CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

Two Interconnected Modes: Speech and Gesture

A graduate project submitted in partial fulfillment of the requirements

For the degree of Master of Arts in Linguistics, General

By

Paige Jordan Reimer

December 2020

The graduate project of Paige Reimer is approved:

Dr. Kenneth V. Luna	Date
Dr. David J. Medeiros	Date
Dr. Laura Ruth-Hirrel, Chair	Date

Table of Contents

List of Images

Abstract

Two Interconnected Modes: Speech and Gesture

By

Paige Reimer

Master of Arts in Linguistics, General

Linguistic research on co-speech gesture has established that gesture is a part of the speech production process, as gestures express semantic information, and can perform pragmatic functions, such as speech acts. Some studies have explored the idea that speech and gesture are two distinct modes that complement each other in their roles of the representation of semantic categories (McNeill 1985; Beattie and Holler 2002). This paper explored the interaction between the speech and gesture modes by examining a semantic category not considered by previous studies, shape. Two participants were recorded using audiovisual equipment speaking together for over an hour and a half in a naturally occurring conversation. Their gestures were identified and analyzed in relation to their cooccurring speech and with regard to criteria set forth by

previous research. The gestures focused on were the iconic gestures, but the other gesture types, metaphoric, deictic, and beats, are also noted. There were few examples for the semantic categories of shape but those that did occur showed shape could be represented by the gesture aspect on a basis of importance to the narrative. As this research continues to support the idea of speech and gesture as two complementary modes divided in their work, it would do well to either expand this examination of shape or continue to analyze the other semantic categories not yet touched upon in other research.

I. Introduction

It has been proposed that speech and gesture are two distinct, but complementary modes of thought due to research that suggested they develop internally together, they are triggered by the same stimulus simultaneously, and convey different kinds of information in distinct ways (McNeill 1985). According to McNeill, speech represents the linear side of thought while gesture represents the imagistic side.

Other research has looked more specifically at how semantic information is represented across the two modes. With McNeill's ideas as a foundation, there have been studies on whether gesture use does lead to higher comprehension rates (Beattie and Shovelton 1999), as well as studies on what semantic categories may be represented by each mode (Beattie and Holler 2002; Beatie and Shovelton 2006). Research concentrated on gestures as an asset to comprehension have shown that comprehension rates are higher when gestures are viewed with their accompanying speech, (Beattie and Shovelton 1999; Cassell 1999) supporting the necessity for further study of gestures, and how the gesture and speech modes interact. Research focused on how the representation of the semantic categories of relative position and relative size of a given referent in a narrative are divided between the speech and gesture modes, (Beattie and Holler 2002). There has yet to be research or studies done on other semantic categories to see if a divide holds, such a result would support the idea that speech and gesture form an integrated system of communication, each expressing semantic properties emerging from the same conceptualization.

This research examined the semantic category of shape, how the physical form of an entity is represented and observe whether the representations are partitioned between speech and gesture. This study focused on the semantic properties represented in iconic gestures. Iconic gestures refer to the same concrete idea as the co-occurring speech refers to but may provide additional information not expressed in that speech (McNeill 1992). Generally, it is observed that shape, as a semantic category, may be represented by gesture depending on its level of importance within the discourse.

II. Review of the Literature

McNeill (1992; 2008) has written that speech and gesture combine to convey thought, each mode performs its own functions, the two modes supporting each other. He argues that gestures are a part of language itself, the output of an integral part of the processing of language. McNeill states "gestures synchronize with speech at the point where speech and gesture co-expressively embody a single underlying meaning," (McNeill, 2005, p.1) to show that language requires two simultaneous modes to express thought. In a single instance where the gesture co-occurs with the speech, together they are co-expressive of the same thought unit. McNeill (2008) goes on to describe how gesture and speech differ: where gesture is global, the meaning of the parts determined by the meaning of the whole, speech is segmented, the meaning of the whole being determined by the parts. Where gestures are synthetic, speech is analytic – one gesture can compromise the meaning of an entire sentence while each segment of speech must be translated for meaning in order to combine the meaning into a whole (McNeill 2008). These contrasting modes of structuring meaning coexist and share the responsibilities for expressing meaning without seeming to overlap in function.

