University of Alberta

Distracting the imagination: Does visuospatial or auditory interference influence gesture and speech during narrative production?

by

Lisa Jean Smithson

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Psychology

Fall, 2011
Edmonton, Alberta

ABSTRACT

The relationship between imagery, iconic gesture production, and speech was assessed among 120 participants in a narrative task. Study 1 indicated no significant relationship between visuospatial working memory capacity and iconic gesture production, and demonstrated that when visual interference is eliminated, iconic gesture rate decreases. Study 2 provided evidence that as visual interference increases in difficulty, participants use iconic gestures to a greater extent. Study 3 provided suggestive evidence that as auditory interference increases in difficulty, participants use iconic gestures to a greater extent. With respect to speech production, strong associations between narrative length and iconic gesture production were demonstrated in every condition except for when visual perceptual interference was eliminated. These results were interpreted within a framework of embodied cognition wherein iconic gesture production and imagery interact bi-directionally to facilitate the activation of imagistic representations and speech production.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction and Literature Review

"Cognition is for action: Only organisms that can orchestrate action have nervous systems, and the basic function of a nervous system is to guide action" (Glenberg, 2008, p. 43).

Theory of Embodied Cognition

Proponents of the embodied cognition perspective assert that cognition is strongly grounded in our physical interactions within the world (Wilson, 2002). Furthermore, our sensory-motor systems are the product of the co-evolution of the mind with the body (Semin & Smith, 2008). As opposed to the cognitive science perspective, which often views perceptual and motor systems as being peripheral devices for receiving input and generating output, the embodied perspective views perceptual and motor systems as being central to cognitive processes (Wilson, 2002).

Our perceptual and motor systems are activated when we interact within our environment, however they are also activated when we merely think about a concept in the absence of tangible and related perceptual input (Barsalou, 1999). One domain that relies heavily upon the activation of perceptual and motor systems in the absence of relevant perceptual input is storytelling. Our interpersonal and intrapersonal experiences within the world are strongly influenced by our ability to produce and understand stories (Mar, 2004). According to Mar (2004, p.1414), storytelling is "a native element of our social interactions" which likely influences both memory and self-construction (Miller, 1995).

In a study by Mar (2004), the neurological underpinnings of narrative production and comprehension were examined. Findings indicated that throughout both story production and comprehension, areas of the brain associated with visual and spatial imagery were activated. Imagery is strongly associated with memory and perception (Hostetter & Alibali, 2008) and therefore, narrative discourse may be understood within a framework of embodied cognition.

Perceptual and motor systems seem to play an influential role in the development and processing of narratives. Memory, gesture, and language all seem to exert influence upon the structure and content of narratives and each of these influences has been investigated with respect to the embodied perspective. These relationships will be addressed throughout the following paragraphs and subsequently an argument for investigating the relation between these components of narrative production will be proposed.

Visuospatial Working Memory from an Embodied Perspective

Visuospatial working memory may be the memory storage system responsible for much of the cognitive processing that allows cognition to be embodied. Visual and spatial information that enters our peripheral perceptual systems is stored and manipulated using visuospatial working memory (Baddeley, 2003). Long-term knowledge is also manipulated and stored in visuospatial working memory and is very useful when we are performing mental imagery tasks (Baddeley & Andrade, 2000). The efficiency of our memory is contingent upon how easy it is for us to visualize a concept (Paivio, 1965). Just as thinking about a concept leads us to enact perceptual and motor simulations of properties associated with that concept, visuospatial working memory can create and manipulate representations even when the corresponding visuospatial perceptual experience is absent (Cattaneo, Fastame, Vecchi, & Cornoldi, 2006). These representations are accessed during narrative production.

Gestures from an Embodied Perspective

Our ability to produce narratives is also determined by our use of gestures and language (Colletta, Pellenq, & Guidetti, 2010). Gestures are hand and arm movements thought to be strongly associated with spatial cognition. Adults and older children use co-speech gestures within their narratives to represent events and characters' attitudes; to synchronize communication with their listener (Colletta *et al*., 2010); and to provide narrative structure (Demir, 2009).

Iconic gestures represent shapes, objects, and functions of a referent (McNeill, 1992). For example, holding the hands forwards with thumbs forming a cross and fingers waving up and down might symbolize a butterfly flapping. It has been suggested that iconic gestures are grounded in visuospatial working memory (Trafton, Trickett, Stitzlein, Saner, Schunn, Kirschenbaum, 2006). Evidence for this claim has emerged from studies assessing the effects of gesture restriction. Gesture restriction has been shown to lead to a decrease in imagery (Rimé, Shiaratura, Hupet, & Ghysselinckx, 1984), and an increase speech dysfluencies which are most pronounced when a speaker is communicating spatial information (Graham & Heywood, 1975; Rauscher, Krauss, & Chen, 1996). It has been argued that gestures maintain spatial imagery (Wesp, Hesse, Keutmann, & Wheaton, 2001) and that when people gesture they may rely on spatial and imagistic representations stored in working memory. When gestures are restricted, our ability to communicate information being held in visuospatial working memory may be limited. Gestures therefore seem to be particularly useful at representing spatial and motoric information (Alibali, 2005).

Iconic gestures are also thought to be the category of gesture most strongly linked to language meaning (Krauss, Chen & Gottesman, 2000). Various functions of iconic gestures have been proposed for the speaker and for the listener. Gestures may help the listener to understand the speakers' communicative message (Goldin-Meadow, 2003; Beattie & Shovelton, 2000) and may help the speaker by facilitating speech production (Krauss & Morsella, 2004), by lightening the load on speakers' working memory (Wagner, Nusbaum, Goldin-Meadow, 2004), or by facilitating the packaging of information that will be communicated (Alibali, Kita, & Young 2000).

Language from an Embodied Perspective

Over the course of development, gesture and speech become integrated semantically and temporally just before the onset of two-word speech (Butcher & Goldin-Meadow, 2000). If gesture production is grounded in embodied cognition, then we would expect language production to also be grounded in embodied cognition because of the strong coupling between gesture and speech in early development. Researchers have collected support for this position. For example, when reading a word that is associated with a motion, the area of the motor cortex that is connected with that action becomes activated (Pulvermuller, 2005). In addition, when appropriate body movements are imagined, metaphorical phrases can be understood more effectively (Wilson & Gibbs, 2007). Language contributes to the quality of narratives in various ways since narrative production requires the simultaneous use of grammar, vocabulary, episodic structure, cohesive devices, and propositional content (Ukrainetz, Justice, Kaderavek, Eisenberg, Gillam, & Harm, 2005).

Competing frameworks

Theories of embodied cognition assert that our cognitive representations are inextricably linked to our perceptual and motor modalities and these are paramount for how we interact within the world (Markman & Brendl, 2005). When relaying a narrative, we rely extensively upon visuospatial working memory for imagery. Gesture production and language production are strongly associated with imagery as well (Hostetter & Alibali, 2008). Although functional roles for visuospatial working memory, gesture production, and language during a narrative task have been addressed, it remains unclear how visuospatial working memory, gesture production and language interact during this task.

One theory has been proposed in order to explain how gestures, imagery, and language are related. This framework is called Gesture as Simulated Action (GSA) and it asserts that "gestures emerge from the perceptual and motor simulations that underlie embodied language and mental imagery" (p. 495). The simulation of actions requires the activation of brain areas that are involved in planning physical actions. The simulation of perceptions requires the activation of brain areas that are involved in perceiving physical objects. All of our simulations are the result of the activation of premotor action states and this activation has the potential to lead to manifest action by spreading to motor areas. When spreading activation occurs, a gesture is produced. According to the GSA theory, there should be a correspondence between action simulations and gesture production (Hostetter & Alibali, 2008). Therefore, according to the GSA theory, a positive correlation would be expected between visuospatial cognitive resources and gesture production. However evidence used to support this claim has been largely inconclusive.

According to Wagner *et al.* (2004), although it makes intuitive sense that gesture would be an ideal means to externalize information encoded visually and spatially, models which present visuospatial working memory as the foundation of iconic gesture production do not adequately address why people also produce gestures when the content of their speech is non-spatial. In an attempt to determine whether the representations linked to gesture production were more reliant upon visuospatial or verbal working memory systems, Wagner *et al.* used a dual-task paradigm wherein participants were asked either to remember a sequence of letters (taxing verbal working memory) or to remember a visual pattern (taxing visuospatial working memory) while they explained math problems. When there is interference in dual-task paradigms, it is assumed that the two tasks are dependent upon similar cognitive processes. Participants were told either to gesture or not to gesture throughout the dual-task. The findings indicated that those who gestured remembered more in both conditions (Wagner *et al.*, 2004). Therefore, although there seems to be a great deal of agreement regarding the relationship between gesture production and spatial cognition (Trafton *et al*., 2006), the assertion that visuospatial working memory resources are predictive of iconic gesture production remains tenuous.

