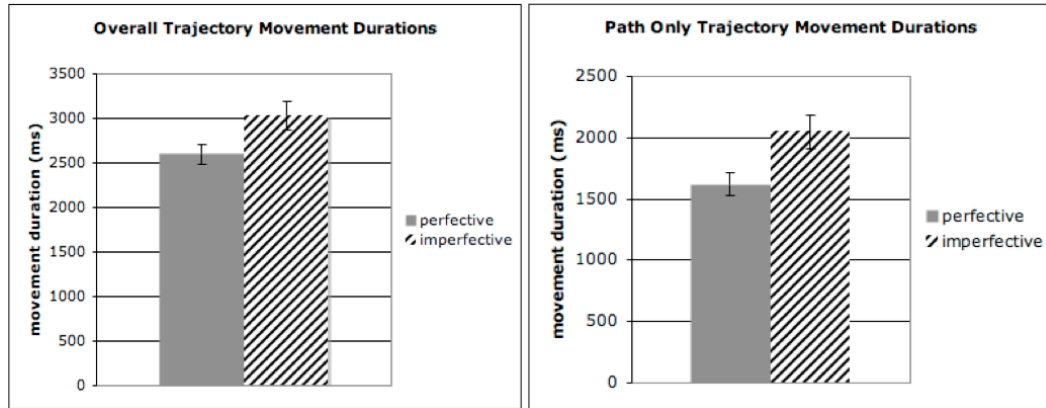


Figure 9. Movement duration differences for the overall trajectory, left panel, and in the region of the visual scene corresponding to the path, right panel, for Experiment 2.

The pattern of results was similar when looking at movement durations in the specific region of the screen corresponding to the visually presented path. The interaction of tense and aspect was not significant, $p > .9$. Also, the main effect of tense was not significant, $p > .9$. However, there is a significant main effect of aspect for the movement duration differences corresponding to the visually presented path, $F(1,62) = 7.849, p < .05$. See Figure 9 (right panel). Movement durations were longer in response to sentences containing imperfective verbs, $M = 2050$ ms, $SD = 1513$, than to sentences containing perfective verbs, $M = 1620$ ms, $SD = 1001$.

Discussion

The results of Experiment 2 dovetail with the results of Experiment 1 and contribute further to our understanding of how aspect is processed. In Experiment 2, we examined not only the processing differences associated with comprehending past tense sentences but also those entailed by comprehending future tense sentences. First, drop location differences emerged as a function only of tense (past or future), with the drop locations of future tense sentences being further from the destination than those of the past tense sentences. This is compatible with the results of Experiment 1, where we found that aspect caused no differences in drop location. Future tense sentences elicit drop locations further from the landmark, suggesting increased uncertainty as to

the completed-ness of an action that has not yet occurred.

This increased attention to the location of a future action is also reflected in the analysis of spatial differences between different verbal aspects. In the x-coordinates, the main effect of tense for a large number of timesteps indicates that processing future tense sentences entails a greater amount of attention to the ongoing uncompleted action of the future-tense event, with participants adhering more closely to the path. In the y-coordinates, a main effect of aspect implies that sentences containing imperfective verbs (*was walking* or *will be walking*) direct attention to the location of the ongoing action, namely the path. Tense also influences these spatial differences; sentences containing future tense verbs are moved into the scene more closely to the path, as this will be the location of the future action. Hence, the influences of aspect and tense additively result in the different movement trajectories produced in response to each of these conditions.

Our movement duration differences also support the results of Experiment 1. A main effect of aspect in both the overall trajectory and in the region of the screen corresponding to the path implies that aspect drives these differential motor outputs. After reading a perfective sentence, movement durations are shorter than after reading an imperfective sentence, regardless of tense.

Experiment 3

In Experiment 3, we sought to expand on the results of Experiment 1 by investigating the way verbal aspect interacts with terrain descriptions. Previous research has shown that context descriptions interact with fictive, or implicit, motion verbs to produce differences in reaction times (Matlock, 2004) and eye-movement patterns (Richardson & Matlock, 2007). In one experiment, when participants first heard context sentences describing a terrain as difficult, inspection times and eye-movement scanning along the path were increased as opposed to when participants heard the terrain described as

easy (Richardson & Matlock, 2007). These results support earlier reaction time results, where participants read narratives describing a terrain, and then made decisions about whether a fictive motion sentence was related to the preceding context (Matlock, 2004). Participants were slower to respond to fictive motion sentences when they had read a context sentence describing difficult terrain than when they had read a context describing easy terrain. While these results demonstrate that context interacts with implicit motion verbs, the impact of such contextual descriptions have not been explored in relation to verbal aspect. Hence, Experiment 3 investigates how contextual information interacts with aspectual forms of explicit motion verbs.

Method

Participants

Sixty-four undergraduates at Cornell University participated in the experiment for extra credit in psychology courses. All were right handed and native speakers of American English.

Materials and Procedures

Twelve sentences were created by adapting the stimuli of Experiment 1. In addition, two linguistic contexts for each stimulus were created. Hence, four versions of each of the 12 experimental items were created, as shown in Table 3 below: (3a) rough path description, simple past sentence, (3b) rough path description, past progressive sentence, (3c) smooth path description, simple past sentence (3d), smooth path description, past progressive sentence.

Table 3: Target stimuli accompanying visual scene for Experiment 3.

	Simple past sentence	Past progressive sentence
Rough context description	3a) <i>The road to the university was rocky and bumpy.</i> / David <u>walked</u> to the university where he sat in class.	3b) <i>The road to the university was rocky and bumpy.</i> / David <u>was walking</u> to the university where he sat in class.
Smooth context description	3c) <i>The road to the university was level and clear.</i> / David <u>walked</u> to the university where he sat in class.	3d) <i>The road to the university was level and clear.</i> / David <u>was walking</u> to the university where he sat in class.

