2018

Referential Indexing: The Role of Space in Reference Processing

Cameron Smith

Recommended Citation

Smith, C.(2018). *Referential Indexing: The Role of Space in Reference Processing.* (Master's thesis). Retrieved from https://scholarcommons.sc.edu/etd/5503

Referential indexing: The role of space in reference processing

by

Cameron Smith

Bachelor of Science
Mercer University, 2014

Bachelor of Arts
Mercer University, 2014

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Arts in

Experimental Psychology

College of Arts and Sciences

University of South Carolina

2018

Accepted by:

Amit Almor, Director of Thesis

Robin K. Morris, Reader

Douglas H. Wedell, Reader

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

I thank the many people who have helped me complete the work for this thesis, particularly my advisor, Amit Almor, for his input, encouragement, and patience throughout the thesis process. I would also like to thank my readers, Robin Morris and Doug Wedell. In addition, I would like to thank the members of the aLab who have provided support in data collection. Finally, I would like to thank all who have participated in this research.

ABSTRACT

A number of studies have suggested that spatial representations are used during language comprehension in a number of different paradigms (e.g. Ferreira et al., 2008; Gunter et al., 2015). These studies have largely focused on the visual modality, suggesting that people may use spatial cues in the visual environment to aid them in language comprehension. Importantly, the primary modality of spoken language is audition, which raises the question of whether the reliance on spatial information is specific to referents that are visible or at least were visible at some point. Thus, in this study, we examined whether spatial representations are used in language comprehension in the auditory modality. In support of this, we show that participants retain auditory non-visual spatial information to encode referents and activate this information in a cross-modal reading task. Our results are consistent with findings from studies of sign language and manual gesture that suggest that spatial representation may influence the reference process (e.g. Emmorey, 1996; Gunter et al., 2015). These results thus represent an important first step in understanding how spatial representation may influence spoken language comprehension.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Conversations frequently involve communicating with and about more than one individual. Therefore, to communicate effectively, interlocutors often must keep track of multiple referents that have been mentioned in conversation. Keeping track of these referents within and across conversations can be challenging and likely requires considerable mental resources and may possibly involve non-linguistic mechanisms. In particular, some research suggests that spatial representation plays a role in reference tracking. This suggestion is based on a number of key findings: 1) sign languages exclusively use physical and spatial mediums to communicate and provide referential information (e.g. Emmorey, 1996; Senghas, 2011), 2) co-speech gestures have been shown to influence reference processing (e.g. Gunter, Weinbrenner, & Holle, 2015), and 3) individuals look at locations where items *used* to be when they are referred to, even when these items are no longer in the physical environment (a phenomenon known as "looking at nothing"; e.g. Ferreira, Apel, & Henderson, 2008; Richardson & Dale, 2005; Staudte & Altmann, 2017). The present paper aims to extend this body of research, which is largely based on evidence from the visual modality, to spatial representations more generally and specifically those based on acoustic information. This introductory chapter will examine each of these three areas of research in more detail. The following sections will highlight the interaction of spatial representation and language processing to illustrate how integral spatial representations are to various aspects of visually-mediated language processing.

1.1 SIGN LANGUAGES

Sign languages inherently make use of spatial representation throughout the entire language process, from influencing the interpretation of the specific signs made to being used to anchor specific individuals to areas in space (Emmorey, 1996). Liddell (1995; 2003) argues that there are three types of space that signers use while communicating: Real Space, Surrogate Space, and Token Space. Real Space is a representation of the physical space surrounding the signer and is shared among signers and non-signers alike; this can be used to accomplish reference through deictic pointing and identifying referents around the interlocutor. Surrogate Space can be used to communicate about referents that are not physically present by using a mental space that consists of "surrogates" that can be referred to through pointing. Finally, Token Space involves a smaller mental space that is equivalent to the size of the individual's signing space (i.e., the peripersonal space around the signer in which signs are produced). From this alone, we can see that signers make use of incredibly complex spatial representations, involving a minimum of three levels of spatial representation that interact with one another. Importantly, this is only one way in which signers make use of spatial representation.

Sign languages also allow individuals to communicate information about space in a much different fashion that users of spoken language (Emmorey, 1996). For example, spoken language is limited by the constraints imposed by the modality of communication: speech. One such constraint is the limit on how much information can be communicated from a speaker to a comprehender at any given time. This entails that, in spoken language, all types of information are communicated in a serial sequential manner. In contrast, sign languages, can communicate spatial information at the same time as other information. For example, a signer can create the sign for a single object and can place this sign at a location within signing space that is further away from the signer's other hand to indicate the distance between the two objects being discussed. This represents another key difference in the use of space in signed and spoken languages: sign languages do not only make use of multiple layers of spatial representation, but also actively communicate spatial information *at the same time* as other information. These key differences in use of spatial representation highlight the differences in use of space in sign language relative to spoken language.

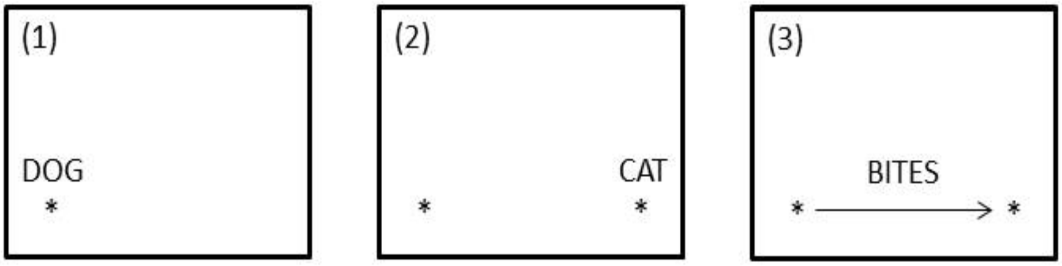

Figure 1.1. *Signing verb agreement. In Panel (1), the sign for 'DOG' is made on the left side of space. In Panel (2), the sign for 'CAT' is made on the right side of space. In Panel (3), the sign for 'BITES' is made in a motion that connects the index for 'DOG' to the index of 'CAT,' meaning "The dog bites the cat" (example from Emmorey, 1996).*

Sign languages can make use of space in other ways as well that are more related to specific linguistic process. Most relevant here, in American Sign Language (ASL) and other sign languages, signers can anchor individual referents to specific areas of space through a process known as indexing. Through this process, the signer may produce the sign of an entity or individual a single time, then by pointing to a location in space around them, the entity is "indexed." Then, later in discourse, the signer need only point to this area in space to refer to that entity (Emmorey, 1996). Further, in many sign languages, including American and Nicaraguan Sign Language (NSL), spatial modulations of individual signs can influence the grammar of the sentence. If an interlocutor indexes two individuals to separate locations in space, a verb sign can be made from one index to another, indicating that the first index was the agent of a sentence, while the second was the patient (Emmorey, 1996; Senghas, 2003; see Figure 1.1; see also Liddell, 2003 for an alternate view on verb agreement in sign languages). Further, spatial modulations of NSL can also provide *who* and *where* information regarding the sentence (Senghas, 2011). Given these findings, it is clear that spatial representations play a central role in tracking reference in sign languages.

It should be noted that the role of spatial representations in sign language extends beyond reference tracking and the other connections between space and language that have been noted for spoken language (metaphor and language of space). Indeed, some studies have demonstrated the influence of the use of sign language on spatial representation extra-linguistically (Emmorey, Kosslyn, & Bellugi, 1993). By comparing deaf signers, hearing signers, and hearing non-signers on measures of image maintenance, image generation, and mental rotation, Emmorey and colleagues (1993) showed that while spatial representation might affect language, language might affect spatial cognition as well. Their results indicated that deaf signers performed faster on tasks involving the generation of both simple and complex images compared to hearing signers, who, in turn, performed better than non-signers; both deaf and hearing signers performed better than non-signers on tasks involving mental rotation, while no differences existed on tasks involving image maintenance tasks (Emmorey et al., 1993).