The assumption that visuospatial working memory resources would be predictive of iconic gesture production is problematic since it does not take into account external influences that can alter available visuospatial working memory resources. One external factor that may influence visuospatial working memory resources is task difficulty. Therefore, rather than internal cognitive resources guiding gesture production, it may be the case that gesture production (or the lack thereof) may influence our ability to activate visuospatial cognitive resources, especially when task difficulty increases. The mechanisms underlying this possible influence are illustrated in Figure 1.

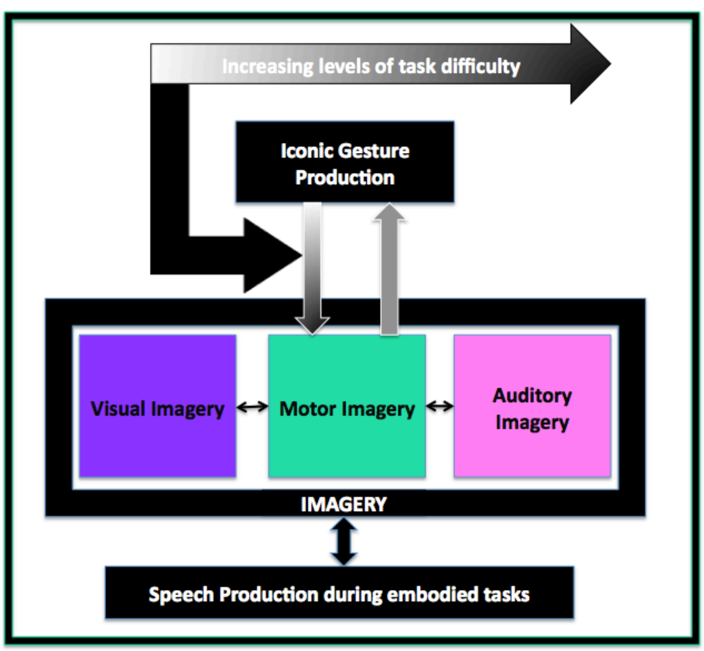

Figure 1: *The gestural facilitation of imagery theory (GFI)*

According to this theory, gestural production facilitates the activation of visuospatial cognitive resources and therefore, of imagistic representations. As such, the model is called the gestural facilitation of imagery theory (GFI). This theory asserts that individuals use gesture in order to activate imagery as a function of task difficulty. Imagery interference increases task difficulty resulting in an increase in iconic gesture production in order to facilitate the activation of imagistic representations. Speech production during embodied tasks is associated with imagery (Wilson & Gibbs, 2007). According to the GFI theory, speech production is indirectly associated with iconic gesture production. When task difficulty increases, and gesture production increases, the relation between language production and iconic gesture is strengthened.

In order to test the relative strengths of the GSA and GFI models, gesture or visuospatial memory must be manipulated systematically and both verbal and visuospatial working memory capacities must be taken into account. Previous studies have relied extensively on gesture restriction in order to elucidate the relation between gesture production and imagery. These studies have provided suggestive evidence that restricting gesture results in a reduction of imagery (Rimé *et al*., 1984). However, despite these suggestive findings, this method of manipulation may put an additional cognitive load upon participants as they must monitor their movements or be restricted unnaturally. Gesture restriction may also lead participants to compensate for this cognitive load by using compensatory body movement and this compensatory movement may distort our interpretation of the relation between gesture production and imagery.

It is possible to manipulate visuospatial working memory interference by manipulating the degree of competing visuospatial perceptual information. Since it is often the case that no perceptually relevant input is available during narrative production, this task requires the activation of brain areas associated with visual and spatial imagery (Mar, 2004). According to Ganis, Thompson and Kosslyn (2004), the brain areas activated during visual mental imagery and vision overlap significantly (up to 90%). The functional cerebral distance model (Kinsbourne, 2006, p.4) posits that "when adjacent, interconnected areas of the brain accommodate two independent activities; they will be advantaged in the performance of the dual task if the actions are congruent, and disadvantaged if they are incongruent". Visually present surroundings cannot be manipulated to parallel that which is created through imagery, however the elimination of perceptually present visual input precludes the possibility of incongruence and may offer the closest experimental approximation of congruence. Therefore, according to the functional cerebral distance model, a reduction in the level of visual input that is perceptually present should facilitate imagery.

Since it has been argued that language is embodied, it is important to take speech production into consideration when assessing the effects of visual perceptual manipulation. Narrative length is thought to be "a basic index of discourse competency" (Minami, 2008, p.89) and should be considered when investigating the association between gesture production, language, and memory. Manipulation of visual perceptual input in combination with the assessment of working memory capacity, gesture production, and language production may clarify the nature of the relationship between these factors during a narrative task.

CHAPTER II

Introduction to Study 1

The purpose of this study was two-fold. The primary purpose of this study was to investigate whether individuals with superior visuospatial skills would gesture more than individuals with poorer visuospatial skills, a relation that has been suggested by proponents of the GSA framework. In order to do this effectively, memory capacities were measured independently of, and then compared to gesture production in a task that remained constant among all participants. According to the GFI theory of gesture production, iconic gestures should be negatively associated with visuospatial working memory. Individuals who have low levels of visuospatial working memory resources were predicted to use iconic gestures to a greater extent than individuals with high levels of visuospatial working memory resources in order to activate the imagistic representations required for their narrative production.

The lengths of participant stories were assessed in relation to both gesture production and visuospatial working memory. A significant relationship between story length and iconic gesture production would suggest a strong coupling between these two modalities of communication during narrative production. A significant relationship with visuospatial working memory resources and story length would be suggestive of the embodiment of language during narrative production.

The secondary purpose of this study was to investigate how the manipulation of visual perceptual input might alter gesture production. In addition, it was of interest to determine whether any effects of visual interference or lack thereof were mediated by visuospatial working memory capacity. The GSA theory maintains that the simulation of actions and perceptions is contingent upon motor and visual imagery and that simulations result from the activation of premotor action states and these give rise to gestures. Gesture production should therefore increase with ease of mental imagery. More specifically if this theory is correct then when all competing visual perceptual information is blocked from view, individuals should gesture more since they should be better able to activate imagery than when their visual perceptual field contains competing perceptual information. According to the GFI theory, gesture production should increase when individuals struggle to activate imagistic representations. This would occur when visual perceptual information competes with visual imagery. Therefore, when competing visual perceptual information is blocked from view, individuals should gesture less since imagery should be more easily accessible.

The GSA and GFI theories do not make specific claims with respect to the relationship between narrative length and gesture production. However, according to the GFI theory, when task difficulty increases, gesture production increases and the relation between language production and iconic gesture rate becomes stronger. Therefore, a reduction in visual interference should reduce the association between iconic gesture production and narrative length.

A narrative production task was used in order to assess gesture and language production in this study and a working memory battery was used in order to test memory capacity. Participants were randomly assigned to two experimental conditions for the story-retelling component of the experiment: (1) Participants relayed their narratives to a listener sitting in front of them. This was labelled the NORMAL condition (2) Participants wore video glasses which completely obscured their vision as they relayed their narratives to a listener sitting in front of them. Nothing was projected onto the screen of the video glasses and as a result participants were only able to see darkness. This was labelled the "video glasses with no distractor" (VGND) condition. The purpose of this condition was to eliminate visual perceptual interference in order to facilitate imagery.

Method

Participants

Forty English undergraduate students from the University of Alberta were recruited to participate in this study. All participants were recruited from the Psychology Subject Pool at the University of Alberta. This research pool is comprised of undergraduate students enrolled in a first year psychology course. In order to participate in this study the students had to be monolingual. Students were considered to be monolingual even if they had studied a foreign language for a year or if they had non-fluent knowledge of another language.

Twenty participants were included in the NORMAL condition (6 males and 14 females). Participants in the NORMAL condition ranged in age from 17-24 years with a mean age of 19.05 years (SD=1.93). Twenty participants were included in the VGND condition (7 males and 13 females). Participants in the VGND condition ranged in age from 17-26 with a mean age of 19.05 years (SD=2.11).

In order to ensure the comparability of participants in the two experimental groups, t-tests were conducted upon all measures that were independent of the experimental manipulation. Participants in the experimental conditions did not differ significantly according to age ($t(38) = 0.000$, $p=1.000$), vocabulary ($t(38)=-1.705$, $p=0.096$), visuospatial short-term memory ($t(38)=-0.760$, $p=0.452$), visuospatial working memory ($t(38)=-1.019$, $p=0.315$), verbal short term memory ($t(38)= 0.392$, $p=0.697$) or verbal working memory ($t(38)=-0.525$, $p=0.603$).