Gesture Classifications

McNeill (1992) identifies several types of gestures and makes the distinction between imagistic and non-imagistic gestures, imagistic gestures are interpreted as conveying an image of some kind, which can be anything from features of an object, an action, or an event. Non-imagistic gestures comprise simpler gestures such as those that point or simple rhythmic movements that serve pragmatic functions (McNeill 1992). Iconic and metaphoric gestures are considered imagistic gestures while deictic gestures and beats are classified as non-imagistic gestures (McNeill 1992). Iconic gestures are defined by McNeill (1992) as gestures that can express additional (complementary) or the same (co-expressive) semantic information about the same event that the co-occurring speech is expressing semantic information about. They can be small and simplistic but more often they are large and complex, and typically performed in the central gesture space. Where iconic gestures represent concrete characters and objects in actions and events, metaphoric gestures represent abstract ideas, using the base of the gesture to refer to the abstract concept being discussed in the speech (McNeill 1992). Deictic gestures are

pointing movements that typically point to a place in the gesture space where a referent has been placed, while beats are pragmatic discourse markers that typically consist of two small, low energy, rapid movements of the fingers or hand, and are performed wherever the hands may be at the time, including resting positions (McNeill 1992).

Semantic Importance of Gesture

Cassell, McNeill, and McCullough (1999) argue that the semantic relationship between speech and gesture is taken into account by listeners, as listeners absorbed and were able to recount semantic information represented by gestures whether or not the information was also expressed in the co-occurring speech. The listener uses both speech and gesture to understand the narrative conveyed by the speaker (Cassell 1999). Their study shows that listeners gain semantic information from gestures, proving their claim that speech and gesture are integral parts of understanding. Gestures present one unified representation and can be processed as one unified representation (Cassell 1999). This study supports the idea that comprehension is greater when gestures are viewed with speech, as speech and gesture are shown to combine to form one unified meaning. Similarly, Beattie and Shovelton (1999) also finds that there is a slightly higher comprehension rate when gestures are viewed, compared to when speech is given without the accompanying gestures. Both studies provide a significant confirmation of the semantic function of iconic gestures.

As previously stated, iconic gestures can be co-expressive or complementary. They can either provide semantic information about the same aspect or a different aspect of a narrative, compared to the aspect represented by the co-occurring speech. An example of a co-expressive gesture would be one where the speaker says “he lunges for him” as his right hand that represents a cat moves quickly toward the left hand that represents a bird (Cassell 1999). The speech and gesture are referring to the same aspect, the action, but the work of representing the whole continues to be divided as the gesture showed the imagistic side and removed any ambiguity from the speech of who was lunging at who. An example of a complementary gesture would be one where the speaker says “she chases him out again” while making a holding motion and swinging that motion back and forth (McNeill 1992). While the speech presents the aspect of action taking place, the gesture expressed the manner in which the action is performed, namely

the motion is referring to the use of an umbrella as a weapon. While iconic gestures can be complementary or co-expressive, the harmonious function of the two modes remains intact, as information provided is never redundant, even from a co-expressive gesture there is still information to be absorbed from the gesture.

Complementary Modes

Beattie and Holler (2002) shows support for the theory put forward by McNeil and set out to demonstrate that speech and gesture were in fact complementary but separate modes, and their study focuses on whether there were observable divisions between the semantic features that were represented by speech or gesture. They choose to study two specific semantic features at length, relative position and relative size; however, they recognized that the categories were too broad and subsequently divided each into subcategories that accounted for each different observable kind of aspect. They collected their data by showing participants two Tom and Jerry cartoons and asked them to recount the narrative to the listener, who could not view the cartoon. The participants were video recorded, and their narratives were analyzed for speech and gesture.

Relative Position Semantic Feature

Relative position is broken down into four subcategories, agent-object, agent-instrument, object-instrument, and object-surrounding space, while relative size is broken down into agent, object, and instrument. The aspects of relative position merely represent where the two items are in relation to each other, an agent-object aspect would demonstrate what position the object is in relation to the agent –is the object behind the agent, to the left of them, above them, etc. The concept is the same for each aspect of relative position.