Sentences were recorded using Mac-based sound software. Each of the 12 experimental items was spliced to produce both a past progressive and a simple past version, ensuring that the prosody of the targets was otherwise identical. Similarly, the context description was spliced onto the beginning of each of these target sentences. A pause of one second separated the offset of each context sentence from the onset of the target sentence. The experimental items were counterbalanced across four presentation lists. Each list contained three instances of each condition, so that all participants heard all 12 target sentences, but only heard one version of each.

Corresponding visual scenes were created for each target sentence frame. These visual scenes were identical in their specifications as the visual stimuli of Experiment 1. See Figure 4.

Additionally, 12 filler items were created to prevent participants from developing strategies specific to the experimental items. The fillers were of the same form as the target sentences: each contained a context description and either a past progressive or simple past verb. These filler trials varied from the target trials such that the first sentence provided no information about the path (i.e., “The weather in the

valley was warm and humid”) and that the second sentence described no movement along the path (i.e., “Janet swam in the pool and then dried in the sun,”). These filler items were accompanied by 12 filler scenes, created using a short path beginning on the right side of the screen and slanting to the top, center of the screen. Besides the direction of the path, each filler scene was quite similar to the target scenes. The procedure was identical to that of Experiment 1.

Results

Mouse movements were recorded during the grab-click, transferal, and drop-click of the character in the experimental trials. Prior to the analyses, the data were screened to remove extremely long trials. Movement durations of 20 seconds or more were removed because they constituted an unusually long time for a mouse-movement. Using this criteria, only three trials (less than 0.4%) of trajectories, were excluded.

Drop Locations

In plotting the drop location in each of the four conditions, the current results demonstrate a similar trend to those observed in earlier offline experiments. See Figure 10. There was not a significant interaction of terrain description and verb aspect (p 's $> .5$). However, there was a main effect of verb aspect when comparing the average drop x-coordinate, $F(1,62) = 8.462, p < 0.05$, with the average drop x-coordinate being further left (closer to the path) when participants heard past progressive sentences ($M = 477$ x-pixel, $SD = 69$) than when they heard simple past sentences ($M = 495$ x-pixel, $SD = 62$). Similarly, there was a main effect of aspect when comparing the average drop y-coordinate, $F(1,62) = 6.048, p < 0.05$, with the average drop y-coordinate being lower (further from the destination) when participants heard past progressive sentences ($M = 219, SD = 37$) than when they heard simple past sentences ($M = 211, SD = 41$).

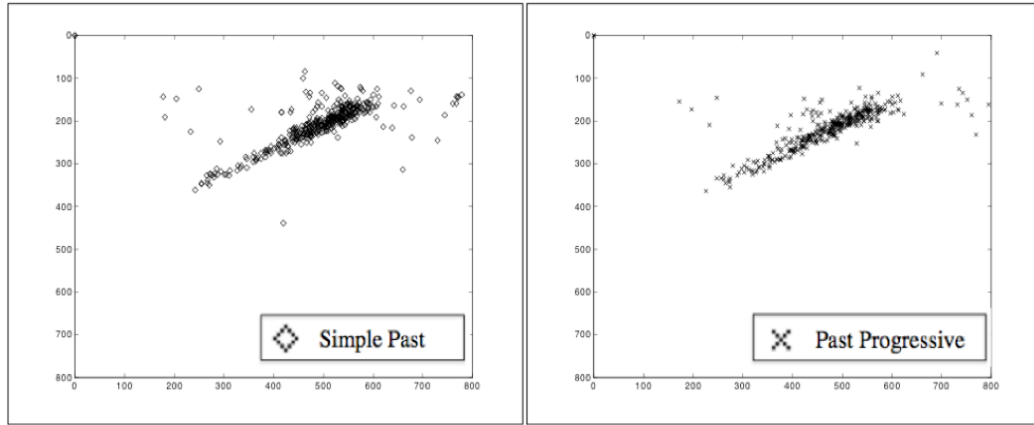


Figure 10. Drop locations in response to simple past sentences, left panel, and past progressive sentences, right panel, for Experiment 3.

Spatial Attraction

Figure 11 shows the average time-normalized trajectories in each of the four conditions. The mean simple past and past progressive trajectories at each of the 101 time-steps in the left panel of Figure 11 illustrate that in the rough terrain context, the average past progressive trajectory curved more toward the path than the average trajectory elicited by the simple past sentences, but only near the end of the trajectory. However, in the smooth terrain description, (Figure 11, right panel), there appears to be greater attraction toward the path across a greater portion of the trajectory for the past progressive sentences.

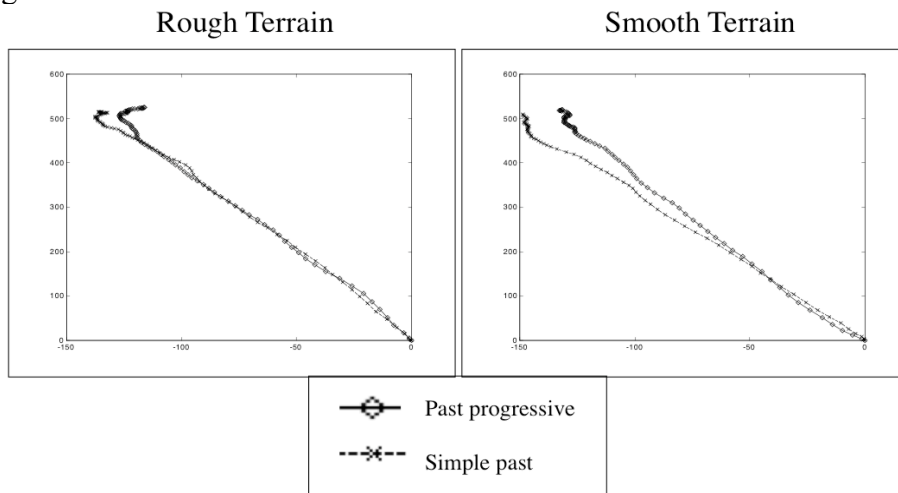


Figure 11. Average time-normalized simple past and past progressive sentence trajectories in rough and smooth terrain contexts, for Experiment 3.