Materials

Automated Working Memory Assessment

A 4-subtest working memory battery called the Automated Working Memory Assessment Short-Form (AWMA-S), was used to evaluate the adults' visuospatial short-term memory, visuospatial working memory, verbal short-term memory and verbal working memory. The AWMA is a standardized, computerized testing assessment. The assessment scoring is automated and the testing sequence is pre-set.

Visuospatial short-term memory was assessed using a dot matrix task. Participants were shown the position of a red dot in a series of 4x4 grids and had to recall the position by pointing to the squares on the computer screen that contained the red dot. The participants were asked to point to the squares in the same order that the dots appeared. The test becomes progressively difficult as the number of dots to be remembered increases.

Visuospatial working memory was assessed using a spatial span task. The participants were asked to view a screen with two shapes. The shape on the right side of the screen had a red dot on it. The participants were asked to identify whether the shape on the right was the same or opposite to the shape presented on the left side of the screen. The shape with the red dot could also be rotated. At the end of the trials, participants were asked to recall the location of each red dot on the shape in the exact sequence that it was presented by pointing to an image with three dots in the form of a triangle (which represented areas on the screen where the red dot had been located). Both of the shapes remained on the screen until the participants identified whether the shape on the right was in the same orientation as the shape on the left. In addition, the points in the form of a triangle remained on the screen as the adults pointed to the areas where the red dots had been located.

Verbal short-term memory was assessed using a digit recall task. Participants heard a sequence of digits and were asked to recall the digits in the correct order. The test becomes progressively more difficult as the digit span increases on subsequent trials.

Verbal working memory was assessed using a listening recall task. Participants heard a series of spoken sentences. They were first asked to identify the sentence as being true or false, and they were subsequently asked to recall the last word of each sentence in the correct sequence. The task increases in difficulty as more sentences are added.

Arrow keys on the computer keyboard were used by the observer in order to indicate the participants' response.

Vocabulary Test

The 'Peabody Picture Vocabulary Test – Third edition' (PPVT-III) was used to assess English language proficiency (Dunn & Dunn, 1997).

During the administration of the test, the examiner said a word and showed the participant a set of four pictures. Each of the four pictures was black and white and had a corresponding number. The test taker was asked to point to the picture that best represented the word's meaning or to inform the examiner of the number corresponding to the picture. The test was scored according to the experimenter's manual (Dunn & Dunn, 1997); that is scores were determined by subtracting the number of incorrect responses from the highest correct item number. The participants' raw scores were used in the analyses.

Pink Panther Cartoons

Two segments of Pink Panther cartoons (one entitled 'In the Pink of the Night', and the other 'Jet Pink'), were shown to the adults. The cartoons are approximately eight minutes in length in total. In the first video, the Pink Panther is being woken up by a cuckoo bird. The Pink Panther tries desperately to silence the cuckoo bird. Eventually the Pink Panther ends up becoming friends with the bird. In the second video, the Pink Panther decides that he wants to be a famous pilot. He gets into an airfield for military jet airplanes and proceeds to fly into the atmosphere and around a city until finally, he gets ejected from the plane.

Procedure

Participants were fully informed about the purpose of the study. They were aware that their speech and gestures were being observed.

The first component of the session was the PPVT-III. Following the vocabulary test, all participants watched two short cartoons and subsequently recalled the cartoons in narrative form while being videotaped. Subsequently, participants' memory was assessed using the Automated Working Memory Assessment, a standardized, computerized battery.

Speech and gesture production from the retelling were coded and gesture rate was calculated [(#of iconic gestures/# of words) x 100]. The gesture rate was multiplied by 100 for ease of interpretability.

Transcription and coding of speech

The videos were transcribed in orthographic words by a native English speaker. The number of word tokens produced was assessed using CHAT transcription software (MacWhinney, 2000).

Coding gestures

Gestures were coded according to the gesture classification system proposed by McNeill (1992). Four types of gesture were coded during the analysis of the videos: iconic, deictic, conventional, and beat. Iconic gestures were the only type reported for the purpose of this study since this is the type of gesture thought to be most strongly associated with language meaning (Krauss *et al*., 2000).

Results

NORMAL condition

In the NORMAL condition, the mean PPVT score was 180.00 (SD=5.96). With respect to the visuospatial memory measures, participants scored a mean of 30.00 (SD=6.10) on the visuospatial short-term memory assessment and 25.40

(SD=8.06) on the visuospatial working memory assessment. With respect to the verbal short-term memory measures, participants scored a mean of 37.55 (SD=5.61) on the verbal short-term memory assessment, and 17.15 (SD=3.57) on the verbal working memory assessment.

Simple correlations between iconic gesture production and memory scores revealed no significant associations (Table 1).

Table 1: *Simple correlations between iconic gesture production and memory scores in the NORMAL condition*

Relationship	Correlation
Iconic gesture production and visuospatial short-term memory	r=-0.120 (p=0.615)
Iconic gesture production and visuospatial working memory	r=-0.058 (p=0.807)
Iconic gesture production and verbal short-term memory	r=-0.008 (p=0.972)
Iconic gesture production and verbal working memory	r=0.267 (p=0.255)

According to Hostetter and Alibali (2007) both spatial skill and verbal skills should be taken into account when assessing the relation between gesture production and spatial cognition. In order to further investigate the relationship between visuospatial memory resources and iconic gesture production, partial correlations were conducted which statistically controlled for the influence of verbal memory resources. When statistically controlling for verbal short-term memory, the correlation between iconic gesture production and visuospatial short-term memory was not significant (r=-0.121, p=0.623). When statistically controlling for verbal working-memory, the correlation between iconic gesture production and visuospatial working-memory was also not significant (r=-0.304, p=0.205).

Simple correlations between narrative length and memory scores revealed no significant associations (Table 2).

Table 2: *Simple correlations between narrative length and the memory scores in the NORMAL condition*

Relationship	Correlation
Narrative length and visuospatial short-term memory	r=0.053 (p=0.825)
Narrative length and visuospatial working memory	r=0.108 (p=0.651)
Narrative length and verbal short-term memory	r=-0.297 (p=0.203)
Narrative length and verbal working memory	r=0.344 (p=0.138)

The relationship between iconic gesture production and narrative length was assessed. Iconic gesture rate was positively associated with narrative length in word tokens (r=0.451, p=0.046).

VGND condition

In the VGND condition, the mean PPVT score was 182.95 (SD=4.94). With respect to the visuospatial memory measures, participants scored an average of 31.25 (SD=4.12) on the visuospatial short-term memory assessment and 27.70 (SD=6.09) on the visuospatial working memory assessment. With respect to the verbal memory measures, participants scored an average of 36.90 (SD=4.84) on the verbal short-term memory assessment and 17.75 (SD=3.65) on the verbal working memory assessment.

Simple correlations between iconic gesture production and memory scores in the VGND condition revealed a significant negative association between iconic gesture production and visuospatial working memory (Table 3).

Table 3: *Simple correlations between iconic gesture production and the memory scores in the VGND condition*

Relationship	Correlation
Iconic gesture production and visuospatial short-term memory	r=-0.315 (p=0.176)
Iconic gesture production and visuospatial working memory	r=-0.530* (p=0.016)
Iconic gesture production and verbal short-term memory	r=-0.034 (p=0.887)
Iconic gesture production and verbal working memory	r=-0.148 (p=0.534)

When statistically controlling for verbal short-term memory, the correlation between iconic gesture production and visuospatial short-term memory was not significant (r=-0.343, p=0.151). When statistically controlling for verbal working-memory, the correlation between iconic gesture production and visuospatial working-memory remained significant (r=-0.515, p=0.024).

Simple correlations between narrative length and memory scores revealed no significant associations (Table 4).

Table 4: *Simple correlations between narrative length and the memory scores in the VGND condition*

Relationship	Correlation
Narrative length and visuospatial short-term memory	r=0.109 (p=0.646)
Narrative length and visuospatial working memory	r=0.042 (p=0.861)
Narrative length and verbal short-term memory	r=-0.100 (p=0.674)
Narrative length and verbal working memory	r=-0.096 (p=0.686)

The relationship between iconic gesture production and narrative length was assessed. Iconic gesture rate was not associated with narrative length in word tokens (r=0.183, p=0.439).

Comparing participants in the NORMAL condition to the VGND condition

In order to determine whether the complete reduction in visuospatial interference would affect gesture production, a comparison between participants in the NORMAL condition and the VGND condition was conducted (Figure 2). A two-tailed, independent-samples t-test revealed a significant difference in gesture rate (t(38) = 2.373, p=0.023). Participants in the NORMAL condition gestured more (M=4.47, SD=4.44, N=20) than those in the VGND condition (M=1.73, SD=2.64, N=20).