Beattie and Holler (2002) provide an example that contains the relative position aspects agent-object and agent-instrument simultaneously in the gesture. The speaker states “the mouse is holding some pointy thing, and he’s gonna sort of spike him.” The speaker, beginning with ‘gonna’ and ending with ‘him’, brings both hands in front of him at stomach height, clenches both into fists with the left right behind the right, and pushes them forward twice (Beattie 2002). There is no information in the speech that represents the position between the agent and the object, nor the agent and the instrument, but there is such information in the gesture. While the speech states who is holding what, the

gesture shows the listener that the agent is holding the instrument in front of them and that they are thrusting forward towards the object. This demonstrates that both the object and the instrument in front of the agent, where the instrument is between them, and the agent is acting upon the object using the instrument. This example was a general representation of a whole that showed agent-object and agent-instrument relative position aspects were entirely represented by gesture. In comparison the following examples will show that in regard to the relative position aspects object-instrument and object-surrounding space, it is the speech that is entirely responsible for their representation.

A general example of the relative position object-instrument aspect is where the speaker says, “the spike is struck in Tom’s tail,” as their right hand becomes a fist at stomach height and then is forcefully moved down before abruptly stopping at knee height (Beattie 2002). Here it is the speech that contains the information about the position of the object, the tail, and the instrument, the spike, conveying the spike has intercepted the tail. The gesture, however, contains only the information of the direction of the instrument, not what the instrument strikes. A similar example for the relative position aspect object-surrounding space is where the speaker states “the dog’s picked him off and he’s just holding him so that his feet don’t touch the ground,” as his right hand in a fist, rises to shoulder height with the arm fully extended (Beattie 2002). The speech informs the listener that the object is off the ground, suspended in the air while the gesture merely represents the holding of something, not necessarily where it is in the surrounding area. This is because the speaker remains seated as they perform their gesture, therefore it is difficult for gesture to demonstrate real space representations of the scene, as the speaker is gesturing in the narrative space, which is typically directly in front of their chest. As you can see, it is important to divide broad categories such as relative position into its discernable aspects, the study showed that half of the subcategories were represented solely by gesture and the other half solely represented by speech. They showed that there is a division of which aspects of relative position are represented by speech versus gesture, the agent-object and agent-instrument aspects being solely represented by gesture, and object-instrument and object-surrounding space being solely represented by speech.

Relative Size Semantic Feature

The relative size aspects do not necessarily have to be in relation to another item. The relative size of an agent, object, or instrument can be obtained by naming them through proper names, pronouns, nouns in general, etc. After the cognitive concept of the object or instrument is obtained through their name, the listener can infer size through common background knowledge and then they can relate that size to other items in the narrative.

Regarding the relative size subcategories, the next few examples will demonstrate the relative size of an agent, an object, and an instrument respectively; and provide the general representation for instances in Beattie and Holler's study. The example for agent begins with the speaker says "Jerry's gonna put this cue through his tail," while they raise their left hand up to chest height, make a fist which moves down forcefully before stopping abruptly above the speaker's lap (Beattie 2002). The speech does not denote any explicit size, but the size of the agent is represented linguistically by naming the character's identity in the context of the cartoon. A listener's background knowledge informs them that 'Jerry' is a mouse and therefore small in size relative to his surroundings, items, and other characters. The work performed by the gesture is the previously explained agent-instrument relative position aspect, as it shows that the agent is holding the instrument in front of them. A similar example is used for the relative size of an object, where the speaker states "the warden is carrying him by... by the scruff of his neck," and the speaker's hand rises in front and to the right of their body and makes a fist as they move it quickly up to shoulder height (Beattie 2002). The speech once again gives the size of the object through linguistic representation but this time the object is represented by a pronoun instead of a proper name. When the speaker uses 'him' and 'his' the listener knows from the context of the conversation they are referring to Tom, and the background knowledge of the cartoon informs them that Tom is a cat, which provides his size relative to the size of the warden, an adult human. Meanwhile the gesture provides the agent-object aspect of relative position, where the speaker as the warden's hand and arm shows the motion of picking up Tom the cat from in front and slightly to the right of them. For the relative size of an instrument, the linguistic representation continues, as the speaker says "Jerry will prod him with... with a pen or something that he's got, with a sharp instrument, either in his toe, or maybe in his tail, or

poke him in the bum or something," while their right hand is a fist, swung up vertically in front of them, and then brought down again where it abruptly stops (Beattie 2002). The speech identifies the instrument as a pen, or pen-like, which gives its relative size compared to Jerry, and as the listener's background knowledge informed them of the size of Jerry, the cognitive concept of 'pen' informs them similarly. Once again, the gesture gives a relative position aspect, this time of agent-instrument, and where the agent is holding the instrument in front of them.