Both the studies of American Sign Language and Nicaraguan Sign Language are informative in that they illustrate the strong interactions between sign language and spatial representation, be it grammatical (Emmorey, 1996; Senghas 2003; 2011), or more general (Emmorey et al., 1993). The evidence that spatial representation and language interact uniquely in users of sign languages raises the question of whether the spatial aspects of reference tracking seen in signers also hold true for users of spoken language. Interestingly, Almor and colleagues (2007) found that reference tracking during reading involves brain circuits related to processing spatial information, which suggests that some instances of non-signed languages do involve the use of spatial representations.

1.2 CO-SPEECH GESTURE

While sign languages are entirely based on overtly spatial hand movements, spoken languages also make use of hand movements in a different capacity, as co-speech gesture. While gestures may not appear as important to spoken language given that they are primarily an accompaniment to spoken language, they reveal much about ongoing spoken language (McNeil and Pedelty, 1995). In fact, spoken language rarely occurs without the use of co-speech gesture (Guellaï, Langus, & Nespor, 2014). Broadly speaking, co-speech gestures can be broken down into four main types: iconic, metaphoric, beat, and deictic (McNeil & Pedelty, 1995). Iconic gestures describe concrete aspects of the accompanying spoken language. Metaphoric gestures convey abstract or conceptual relationships. Beat gestures mark boundaries in discourse (or syntax). Finally, deictic gestures are used to point at locations in order to refer to entities in a discourse. Importantly, all of these types of co-speech gesture are involved in both language production and comprehension. While there has been much research on co-speech gesture and its role in language production and comprehension, the discussion of gesture here is limited to its relationship to reference.

A number of studies have examined the role of gesture in reference processing in a variety of languages (e.g. Hemforth et al., 2012; Konieczny et al., 2010; So et al., 2009). Hemforth et al. (2010) and Konieczny et al. (2012) investigated the relationship between reference processing and pointing gestures in French and German. After reading a two-sentence discourse, participants had to judge the plausibility of the sentences. To do this, participants had to move their hand to the left or the right in order to press a button to make the judgment. Both studies found that judgments following a reference to the subject of the first sentence were made faster with a leftward motion while judgments following reference to the object of the first sentence were made faster with a rightward one. These results were interpreted as evidence that discourse entities were represented spatially with subjects stored on the left and objects stored on the right. Gestures (in this case, pointing) were facilitated when the direction and location of the referential entity in the mental representation were compatible.

Other studies have examined gestures in response to spoken language production as opposed to comprehension. So and colleagues (2009) reported that when discussing a referent, participants' gestures matched in locations to the location of the referent. Importantly, these gestures only occurred when the referent was also being specified in speech and did not occur to compensate for speech that did not specify a referent. Due to this, So and colleagues (2009) suggested that co-speech gesture has little communicative value, but is nevertheless linked to speech at the discourse level. Thus, a gesture is not sufficient for meaningful communication but is nevertheless linked to spoken language production and comprehension. The fact that this coupling between gesture and language occurs in the context of making reference suggests that spatial indexing may occur in spoken language production.

While So et al. (2009) focuses on co-speech gesture in general, there have been studies that have examined specific types of gesture (independent of the types proposed by McNeil and Pedelty, 1995). One of these gestures that has received attention in the literature is abstract pointing (Gunter et al., 2015; So et al., 2009). Abstract pointing is a pointing gesture directed towards empty space, as opposed to pointing at a specific object or person. So et al. (2009) argue that abstract pointing is redundant, as it typically provides referential information that is already specified in speech. However, despite this claim, other findings suggest that abstract pointing can be used similarly to indexing in sign languages, establishing a referent in space (Gunter et al., 2015). The study conducted by Gunter and colleagues (2015) focused on examining gesture-speech match or mismatch and the consequential effects on discourse comprehension. Participants viewed recordings of individuals discussing various referents and associating them with different pointing gestures to locations in space; later videos depicted these referents being discussed with pointing gestures that matched or mismatched the previous information regarding each referent (e.g. in the first video, the individual would point in one direction while discussing referent A but would point the opposite direction while discussing referent A in the second video). The results indicate that when pointing did not match the utterance, there was a deficit to language comprehension, suggesting that abstract pointing is involved in referential processing and that at least some gestures have greater communicative functions than proposed by So and colleagues (2009).

It should come as no surprise, particularly given McNeil and Pedelty's (1994) classification of gestures, that co-speech gestures are used for a wide variety of linguistic functions. Gesture is used to divide discourse (Guellaï et al., 2014), coordinate social interaction (Haviland, 2000), and convey other linguistic information as well, indicating that gesture is involved in language processing at-large, not just reference processing (e.g. Goldin-Meadow & Singer, 2003). More specifically, Guellaï and colleagues (2014) highlighted that the prosody of speech is not modality specific and can be perceived in spontaneous gestures that accompany speech. These authors argued for two claims regarding the prosodic nature of co-speech gesture. First, participants were able to perceive congruency between unintelligible speech and co-speech gesture, indicating that co-speech gesture accompanies vocalization naturally. Second, participants preferentially chose the meaning indicated by gesture in ambiguous sentences where mismatched gesture and prosody led to different meanings. These findings, in conjunction with those from So et al. (2009) illustrate that speech and gesture form a single communication system where gesture serves not only as a referential guide, but also as a prosodic cue in addition to multiple other functions. In fact, gesture has been shown to enable easier processing of syntactically complex sentences, indicating that gesture can serve an interpretative function beyond those discussed already (Holle, Obermeier, Schmidt-Kassow, Friederici, Ward, & Gunter, 2012).

Together, the studies of sign language and gesture show that spatial representations are used through the language process in a variety of functions in language production and comprehension. In sign languages, spatial representation plays a crucial role in almost every aspect of language processing out of necessity whereas research on co-speech gesture indicates that spatial representations also play a diverse role in spoken language processing, including, but not limited to, referential processing (Guellaï et al., 2014; Gunter et al., 2015; Hemforth et al., 2010; So et al., 2009).

1.3 “LOOKING-AT-NOTHING”

Sign languages and co-speech gesture are both physical indications of reference processing involving spatial representation to various degrees, there are other studies of spoken language processing, independent of gesture, that have indicated that speakers use spatial representations. Eye-tracking studies using the visual world paradigm have shown that speakers and listeners will direct their eye movements towards relevant visual information in their environment when it is mentioned (Richardson & Dale, 2005). The extent to which two interlocutors fixate on similar visual information during conversation is directly associated with the success of comprehension by both. The association between eye fixations and successful language comprehension suggests that spatial representation of the environment is important to discourse comprehension, as increased similarity in interlocutors’ representation of the environment leads to more successful comprehension.