To statistically assess spatial differences, the trajectories were time normalized, and a series of t-tests was conducted on each of 101 time-steps to compare spatial differences across participants and across conditions, as in Experiments 1 and 2. Again, an observed divergence was not considered significant unless the divergence elicited p -values $< .05$ for at least eight consecutive time-steps. The x- and the y-coordinates were tested separately.

In both the x-and y-coordinates, the interaction of context and aspect was not significant, all p 's $>.05$. In the rough context description condition, there was significant divergence of the past progressive x-coordinates away from the simple past x-coordinates and toward the path between time-steps 89 and 101, p 's $<.05$, and no significant divergence in the y-coordinates. This difference is similar to the observed differences in drop locations for past progressive and simple past sentences. Hence, this significant divergence so late in the time-normalized trajectories may simply be an artifact of drop locations.

In the smooth context description, there were significant divergences of the past progressive x-coordinates away from the simple past x-coordinates towards the path between time steps 48 and 60, p 's $<.05$, and again between time steps 65 and 89, p 's $< .05$. There was also significant divergence of the average past progressive y-coordinates away from the average simple past y-coordinates and towards the path between time steps 89 and 101. Again, this divergence late in the trajectory may be an artifact of the drop locations in each condition.

To examine the time-sensitive differences in spatial attraction to the visual scene's path across conditions, we looked separately at the average x- and y-coordinates within eight 250ms time-bins of the movement duration. As in Experiment 2, there was no significant interaction between aspect and terrain, p 's $>.1$, or main effect of either variable, p 's >2 .

Movement Durations

Finally, we concluded our investigation of the online processing of context and verbal aspect by looking at the temporal dynamics of the movement of the character into the visual scene. There was no significant interaction of context and aspect when comparing overall movement durations, $p's > .2$. However, there was a significant interaction of context and aspect on movement durations within in the specific region of the screen corresponding to the depicted path, $F(1, 63) = 4.6, p < .05$. See Figure 12.

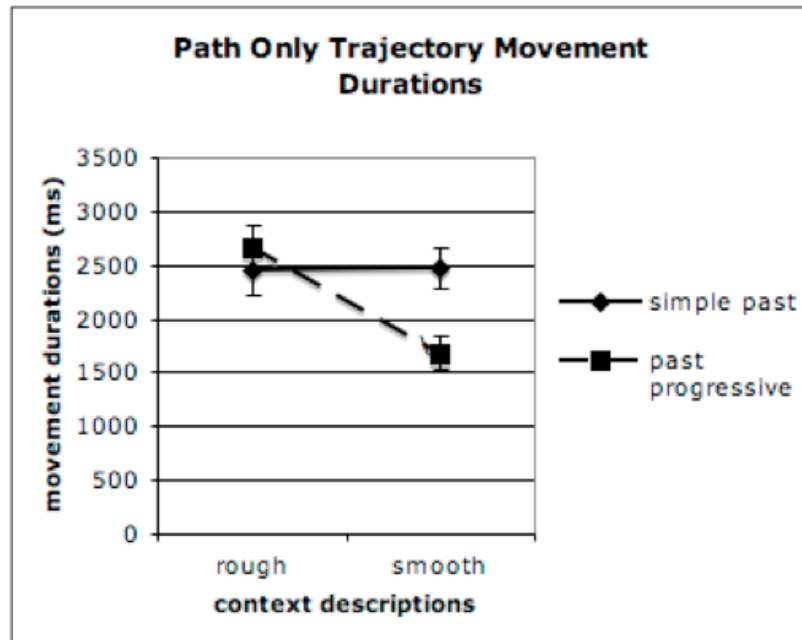


Figure 12. Movement duration differences in the region of the visual scene corresponding to the path, for Experiment 3.

In the region of the path, the average movement duration for simple past sentences was not significantly different when the context was first described as rough ($M= 2448$ ms, $SD=1849$) or smooth ($M=2479$ ms, $SD= 1527$). By contrast, with past progressive sentences, the average movement duration in the region of the path was slower when the context was first described as rough ($M = 2668$ ms, $SD=1680$) than when it was described as smooth ($M=2122$ ms, $SD=1240$). As seen in Richardson and Matlock (2007), when language directs attention to a mental simulation of motion in a

region of the screen (either by fictive motion verbs in their case, or by past progressive aspect in the present case), one sees robust effects of context stories describing how that region would affect that motion. Because simple past sentences focus attention on completed action, context descriptions of the traversal region do not significantly impact the movement dynamics. On the other hand, because past progressive sentences encourage attention to the ongoing-ness of the action in the specific region of that movement, context descriptions of that path do influence processing.

Discussion

In addition to corroborating previous work on grammatical aspect, these data also reveal new insights about processing through the examination of continuous motor output in response to aspectual and contextual differences. First, drop locations reliably differed between aspectual forms, with the past progressive condition eliciting drop locations closer to the path, and the simple past condition eliciting drop locations closer to the destination. This tendency to drop a character closer to the path in the past progressive condition and closer to the destination in the simple past condition, replicates previous evidence that the ongoing nature implied by a past progressive sentence draws attention to the middle portion of the path, whereas there is a tendency to focus attention on the destination in response to simple past sentences. However, we did not find this trend in Experiments 1 and 2. This may be because the contextual description of the path in Experiment 3 made the visually presented path more salient. By doing so, these context descriptions may constrain the locations for these drop points, thus decreasing variability and allowing us to detect statistically significant differences.

Similarly, contextual descriptions interacted significantly with verbal aspect in movement durations, but only in the region of the screen depicting the path. Contextual descriptions did not significantly modulate simple past movement

durations, because of the simple past's emphasis on the completed action. However past progressive movement durations were significantly faster when preceded by an easy terrain description than when preceded by a rough terrain description. That we do not find differences in the overall movement duration trajectories may again be due to the contextual description of the path making that region of the screen more salient.