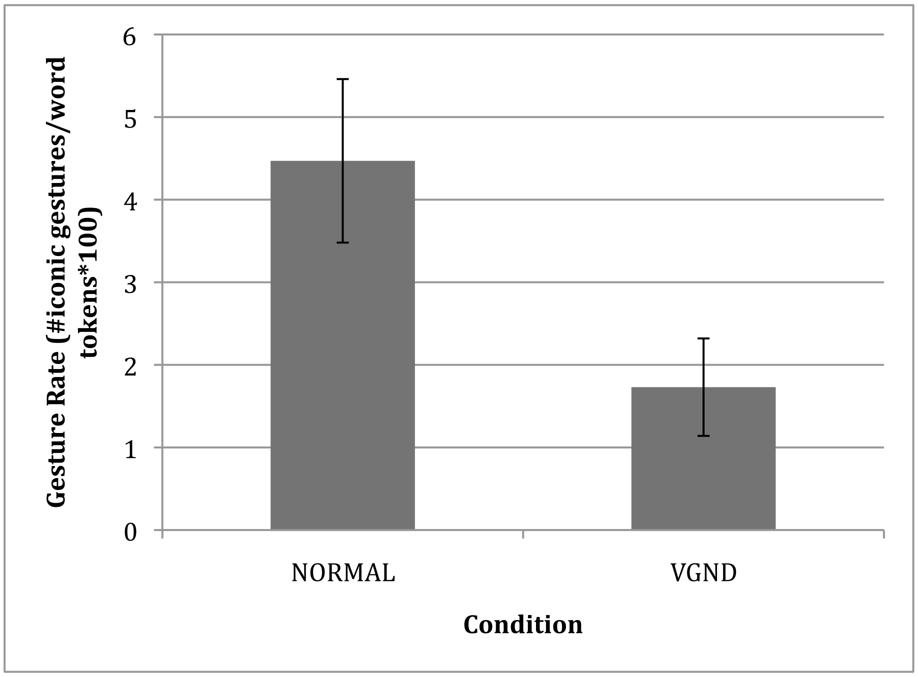

Figure 2: *Gesture rates among participants in the NORMAL and VGND conditions. Error bars show the standard error of the mean.*

Comparisons among these groups were further conducted in order to assess whether participants differed according the quality of their verbal narratives (Table 5). Results indicated that participants in these conditions did not differ with respect to the length of their stories ($t(38)=-0.662$, $p=0.512$).

Table 5: *Information concerning verbal and gestural measures of narrative production for participants in the NORMAL and VGND conditions*

Condition	Iconic Gesture Rate	Word Tokens
NORMAL (N=20)	4.47 (4.44)	466.80 (283.82)
VGND (N=20)	1.73 (2.64)	531.45 (331.77)

Discussion

The primary purpose of this study was to clarify the nature of the relationship between iconic gesture production and visuospatial working memory capacity. According to the GSA framework, individuals with high levels of visuospatial cognitive resources should gesture more than individuals with low visuospatial cognitive resources. If the ability to store and manipulate visuospatial information is predictive of iconic gesture production, individuals with superior visuospatial memory capacities would be expected to gesture more. However, this was not the case in either condition assessed. In fact, participants in the VGND condition who had greater visuospatial working memory resources were found to gesture less. It could be argued that in order for the positive association between iconic gesture production and visuospatial memory resources to become apparent, the ability to store and manipulate verbal information would have to be taken into account. For example, it has been argued that both our visuospatial and verbal skills together influence iconic gesture rates (Hostetter & Alibali, 2007). However, even when statistically controlling for verbal memory resources, there remained a negative correlation between iconic gesture production and visuospatial working memory in the VGND condition, and no significant relationships among iconic gesture production and visuospatial working memory capacity among participants in the NORMAL condition. This result may indicate that when the task difficulty is reduced through the reduction of competing visuospatial interference, then only individuals with poor visuospatial working memory resources depend upon gesture production to facilitate the production of their narrative utterances.

Therefore when task difficulty is sufficiently decreased, visuospatial working memory resources may be predictive of ionic gesture production. This evidence contrasts with the assumption of the GSA framework that individuals with stronger spatial skills should gesture more than those with weaker spatial skills (Hostetter & Alibali, 2008). Instead, this evidence lends support to the GFI theory suggesting a negative association between gesture production and visuospatial working memory measures.

In the NORMAL condition a positive association was found between iconic gesture production and narrative length suggesting a strong coupling between these two modalities of communication in this condition. According to the GFI framework, as task difficulty decreases the coupling between gesture production and language is attenuated, which may explain why participants in the VGND condition did not display a strong association between these measures. In both conditions, no relationship was found between narrative length and visuospatial working memory resources, suggesting that language production used during narrative production, although thought to be embodied (Wilson & Gibbs, 2007), may not be directly reliant visuospatial memory resources.

The secondary purpose of this study was to investigate the effects of visual perceptual manipulation upon iconic gesture production and speech production. It was hypothesized that if imagery leads to gesture production, then visual perceptual input should interfere with visuospatial working memory and that this should lead to a reduction iconic gesture rate. It was hypothesized that if gesture production facilitates imagery activation, then visual perceptual input should interfere with visuospatial working memory leading to an increase in iconic gesture production. In support of the GFI theory, individuals who wore the video glasses with no image projected onto the glasses gestured significantly less than individuals telling the narrative without them. Participants in these conditions did not differ with respect to the length of their narratives.

The findings from this study seem to echo those in a study by Wesp, Wheaton, and Wheatley (1996) wherein participants were asked to describe a route either with their eyes opened or closed. Participants gestured less when they had their eyes closed compared to when they were open. The findings from Wesp *et al.* (1996) and from the current study are suggestive of a relationship wherein gesture production increases as a result of visual interference rather than being inhibited by visuospatial interference and therefore lends support to the GFI theory.

However, it is important to interpret these findings with caution since gestures serve many communicative functions and as a result, listener visibility may have influenced gesture production in both cases. In a study by Alibali and Heath (2001), it was demonstrated that in a face-to-face condition, participants used more iconic gestures than when listeners could not be seen. Similar to the results from Alibali and Heath (2001), in the present study, the comparison between iconic gesture production in the face-to-face condition (NORMAL) and the blank video glasses condition (VGND) revealed a significant difference wherein participants gestured more in the face-to-face condition. Although the blank video glasses condition may have afforded greater use of imagery since no visual perceptual input was interfering with imagery, it may have also influenced gesture production since the listener was not visible in this experimental condition. The relative influences of perceptual facilitation and communicative functions cannot be disentangled in the current study and therefore this issue will not be addressed further here.

CHAPTER III

Introduction to Study 2

Since gestures serve many communicative functions and listener visibility may influence gesture production, the results of Study 1 cannot conclusively be attributed to either the effects of listener visibility or perceptual facilitation. The purpose of Study 2 was to determine how gesture production is affected when visual interference is manipulated while holding listener visibility constant.

If imagery leads to gesture production, as is claimed by the GSA framework, then visual perceptual input should interfere with visuospatial working memory and this should lead to a reduction in iconic gesture production. The presentation of a more difficult visuospatial distractor should result in lower gesture rates. In addition, individuals with stronger visuospatial working memory capacities should be less affected by visuospatial interference and therefore a positive relationship between visuospatial working memory and iconic gesture production would be expected.

If gesture production facilitates imagery activation, as is claimed by the GFI theory, then visual perceptual input should interfere with visuospatial working memory leading to an increase in iconic gesture production. Therefore, when a more difficult visuospatial distractor is presented, gesture rates should be higher as compared to when easier stimuli are presented. It was predicted that individuals with higher visuospatial working memory scores would use iconic gestures less in both conditions. In other words a negative correlation between iconic gesture and visuospatial working memory scores was expected.

A significant relationship between story length and iconic gesture production would suggest a strong coupling between these two modalities of communication during narrative production and a significant relationship with visuospatial working memory resources and story length would be suggestive of the embodiment of language during narrative production. In addition, in accordance with the GFI model, it was predicted that an increase in visuospatial distraction would strengthen the relationship between iconic gesture production and speech.

Methods were identical to Study 1 with the exception that participants were randomly assigned to two different experimental conditions for the story-retelling component of the experiment: (1) Participants wore video glasses with a “simple visual distractor” (SVD) projected onto the screen. The SVD distractor consisted of a red dot moving through a series of parallel lines (2) Participants wore video glasses with a “complex visual distractor” (CVD) projected onto the screen. The CVD distractor consisted of an array of shapes and colors which morphed kaleidoscopically.

Method

Participants

Forty participants were included in this study. Participants included 6 males and 14 females in the SVD condition ranging in age from 17 to 22 with an average of 18.70 years (SD=1.26). Participants included 7 males and 13 females in the CVD condition ranging in age from 17 to 22 with an average age of 19.15 years (SD=1.46).