Relative Size Exception. There was one exception in Beattie and Holler's study where a gesture seemingly represented information about size, otherwise size was solely represented by speech. This instance was when the speaker says, "but that spike is still in his tail," while the speaker's right hand with index finger extended rises in front of them and moves from lap height to shoulder height and down again immediately following the same path back to lap height (Beattie 2002). They believed that this motion could be representing the dimensions of the spike and therefore giving information about the instrument's size, but the gesture continues for when the hand with the index finger still extended reaches the original starting point, it is moved around the body until it reaches behind the speaker and points at the base of their back (Beattie 2002). It was imperative they resolve this exception in order to support the theory of two complementary but separate modes, for if they could not show a division of labor in representation between speech and gesture, they could not uphold the basis of that theory. They original tried to account for the exception by explaining the viewpoints gestures take, and how this particular gesture was a duel viewpoint gesture (Beattie 2002), but after more investigation a new article came out with a more concise justification.

Relative Size Reclassification. Beattie and Shovelton (2006) found size information is reliably represented by imagistic gestures and set out to determine how important the instances of size represented by gesture were compared to those represented by speech. They determined that gestures receive the most significant aspects of the information being conveyed in a narrative and represent them accordingly, and that they did so in an acceptably unambiguous manner where listeners received the information (Beattie and Shovelton 2006). They observed that the level of importance of the size information of any given item would determine whether it was represented by

speech or gesture, no matter if the item in question was an agent, object, or instrument. It becomes apparent that relative size divided into agent, object, and instrument aspects would not yield a clean division of the representation between speech and gesture. Instead dividing relative size on the basis of relative importance to the topic or narrative being discussed better determined by which mode size was represented. They have a clear example where the speaker states "he just makes a ball out of the two," as they bring their flat hands wide apart, with the palms facing toward each other, positioned at 2 and 8 in front of them, forming a round shape (Beattie and Shovelton 2006). This ball's size was determined to be among the most important in the narrative by five observers, and the gesture effectively and efficiently represented the object's size. The size of the items of lesser importance were represented in the speech through the ways that were demonstrated by previous size examples from Beattie and Holler. This subsequent reclassification of relative size in Beattie and Shovelton's 2006 research demonstrates how important it is to determine the correct division of aspects for the broad semantic feature categories. Each gesture and speech segment can represent multiple semantic features at once. Until such a time when all semantic features are studied individually there will not be a clear answer when attempting to determine definitive conclusions concerning the division of labor between the speech and gesture modes.

III. Current Study

Beattie and Holler (2002) demonstrated a way forward in determining what representational work is done by speech versus what is done by gesture by taking semantic features and researching them individually. They began with the study of two semantic features, relative position and relative size, and by doing so have shown that a divide between representation by speech and gesture can be determined. This paper contributes to this area of research through the observation of a third semantic feature, shape, in order to determine whether the representation of information is divided between speech and gesture; if so, what motivates shape information to be represented by speech or by gesture.

These are the research questions this paper will attempt to comment on. How will other semantic features, such as shape (etc.), break down to show the same divide between gesture and speech that can be observed between relative position and relative size? How will observations of previously studied semantic categories, such as relative position and relative size, compare to previous results? In what ways will the continuation of the study of semantic features prove or disprove the idea that gesture and speech modes present as complementary?

IV. Methodology

4.1 Participants

The two participants in this analysis are friends of the author - and friends themselves. They are therefore familiar with each other and comfortable conversing with one another. They received a meal each in exchange for their consent to record their natural conversation for a predetermined amount of time. When the initial recording failed partway through, both participants consented to be filmed again for the previously established amount of time. Both participants were in their twenties at the time of filming. The native language of both participants is English. They were not informed of the topic of the study in order to produce the most natural data.