Finally, while the coarse measure of raw-time spatial attraction to the path did not reveal statistically significant results, there was a significant spatial divergence of the past progressive trajectory away from the simple past trajectory and toward the destination in both contextual descriptions. Divergences late in the trajectory may be a result of differences in drop location, but divergences across the trajectories after the smooth context description provide further evidence for processing differences between these two aspectual forms. While these results are encouraging, they are not as convincing as the path-movement duration results (Figure 12). It is curious that the spatial attraction differences were detected in the smooth context description but not as robustly in the rough context description. Perhaps the visual stimuli used to depict the path simply did not appear to afford difficult or uneven travel, and the incongruence in the linguistic description and the visual appearance of the path hindered the emergence of full spatial differences in this context description. Future work will investigate this possibility.

General Discussion

The results of the three experiments reported here are in line with previous research using mouse-tracking (Anderson, et al., 2008), narrative reading (Madden & Zwaan, 2003), surveys (Fausey & Matlock, in press; Matlock, in press), and offline judgment tasks (Matlock, et al., 2007). Experiment 1 provides new evidence that different grammatical forms influence the processing of event descriptions, with the simple past (e.g., *walked*) focusing attention on the end of the path and the location of the

completed action, and past progressive (e.g., *was walking*) focusing attention to the “middle” of the event and the spatial region of that ongoing action. Experiment 2 provides evidence that other tenses additively influence the processing of event descriptions. These data show that the open-ended nature of future tenses, indicating action events that have not yet occurred, and imperfective aspect additively increase attention to the location of the verb’s action: the path. Experiment 3, in addition to corroborating previous work on grammatical aspect, reveals new insights about their processing by the examining continuous motor output in response to aspectual and contextual differences. These data show that context descriptions of the location of an ongoing action interact with verbal aspect to differentially influence processing of these sentences.

Despite the promise of these findings, more research is needed to fully understand the processing of verbal aspect. For example, these results do not speak to the processing of action verbs that do not explicitly code a destination. Also, our results do not shed light on the processing of perfect sentences (e.g., *John has closed the door*). However, recent empirical evidence suggests that perfect and imperfective sentences are processed differently, producing systematic differences in motor output (Bergen & Wheeler, 2010). Using the Action Compatibility Effect methodology (Glenberg & Kashchak, 2002), the authors demonstrate that imperfective sentences about hand motions facilitate and speed manual responses to those sentences, while perfect sentences do not. These data suggest that perfect sentences do not emphasize the ongoing action of the verb and are not processed in the same way as imperfective sentences are.

Another outstanding issue is the exact raw time course of processing different aspectual forms. As reported in our experiments here and from reaction time experiments (Madden & Zwaan, 2003), the exact time course of verbal aspect is

elusive because the continuous or progressive aspectual forms indicate an action that is ongoing. As such, perceptual simulations corresponding to the diffuse number of alternatives corresponding to this ongoing action may be difficult to capture in raw time. Future research may need to employ other methodologies, such as eye-tracking, to more thoroughly understand the time course of comprehending imperfective verbal aspect. For example, mouse-tracking is not as immediate as the eye-tracking, typically lagging a few hundred milliseconds behind an eye-movement. Therefore, eye-tracking provides more immediate information about intermediate stages of processing.

The current research has implications for several areas of research. First, grammatical aspect has been considered to provide minimal or secondary semantic information. Critically, however, our results indicate that aspect can significantly influence on-line processing. Second, this work also investigates grammatical aspect using a novel approach, allowing for the examination of more fine-grained temporal information, which complements the existing reaction time data. Finally, our results provide evidence to support cognitive linguists' claims regarding meaning as a conceptualization of linguistic descriptions, and the idea that aspect, like many domains of language, involves dynamic conceptualization (Langacker, 1987; Talmy, 2000).

More broadly, this work resonates with embodied cognition work on perceptual simulation and language understanding (Barsalou, 1999). It dovetails with the methodological advances of Balota and Abrams (1995) by providing new evidence from the temporal dynamics of a response after the response has been initiated, and by demonstrating that the motor system is not a robot-like automaton triggered by completed cognitive processes. Rather, motor processes are co-extensive with cognitive processes during perceptual/cognitive tasks (e.g., Balota & Abrams, 1995; Gold & Shadlen, 2000; Spivey et al., 2005). This work also comports with our

understanding of how mental models and visual information are coordinated in motor output. Similar to the way understanding spatial events is created and observed through tracking eye movements (Richardson & Matlock, 2007; Spivey & Geng, 2001), this work demonstrates that event understanding takes place differently as a function of changes in verbal aspect. Finally, this work explores a new way that language may influence thought in real time. Specifically, our results suggest that differences in underlying perceptual simulations, resulting from these differences in the dynamics of a motor response, may account for observed processing differences in comprehending sentences that employ different verbal aspects.

CHAPTER 5

CONCLUSIONS AND DISCUSSION

Taken together the previous chapters provide support for embodied cognition, and taken individually, each contributes to different specific areas of research. Chapter 2 contributes to language-action interactions by summarizing empirical data on such interactions and by placing this review of more recent research within a historical context of data. Also, this chapter highlights both the bi-directional relationship between language and action, as well as pointing out some of the outstanding questions that need to be addressed. Chapter 3 contributes to children's sentence processing research by examining individual differences within a single age group. These results suggest that one of the contributing factors that facilitates the transition from child-like to more adult-like use of visual context is linguistic experience. Chapter 4 contributes new empirical data to our understanding of verbal aspect. These data show that different aspectual forms differentially influence motor output as a result of processing differences associated with the perfective and imperfective forms. Further, this work extends our understanding of these processing differences to the future tense and to their interaction with contextual information. While each of the preceding chapters makes contributions to individual areas of research, they also as a whole make contributions to our understanding of embodied language processing more generally.