In order to ensure the comparability of participants in the two experimental groups, t-tests were conducted upon all measures that were independent of the experimental manipulation. Participants in the experimental conditions did not differ significantly according to age ($t(38) = -1.043$, $p=0.304$), vocabulary ($t(38)=-0.749$, $p=0.458$), visuospatial short-term memory ($t(38)=-0.924$, $p=0.362$), visuospatial working memory ($t(38)=-1.807$, $p=0.079$) or verbal working memory ($t(38)=-1.520$, $p=0.137$). Participants in the experimental conditions did, however differ with respect to their verbal short-term memory ($t(38)=-3.421$, $p=0.002$).

Procedure

To ensure that our appraisals of stimuli difficulty were in fact correct, ten participants from each of the visual distractor conditions were asked to rate the difficulty of producing a narrative with the relevant stimuli present. Participants verbally rated the level of difficulty on a scale ranging from 1 to 10 (with 10 being very difficult). The average rating of difficulty in the SVD condition was 4.05 ($SD=2.36$) and the average rating of difficulty in the CVD condition was 6.65 ($SD=1.25$). A two-tailed, independent-samples t-test revealed that participants in the SVD condition found it less difficult to tell the story with the distractor present than participants in the CVD condition ($t(18)=-3.077$, $p=0.006$).

Results

SVD condition

In the SVD condition, the mean PPVT score was 179.65 ($SD=11.98$). With respect to the visuospatial memory measures, participants scored an average of 31.10 ($SD=4.01$) on the visuospatial short-term memory assessment and 25.35

(SD=5.86) on the visuospatial working memory assessment. With respect to the verbal memory measures, participants scored an average of 35.20 (SD=4.74) on the verbal short-term memory assessment and 16.60 (SD=4.03) on the verbal working memory assessment.

Simple correlations between iconic gesture production and memory scores in the SVD condition revealed no significant associations (Table 6).

Table 6: *Simple correlations between iconic gesture production and the memory scores in the SVD condition*

Relationship	Correlation
Iconic gesture production and visuospatial short-term memory	r=0.162 (p=0.495)
Iconic gesture production and visuospatial working memory	r=-0.052 (p=0.828)
Iconic gesture production and verbal short-term memory	r=-0.308 (p=0.187)
Iconic gesture production and verbal working memory	r=0.007 (p=0.975)

Simple correlations between narrative length and memory scores revealed no significant associations (Table 7).

Table 7: *Simple correlations between narrative length and the memory scores in the SVD condition*

Relationship	Correlation
Narrative length and visuospatial short-term memory	r=0.390 (p=0.089)
Narrative length and visuospatial working memory	r=0.175 (p=0.460)
Narrative length and verbal short-term memory	r=-0.059 (p=0.806)
Narrative length and verbal working memory	r=0.323 (p=0.165)

The relationship between iconic gesture production and narrative length was assessed. Among participants in the SVD condition, iconic gesture rate was positively associated with word tokens (r=0.469, p=0.037).

CVD condition

In the CVD condition, the mean PPVT score was 182.00 (SD=7.30). With respect to the visuospatial memory measures, participants scored an average of 32.20 (SD=3.50) on the visuospatial short-term memory assessment and 28.70 (SD=5.87) on the visuospatial working memory assessment. With respect to the verbal memory measures, participants scored an average of 40.50 (SD=5.05) on the verbal short-term memory assessment and 18.55 (SD=4.08) on the verbal working memory assessment.

Simple correlations between iconic gesture production and memory scores in the CVD condition revealed no significant associations (Table 8).

Table 8: *Simple correlations between iconic gesture production and the memory scores in the CVD condition*

Relationship	Correlation
Iconic gesture production and visuospatial short-term memory	r=0.284 (p=0.225)
Iconic gesture production and visuospatial working memory	r=-0.059 (p=0.804)
Iconic gesture production and verbal short-term memory	r=-0.072 (p=0.761)
Iconic gesture production and verbal working memory	r=-0.049 (p=0.837)

Simple correlations between narrative length and memory scores revealed no significant associations (Table 9).

Table 9: *Simple correlations between narrative length and the memory scores in the CVD condition*

Relationship	Correlation
Narrative length and visuospatial short-term memory	r=0.261 (p=0.267)
Narrative length and visuospatial working memory	r=0.040 (p=0.867)
Narrative length and verbal short-term memory	r=-0.180 (p=0.448)
Narrative length and verbal working memory	r=-0.025 (p=0.915)

The relationship between iconic gesture production and narrative length was assessed. Among participants in the CVD condition, iconic gesture rate was positively associated with narrative length in word tokens (r=0.451, p=0.046).

Comparing participants in the SVD condition to the CVD condition

It was of interest to determine whether the different visuospatial distractors had a different effect upon iconic gesture production. The SVD and CVD conditions were therefore compared (Figure 3). A two-tailed, independent-samples t-test showed a significant difference in gesture rate ($t(38) = -2.291$, $p=0.028$). Participants in the CVD condition gestured more ($M=2.52$, $SD=2.75$) than the participants in the SVD condition ($M=0.81$, $SD=1.88$).

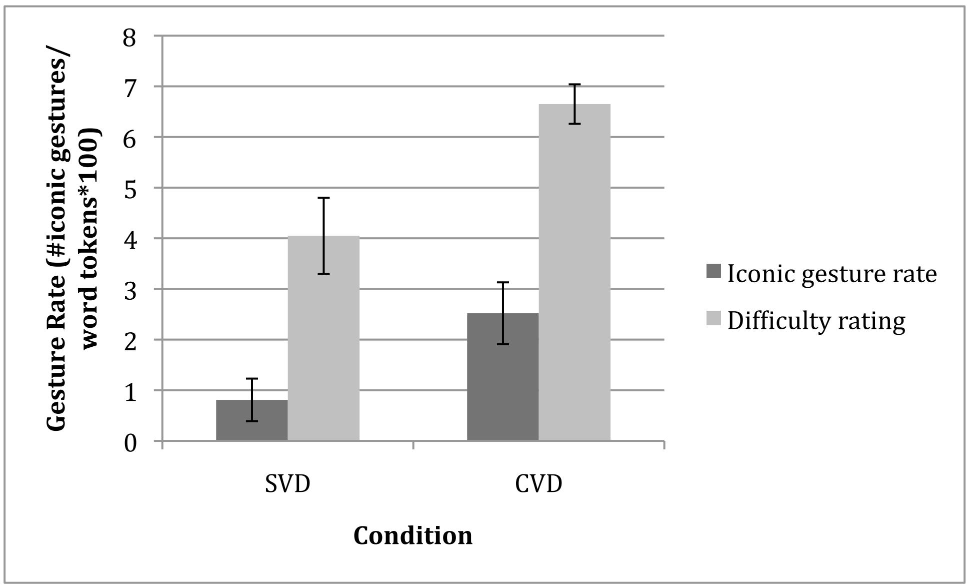

Figure 3: *Gesture rates among participants in the SVD and CVD conditions*

Table 10 shows the averages and standard deviations of the verbal and gestural measures of the narratives produced by participants. Comparisons assessing the narrative quality indicated that participants in the SVD condition told shorter stories ($t(38)=-2.363$, $p=0.023$) than participants in the CVD condition.

Table 10: *Information concerning verbal and gestural measures of narrative production for participants in the SVD and CVD conditions*

Condition	Iconic Gesture Rate	Word Tokens
SVD (N=20)	0.81 (1.88)	319.25 (240.16)
CVD (N=20)	2.52 (2.75)	502.60 (250.51)

Visual perceptual manipulation or verbal short-term memory effects?

As was mentioned earlier, the primary purpose of this study was to determine whether visual perceptual input might have an influence upon iconic gesture production. It was determined that participants in the SVD and CVD conditions differed significantly with respect to their gesture production, however they also differed significantly with respect to their baseline levels of verbal short-term memory scores. Further analyses were therefore required in order to disentangle the relative contributing effects of the differing levels of verbal short-term memory and the manipulation of the visual perceptual input respectively. In order to assess this, a univariate analysis of covariance was conducted. The results revealed that iconic gesture rates remained significantly different when controlling for the effect of verbal short-term memory [$F(1,37)=6.195$, $p<0.05$]. This result lends support to the claim that the visual perceptual input is in fact influencing iconic gesture production.

Discussion

In the visual interference conditions, participants could not see the listener and therefore, the perceptual influences of the visual images presented on the video glasses on gesture production can be examined in isolation from the communicative influences of listener visibility.

It was hypothesized that if imagery leads to iconic gesture production, then the interference from the more difficult visual distractor should inhibit gesture production to a greater extent than the interference from the easier visual distractor. It was also hypothesized that if iconic gesture production is used to facilitate the activation of imagery, then individuals should use more iconic gestures with greater perceived visual distractor difficulty.