While this alone indicates that spatial representation is involved in language processing in the presence of a relevant visual environment, the use of spatial representation can be seen more directly through a phenomenon known as “looking-at-nothing” (Ferreira, Henderson, & Apel, 2008). This finding also comes from visual world studies where participants view a number of items while engaging in a discourse. Halfway through the experiment, the items are removed from the screen while the participants are still engaged in discourse. After this removal of items, participants continued to look at the areas where items had been located when discussing those items. Ferreira and colleagues propose that speakers use eye movements as a memory cue allowing them to better comprehend the information being discussed. Other studies have supported Ferreira et al.'s (2008) view that "looks-at-nothing" serve as memory retrieval cues, though subsequent studies have highlighted that memory for visual features *and* spoken information is activated by these cues (Hollingworth, 2009; Schold, Mehlhorn, & Krems, 2016). When both an object location and object reference were given, if the location is congruent there is facilitation in responding relative to when the given location is incongruent (Johansson & Johansson, 2014; Scholz et al., 2016). However, until recently, no study had examined whether "looks-at-nothing" when only the location could be recalled facilitated retrieval of the identity of the object at that location; that is to say, no study had examined whether eye movements to *where* an object had previously been located facilitated (or did not facilitate) retrieval of *what* had been in that location (Staudte & Altmann, 2017). Staudte and Altmann tested this by having participants memorize sequences of letters presented onscreen and to verify whether their location or sequence was correct at test. Disrupting looks to a specific location did not impact recall of *what* was in that location while anticipatory looks towards a location did not facilitate recall of *what*. However, unlike most other studies, this study did not employ a traditional visual world paradigm or have participants engage in a discourse and, instead, used grids of letters without a language component. Thus, while Staudte and Altmann indicate a scenario in which "looking-at-nothing" does not seem to facilitate memory of what the target's identity is, it may be that engaging in discourses about specific items in the visual environment does so.

At a more basic level, this phenomenon appears to indicate that speakers have a faithful spatial representation of their environment. Due to this, despite the absence of visual information critical to their discourse, individuals can still use this spatial representation as a referential aid while engaging in a discourse. Remarkably, this phenomenon appears similar to two different aspects of sign languages. First, this appears similar to the ideas of Real Space and Surrogate Space proposed by Liddell (1995; 2003). It is similar to Real Space in the sense that speakers appear to faithfully maintain a spatial representation of their environment for use in communication with others. The relation to Surrogate Space is more intriguing. Recall that Liddell (1995; 2003) suggests that Surrogate Space is *nonphysical* and is used by signers to refer to items not present in Real Space. Thus, "looking-at-nothing" appears to be similar to Surrogate Space insofar as participants in a discourse use eye movements to aid them in production and comprehension of information about items no longer in the physical environment. Second, by looking at areas where the items being discussed *used* to be, speakers engaged in a discourse appear to spatially "index" relevant visual stimuli in their environment, similar to the indexing of referential entities in space done in signed discourse (Emmorey, 1996).

1.4 PRESENT STUDY

The evidence reviewed so far suggests that language processing involves spatial representations in both signers and speakers of spoken language generally, and that the use of these representation extends to reference tracking in signers and potentially speakers of spoken language. Importantly, all the reviewed research investigating the involvement of spatial representations have only examined visually conveyed spatial information: signs and gestures accompanying signed and spoken language, and images or letters in a visual world studies of spoken language (in the case of “looking-at-nothing”). This represents an important limitation in the understanding of the role of spatial representations in spoken language in which the primary modality is *audition* and not vision. To address this limitation, the present experiments aim to determine whether spatial representations are involved in reference tracking during spoken language comprehension even for referents that were not first seen at a specific location. Evidence showing such involvement will support a general role for spatial representation in reference tracking in language processing. A lack of support will suggest that the use of spatial representations in language and reference processing is modality specific. As such, the present studies aim to investigate whether spatial representation plays a role in reference tracking in spoken and written language, by using spatial separation of auditory stimuli to mimic the process of indexing in spoken language similar to the use of gesture and eye-movements in other paradigms (Ferreira et al., 2008; Gunter et al., 2015). The first experiment tested whether people keep track of acoustically conveyed spatial location of possible referents. The second experiment tested whether this information is activated during reference tracking.

CHAPTER II

EXPERIMENT ONE

Experiment One (E1) investigated whether a process similar to indexing in sign language could occur during spoken language processing. Specifically, this experiment tested whether comprehenders would keep track of spatial information associated with potential referents, even when that information is not conveyed visually, and when it does not play any overt role in the task they are performing. To accomplish this, we had participants listen to pairs of short introductions by different speakers heard from different directions. The spatial separation of these referents served an analogous purpose to entities being indexed in space in sign language. In order to determine whether spatial information was automatically encoded, participants then responded to information regarding the referents from a matching or mismatching spatial directions. If participants were indeed coding the spatial information about the referents, we expected to see effects of the direction match on their responses.

2.1 METHOD

Participants

Fifty-two participants (13 male, 39 female, *M* age = 20.21) took part in this study for extra credit towards a psychology course. Of these 52 participants, four were excluded: three were excluded due to being nonnative speakers of English and the fourth was excluded due to equipment malfunction. Thus, data for 48 participants (12 male, 36 female, *M* age = 20.68) were included in the analysis. All included participants were right-handed, native speakers of English wish normal to corrected-to-normal visual acuity (self-reported). Participants gave their informed consent to participate in this research as approved by the Institutional Review Board of the University of South Carolina.

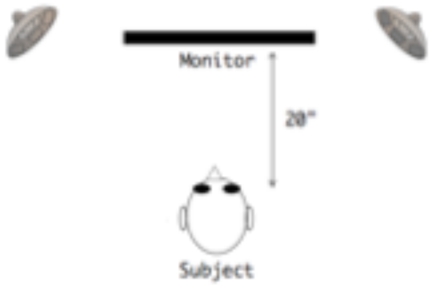

Figure 2.1. *Schematic of the speaker set-up used in both experiments.*

Apparatus

Auditory stimuli were presented through two speakers, located 45 degrees to the left and right of the participant. Each speaker was controlled via an independent sound channel, enabling the individual control of each speaker. Each auditory stimulus was presented via QLab 3, a software enabling the control and manipulation of auditory stimuli (Figure 53, LLC., Baltimore, MD).

Responses were collected through E-Prime 2.0 software on a computer monitor with a square 15" screen with a resolution of 800 x 600 (Psychology Software Tools, Pittsburgh, PA).

Auditory stimuli

Three hundred stimuli items were created. Each stimulus was an introduction of a character, including their name, job title, and place of work (e.g. "Hi, I'm John and I'm a lawyer at Smith, Jones, and Jones, LLP."). Half of the characters were male and half female. Two male and two female native English speakers where then recorded reading one-quarter of the stimuli each for a total of 75 recordings per individual. Stimuli were then paired so that each stimulus pair consisted of two unique voices of either the same or different genders. In addition, each voice was recorded speaking each name separately from the sentences. Auditory stimuli were created in Audacity and were manipulated through PRAAT software so that each sentence was set at 55 dB (Boersma, 2001).

Procedure

The experiment consisted of 75 trials – three practice trials for the participants to familiarize themselves with the procedure of the experiment and 72 experimental trials (see Table 2.1 for sample items). Each trial consisted of a fixation cross lasting one second followed by the presentation of two auditory stimuli. Each auditory stimulus was presented from either the speaker 45 degrees to the right of the participant or 45 degrees to the left of the participant. Each trial consisted of a single introduction from each direction followed a question that appeared onscreen pertaining to one of the two referents (e.g. "Who was the lawyer?" or "Who works in New York City?"). Each question appeared onscreen for 3 seconds. Following the presentation of the question, the name of one of the referents was played from either the left or the right speaker. The correct answer to this question is henceforth referred to as the "target" and the referent that was played after the question is referred to as the "probe." Participants were instructed to press the spacebar if the probe was the correct answer (i.e., if the probe matched the target) to the question they had read or to press the right shift button if the probe was the incorrect answer. Spatially congruent trials were trials in which both the probe and the target appeared from the same direction, while spatially incongruent trials were those in which the probe and target appears from different directions. Participants were given 5000 ms to respond to the name before moving on to the next trial.