4.2 Data Collection

The recordings were audio-visual and therefore recorded both their voices as well as the majority of their person; at any given time between the two recordings, only the calves are partially out of frame. All the hand gestures were identified and counted and the speech accompanying them analyzing for semantic content. The first recording was made at Erica's home on an iPad positioned on a dresser aimed at a bed where the participants sat and conversed. The recording captures only twenty-two minutes which was not enough to capture sufficient data. The second recording was made at Matt's home, on an iPhone positioned on a table aimed at a couch where the participants sat to speak to one another. This recording successfully recorded an hour and twenty minutes and seven seconds. The participants engaged in natural talk, having a natural dialogue or what is called naturally occurring conversation. The study benefits from analyzing natural conversation because only through studying naturally produced speech and gesture can researchers discover how they come together to create meaning. It allows researchers to put in context the speech and gestures they observe and thus provides the opportunity to potentially observe and describe the processes taking place (Eggins 2000). The participants were instructed to partially face the camera and partially face each other in each video, in order to see each other while conversating and so the camera was able to record their gestures.

4.3 Coding of Gestures

All the hand gestures in the videos were identified according to McNeil's gesture classifications. The gestures were determined to be either imagistic or non-imagistic, and the non-imagistic, deictic gestures and beats, were counted, but not analyzed for meaning as they are unrelated to the aims of the current study. The imagistic gestures were counted, and each was determined whether it had an observable semantic feature it was representing. Any gesture that conveyed information about shape was pulled out, examined, and compared to the speech.

Imagistic		**Non-Imagistic**	
Iconic	Represent aspects of speech, additional or co-expressive, concrete	Deictic	Pointing gestures, often point to referent gesture space
Metaphoric	Creates a base for a spoken concept, abstract	Beats	Small, low energy, rapid, repetitious movements of fingers or hands in time with speech

In keeping with the same criteria as Beattie and Holler, the event or object each imagistic gesture represented was identified and the gestures examined for their semantic content. Specifically, the semantic feature they may be representing, and they were compared to the speech they accompanied. The semantic information being represented by the gesture was determined by examining the speech and the hand shape of the stroke – particularly the positions of the fingers and palms relative to each other.

4.4 Coding of Speech

The speech throughout the videos was reviewed and all referring expressions with physical objects that covertly expressed shape information were counted. These speech items covertly include object's shape based on background knowledge. The speech that cooccurred with the gestures was analyzed for their semantic content, i.e. what semantic feature it may be representing, and the beginning and end of each of the gestures was identified where it lined up with the co-occurring speech down to the exact frame,

providing temporal relation. The speech was evaluated for its semantic meaning compared to the semantic meaning of the accompanying gesture and analyzed for repetition or standalone aspects that were represented exclusively by one or the other, giving semantic relation and function. The utterances that occurred before and after the co-occurring speech was given to provide the speech and gestures more semantic context. As with the Beattie and Holler study, the utterances and the semantic units were taken into consideration in that the same utterance that accompanied the gesture was considered, as well as any neighboring utterances with relevant verbal information so long as they came directly before or after the utterance with the lexical affiliate and didn't contain a new iconic gesture. The motivations behind the speech were speculated on to draw conclusions. Any spoken expression that conveyed information about shape was identified and checked to see if it occurred with a gesture. If a gesture accompanied, it was examined for whether it also conveyed information about shape.

V. Results

Table 1 shows the number of each type of gesture identified in both of the recordings. The participants used thirteen iconic gestures, ten metaphoric gesture, eight deictic gestures, and seven beats gestures. Any other visible movement was too ambiguous to be determined and are assumed to be non-gestures such as self-touching and object manipulations.

Table 1. *The number of identifiable gestures.*

Imagistic		**Non-Imagistic**	
Iconic	13	Deictic	8
Metaphoric	10	Beats	7

Table 2 shows the number of iconic gestures whose semantic information conveyed shape, relative position, or relative size. Across the two recordings: three iconic gestures convey shape information, five convey information about relative position, and two convey information about relative size. Non-imagistic gestures are not semantic in nature, as previously stated non-imagistic gestures tend to serve pragmatic functions.

Table 3. *The number of imagistic gestures that convey semantic information.*

	Shape	**Relative Position**	**Relative Size**
Iconic	3	5	1

Table 3 shows the number of speech utterances that express a physical object where shape information could be inferred. It also shows the only time that an utterance such as this occurs with a gesture that also expressed shape information. There were forty-one instances of utterances of physical objects that covertly expressed shape information, and only one where the utterance had a gesture accompany it that also conveyed shape information.

	Utterances	**Utterances w/ Gesture**
Speech Covertly Expresses Shape Information	41	1

There were no instances where shape information was found to be overtly conveyed by the speech. The recording was reviewed multiple times, but no speech concerning shape occurred – not even words such as ‘circular’, ‘round’, ‘rectangular’, ‘square’ etc.