Instead, the results support the GFI theory by revealing that with an increase in perceived task difficulty, gesture production increased significantly (it is assumed that task difficulty should be positively associated with the degree of visual interference and therefore negatively associated with the production of imagery). Gesture production seems to increase as a function of visual-specific task difficulty. Rather than lending support to the GSA theory, it seems that gestures enhance the activation of spatial imagery and this may facilitate narrative production.

With respect to language production, participants in both conditions demonstrated a strong positive association between narrative length and iconic gesture production. This provides further evidence for the strong association between gesture and speech. Furthermore, individuals in the less difficult condition relayed shorter stories than those in the more difficult condition. Therefore, even when task difficulty is manipulated, the strong association between iconic gesture production and narrative length remains. The use of iconic gestures during difficult tasks may facilitate lexical access for narrative length. Support for the position that gestures can facilitate lexical retrieval has been suggested in previous literature (Krauss *et al*., 2000).

The results of this study suggest that visual perceptual input interferes with the use of imagery and that iconic gesture production is associated with both the activation of imagistic representations and speech.

CHAPTER IV

Introduction to Study 3

The majority of research conducted addressing imagery focuses on visual and spatial imagery and rarely addresses auditory imagery (Hubbard, 2010). Auditory imagery is defined as the “introspective persistence of an auditory experience, including one constructed from components drawn from long-term memory, in the absence of direct sensory instigation of that experience” (Intons-Peterson, 1992, p. 46). This form of imagery relies on many of the same neural mechanisms as auditory perception (Hubbard, 2010).

Participants in Studies 1 and 2 were asked to watch two cartoons with sound. Although no words were spoken throughout the videos, sound effects occur throughout the films. Presumably, when participants activate their memories for the videos that they must recall, they activate mental imagery that encompasses the sensory modalities that were perceptually activated when the videos were originally watched. These modalities would include auditory and visual perceptual systems. According to Schifferstein (2008-9) imagery is subjectively appraised as being most vivid for visual and auditory systems.

If gestures truly are specifically tied to visual and spatial imagery as many researchers have claimed rather than being tied to imagery in general (encompassing auditory imagery as well), then we would expect an auditory distractor to have no effect upon gesture production. However, if gestures are grounded in imagery in a broader sense, then we would expect gesture production to be similarly affected by an auditory distractor as it was by a visual distractor.

With respect to narrative length, auditory interference has been found to disrupt the functioning of verbal working memory (Baddeley & Andrade, 2000). Since verbal working memory is responsible for holding both “speech based information” (Baddeley & Andrade, 2000, p.127) and non-speech sounds, an increase in the difficulty of auditory interference may lead participants to tell shorter stories. As in Studies 1 and 2, it was hypothesized that participants would demonstrate a strong association between narrative length and gesture production when task difficulty increased.

Participants were randomly assigned to two experimental conditions for the story-retelling component of the experiment: (1) Participants heard a “simple auditory distractor” (SA) while relaying their narratives. This distractor consisted of a simple beeping noise which occurred rhythmically (2) Participants heard a “complex auditory distractor” (CA) while relaying their narratives. This distractor consisted of a beeping noise which did not occur rhythmically but rather occurred in irregular patterns and intervals.

Method

Participants

A total of 40 participants were included in this study. Participants included 8 males and 12 females in the SA condition ranging in age from 18 to 22 with an average age of 19.45 years (SD=1.19). Participants included 6 males and 14 females in the CA condition ranging in age from 18 to 20 with an average age of 19.10 years (SD=0.79).

In order to ensure the comparability of participants in the two auditory experimental conditions, t-tests were conducted upon all measures that were independent of the experimental manipulation. Participants in the SA and CA conditions did not differ with respect to age (t(38)=1.096, p=0.280), vocabulary (t(38)=-1.005, p=0.321), visuospatial short-term memory (t(38)=-0.509, p=0.614), visuospatial working memory (t(38)=-1.104, p=0.277), verbal short term memory (t(38)= -1.012, p=0.318) or verbal working memory (t(38)=-0.675, p=0.504).

Procedure

To ensure that appraisals of the stimuli difficulty were correct, participants were asked to gauge the difficulty of telling the story with the distractor present. Twenty participants from the SA condition and twenty participants from the CA condition were asked to rate the difficulty of telling the narrative with the sound in the background on a scale of 1-10 (with 10 being very difficult). A two-tailed, independent-samples t-test was conducted which showed that participant ratings of difficulty differed significantly according to condition (t(38)=-1.677, p=0.102). Since the sound stimulus used in the CA condition was more sporadic and less predictable than the sound stimulus used in the SA condition, a one-tailed t-test was assessed. The critical t-value with 40 degrees of freedom with an alpha level of 0.05 is 1.684. Therefore, the sound stimulus in the CA condition was rated as marginally more difficult than the sound stimulus used in the SA condition.

In order to demonstrate the comparability of the visual and auditory distractors in terms of cognitive load, it was necessary to determine that the difficulty ratings were comparable. Since the SVD visual condition and the SA auditory condition were rated as being less difficult distractors for the visual and auditory conditions respectively, their ratings of difficulty were compared. Participants in the SVD condition provided ratings of difficulty with an average of 4.05 (SD=2.36), whereas participants in the SA condition provided ratings of difficulty with an average of 4.70 (SD=2.41). A two-tailed, independent-samples t-test revealed that ratings of difficulty ($t(38)=-0.700$, $p=0.490$) were not significantly different. The CVD visual condition and CA auditory condition were also compared. Participants in the CVD condition provided ratings of difficulty with an average of 6.65 (SD=1.25), whereas participants in the CA condition provided ratings of difficulty with an average of 5.93 (SD=2.20). The results revealed that ratings of difficulty were not significantly different ($t(38)=0.961$, $p=0.345$). Additionally, overall ratings for difficulty of the visual and auditory distractors were compared. The results demonstrated no significant difference in difficulty between the visual and auditory conditions ($t(58)=0.059$, $p=0.953$).

Results

SA condition

In the SA condition, the mean PPVT score was 180.25 (SD=8.53). With respect to the visuospatial memory measures, participants scored an average of 30.15 (SD=4.59) on the visuospatial short-term memory assessment and 23.85 (SD=5.54) on the visuospatial working memory assessment. With respect to the verbal memory measures, participants scored an average of 35.65 (SD=5.67) on the verbal short-term memory assessment and 17.20 (SD=2.73) on the verbal working memory assessment.

Simple correlations between iconic gesture production and memory scores in the SA condition revealed no associations (Table 11).

Table 11: *Simple correlations between iconic gesture production and the memory scores in the SA condition*

Relationship	Correlation
Iconic gesture production and visuospatial short-term memory	r=0.032 (p=0.893)
Iconic gesture production and visuospatial working memory	r=0.256 (p=0.275)
Iconic gesture production and verbal short-term memory	r=-0.184 (p=0.438)
Iconic gesture production and verbal working memory	r=0.120 (p=0.613)

Simple correlations between narrative length and memory scores revealed no significant associations (Table 12). However, it should be noted there was a marginally significant association between narrative length and visuospatial working memory scores.

Table 12: *Simple correlations between narrative length and the memory scores in the SA condition*

Relationship	Correlation
Narrative length and visuospatial short-term memory	r=0.277 (p=0.237)
Narrative length and visuospatial working memory	r=0.435 (p=0.055)
Narrative length and verbal short-term memory	r=-0.253 (p=0.283)
Narrative length and verbal working memory	r=0.315 (p=0.176)

The relationship between iconic gesture production and narrative length was assessed. Among participants in the SA condition, iconic gesture rate was positively associated with narrative length in word tokens (r=0.472, p=0.036).

CA condition

In the CA condition, the mean PPVT score was 182.70 (SD=6.80). With respect to the visuospatial memory measures, participants scored an average of 30.85 (SD=4.09) on the visuospatial short-term memory assessment and 25.75 (SD=5.34) on the visuospatial working memory assessment. With respect to the verbal memory measures, participants scored an average of 37.65 (SD=6.78) on the verbal short-term memory assessment and 18.00 (SD=4.54) on the verbal working memory assessment.

Simple correlations between iconic gesture production and memory scores in the CA condition revealed no significant associations (Table 13).

Table 13: *Simple correlations between iconic gesture production and the memory scores in the CA condition*

Relationship	Correlation
Iconic gesture production and visuospatial short-term memory	r=-0.130 (p=0.585)
Iconic gesture production and visuospatial working memory	r=-0.365 (p=0.113)
Iconic gesture production and verbal short-term memory	r=-0.173 (p=0.465)
Iconic gesture production and verbal working memory	r=-0.089 (p=0.710)

Simple correlations between narrative length and memory scores revealed no significant associations (Table 14).