The order of presentation of the auditory stimuli (right-left or left-right), the genders of the referents (same-gender or mixed-gender), the name played after the question (correct or incorrect answer), and direction of the post-question name (left or right) were counterbalanced to prevent any effect of specific ordering of stimuli. See Figure 2.2 for a schematic of an experimental trial.

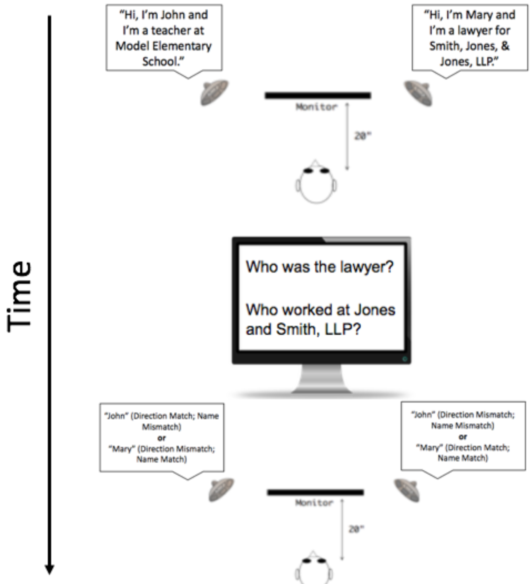

Figure 2.2. *Schematic of a typical trial in Experiment One. Participants heard a two introductions, read a question on-screen, and responded to an auditor probe after reading the question.*

2.2 RESULTS

For this experiment, we were interested in the response times to the single name presentation. Before conducting any data analyses, we removed trials to which the participant did not respond, which affected 1.3% of the data. We removed trials in which participants answered the question wrong, affecting a further 7.5% of the data. In addition, we reasoned that participants were not able to fully process and generate responses faster than 300ms due to the nature of the task (e.g. a judgment regarding a presented answer). Thus, we removed responses that occurred in less than 300ms, removing an additional 1.8% of the data. In all, 10.6% of the data were removed prior to analysis. In addition, outliers were then identified by an adjusted boxplot rule, based on the upper and lower quartiles, along with a robust skewness estimator, separately for each subject. An additional 5.8% of the data were removed as outliers. Following this preprocessing, log-transformed times were than analyzed.

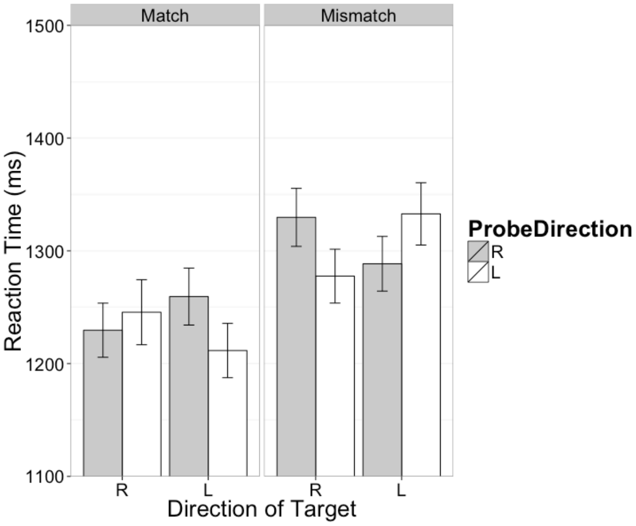

Figure 2.3. *Graphical results of the three-way interaction between target direction, probe direction, and name condition; error bars represent SE. This interaction is driven by a two way interaction between target direction and probe direction present in the name mismatch condition but not in the name match condition.*

Data analyses were conducted in the lme4 package (Bates, Maechler, and Bolker, 2011) in R (v.3.2.3; R Development Core Team, 2015), using mixed-effects modeling with the direction of the target introduction (Target; Left vs. Right), the direction of the probe (Probe; Left vs. Right), the order of introductions (Order; Left-to-right vs. Right-to-left), and whether the name of the target referent matched the name of the probe referent (Name Condition; Match vs Mismatch) as fixed effects. Following Baayen, Davidson, and Bates (2008), we first fit the maximal model including all interaction terms and then tried to eliminate each coefficient; if this elimination did not result in a significant loss of model fit, we then tried to eliminate the subsequent coefficient. We included the maximal structure of by-participant and by-item random intercepts and slopes that allowed the models to converge, as suggested by (Baayen et al., 2008; Barr, Levy, Scheepers, & Tily, 2013). We report the coefficients of the final models, along with random effects structures.

Table 2.1

Coefficient estimates for the reduced model including probe direction, target direction, and name condition as fixed effects. Dummy coding was used with Left as the baseline levels for probe direction and target direction, and Match for name condition.

Coefficient Estimates

β	*Condition*	*Est.*	*Std. Error*	*t*	*p* <
β_0	(Intercept)	7.015	0.047	148.05	0.001***
β_1	Right (Probe)	0.001	0.023	0.035	=0.972
β_2	Right (Target)	0.033	0.022	1.547	=0.121
β_3	Mismatch	0.103	0.025	4.192	0.001***
β_4	Right(Probe) x Right(Target)	-0.046	0.032	-1.456	=0.145
β_5	Right(Probe) x Mismatch	-0.052	0.033	-1.611	=0.107
β_6	Right(Target) x Mismatch	-0.061	0.032	-1.922	=0.055
β_7	Right(Probe) x Right(Target) x Mismatch	0.117	0.046	2.538	0.05*

Following this procedure, the four-way interaction between target direction, probe direction, order of presentation, and name condition was nonsignificant. We elected to fit a model with only target direction, probe direction, and name condition as fixed effects to determine if order influenced responses to probes. After removing the order variable from our analysis, we compared this reduced model to the full model, including order as a fixed effect. The model with three fixed effects was not significantly different from the model with four, indicating that the order of referent introduction did not affect model fit,

$\chi^2(10) = 20.36$, $p = 0.12$[1]**.** Model results for the reduced model are reported in Table 2.1. The three-way interaction between target direction, probe direction, and name condition was significant $\chi^2(1) = 6.43$, $p < 0.05$ in the reduced model (see Figure 2.3)[2].

In order to gain a better understanding of the three-way interaction, we conducted interaction contrasts. Interaction contrasts revealed that the three-way interaction was driven be a two-way interaction between probe direction and target direction that was present in the present in the name mismatch condition ($\chi^2(1) = 5.27$, $p < 0.05$), but not in the name match condition ($\chi^2(1) = 1.66$, p = N.S.). Coefficients for these models are presented in Table 2.2 and Table 2.3 respectively.

Table 2.2

Coefficient estimates for the model of the two-way interaction between target direction and probe direction for the name mismatch condition. Dummy coding was used with Left as the baseline levels for probe direction and target direction.

Coefficient Estimates

β	*Condition*	*Est.*	*Std. Error*	*t*	*p <*
β_0	(Intercept)	7.119	0.045	157.88	0.001***
β_1	Right(Probe)	-0.051	0.023	-2.224	0.05*
β_2	Right(Target)	-0.029	0.023	-1.268	=0.205
β_3	Right(Probe) x Right(Target)	0.074	0.032	2.302	0.05*

2.3 DISCUSSION

As expected, participants responded faster overall for probe referents that matched target referents than those that did not. This is an intuitive finding and is supported by the interpretation that upon seeing the question onscreen, participants likely

---

[1] Random intercepts and slopes were included for probe direction, target direction, and name condition by subjects and items.