The first such contribution to embodied cognition claims is methodological. Chapter 2 introduces the use of mouse-tracking as a continuous motor output index of underlying cognitive processing that lends itself readily to investigating embodied cognition claims. Chapter 3 uses this method with a new population of participants, specifically young children. The results reported in Chapter 3 suggest that mouse-

tracking can reliably be used to investigate the underlying cognitive processing of five-year-olds. Chapter 4 again extends the methodology beyond the two- and three-forced choice visual scenes used in previous research. Similarly, these data provide a new way of investigating verbal aspect, which has proven difficult to examine (Madden & Zwaan, 2003). Although the diffuse number of perceptual simulations corresponding to past progressive sentences have been difficult to detect with reaction time tasks, mouse-tracking easily allows researchers to examine their processing and how this differs from processing simple past sentences. Using mouse-tracking to investigate verbal aspect opens the door to new ways of using mouse-tracking to directly index motor output differences in sentence comprehension.

The second contribution this dissertation makes is to the theory of embodied cognition, especially in Chapters 3 and 4. Chapter 3 provides evidence that while children are not adult-like in their use of visual scene information to disambiguate a garden-path sentence, they are adult-like in another way. Specifically, children show the same graded spatial attraction to the incorrect destination, indicating a garden-path effect. The cognitive task of sentence processing leaks immediately into the motor output produced by both children and adults. This implies that even though children are not like adults in their use of visual scene information, they are very much like adults in the way language immediately influences continuous motor output. Similarly, Chapter 4 provides evidence from an area of language processing that previously has been difficult to explore. Capturing processing differences created by comprehending sentences containing different aspectual forms has been elusive. However, these data suggest that not only do sentence- and word-comprehension automatically bleed into motor output, but grammatical aspect does as well. Hence, even at the level subtle temporal information, language comprehension immediately influences action.

Despite the contributions of this research, there are still several questions that remain. First, there are questions regarding the methodology. Many researchers interpret mouse-tracking data as an indicator of graded underlying processing. However, is it possible that the cognitive system produces discrete outputs, which the motor system then averages to create the continuous trajectories? Second, there are questions regarding the nature of the interaction between language and action. While the new empirical data reported in Chapters 3 and 4 provide evidence for language influencing action, they do not speak to whether this interaction is bidirectional. Chapter 2 suggests that without evidence for such bi-directional influences, it may be possible that cognitive modules are doing the “heavy-lifting” of cognition, and that their activity subsequently leaks into the motor areas of the brain. This description suggests that motor and language interactions are trivial. In the remainder of this chapter, I will address these two questions individually.

Methodological concerns

Researchers using mouse-tracking assume that the graded spatial differences in the mouse-movement trajectories index underlying continuous cognitive processing. The graded spatial attraction of these hand movements provides evidence both of the continuous uptake and integration of visual and linguistic information and of the dynamic competition between partially active alternatives made salient by this integration (Spivey, Grosjean, & Knoblich, 2005). The continuous mouse-tracking results of word identification and sentence processing tasks are taken as support for continuous parallel processing in cognition, in a similar way to how eye-tracking results provide such support (Magnuson, 2005).

However, it has recently been suggested that the signature curvatures in these mouse-movement trajectories can actually be explained, in principle, by a model in which cognitive processing is discrete and serial (considering a single symbolic

representation at a time), and the motor output is produced by a continuous parallel processing system (van der Wel, Eder, Mitchel, Walsh, & Rosenbaum, 2009). In this model, two motor movements corresponding to a strategic upward movement and then a perceptual decision movement are asynchronously averaged to produce a smoothly curved motor output (Henis & Flash, 1995). This distinction between perceptual processing and action planning provides an existence proof in which motor output may be continuous, but the underlying cognitive decisions are serial. This model then creates problems for theories of embodied cognition that propose that cognition is dynamically coupled with action.

It seems unlikely though that one neural system (cognition) behaves in one way (i.e., using discrete representations in sequence) and then feeds into a second system (action) that behaves in a qualitatively different way (i.e., using continuous representations in parallel). In their reply to van der Wel et al. (2009), Spivey and colleagues used the same equations that van der Wel and colleagues (2009) used for their model, adding a mechanism of dynamic competition between the multiple simultaneous cognitive representations that drive those motor commands (Spivey, Dale, Grosjean, & Knoblich, 2010). As there is nothing uniquely serial about the equations used by Henis and Flash (1995), the results of Spivey et al.'s model provide evidence that the perceptual and motor decisions can both be made in a continuous, parallel fashion. For example, cognitive representations for two response options initiate motor commands for both potential reach locations (Cisek and Kalaska, 2005), and the averaging weights for those two motor commands start out equal. This instigates motor output that is initially aimed at the midpoint between the two potential reach locations. As one cognitive representation receives increasing perceptual support, its weight ramps up, while the weight for the other cognitive representation ramps down. These changing weights are used to produce a dynamically averaged

motor movement that smoothly curves in a manner identical to the human data. Hence, a dynamic and continuous cognitive task flows smoothly into a dynamic and continuous motor output.

Unanswered theoretical questions

Does action influence language processing similar to the way language influences action? Chapter 2 suggests that it may be possible to reconcile findings that demonstrate that language influences action by claiming that language is central. In such an explanation, the activation in language areas spreads to supplementary areas, such as motor areas, giving rise to results that look like language influencing action. In other words, language could be processed in an encapsulated fashion, and this information then spreads into motor output. This account suggests that the activity observed in motor areas during language processing is peripheral and redundant to the cognitive task (Mahon & Caramazza, 2008). The empirical results reported in Chapters 3 and 4 are vulnerable to this counterargument. While both children's sentence processing of garden-path sentences and adults' processing of verbal aspect influences motor output, the data reported here do not provide evidence for the bi-directional influence of action on language.

As described in Chapter 2, there is evidence that motor activity influences language processing. These data come from TMS data (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005) and from language production research, demonstrating that gestures influences language. For example, speakers continue to use gesture when their listener cannot see the gesture (Iverson & Goldin-Meadow, 1998, 2001), implying that gestures not only help the listener but also assist the speaker (Goldin-Meadow & Sandhofer, 1999). Similarly, when arbitrary movements accompany speech, fewer verbal disfluencies are observed when that movement matches the implied direction of the speech (Casasanto & Lázaro, 2007), which suggests that

action differentially influences language production.