Table 14: *Simple correlations between narrative length and the memory scores in the CA condition*

Relationship	Correlation
Narrative length and visuospatial short-term memory	r=-0.158 (p=0.506)
Narrative length and visuospatial working memory	r=0.204 (p=0.387)
Narrative length and verbal short-term memory	r=-0.228 (p=0.333)
Narrative length and verbal working memory	r=0.037 (p=0.878)

The relationship between iconic gesture production and narrative length was assessed. Among participants in the CA condition, iconic gesture rate was positively associated with narrative length in word tokens (r=0.471, p=0.036).

Comparing participants in the SA condition to the CA condition

Iconic gesture production in the SA condition versus the CA condition was compared. Participants in the SA condition had a mean iconic gesture rate of 5.28 (SD=4.14) and participants in the CA condition had a mean iconic gesture rate of 6.96 (SD=3.94). The results revealed no significant difference (t(38)=-1.314, p=0.197).

Table 15 summarizes the averages and standard deviations for verbal and gestural measures of narrative production. Comparisons of narrative quality revealed no significant differences in terms of the length of the stories (t(38)=-0.826, p=0.414).

Table 15: *Information concerning verbal and gestural measures of narrative production for participants in the SA and CA conditions*

Condition	Iconic Gesture Rate	Word Tokens
SA (N=20)	5.28 (4.14)	418.70 (198.38)
CA (N=20)	6.96 (3.94)	475.05 (231.84)

Discussion

The simple visual and auditory distractors did not differ with respect to perceived task difficulty nor did the complex visual and complex auditory distractors differ with respect to perceived task difficulty. If it is the case that gesture production is specifically associated with visuospatial imagery rather than being associated with imagery in general then we would expect a different pattern of results in comparison to Study 2. Namely we would expect no difference in gesture production with an increase in difficulty. Although no significant difference with respect to iconic gesture production was found among participants in the SA and CA groups, a trend emerged wherein as perceived ratings of task difficulty increase, gesture rates also increase.

With respect to measures of narrative production, participants in both conditions demonstrated a strong positive association between iconic gesture production and narrative length. This suggests a very strong positive association between iconic gesture production and language in these conditions.

CHAPTER V

General Discussion

Narrative productions were used in this study to assess gesture production and language production with respect to working memory resources. The overarching purpose of these studies was to clarify the relationship between visuospatial working memory resources, iconic gesture production and narrative length by testing two competing theories concerning the relationship between these factors.

Evaluating the GSA and GFI Theories

The GSA framework is based upon an embodied cognition perspective and argues that “gestures emerge from the perceptual and motor simulations that underlie embodied language and mental imagery” (Hostetter & Alibali, 2008, p.495). According to this theory, individuals who have stronger visuospatial cognitive resources should gesture to a greater extent than individuals who have poorer visuospatial cognitive resources. Study 1 did not support this claim. Further, according to this framework, it would be argued that as the level of interference is reduced, gesture production should increase since this would presumably facilitate perceptual and motor simulations. However, the results from Study 1 did not support this claim. Since Study 1 confounded visual perceptual influences with listener effects, a clear comparison between the VGND and NORMAL conditions is not possible and as a result, the evidence from this study is largely speculative.

According to the GSA framework, more distracting visuospatial stimuli would be expected to interfere with our perceptual and motor simulations to a greater extent and therefore, in Study 2 it was argued that individuals in the CVD condition should gesture less than those in the SVD condition. Once again this prediction was not supported by our results. Instead, participants in the CVD condition gestured significantly more than those in the SVD condition.

The results from these studies lend support instead to the GFI theory of gesture production. According to this theory individuals use gesture in order to activate imagery as a function of task difficulty. For example, in Study 1, participants gestured less when fewer perceptual distractors were present, in Study 2, participants gestured more when the visual distractor was rated as being more difficult, and in Study 3, participants showed a trend to gesture more when the auditory distractor was rated as being more difficult.

In summary, there are three aspects of the GSA framework that are largely incompatible with the results from the current studies.

Bidirectional Relationship between Gesture and Imagery

1) The GSA framework argues that imagery gives rise to iconic gesture production. However, a positive correlation between visuospatial working memory and iconic gesture production in the NORMAL condition was not found. Furthermore, none of the experimental conditions illustrated a positive association between visuospatial working memory and iconic gesture production. Although it may be the case that imagery is necessary for gesture production, participants gestured to a greater extent when the difficulty of the visual distractor was

increased, and since this presumably inhibited the extent to which they could use their imagery resources effectively, it may also be the case that iconic gesture production gives rise to imagery. This may facilitate the generation of imagistic representations when imagery is disrupted with competing visual information. It would be useful for this model to take into account the bi-directional relationship between imagery and gesture production.

Task Difficulty and Gesture Production

2) In order to explain why task difficulty may result in an increase in gesture production, it has been argued that it requires cognitive effort to inhibit spreading activation from occurring and that as task difficulty increases, "the effort of maintaining a high threshold may be abandoned or lessened so those resources can be devoted to the current task" (Hostetter & Alibali, 2008, p.505). However, why individuals might be inclined to inhibit their gestures is not addressed. Furthermore, since much of human communication is reliant upon imagery activation, task difficulty may always be confounded with imagery difficulty. The GSA framework argues that task difficulty would lead to an increase in gesture production, whereas imagery difficulty would lead to a decrease in gesture production. These competing predictions support a framework that is not easily falsified and more importantly, may not be a sensible approach to understanding the relationship between iconic gesture production, imagery, and task difficulty.

Therefore, although the GSA model may be correct in its assumption that as difficulty increases, gesture production increases, it is not thought to be the case that this is the result of a decreased level of resources available to inhibit gestures but rather because as difficulty increases, the functional roles of iconic gesture production become more useful and are used for the purposes of activating imagery and facilitating speech production. The finding that gesture production increases with an increase in perceived task difficulty lends support to the functional roles of iconic gesture production.

Cross-modal Imagistic Interactions

3) Only visual and motoric imagery are included within the GSA framework as influencing perceptual and action simulations, however, the results from Study 3 suggest that auditory imagery may have an effect upon gesture production as well. Imagery from all perceptual systems interacts so completely that limiting our understanding of gestural production to the influence of solely motor and visual imagery may not take into account other important imagery influences. According to Shimojo and Shams (2001, p.505) "cross-model interactions are the rule and not the exception in perception". Research has demonstrated that when sound is present during the perception of a visual stimulus the perception of the intensity of the visual stimulus is increased. In addition, the quality of visual perception is transformed when auditory stimuli is present (Shimojo & Shams, 2001). The cross-modal interactions that occur during perception may be mirrored by similar interactions in memory retrieval. It is argued that all forms of imagery combine together during perception and memory retrieval and as a result, iconic gesture should be understood in relation to all forms of task relevant imagery.

The GFI model takes into account the three aforementioned problematic aspects of the GSA framework (see Figure 1). It is thought that when participants relay a narrative, they rely extensively upon imagery. Additionally, the perceptual systems which are activated when the information is encoded will be activated in imagery when the information is retrieved. Depictive hand movements are themselves motor actions, and evidence has accumulated to support a strong link between motor imagery and gesture production. People tend to gesture more when they describe patterns that are physically made as opposed to only being viewed, when they mentally rotate shapes as opposed to imagining them stationary, and when they describe nouns having strong associations with actions (Hostetter & Alibali, 2008). These findings suggest a very close link between motor imagery and gesture production. However, as mentioned before, it does not appear to be the case that motor imagery is exclusively associated with iconic gesture production. As a result of cross-modal interactions, it is thought that all sources of sensory imagery activated during perception will influence one another during recall. For the current task, participants were asked to watch and recall cartoons where visual and auditory perceptual information was available to them. It is therefore thought that within the current study, visual imagery and auditory imagery interacted with motor imagery and together influenced gesture and speech production.

Any interference to task relevant imagery, (in this case whether it be auditory, visual, or motor) is thought to reduce the overall efficiency of imagery and therefore leads to a greater sense of perceived difficulty (since narrative production is so heavily reliant upon imagery access). As imagery interference increases, perceived task difficulty increases, and the need for gestures to facilitate imagery increases. As a result gesture rate increases overall. Additionally, this gesture rate increase affects language production since language is also reliant upon imagery. More specifically, as task difficulty increases, this leads to an increase in gesture rate to facilitate imagery activation which then influences language production, resulting in a stronger association between gesture and speech when task difficulty increases. Evidence for this component of the model has been gleaned from the three studies presented here.