[2] Random intercepts and slopes were included for order, probe direction, target direction, and name condition by subjects and for probe direction, target direction, and name condition by items.

had the correct answer in mind, thus incurring a processing cost when the expected response (e.g. the correct answer) was not heard. Importantly, responses to an unexpected response (e.g. the incorrect answer) were modulated by the spatial locations of the target and the probe. Thus, the results of this experiment provide support of our central hypothesis: that spatial indices are maintained during spoken language comprehension. However, Experiment One does not offer any clarity on whether these indices are used in reference tracking during language comprehension or are merely activated by specific language comprehension tasks. The next experiment aimed to test whether these spatial indices are accessed during reference processing during spoken language comprehension.

Table 2.3

Coefficient estimates for the model of the two-way interaction between target direction and probe direction for the name match condition. Dummy coding was used with Left as the baseline levels for probe direction is right and for target direction is right.

Coefficient Estimates

β	*Condition*	*Est.*	*Std. Error*	*t*	*p <*
β_0	(Intercept)	7.017	0.049	143.19	0.001***
β_1	Right(Probe)	-0.001	0.024	-0.536	=0.957
β_2	Right(Target)	0.033	0.024	1.353	=0.175
β_3	Right(Probe) x Right(Target)	-0.042	0.033	-1.291	=0.198

CHAPTER III

EXPERIMENT TWO

Given that E1 showed that participants keep track of acoustic spatial information associated with referents, Experiment Two (E2) tested whether this information is used during the processing of linguistic descriptions of these referents. Thus, instead of having participants validate the response to a question about the referents, following the introductions, participants read onscreen a sentence presented word by word in rapid serial visual presentation (RSVP) in which one of the two referents was mentioned. During the sentence, one of the two referent names was played at one of three points: before the sentential reference, directly after the sentential reference, or after the verb phrase and participant had to simply press a button when they heard the name. We hypothesized that we would find a referent match effect, indicating that participants responded faster when the referent heard during the sentence matched the sentential reference. We also hypothesized that this referent match effect would be *faster* when the referent was played from the *matching* direction. Finally, we hypothesized that the nature of this effect would be different directly after the sentential reference while referential information is being processed compared to after the verb phrase when referential processing has largely finished. We chose three probe positions (before the sentential reference, after the reference, and after the verb phrase). We hypothesized that there would be no effect before the reference in the sentence, given that the referential information would not have been read yet; we further hypothesized that there would be a referent-direction match interaction effect following the reference and that the effect may disappear following the verb phrase.

3.1 METHOD

Participants

Forty-four participants (6 male, 36 female, *M* age = 20.64) took part in this study for extra credit towards a psychology course. All participants were right-handed, native speakers of English wish normal to corrected-to-normal visual acuity (self-reported). Participants gave their informed consent as approved by the Institutional Review Board of the University of South Carolina.

Apparatus

An identical apparatus to that of E1 was used for this experiment.

Auditory Stimuli

The auditory stimuli from E1 were used for this experiment.

Reading task

The reading task consisted of a single sentence presented in a word-by-word manner at the center of the screen for 500 ms per word. The sentence pertained to one of the two referents introduced in the auditory portion of the experiment. The sentence was constructed so that the content did not pertain to any of the introductory statements; thus, either referent could appear in the sentence. In each sentence, the reference was the job of the referential entity (e.g. "the doctor"). Each sentence was the same length and the reference occurred at the exact same point in each sentence so that length was not a confound[3]. The chosen referent for each sentence was counterbalanced across participants.

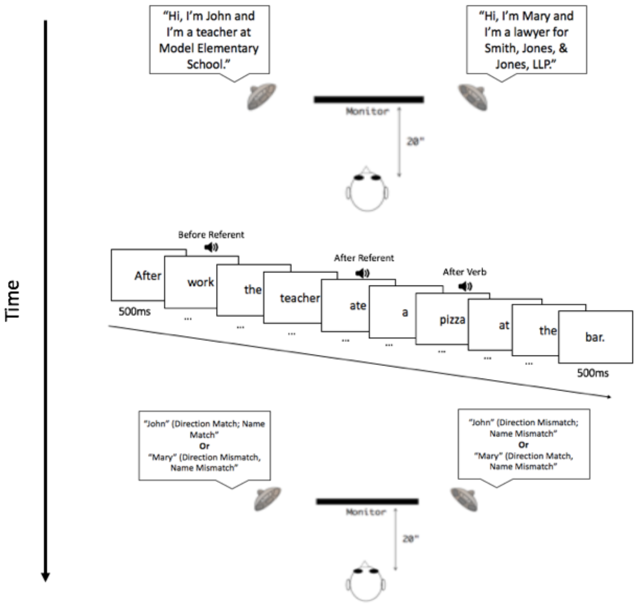

Figure 3.1. *Schematic of a typical trial in Experiment Two. Participants heard two introductions, read a sentence regarding one of the referents word-by-word onscreen and heard a probe in one of three locations during the sentence.*

Procedure

The experiment consisted of 75 trials – three practice trials that familiarized the participants with the reading task and 72 experimental trials. Experimental trials consisted of two referents introducing themselves from the speakers 45 degrees to the right and to the left of the participant followed by the reading task, consisting of one sentence. During the word immediately preceding the reference in the sentence, immediately after the reference, or immediately after the verb, an auditory probe corresponding to one of the two potential referents names was played to which the participants were required to respond. Participants were required to press the spacebar in

---

[3] Example sentence: "After work, John ran two miles at the gym."

response to the probe. Following the sentence, a comprehension question appeared onscreen that pertained to the passage that had just been read (see Figure 3.1 for a schematic of a sample trial). Participants were instructed to respond to this question with a yes or no answer by pressing the spacebar or right shift button respectively. The comprehension questions were used to ensure that the participant was processing the sentences onscreen while responding to the probe and not just responding to the probe alone.

3.2 RESULTS

For this experiment, we were interested in response times to the probe played during the sentence; specifically, the responses at points after the reference in the sentence had been read. Before conducting any data analyses, we removed trials in which the comprehension question regarding the two sentences was answered incorrectly, affecting 20.8% of responses. In addition, we reasoned that participants were unable to respond to the probe faster than 300 ms and that responses slower than 6000ms reflected failures to attend to the task. Responses that did not fall in this range were removed from the dataset as well and we excluded outliers based on the adjusted boxplot rule used in E1; this resulted in a further 5.9% of responses being removed from the dataset. In all, 26.7% of responses were removed and log-transformed reaction times were analyzed.

Data analyses were conducted in the lme4 package (Bates, Maechler, and Bolker, 2011) in R (v.3.2.3; R Development Core Team, 2015), using mixed-effects modeling with wether the direction of the probe matched that character's introduction (Direction Match; Match vs. Mismatch), the speaker the probe was played from (Probe Speaker; Left vs. Right), whether the probe name matched the sentential referent (Name

Match; Match vs. Mismatch), the point at which the probe was played in the sentence (Probe Condition; Before Referent, After Referent, After Verb Phrase), and the spatial order of the introductory statements (Order; Left-to-right vs. Right-to-left) as fixed effects, as well as all possible interactions between the five factors. Following the same procedure as E1, we first fit the maximal model including all interaction terms and then tried to eliminate each coefficient one-by-one (Baayen et al., 2008). We included the maximal structure of by-participant and by-item random intercepts and slopes that allowed the models to converge (Baayen et al., 2008; Barr, Levy, Scheepers, & Tily, 2013). We report the coefficients of the final models, along with random effects structures.