There is also evidence that gestures impact children's cognitive processing. For example, gesture seems to aid children's mathematical problem solving (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). More recent evidence shows that gestures also reduce cognitive load during the explanation of conservation tasks, even when referring to objects that are not present (Ping & Goldin-Meadow, 2010). This evidence suggests that action, specifically gestures, facilitates children's cognitive processing by decreasing the cognitive load of such tasks. Hence, taken with the research presented in Chapter 3, the emerging consensus is that the interaction between cognition and action is bi-directional, even in children.

Similarly, the data presented in Chapter 4 provide evidence that processing sentences containing different verbal aspects influences motor output. However, these data do not provide any evidence of action influencing the processing of different verbal aspects. As noted earlier, verbal aspect has been difficult to investigate in previous research, because imperfective sentences (i.e., *David was going to the store*) imply an on-going action of the verb. As such, the underlying perceptual simulations can correspond to many different stages of incomplete or ongoing action. Because of this, investigations of verbal aspect have been sparse, and at this time, there is no empirical data regarding the influence of action on comprehending these types of sentences.

In order to address this, future work is planned to investigate whether motor movements influence the comprehension of verbal aspect. Specifically, this work will perturb the motor output and investigate the impact of these perturbations on the comprehension of sentences like those used in Experiment 3 of Chapter 4 (i.e., *The road to the university was rocky and bumpy. David was running to the university for class*). The proposed experiment will again use both past progressive and simple past

target sentences, preceded by easy terrain descriptions (*The road to the university was level and clear*) or rough terrain descriptions (*The road to the university was rocky and bumpy*). Unlike Experiment 3, on half of the trials, the motor output will be perturbed as the participant moves the character into the scene. In order to gauge the impact of such motor perturbations, comprehension questions about the sentences will follow each trial. If motor movements influence the processing of these sentences, then disrupting the motor output should disrupt language comprehension more when the context sentence describes smooth travel than when it describes bumpy travel. Also, these perturbations should be least disruptive in the past progressive, difficult terrain description. If perceptual simulations of rough, ongoing actions are created in response to this condition, as is suggested by the results obtained in Experiment 3, then bumpy motor movements (disrupting the motor output) should be less disruptive to comprehending these sentences. Such differential results would suggest that motor output influences processing of sentences containing different aspectual forms and would support the hypothesis that language and motor interactions truly are bi-directional.

Even once these questions are addressed, many questions about embodied cognition will remain. First, as mentioned in Chapter 2, embodied cognition cannot easily account for the way participants seem to embody depictions of their names on a computer screen (Markman & Brendl, 2005). Second, embodied cognition currently cannot explain negation without appealing to logical symbols and rules (Kaup, Yaxley, Madden, Zwaan, & Lüdtkke, 2006) although strides are being made to account for the processing of negated sentences using sensorimotor primitives (Anderson, Huette, Matlock, & Spivey, 2010; Kaup, Lüdtkke, & Zwaan, 2006). Third, opponents of embodied cognition suggest that conceptual processing of abstract ideas, like justice, cannot be explained sufficiently by strong embodied theories (Mahon &

Caramazza, 2008). As such, the research presented in this dissertation makes advances to embodied cognition while suggesting the need for future research in order to more concretely flesh out the predictions of embodied cognition.

REFERENCES

- Abrams, R.A. & Balota, D.A. (1991). Mental chronometry: Beyond reaction time. *Psychological Science, 2*, 153-157.
- Ahmed, A. & Ruffman, T. (1998). Why do infants make A not B errors in a search task, yet show memory for location of hidden objects in nonsearch task? *Developmental Psychology, 34*, 441-453.
- Alloppenna, P. D, Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition: evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419-439.
- Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191-238.
- Anderson, S.E., Farmer, T.A., Goldstein, M., Schwade, J., & Spivey, M. (to appear). Individual differences in measures of linguistic experience account for variability in the sentence processing skill of five-year-olds. In I. Arnon & E.V. Clark (Eds.), *How children make linguistic generalizations: Experience and variation in first language acquisition*. Trends in Language Acquisition Research: John Benjamins Publishing Company.
- Anderson, S. E., Huette, S., Matlock, T., & Spivey, M. J. (2010). On the temporal dynamics of negated perceptual simulations. In F. Parrill, M. Turner, & V. Tobin (Eds.), *Meaning, form, and body* (pp. 1-20). Stanford: CSLI Publications.
- Anderson, S. E. . Matlock, T., Fausey, C., & Spivey, M.J. (2008). On the path to understanding on-line processing of grammatical aspect. *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 143-148), Mahwah, NJ: Lawrence Erlbaum Associates.
- Barsalou, L. (1999). Language comprehension: Archival memory or preparation for

- situated action? *Discourse Processes*, 28, 61-80.
- Barsalou, L.W. (1999b). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Balota, A.D. & Abrams, R.A. (1995). Mental chronometry: Beyond onset latencies in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1289-1302.
- Barsalou, L.W. (1999a). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes*, 28, 61-80.
- Barsalou, L.W. (1999b). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Barsalou, L. W., Simmons, W., K., Barbey, A., K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7, 84-91.
- Barsalou, L.W., Solomon, K.O., & Wu, L.L. (1999). Perceptual simulation in conceptual tasks. In M.K. Hiraga, C. Sinha, & S. Wilcox (Eds.), *Cultural, typological, and psychological perspectives in cognitive linguistics: The proceedings of the 4th conference of the international cognitive linguistics association, Vol. 3*. Amsterdam: John Benjamins, 209-228.
- Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and Language*, 112, 150-158
- Bertenthal, B.I. (1996). Origins and early development of perception, action, and representation. *Annual Review of Psychology*, 47, 431-459.
- 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.