VI. Discussion

In this section there will be attempts to interpret the results and speculate on their relation to the research questions at hand. There will be a general discussion of the results, and an in-depth discussion of three instances where the gestures expressed information about shape in order to determine if how the representation of shape information is expressed through speech and gesture is divided. Then a short analysis of the previously studied semantic categories, relative position and relative size, as well as others.

The speech and the gesture of the following examples will be analyzed for their semantic content concerning shape. There may be other semantic information that is expressed, but for a small inquiry such as this, there must be a limit to how many variables to focus on at a time. Other semantic information may be noted, but it is whether shape is expressed by speech or gesture that will be the main analysis going forward. As stated previously there is a focus on iconic gestures as they are the type of gestures found expressing the relevant semantic information.

In the parent study, relative position was broken down into agent/object, agent/instrument, object/object, object/instrument, and object/surrounding space. The first three aspects were exclusively represented by gestures, the following two solely represented in speech. The relative size semantic category was determined to be represented by gesture or speech depending on the importance of the information regarding the narrative. The data showed that the division of representation regarding shape information likely followed a similar path to that of relative size than that of relative position as there appears to be some narrative related aspect that triggers the shape information to be expressed by gesture. Each instance of conversation has been given context as to the topic, an image of the gesture itself, a description of the speech relative to the gesture, and then the speech and gesture are analyzed for semantic content and speculated upon.

Gestures

Macaron Ice Cream Sandwich Gesture

In this first gesture example, the speaker identifies the item as a macaron ice cream sandwich, and the accompanying gesture forms a partial shape, the whole of which is easy to infer. The speaker is describing the macaron ice cream sandwiches from a restaurant that the listener has not been to.

Figure 1. The macaron ice cream sandwich gesture.

1. 6:48:28 6:49:06

S: “And they have the little macarons, so you can have a [macaro]n ice cream sandwich. A fresh one.”

The speaker’s hand hovers in an open C-shape gesture: fingers together, thumb out, palm down. They hold this gesture and then returns their hand back to their head and plays with their hair, a self-touch non-gesture.

This gesture denotes the partial shape of the object, the macaron ice cream sandwich, which is not explicitly conveyed in the speech. The speech itself identities the object as a macaron, or more specifically a macaron ice cream sandwich. As the gesture denotes the shape of a half circle, and circles are symmetrical, the other half of the shape is easily filled in. The only referent in the speech the gesture could be related to is the mention of the macaron. Here the speech is covertly conveying shape information, but it is of a prototypical macaron, which by the speaker's clarification of the item as a macaron ice cream sandwich, the macaron being referred to is atypical. A listener with a referent for the word macaron may have an idea of what the object looks like, but the gesture itself gives a certainty to its shape that the speech does not give. Despite the shape being implicit in the meaning of the word, the shape information might be conveyed in the gesture because the speaker is referring to an atypical specimen and therefore likely considers the referent to be unfamiliar to the listener. The shape information may be

necessary to express in a gesture for classification as the traditional macaron is filled with ganache, buttercream, or jam, and this macaron is filled with ice cream.

Dark Chasm Gesture

In the second example, the speaker describes a chasm and gestures the shape of the body of the chasm. A chasm is defined as 'a deep fissure in the earth, rock, or other earthly surface' which conjures a vague mental image of a narrow opening in the ground. The gesture shows the chasm refers to exists at a specific angle whose direction is diagonal, a detail about the shape that is solely expressed through the gesture.

The context for the topic is the participants are discussing the environment of a Dungeons and Dragons session they took part in earlier that week. The chasm is part of a cave the characters entered where the chasm is on one side of the cave and across on the other wall is a ledge. The chasm was wide enough for the characters to fall into, and long enough that it spanned the length of the cave, as far as the characters could see. The gesture arises as the listener expresses a contrary opinion and the speaker affirms their position that the chasm was a dark, confined space.

Figure 2. The dark chasm gesture.

2. 18:20:13

S: "If you guys had jumped down the chasm it would have been so over."

19:04:32

S: "I said the answer lies in darkness, in confined space."

19:06:20

L: "I don't know if a chasm is a confined space."

19:11:05 19:11:17

S: "the chasm is da[rk" /paus]e/ "it's confined space."