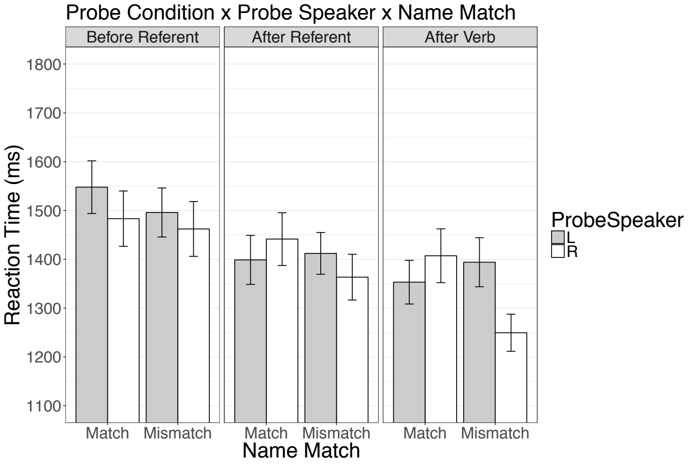

Figure 3.2. *Graphical results of the marginal three-way interaction between probe condition, name match, and probe speaker.*

Given the lack of an effect of Order in E1, we elected to first test whether order should be included in the maximal model. First, we fit a model with the maximum combination of interactions between all five fixed effects as well as a model with only Direction Match, Probe Condition, Name Match, and Probe Speaker as fixed effects.

After removing the order variable from our analysis, we compared this reduced model to the maximal model. The model with four fixed effects was not significantly different from the model with five, indicating that the order of referent introduction did not affect model fit, $\chi^2(16) = 13.364$, $p = 0.64$[4]**.** Model results for the reduced model are reported in Table 3.1.

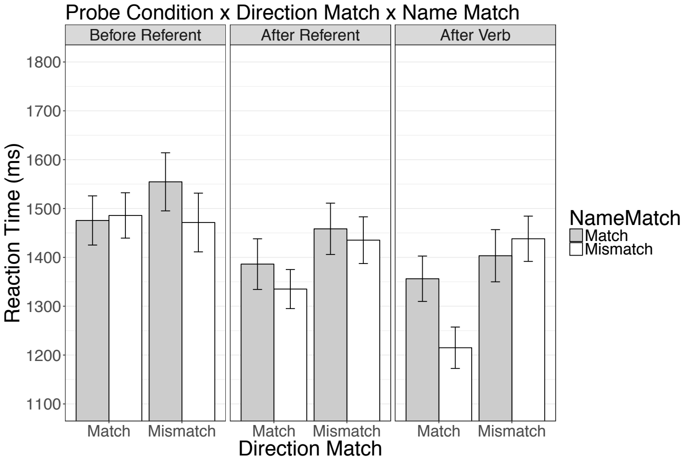

Figure 3.3. *Graphical results of the three-way interaction between probe condition, direction match, and name match.*

Following this procedure, the four-way interaction between direction match, name match, probe condition, and probe speaker was nonsignificant ($\chi^2(1) = 0.229$, $p = 0.63$). The three-way interaction between Direction Match, Probe Condition, and Probe Speaker was nonsignificant, as was the three-way interaction between Direction Match, Name Match, and Probe Speaker. The three-way interaction between Name Match, Probe Condition, and Probe Speaker was marginally significant, $\chi^2(1) = 3.22$, $p = 0.073$. Finally, the three-way interaction between Direction Match, Name Match, and Probe

---

[4] Random intercepts were included for Probe Speaker and Direction Match by subjects and by Probe Speaker by items. Random slopes were included for Probe Speaker both by subjects and by items. These same intercepts and slopes were used for all following models unless otherwise specified.

Condition was significant, $\chi^2(1) = 4.753, p < 0.05$. The model results for the final model including both three-way interactions are reported in Table 3.2. The first three-way interaction is depicted in Figure 3.2 and the second in Figure 3.3.

Table 3.1

Coefficient estimates for the model of the four-way interaction between Direction Match, Name Match, Probe Condition, and Probe Speaker. Dummy coding was used with Match as the baseline levels for Direction Match and Name Match and Left for Probe Speaker.

Coefficient Estimates					
β	*Condition*	*Est.*	*Std. Error*	*t*	$p <$
β_0	(Intercept)	7.231	0.091	79.040	0.001***
β_1	Mismatch(Direction)	0.082	0.085	0.963	=0.335
β_2	Mismatch(Name)	0.060	0.086	0.698	=0.485
β_3	ProbeCondition	-0.077	0.028	-2.703	0.01**
β_4	Right	-0.056	0.088	-0.638	=0.523
β_5	Mismatch(Direction) x Mismatch(Name)	-0.218	0.121	-1.807	0.10•
β_6	Mismatch(Direction) x ProbeCondition	-0.007	0.039	-0.175	=0.861
β_7	Mismatch(Name) x ProbeCondition	-0.026	0.039	-0.660	=0.509
β_8	Mismatch(Direction) x Right	-0.055	0.121	-0.451	=0.652
B_9	Mismatch(Name) x Right	-0.003	0.120	-0.023	=0.982
B_{10}	ProbeCondition x Right	0.056	0.040	1.403	=0.161
B_{11}	Mismatch(Direction) x Mismatch(Name) x ProbeCondition	0.106	0.056	1.908	0.10•
B_{12}	Mismatch(Direction x Mismatch(Name) x Right	0.165	0.171	0.967	=0.334
B_{13}	Mismatch(Direction) x ProbeCondition x Right	-0.014	0.056	-0.254	=0.799
B_{14}	Mismatch(Name) x ProbeCondition x Right	-0.053	0.056	-0.946	=0.344
B_{15}	Mismatch(Direction) x Mismatch(Name) x ProbeCondition x Right	-0.038	0.079	-0.480	=0.632

To better understand the nature of the three-way interaction between probe condition, direction match, and name match, we performed interaction contrasts. Interaction contrasts revealed that a two-way interaction between direction match and name match was present in the "After Verb" condition, $\chi^2(1) = 5.613, p < 0.05$[5]. The

---

[5] Random intercepts and slopes were included for both Direction Match and Name Match by subject and by items for the "After Referent" condition. Random intercepts were included for both Direction Match and Name Match by subjects and by items while random slopes were included for Direction Match by subjects and by items in the "After Verb" analysis. A random intercept was included by subjects and items for the "Before Referent" condition.

model results for this model are reported in Table 3.2. However, there was no significant two-way interaction in the “Before Referent” or “After Referent” conditions; further analyses revealed that there were no main effects of direction match or name match in these conditions.

Table 3.2

Coefficient estimates for the model containing the marginal three-way interaction between Probe Condition, Probe Speaker, and Name Match and the significant three-way interaction between Probe Condition, Direction Match, and Name Match. Dummy coding was used with Match as the baseline levels for Direction Match and Name Match is Match and Left for Probe Speaker.