- Britt, M. A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language*, 33, 251-283.
- 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.
- Caciopoo, J.T., Priester, J.P., & Bernston, G.G. (1993). Rudimentary determination of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65, 5-17.
- Calvert, S., Rideout, V., Woolard, J., Barr, R., & Strouse, G. (2005). Age, ethnicity, and socioeconomic patterns in early computer use: A national survey. *American Behavioral Scientist*, 48, 590-607.
- Calvo-Merino, B., Grezes, D.E., Glaser, J., Passingham, R.E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, 16, 1905-1910.
- Casasanto, D. & Lozano (2007). Embodied language production: Evidence from gesture, speech, disfluency, and motor action. Paper presented at Embodied Sentence Processing Conference: Behavioral, Neuropsychological, and Computational Perspectives, Saerbruecken, Germany.
- Chambers, C., Tanenhaus, M., & Magnuson, J. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 687-696.
- Choi, Y. & Trueswell, J.C. (2010). Children's (in)ability to recover from garden-paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, 106, 41-61.

- Chen, S., & Bargh, J.A. (1999) Consequences of automatic evaluation: Immediate behavior predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25, 215-224.
- Cisek, P. & Kalaska, J.F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron*, 45, 801-814.
- Comrie, B. (1976) *Aspect*. Cambridge: Cambridge University Press.
- Crain, S & Meroni, S. (2002). Children's use of referential context. Presented at the Annual Boston University Conference on Child Language Development, 2002, Boston, MA.
- Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing exemplars. *Memory and Cognition*, 35, 15-28.
- Dietrich, E. & Markman, A. (2003). Discrete thoughts: Why cognition must use discrete representations. *Mind and Language*, 18, 95-119.
- Dijkstra, K., Kaschak, M.P., & Zwaan, R.A. (2006). Body posture facilitates retrieval of autobiographical memories. *Cognition*, 102, 139-149.
- Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58 (5, whole no. 270).
- Dunn, L. M. & Dunn, L.M. (1997) *PPVT-III: Peabody Picture Vocabulary Test*. Circle Pines, MN : American Guidance Service.
- Eberhard, K., Spivey-Knowlton, M., Sedivy, J., & Tanenhaus, M. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24, 409-436.
- Ellis, R., & Tucker, M. (2000). Micro-affordances: The potentiation of components of action by seen objects. *British Journal of Psychology*, 91, 451-471.
- Elman, J. & McClelland, J. (1988). Cognitive penetration of the mechanisms of

- perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Cognition*, 27, 143-165.
- Farmer, T.A., Anderson, S.E., & Spivey, M.J. (2007). Gradiency and visual context in syntactic garden-paths. *Journal of Memory & Language*, 57, 570-595.
- Fausey, C.M., & Matlock, T. (in press). Can grammar win elections? *Political Psychology*.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59, 1-173.
- Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Development Psychology*, 42, 98-116.
- Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge in semantic memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 182-196.
- Fischer, M.H. & Zwaan, R.A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, 61, 825-850.
- Fodor, J. (1975). *The language of thought*. New York: Crowell.
- Frawley, W. (1992). *Linguistic semantics*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Frazier, L. & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.
- Frazier, L. & Fodor, J. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291-325.
- Freeman, J.B., Ambady, N., Rule, N.O., & Johnson, K.L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person

- construal. *Journal of Experimental Psychology: General*, 137, 673-690.
- Gibson, J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.
- Glenberg, A. (1997). What memory is for. *Behavioral and Brain Sciences*, 20, 1-55.
- Glenberg, A.M. & Kashak, M.P. (2002). Grounding language in action. *Psychonomics Bulletin and Review*, 9, 558-565.
- Glucksberg, S. (1964). Functional fixedness: Problem solution as a function of observing responses. *Psychonomic Science*, 1, 117-118.
- Gold, J. & Shadlen, M. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, 404, 390-394.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S.D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, 12, 516-522.
- Goldin-Meadow, S. & Sandhofer, C.M. (1999). Gesture conveys substantive information about a child's thoughts to ordinary listeners. *Developmental Science*, 2, 67-74.
- Grant, E. & Spivey, M. (2003). Eye movements and problems solving: Guiding attention guides thought. *Psychological Science*, 14, 462-466.
- Greenwald, A. (1970). A choice reaction time test of ideomotor theory. *Journal of Experimental Psychology*, 86, 20-25.
- Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in the motor and premotor cortex. *Neuron*, 41, 301-307.
- Hauk, O. & Pulvermuller, F. (2004). Neurophysiological distinction of action words in fronto-central cortex. *Human Brain Mappings*, 21, 191-201.
- Hebb, D. (1968). Concerning imagery. *Psychological Review*, 75, 466-477.
- Henis, E.A., & Flash, T. (1995). Mechanisms underlying the generation of averaged modified trajectories. *Biological Cybernetics*, 72, 407-419.

- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368-1378.
- Houghton, F. & Tipper, S.P. (1996). Inhibitory mechanisms of neural and cognitive control: Application to selective attention and sequential action. *Brain and Cognition, 30*, 20-43.
- Huette, S. & McMurray, B. (2010). Continuous dynamics of color categorization. *Psychonomic Bulletin & Review, 17*, 348-354.
- Hurtado, N., Marchman, V.A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*, F31-F39.
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337-374.
- Iverson, J.M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature, 396*, 228.
- Iverson, J.M., & Goldin-Meadow, S. (2001). The resilience of gesture in talk: Gesture in blind speakers and listeners. *Developmental Science, 4*, 416-422.
- Kaden, S.E., Wapner, S., & Werner H. (1955). Studies in physiognomic perception: II. Effect of directional dynamics of pictured object and words on the position of the apparent horizon. *Journal of Psychology, 39*, 61-70.
- Kaup, B., Lüdtke, J., & Zwaan, R.A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics, 38*, 1033-1050.
- Kaup, B., Yaxley, R.H., Madden, C.J., Zwaan, R.A., & Lüdtke, J. (2006). Experiential simulations of negated text information. *The Quarterly Journal of Experimental Psychology, 60*, 976-990.