Ultimately, iconic gesture production may directly facilitate imagery, may indirectly facilitate imagery, or may simultaneously directly and indirectly facilitate imagery. Future studies are required in order to investigate the nature of this relationship. In order to investigate this relationship behaviorally, a distractor would have to be imposed which would not be thought to interfere with imagery. However, all sensory stimuli that can be perceived are thought to influence imagery. Even distractors that are not consciously perceived are thought to influence imagery (Poetzl, 1960). Therefore any distractor that is consciously or unconsciously identifiable would be expected to interfere with imagery. Since the nature of the relation between gesture and imagery seems to be difficult to obtain with the use of behavioural data, it is suggested that neural imaging studies be conducted in order to detect the localization of interference effects and the effectiveness of gesture production in reducing the cognitive load or enhancing imagery.

Limitations

Narrative production is a task that requires the use of mental imagery. Since a cognitive load of any sort may detract from the cognitive resources which could be assembled to perform a narrative task effectively, it may be argued that increases in the difficulty of visual or auditory stimuli did not result in a specific imagery interference effect, but rather that they resulted in an increased cognitive load, influencing mental imagery only indirectly. By imposing a greater cognitive load via visual and auditory distractors, this may have led to a general reduction in cognitive processing available for the task in general rather than specifically interfering with imagery and in either case, this led to an increase in gesture production.

Additionally, a compensatory role of language was not thoroughly investigated since only the number of words, rather than the content of the words was assessed.

Future directions

This study provides evidence supporting iconic gesture productions' facilitating role in imagery activation during a narrative production task when perceived difficulty increases. It also supports the position that visual and auditory distractors may interfere with imagery during recall and that gesture facilitates imagery activation. Future studies are required to more specifically assess the nature of the relationship between imagery and iconic gesture production.

In order to determine how closely iconic gestures are linked with the different components of imagery it would be important to conduct studies which systematically manipulate aspects of perceptual input during encoding, and sensory distractors during retrieval. For example, a study could be conducted wherein the presence of relevant sound during encoding and the presence of a distractor during recall could be manipulated. A finding wherein participants in the sound-on with auditory-distraction-present condition display increased gestural productions (as compared to participants in the other three conditions) would lend more conclusive support to the link between auditory imagery and gesture production.

In addition it is of interest to garner evidence supporting the position that motoric imagery is in fact more closely linked to iconic gesture production than other forms of imagery. A task where participant imitation of the actions of the protagonist in the cartoon during encoding and foot tapping during retrieval could be manipulated. A finding wherein participants in the imitation and motoric distractor condition displayed increased gestural productions (as compared to participants in the other three conditions) would lend support to the link between motoric imagery and gesture production.

If difficulty ratings were comparable in these designs then a comparison of the influence of the manipulations upon gesture production may lend support to one type of interference being more influential with respect to gesture production than another.

Conclusion

This series of studies has contributed to literature in three different respects. Firstly, it has provided evidence that individuals' visuospatial cognitive resources are not necessarily predictive of iconic gesture production. Secondly, individuals tend to use iconic gestures to a greater extent when perceived difficulty of visual or auditory interference increases. Finally, auditory interference was shown to influence iconic gesture production. This is a finding that has not been previously investigated and may be a fruitful avenue for future inquiry.

Overall this study supported a model including a bi-directional relationship between iconic gesture production and imagery. These findings do not counter an embodied cognition perspective of gesture production. On the contrary, they can be well understood within an embodied framework. The embodied perspective argues that our cognition is strongly grounded in our physical interactions within the world and that our perceptual and motor systems are central to cognitive processes (Wilson, 2002). The finding that gesture production may enhance imagery activation, through motoric movements that access imagery resources, parallels this claim.

According to Glenberg (2008, p.43), "Cognition is for action". The results of the current study lend support to this claim with the addendum that when it comes to gesture production, *action is for cognition*.

REFERENCES

Alibali, M.W., Kita, S., Young, A.J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and cognitive processes, 15 (*6), 593-613.

Alibali, M.W., Heath, D.C. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188.

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. Spatial Cognition & Computation, 5, 307-331.

Baddeley, A.D., Andrade, J. (2000). Working memory and the vividness of imager. *Journal of Experimental Psychology: General, 129*, (1), 126-145.

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

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-609.

Beattie, G.W., Shovelton, H.K. (2000). Iconic hand gestures and the predictability of words in context in spontaneous speech. *British Journal of Psychology*, 91, 473-492.

Bucci, Wilma & Norbert Freedman (1978). Language and hand: The dimension of referential competence. *Journal of Personality*, 46, 544–622.

Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235-258). Cambridge: Cambridge University Press.

Cattaneo, Z., Fastame, M. C., Vecchi, T., & Cornoldi, C. (2006). Working memory, imagery and visuo-spatial mechanisms. Amsterdam, Netherlands: John Benjamins Publishing Company.

Colletta, J.M., Pellenq, C., Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication, 52*, 565-576.

Demir, O.E. (2009). A tale of two hands: Development of narrative structure in children's speech and gesture and its relation to later reading skill. Dissertation. The University of Chicago.

Dunn, L.M. & Dunn, L.M. (1997). *Examiner's manual for the PPVT-III: Peabody Picture Vocabulary Test-Third Edition.* Circle Pines, MN: American Guidance Service.

Ganis, G., Thompson, W.L. & Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research, 20*, 226-241.

Glenberg, A.M. (2008). Toward the integration of bodily states, language, and action. In G.R. Semin & E.R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches,* (pp. 43-70). Cambridge: Cambridge University Press.

Goldin-Meadow, S. (2003). Hearing gesture: How our hands help us think. Cambridge, MA: Harvard University Press.

Graham, J. A., & Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology* , *5*, 185-195.

Hostetter, A.B., Alibalil, M.W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture, 7* (1), 73-95.

Hostetter, A.B., Alibali, M.W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15* (3), 495-514.

Hubbard, T.L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin, 136* (2), 302-329).

Intons-Peterson, M. J. (1992). Components of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 45–71). Hillsdale, NJ: Erlbaum.

Kinsbourne, M. (2006). Gestures as embodied cognition: A neurodevelopmental interpretation. *Gesture, 6* (2), 205-214.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). New York: Cambridge University Press.

Krauss, R.M., Morsella, E. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology, 117,* 411-424.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates

Mar, R.A. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia, 42*, 1414-1434.

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science, 16*(1), 6-10.

McNeill, D. (1992). *Hand and mind*. The University of Chicago Press, Chicago and London.

Miller, P. J. (1995). Personal storytelling in everyday life: Social and cultural perspectives. In R. S. Wyer Jr. (Ed.), *Advances in social cognition. Knowledge and memory: The real story* (vol. VIII, pp. 177–184). Hillsdale, NJ: LEA.

Minami, M. (2008). Telling good stories in different languages. *Narrative Inquiry, 18* (1), 83-110.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *Journal of Verbal Learning & Verbal Behavior, 4*, 32-38.

Poetzl, O. (1960). The relationships between experimentally induced dream images and indirect vision. *Psychological Issues, 2* (3), 46-106.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews, 6*, 1–6.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*, 226-231.

Rimé, B., Shiaratura, L., Hupet, M., & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion, 8*, 311–325.

Schifferstein, H.J.J. (2008-9). Comparing mental imagery across the sensory modalities. *Imagination, Cognition and Personality, 28* (4), 371-388.

Semin, G.R. & Smith, E.R. (2008). Introducing embodied grounding. In G.R. Semin & E.R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches,* (pp. 1-5). Cambridge: Cambridge University Press.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology, 11*, 505-509.

Trafton, J.G., Trickett, S.B., Stitzlein, C.A., Saner, L., Schunn, C.D., Kirschenbaum, S.S. (2006). The relationship between spatial transformations and iconic gestures. *Spatial cognition and computation, 6*, (1), 1-29.

Ukrainetz, T.A., Justice, L.M., Kaderavek, J.N., Eisenberg, S.L., Gillam, R.B., Harm, H.M. (2005). The development of expressive elaboration in fictional narratives. *Journal of Speech, Language, and Hearing Research, 48*, 1363-1377.

Wagner, S.M., Nusbaum, H., Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language, 50*, 395-407.

Wesp, R., Wheaton, K., & Wheatley, T. (1996, May). *Reducing visual stimulation may decrease conversational gestures.* Poster presented at the meeting of the American Psychological Society, Washington, DC.

Wesp, R., Hesse, J., Keutmann, D. & Wheaton, K. (2001). Gestures maintain spatial imagery. *American Journal of Psychology*, 114, 591–600.

Wilson, M. (2002). Theoretical and review articles: Six views of embodied cognition. *Psychonomic Bulletin & Review, 9* (4), 625-636.

Wilson, N. L., & Gibbs, R. W., Jr. (2007). Real and imagined body movement primes metaphor comprehension. *Cognitive Science, 31*,721-731.