Coefficient Estimates

β	*Condition*	*Est.*	*Std. Error*	*t*	*p <*
β_0	(Intercept)	7.235	0.086	83.745	0.001***
β_1	Mismatch(Direction)	0.074	0.063	1.174	=0.240
β_2	Mismatch(Name)	0.019	0.074	0.254	=0.799
β_3	ProbeCondition	-0.073	0.025	-2.988	0.01**
β_4	Right	-0.065	0.065	-0.992	=0.321
β_5	Mismatch(Direction) x Mismatch(Name)	-0.134	0.085	-1.567	=0.117
β_6	Mismatch(Direction) x ProbeCondition	-0.014	0.086	-0.497	=0.619
β_7	Mismatch(Direction) x Right	-0.039	0.033	-1.193	=0.233
β_8	Mismatch(Name) x ProbeCondition	-0.017	0.034	-0.489	=0.625
B_9	Mismatch(Name) x Right	0.079	0.087	0.923	=0.355
B_{10}	ProbeCondition x Right	0.049	0.028	1.748	0.10•
B_{11}	Mismatch(Direction) x Mismatch(Name) x ProbeCondition	0.087	0.039	2.186	0.01**
B_{12}	Mismatch(Name) x ProbeCondition x Right	-0071	0.039	-1.799	0.10•

The interaction between direction match and name match in the “After Verb” condition shows that when the direction of the probe matches the direction of the sentential referent’s introduction and the probe is the same referent as the sentential referent, participants responded slower than when the probe was not the same referent as the sentential referent while there was no effect of name match when the direction mismatched that of the target’s introduction. This effect is particularly striking because it is the opposite of the effect found in the mismatch condition in E1.

Table 3.3

Coefficient estimates for the linear mixed effects model containing the significant two-way interaction between Name Match and Direction Match in the "After Verb" condition. Dummy coding was used with Match as baseline levels for Direction Match and Name Match.

Coefficient Estimates

β	*Condition*	*Est.*	*Std. Error*	*t*	*p <*
β_0	(Intercept)	7.061	0.067	104.643	0.001***
β_1	Mismatch(Direction)	0.009	0.042	0.206	=0.837
β_2	Mismatch(Name)	-0.091	0.040	-2.279	0.05*
β_3	Mismatch(Direction) x Mismatch(Name)	0.134	0.056	2.384	0.05*

To better understand the three-way interaction between name match, probe condition, and probe speaker, we conducted interaction contrasts. Interaction contrasts revealed a significant interaction between name match and probe speakers in the "After Verb" condition but not in the "Before Referent" or "After Referent" conditions, $\chi^2(1) = 4.239$, $p < 0.05$[6]. The model results for the significant two-way interaction are reported in Table 3.4. In addition, there were no main effects in either the "Before Referent" or "After Referent" conditions. The interaction between name match and probe speaker in the "After Verb" condition shows that when the probe is not the name of the sentential referent (e.g., "Name Mismatch"), responses were faster to probes played from the right; this effect did not hold in the "Name Match" condition. Critically, these results should be interpreted with caution given that the three-way interaction between probe condition, name match, and probe speaker was marginal, not significant.

---

[6] Random intercepts and slopes were included for both Name Match and Probe Speaker by subject and by items for the "After Referent" and "After Verb" conditions. A random intercept was included by subjects and items as well as a random slop for Name Match by Subjects for the "Before Referent" condition.

Table 3.4

Coefficient estimates for the linear mixed effects model containing the significant two-way interaction between Name Match and Probe Speaker in the "After Verb" condition. Baseline for Name Match is Match and for Probe Speaker is L.

Coefficient Estimates

β	*Condition*	*Est.*	*Std. Error*	*t*	*p <*
β_0	(Intercept)	7.047	0.081	87.376	0.001***
β_1	Mismatch(Name)	0.023	0.040	0.565	=0.837
β_2	Right	0.048	0.047	1.023	=0.306
β_3	Mismatch(Name) x Right	-0.116	0.056	-2.071	0.05*

3.3 DISCUSSION

The results from E2 provide very limited support for the hypothesis that individuals use spatial indices during the reference process. The significant interaction between name match, direction match, and probe condition suggests that spatial representations do interact with reference processing. However, the nature of this interaction did not influence response times in the hypothesized direction. First, we hypothesized that there would be an effect of the spatial information directly after the referent, indicating that participants had indexed a referent in space and used this information to aid in processing the referent; no such interaction was found. Second, we hypothesized, overall, that if the direction of the probe matched the introduction (e.g. "Direction Match") that there would be faster processing for the correct name (e.g. "Name Match); no such interaction was found. However, the presence of the interaction in the "After Verb" condition does illustrate that there *is* an effect of spatial representation but either 1) our task does not reliably elicit this effect across the sentence or 2) the effect of spatial information does not interact during reference processing but instead occurs later in sentence processing. The lateness of this effect is compatible with models of reference processing that suggest that discourse integration, including reference, takes time and does not occur directly upon reading a referent and, instead, is delayed to a later point in the sentence (e.g. Peters, Boiteau, & Almor, 2016). Indeed, an effect of spatial information was only found at the latest probe position tested, suggesting that the referential information may not have been integrated until this point.

Importantly, the direction of the interaction was opposite of what was hypothesized. In the “After Verb” condition, participants responded faster to a mismatching name from the correct direction while the responses to names from the mismatching direction were not different from each other. This may be reflective of more complex integration and comprehension due to the nature of the task (reading a sentence). Indeed, responding faster to a mismatching name may indicate that participants were able to swiftly reject a mismatching name whereas responding to the matching name may require integration of the sentential and auditory reference into a single representation. This is compatible with the Informational Load Hypothesis that suggests that interference in the repeated name penalty is due to the integration of the multiple representations of the same referent created by repeated name use (Almor, 1999; Almor & Nair, 2007). I return to this consideration in the General Discussion.

The overall lack of an effect of the order of referent introduction may be somewhat surprising, given that there are several theoretical accounts of English (and other languages with a left-to-right orthography) that have suggested that speakers of these languages prefer events to unfold in the left-to-right direction (Chatterjee, 2001; Chatterjee, Southwood, & Basilico, 1999; Maas & Russo, 2003). However, both E1 and

E2 failed to find an effect of order, suggesting that this preference may not be a concern for the paradigm used in these experiments.

More generally, the task used in this experiment may have contributed to the nature of the effect of spatial representation. Instead of having participants engage in a judgment task or decision-making task regarding the probe, participants merely had to indicate that they had *detected* a probe. This introduces the possibility that participants only needed to respond as soon as they heard any noise begin to play instead of attending to and processing the referent during the sentence task. However, given that we did find the hypothesized interaction (albeit in a different direction), it appears that participants did attend and process the probe. In addition, the task used in this experiment is a cross-modal task between spoken language processing and visual language processing. Importantly for this study, there is much evidence that the visual modality is "dominant" in cross-modal correspondence (e.g. Shams & Beierholm, 2010). This dominance suggests that visual information is much more likely to influence other modalities during processing whereas the other senses (e.g. audition, touch) are less likely to exert their influence over visual processing. Given that our task was anticipating auditory information to influence a visually-mediated language task, this may be important regarding the nature of the results obtained from E2.

CHAPTER IV

GENERAL DISCUSSION

The main goal of this paper was to explore the possibility that the use of spatial representations to aid language and, more specifically, reference comprehension in discourse is not specific to the visual modality. More specifically, we aimed to test whether spatial “indices” were unique to sign languages or whether they were used during spoken language comprehension. We tested this idea using a novel paradigm that, to our knowledge, has never been used to explore this phenomenon. E1 had participants listen to two individuals introduce themselves from spatially distinct locations and later presented one of these individuals from one of these locations in a verification task. E2 presented two individuals similarly to E1 but participants engaged in a probe detection task while reading a sentence; the probe was played before or after the referent in the sentence. The results from E1 suggest that spatial representations are indeed maintained during spoken language comprehension, while the results from E2 provide limited support for the activation of spatial representation during reference processing. However, the task used in E2 may have contributed to the limited nature of the effect.

In E1, we found that when the probes presented to the participants were the correct answer to the question response times were faster than when the probe was incorrect (e.g. name mismatch). However, only the name mismatch condition revealed an effect for the spatial manipulation: participants were faster at rejecting the incorrect name from the right direction (e.g. hearing the in correct answer from the direction the correct answer was expected) than the incorrect name from the incorrect direction.