The speaker's hands are in a resting position until they come up out in the space in front of them where they place them a couple inches apart, flat palms facing in towards each other, at an angle. Their fingers and thumb are straight with their fingers together and thumbs out, all are flush with the palms; they hold this gesture position until further speech prompts a new gesture.

This gesture provides the shape of the angle at which the chasm is orientated in the earth, the shape of the object, as well as the idea that it is longer than it is wide. This information is not expressed in the speech, and not given by the definition of the referent. The speech continues to describe the chasm as both dark and confined but never conveys information about shape. The listener can clearly see that the shape of the chasm is at an angle. The reason to convey this information in the gesture could be because the listener and speaker have identified a misunderstanding between them of what constitutes a chasm. The speaker is clarifying after the listener has expressed an opposing opinion.

External Battery Gesture

In the last gesture example, the speaker describes an external battery pack they were thinking of purchasing, and the accompanying gestures performs both a semantic and a pragmatic function. The speaker performs a lexical retrieval for the word 'external' with superimposed beats over an iconic gesture that gives the battery's shape.

Figure 3. The external battery gesture.

3. 38:32:18 38:34:08

S: "one of those Pikachu, uh,[batteries, ex]ternal batteries"

The speaker's right hand comes up from the lap first, followed quickly by their left hand, the palms face each other while the middle, ring, and pinky are curled closely to the palms. The thumbs are out in a grip-like position while the index fingers are lightly

curled but separated from the rest of the fingers, held out further away from the palms. They pull the index fingers and thumbs towards and away from each other several times.

The speaker performs five superimposed beats on top of the iconic gesture, as the placement of the thumbs and fingers form a rectangular shape. Beats have been classified as pragmatic discourse markers that typically consist of two small, low energy, rapid movements of the fingers or hand, and are performed wherever the hands may be at the time, which seems to include if the hands happen to be performing a gesture. The beats in this example are performing the pragmatic function of assisting lexical recall, as the speaker searches for the word batteries. In lexical retrieval, often a speaker can picture the item they are trying to name before they recall the word for it. This gesture provides the listener with the shape of the battery that the speaker is remembering to have looked at. The speech does not provide any shape information regarding the referent, as the prototypical battery is small and cylindrical. As clarification seems to be connected to the previous gesture examples, the pragmatic lexical retrieval function of beats is connected to this example.

Further Analysis

With the data available here, it is difficult to say anything definitive about how shape is represented. Speech is certainly capable of representing shape semantic information, as is gesture, demonstrated by these examples, so the question remains of how the work of representing is partitioned between the two.

There were very few examples of iconic gestures in the data set, and only three encoded information about shape. This could be because it may be necessary to have an additional process taking place, clarification or lexical retrieval, for shape information to be expressed by gesture; or because there is an abundance of shape information being expressed covertly by speech. While the intention was to only study shape, the decision was made to look more closely at all the iconic gestures in the data set to see what information was being expressed and whether it was also expressed by speech. This includes the gestures that encoded information about shape, as they may have included other semantic properties in the gesture or speech. The semantic categories the gestures were analyzed for come from Beattie and Shovelton (1999), in order to determine

whether gestures contributed to comprehension, they had to catalogue the types of semantic information encoded in the iconic gestures. The categories included but were said not to be limited to the direction of movement, orientational movement, rotational movement, objects making contact, manner of an action, the speed of an action, the specific identity of an object, the shape of an object, the relative size of objects, the relative position of objects, and the number of objects. The instances of relative position and relative size were compared to previous research.

Across all the iconic gestures, shape was encoded in three, relative position was encoded in five, relative size was encoded in one, and manner was encoded in four. In the iconic gestures' accompanying speech identity was encoded thrice, and relative size was encoded once. It can be said that identifying the object in the speech can give similar semantic information that is being encoded in the gesture, but in each of the examples, the gesture is giving more detail than the identity does, and it cannot be said that the understanding of the referent identified is shared between speaker and listener. Clearly both participants have a different idea of what constitutes a chasm, and that is why there was need for clarification.