- Klatzky, R.L., Pellegrino, J., McCloskey, B.P., & Lederman, S.J. (1993). Cognitive representations of functional interactions with objects. *Memory and Cognition*, *21*, 294-303.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comments on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35-54.
- MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676-703.
- Madden, C.J. & Zwann, R.A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, *31*, 663-672.
- Magliano, J.P. & Schleich, M.C. (2000). Verb aspect and situation models. *Discourse Processes*, *29*, 83-112.
- Magnuson, J. S. (2005). Moving hand reveals dynamics of thought. *Proceedings of the National Academy of Sciences*, *102*, 9995 - 9996.
- 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.
- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9-F16.
- Markman, A., & Brendl, C. (2005). Constraining theories of embodied cognition. *Psychological Science*, *16*, 6-10.
- Marlsen-Wilson, M. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71-102.

- Massaro D. (1999). Speechreading: Illusion or window into pattern recognition. *Trends in Cognitive Sciences*, 3, 310-317.
- Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory and Cognition*, 32, 1389-1400.
- Matlock, T. (in press). The conceptual motivation of aspect. In G. Radden, P. Koch, and K. Panther (Eds.), *Motivation and the Lexicon: Cognitive, Communicative, Perceptual and Socio-Cultural Factors*. Amsterdam: John Benjamins.
- Matlock, T., Fausey, C., Cargill, S., & Spivey, M. (2007, November). On the path toward understanding the dynamics of aspect descriptions in motion events. Paper presented at 48th Annual Meeting of the Psychonomic Society, Long Beach, California.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science*, 19, 22-24.
- Meegan D. & Tipper, S.P. (1998). Reaching into cluttered visual environments: Spatial and temporal influences of distracting objects. *Quarterly Journal of Experimental Psychology*, 51A, 225-249.
- Morrow, D.G. (1985). Prominent characters and events organize narrative understanding. *Journal of Memory and Language*, 24, 304-319.
- Nazir, T.A., Boulenger, V., Roy, A., Silber, B., Jeannerod, M., & Paulignan, Y. (2007). Language-induced motor perturbations during the execution of a reaching movement. *Journal of Cognitive Neuroscience*, 18, 1607-1615.
- Neisser, U. (1967). *Cognitive psychology*. East Norwalk, CT: Appleton-Century-Crofts.
- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive*

- psychology*. San Francisco, California: W.H. Freeman.
- Newell, A., Shaw, J.C., & Simon, H.A. (1958). Elements of a theory of human problem solving. *Psychological Review*, *65*, 151-166.
- Novick, J.M., Trueswell, J.C., Thompson-Schill, S.L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience* *5*, 263-281.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International University Press.
- Ping, R., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science*, *34*, 602-619.
- Pulvermuller, F. (1999). Words in the brain's language, *Behavioral and Brain Sciences*, *22*, 253-279.
- Pulvermuller, F. (2001). Brain reflections of words and their meanings. *Trends in Cognitive Sciences*, *5*, 517-524.
- Pulvermuller, F., Hauk, O., Nikulin, V.V., & Ilmoniemi, R.J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*, 793-797.
- Pylyshyn, Z. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, Mass: MIT Press.
- Richardson, D. C. & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, *102*, 129-138.
- Riskind, J.H. (1983). Nonverbal expressions and the accessibility of life experience memories: A congruence hypothesis. *Social Cognition*, *2*, 62-86.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131-141.

- Sekuler, R., Sekuler, A., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385, 308.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.
- Smith, L.B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A not-B error. *Psychological Review*, 106, 235-260.
- Snedeker, J. & Trueswell, J. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49, 238-299.
- Spivey, M. & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235-241.
- Spivey, M.J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, 102, 10393-10398.
- Spivey, M. J., Dale, R., Knoblich, G., & Grosjean, M. (2010). Do curved reaching movements emerge from competing perceptions? *Journal of Experimental Psychology: Human Perception and Performance*, 36, 251-254.
- Spivey, M. J., & Tanenhaus, M. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1521-1543.
- Spivey, M., Tanenhaus, M., Eberhard, K., & Sedivy, J. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, 45, 447-481.
- Stevens, J., Fonlupt, P., Shiffrar, M., & Decety, J. (2000). New aspects of motion

- perception: Selective neural encoding of apparent human movements.
Neuroreport: An International Journal of Rapid Communication of Neuroscience Research, 11, 109-115.
- Talmy, L. (2000). *Toward a cognitive semantics*. Cambridge, MA: MIT Press.
- Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, 268, 1632-1634.
- Tanenhaus M. & Trueswell, J. (1995). Sentence comprehension. In J. Miller and P. Eimas (Eds.), *Handbook of cognition and perception*. New York: Academic Press.
- Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, 273-273-281.
- Thelen, E., Schoener, G., Scheier, C., & Smith, L. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24, 1-86.
- Thelen, E. & Smith, L.B. (Eds.). (1994). *A dynamical systems approach to development and cognition*. Bradford Books, MIT Press.
- Thomas, L., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin and Review*, 14, 663-668.
- Tipper, SP, Howard LA, & Houghton G. (2000). Behavioural consequences of selection from neural population codes. In S. Monsell and J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 233-245). Cambridge, MA: MIT.

- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89-134.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, 33, 285-318.
- Trueswell, J., Tanenhaus, M., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 528-553.
- Tucker, M. & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 830-846.
- van der Wel, R.P.R.D., Eder, J., Mitchel, A., Walsh, M., & Rosenbaum, D. (2009). Trajectories emerging from discrete versus continuous processing models in phonological competitor tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 588-594.
- Wojnowicz, M.T., Ferguson, M.J., Dale, R., & Spivey, M.J. (2009). The self-organization of explicit attitudes. *Psychological Science*, 20, 1428-1435.
- Zwaan, R.A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B.H. Ross (Ed.), *The psychology of learning and motivation, volume 44* (pp. 35-62). New York: Academic Press.
- Zwaan, R.A. & Taylor, L.J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1-11.
- Zwitserslood, P. (1989). The locus of the effects of sentential-semantic context in

spoken-word processing. *Cognition*, 32, 25-64.