To our knowledge, this is the first time that an influence of spatial representation on reference processing has been reported using the auditory modality as opposed to a visual modality (e.g. Emmorey, 1996; Ferreira et al., 2008; Guellaï et al., 2014). The spatial effect found in E1 (and partially in E2) suggests that individuals keep track of spatial representations to index referents to specific locations in space, similar to reference use in sign languages (Emmorey, 1996; Senghas, 2003; 2011). This matches a host of previous research suggesting that spatial information may influence reference processing in a visual modality (e.g. Ferreira et al., 2008) or during spoken language comprehension accompanied by manual gesture (e.g. Gunter et al., 2015). Importantly, the only spatial aspect of the stimuli used in the experiment was our spatial manipulation – the content of the introductions or sentences was not spatial in nature. This is an important finding as it illustrates that the participants used abstract spatial representations created specifically for each of the referents without any other spatial information contaminating this process.

Importantly, the direction of the interaction in E1 is different than that of E2. In E1, participants were faster at responding to incorrect names presented from the mismatching direction, while in E2 participants responded faster to incorrect probes played from the matching direction while there was no effect in the mismatching direction. This may be due to the fact that in E1, participants could have already had the answer in mind when the probe was played, thus creating a floor-effect in the "name match" condition. In E2, however, participants were likely to integrate the referent read in the sentence with an auditory probe; thus, the integration of these two representations may have had a penalty when coming from the correct direction relative to the incorrect name, which requires no integration. This effect is remarkably similar to the repeated name penalty which has been hypothesized to be due to the integration of multiple representations (Almor, 1999; Almor & Nair, 2007).

Interestingly, the finding that individuals do appear to "index" referents to particular locations in space is similar to both the phenomenon of "looking-at-nothing" (e.g. Ferreira et al., 2008) and the concept of Surrogate Space (e.g. Liddell 1995; 2003). Liddell's initial proposal was that only Real Space was shared by signers and non-signers, whereas Surrogate and Token Space were used exclusively by signers. However, studies of visually-mediated language (e.g. Ferreira et al., 2008; Richardson & Dale, 2005; Staudte & Altmann, 2017) appear to suggest that Surrogate Space may be available to non-signers as well. The present findings appear to support this idea as well, given that if participants were indexing referents in space they were almost certainly doing so in a nonphysical spatial representation, given that no actual individuals were sitting in the room alongside the participants.

In addition to the discussion of spatial representations, there is an entire body of literature investigating the effects of laterality in language that could be important to this and future studies of spatial representation. However, neither E1 or E2 found an effect of order of introduction, suggesting that the general preference for events to unfold in a left-to-right manner in English speakers (and other languages with a left-to-right orthography) may indicate that this preferent is not a concern for the present paradigm (Chatterjee, 2001; Maas & Russo, 2003).

In sum, the previous findings regarding spatial representation in language comprehension from studies of sign languages, eye-tracking, and manual gesture as well as findings regarding laterality on language comprehension all suggest that spatial representation plays a role during language comprehension (Boiteau & Almor, 2016; Emmorey, 1996; Ferreira et al., 2008; Guellaï et al., 2014; Gunter et al., 2015). Our results support these claims from a novel paradigm and suggest that these effects extend beyond the visual modality into the auditory modality as well. However, the effect of spatial representation on language processing in the auditory modality is unclear and further studies should be conducted to elucidate the nature of this effect.

Limitations

In contrast to E1, the effect found in E2 is unclear. The spatial effect was only found after the verb in the sentence but not directly after the referent, as hypothesized. However, our task may have been responsible for the lack of a spatial effect. We employed a reading task following the auditory introduction of two individuals during which a probe detection was required. Detection of probes alone (instead of a judgment task or decision task) may have been insufficient to result in an effect across the sentence and not just at a single position. However, by placing one probe directly after the referent in the sentence, there may not have been enough time for the participant to completely process the referent at the time of the presentation of a probe. Future studies may be able to refine the current methodology and use exclusively auditory stimuli as opposed to using visual and auditory stimuli to minimize cross-modal influences. In addition, further studies using cross-modal stimuli may be able to change the probe detection task into a judgment task to maximize the potential effect of the probe on referential processing.

REFERENCES

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, *106*, 748-765.

Almor, A. & Nair, V.A. (2007). The form of referential expressions in discourse. *Language and Linguistic Compass*, *1*, 84-99.

Almor, A., Smith, D. V., Bonilha, L., Fridriksson, J., & Rorden, C. (2007). What is in a name? Spatial brain circuits are used to track discourse references. *Neuroreport*, *18*, 1215–9.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255-278.

Bates, D., Maechler, M. & Bolker, B. (2011). lme4: linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9-10), 341-345.

Boiteau, T. W., & Almor, A. (2016). Transitivity, space, and hand: The spatial grounding of syntax. *Cognitive Science.* doi: 10.1111/cogs.12355Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*, 1-22.

Chatterjee, A. (2001). Language and space: Some interactions. *TRENDS in Cognitive Sciences, 5(2),* 55-61.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events, and spatial representations. *Neuropsychologia, 37,* 395-402.

Emmorey (1996). The Confluence of Space and Language in Signed Languages. In: Bloom, P. (Ed.). *Language and space.* Cambridge, Mass: MIT Press. Chapter 5. (pp. 171-209).

Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*, *46*, 139-181.

Emmorey, K., McCullough, S., Mehta, S., Ponto, L. L. B., & Grabowski, T. J. (2013). The biology of linguistic expression impacts neural correlates of spatial language. *Journal of Cognitive Neuroscience*, *25*, 517-733.

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *12*(11), 405-410.

Gordon P. C., Grosz B. J., & Gilliom L. A. (!993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*(3), 311-347.

Guellaï, B., Langus, A., & Nespor, M. (2014). Prosody in the hands of the speaker. *Frontiers in Psychology*, *5*, 70.

Gunter, T. C., Weinbrenner, J. E. D., & Holle, H. (2015). Inconsistent use of gesture space during abstract pointing impairs language comprehension. *Frontiers in Psychology*, *6,* 80.

Hemforth, B., Homassel, A., Feldmeth, D.M., & Konieczny, L. (2012). Grounding of anaphora in pointing gestures: Order of mention and prominence. CUNY presentation.

Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., & Gunter, T. C. (2012). Gesture facilitates the syntactic analysis of speech. *Frontiers in Psychology*, *3*, 74.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228-238.

Konieczny, L., Haser, V., Muller, D., Weldle, H., Wolfer, S., & Hemforth, B. (2010). Grounding of anaphora in pointing gestures. CUNY Presentation.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago: University of Chicago Press.

Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology*, 25, 353-382.

Levinson, S. C. (2003). Space in Language and Cognition. Cambridge: Cambridge University Press.

Li, P. & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, *83*, 265-294.

Liddell, S.K. (2003). *Grammar, gesture, and meaning in American Sign Language.* New

York: Cambridge University Press.

Maas, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events: Nature or culture? *Psychological Science, 14(4),* 296-301.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Figure 53, LLC. [QLab 3]. (2016). Retrieved from https://figure53.com/.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*, 39-54.

Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development*, *18*, 511-531.

Senghas, A. (2011). The emergence of two functions for spatial devices in Nicaraguan Sign Language. *Human Development, 53*, 287-302.

So, W. C., Kita, S., Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, *33*, 115-125.

Staudte, M. & Altmann, G.T.M. (2017). Recalling what was where when seeing nothing there. *Psychonomic Bulletin and Review*, *24*, 400.