There is a gesture accompanied by speech that refers to turning down the volume of a television, and in the gesture the pointer and the thumb of the right hand comes up close together and rotated creating a small circle. This action shows the listener that the speaker is turning down the volume using a dial button on the television itself, and from the closeness of the digits on the hand, it can be assumed that the dial button is very small. In comparison, regarding the macaron ice cream sandwich gesture, the speech does provide the size of the macaron, qualifying it as 'little' which matches with the identity given to the object. A macaron is typically small, therefore not only does the semantic information size come from the semantic information identity but also directly from the speech itself. As seen in Beattie and Shovelton (2006) size information most important to the narrative is encoded in the gesture, therefore it can be assumed that since the size information is being encoded in the speech, it is of some lesser importance to information of the shape, which is being expressed by the gesture. It is possible that like size, shape information is represented on a basis of importance to the narrative, where when an

additional process is taken place, like clarification or lexical retrieval, it is encoded in the gesture.

Among the iconic gestures there were three instances of relative position that represented agent-instrument aspect that were solely represented by the gesture. There was a gesture where the speaker identified a great sword and held out a closed fist in front of the body, demonstrating the holding of said sword. Another time the speaker discussed the action of ‘smashing people’ and gesture the holding of both, sword as described above and shield where the opposite forearm was held out in front, side of arm facing out. The speaker moved left to write, indicating a shoving using the shield and the sword arm. Lastly there was an instance where the speaker identified a selfie stick and held a fist above their head and moved it right to left while looking a foot above the fist, miming taking a picture. These three iconic gestures all demonstrate aspect of relative position agent-instrument. This aspect was found by Beattie and Holler (2002) to be solely represented by gesture; a result we see repeated here. The accompanying speech of all three gestures did not express information regarding the relative position of any agent to any of the mentioned instruments. There were no instances where speech and gesture explicitly represented the same semantic category at the same time, which falls in line with the theory that speech and gesture are two distinct but complementary modes.

VII. Conclusion

This paper set out to study the semantic category shape in an attempt to test the idea that speech and gesture are two distinct modes of thought that work complementary with each other. This micro study observed that shape can be expressed by gesture, and it is known that shape can be represented by speech either overtly or covertly. It is possible that future research could see a division of aspects of shape in some way, albeit pertaining to aspects not seen here, but conclusions were not able to be made here. The observations made here, are in line with the previous research on relative position and relative size, where speech and gesture are two distinct modes, each responsible for conveying different types of information. In this case, it could be possible that gesture is representing shape information because it is more relevant or important to the discourse due to the need for clarification or lexical retrieval. It would do to research further to see if shape information is consistently expressed in gesture when there are other discourse functions or pragmatic functions being performed.

The potential for future research is high, as it would be beneficial to study not only to study shape on a larger scale with bigger data sets but to study other semantic categories in depth as well. It is possible that the number of tokens was small due to the possible prerequisite of an additional process of some kind and that shape is often expressed covertly through speech. Further research could help confirm the theory that speech and gesture are two distinct, but complementary modes that work in tandem with each other to process and express internal speech. Further research regarding shape could reveal the motivations to encode the information in gesture.

Bibliography

Beattie, Geoffrey, & Shovelton, Heather. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, *123*(1-2), 1–30. doi: 10.1515/semi.1999.123.1-2.1

Beattie, Geoffrey, & Shovelton, Heather. (2006). When size really matters: How a single semantic feature is represented in the speech and gesture modalities. *Gesture*, *6*(1), 63–84. doi: 10.1075/gest.6.1.04bea

Cassell, Justine, McNeill, David, & McCullough, Karl-Erik. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & cognition*, 7(1), 1-34.

Eggins, Suzanne. (2000). Researching Everyday Talk. In Researching Language in Schools and Communities (pp. 130–151). By Len Unsworth, Cassell.

Holler, Judith, & Beattie, Geoffrey. (2002). A micro-analytic investigation of how iconic gestures and speech represent core semantic features in talk. *Semiotica*, *2002*(142). doi: 10.1515/semi.2002.077

Krauss, Robert M., Chen, Yihsiu, & Gottesman, Rebecca F. (2000). 13 Lexical gestures and lexical access: a process. *Language and gesture*, 2, 261.

Mcneill, David. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350–371. doi: 10.1037/0033-295x.92.3.350

McNeill, David. (1992). Chapter 3: Guide to Gesture Classification, Transcription, and Distribution. In *Hand and Mind: What Gestures Reveal about Thought* (pp. 75–104). Chicago: University of Chicago Press.

McNeill, David. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

McNeill, David. (2008). *Gesture and thought*. University of Chicago press.