THE RELATIONSHIP BETWEEN

REPRESENTATIONAL, BEAT, AND COHESIVE GESTURES

AND SPEECH DISFLUENCY

IN THE FIRST 2012 PRESIDENTIAL DEBATE

by

KATERINA J. PUZINAUSKAS

TONY BUHR, COMMITTEE CHAIR
DANA PATTON
MARCIA HAY-McCUTCHEON
ANGELA BARBER
SPYRIDOULA CHEIMARIOU

A THESIS

TUSCALOOSA, ALABAMA

2017

ABSTRACT

Hand gestures are yoked to speech in a highly-correlated system, often referred to as co-speech (Hostetter, 2011). Both disfluency and gesture have been show to influence audience reaction during political debates (e.g., Pennebaker, Mehl, & Niederhoffer, 2003). This study examined the relation between speech disfluency and hand gesture in the First 2012 Presidential Debate between President Barack Obama and Governor Mitt Romney. Specifically, the frequency and type of speech disfluencies and gesticulations generated by candidates were identified, annotated, and compared. Speech and gesture variables were examined across multiple speech domains, including utterance and speaking turn. Results showed a correlation between the level of speech disfluency and the type of gesture within a set speaking turn for both speakers. Overall, findings suggest 1) associations between speech disfluencies and the five taxonomies of gestures (iconic, metaphoric, deictic, cohesive, and beat, and 2) relationships between variables and audience-reported outcomes for debate success.

DEDICATION

I would like to dedicate this research to President Barack Obama and Governor Mitt

Romney. Thank you for putting on a good show.

LIST OF ABBREVIATIONS AND SYMBOLS

df Degrees of freedom: number of values free to vary after certain

M Mean: the sum of a set of measurements divided by the number of measurements in the given set

SD Standard deviation: the extent of deviation of a population, expressed as a quantity

p Probability: associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value

z Computed value of t test

SE Standard error: the statistical accuracy of an estimate, equal to the standard deviation of the population distribution

$\geq$ Greater than a given value

$\leq$ Less than a given value

ACKNOWLEDGEMENTS

First, I would like to express sincere gratitude to my advisor Dr. Buhr for the continuous support of my research. Without your patience, knowledge, and guidance, none of this would have been possible. To my Thesis Committee, thank you for your resolute assistance and counsel. I would like to thank the incredible researchers and coders who dedicated their time and knowledge to this project, especially Elizabeth Wyatt and Michaela McLaughlin. Finally, I am so grateful for my family and friends who provided unwavering encouragement and understanding throughout this process.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

When people talk, they often make unrehearsed movements of their arms and hands. These movements shift harmoniously with what the speaker is saying to create a cohesive message. Commonly, these hand motions are referred to as gestures. Gestures are defined by A.M. Zeiss as "non-practical, nonlinguistic discrete body movement that arises from the intention to express something, and that encodes some aspects of what is intended to be expressed, and/or how the accompanying speech should be interpreted" (1998). This classification encompasses an immense array of movements of the hands, arms, and body, and includes the multifaceted relationships between speaker and listener; therefore it is unsurprising that the precise classification of gestures and other non-verbal behaviors is still debated (Efron, 1941; Ekman & Friesen, 1969; Zeiss, 1998). David McNeill offers a simpler definition, calling gestures "idiosyncratic spontaneous movements of the hands and arms accompanying speech" (1992, p. 37). Although gestures often present an identical or complementary message to the verbal utterance they accompany, they can also depict information that differs from spoken speech (McNeill). Thus, the study of gestures is integral in understanding gestalt communication between speaker and listener.

Gesture

As a visual medium which complements language, gestures exist in both dimensions of time and space. Gestures follow the sequential timeline of an utterance, called the linear-temporal framework. As physical movements, they also occur within a three-dimensional space

(McNeill, 1992; Kendon, 1980). Spatially, a prototypical gesture has three main phases: the preparation, the stroke, and the retraction (McNeill; Kendon). During the preparation phase, the hands leave the resting position and move out and away from the speaker's body. The stroke is the principal portion of a gesture, and is the only obligatory phase (McNeill; Kendon). During the stroke, the meaning of the gesture is expressed in congruence with the verbal information it references. A typical stroke falls within the boundaries of the waist, shoulders, and elbows (McNeill). Finally, the retraction phase occurs when the hands return to the resting position generally in front of the center of the body. When considering the linear-temporal relationship between gestures and discourse, gestures can be broadly classified into two taxonomies: representational gestures and beat gestures (Hostetter & Alibali, 2007). Representational gestures are hand and arm movements that diametrically reference the semantic content of an utterance. They typically have all three-gesture phases (preparation, stroke, and retraction). Representational gestures can be further dissected into three distinct groupings: iconic, metaphoric, and deictic (Hostetter & Alibali; McNeill). Iconic gestures physically portray tangible actions or objects; metaphoric gestures illustrate abstract ideas; and deictic gestures consist of pointing (generally with the index finger) to a referent, which can be either concrete or conceptual (McNeill; Terken, 1991).

In contrast, beat gestures contain no semantic content (McNeill, 1992). These gestures involve simple, repetitive flicks of the hand that coincide with the melodic line of speech (McNeill; Terken, 1991). Whereas iconic, metaphoric, and deictic gestures have ties to semantic content, beat gestures have no referents; their form remains the same, even as the content of the verbal message shifts (McNeill). Characterizing of over half of all gestures used (McNeill), beat

gestures are theorized to provide emphasis to specific words and phrases. Unlike representational gestures, beats only have two movement phases: in/out or up/down (McNeill).

The final taxonomy of gestures, which ties together thematically-related parts of discourse, is cohesive. Cohesive gestures are comprised of any combination of representational and/or beat gestures, and are often observed within the realm of political speech (McNeill, 1992). Cohesion occurs when a gesture form is repeated in order to call attention to the continuity of a topic within discourse. For the purposes of this thesis, gestural focus was restricted to McNeill's broad gesture taxonomy consisting of representational (iconic, metaphoric, deictic), beat, and cohesive gestures.

Co-Speech: A System of Language and Gesture

Gesturing is ubiquitous in human communication; people of all ages, ethnicities, and cultures gesture (Alibali, 2005; Driskell & Radtke, 2003; Feyereisen & de Lannoy, 199; Goldin-Meadow, 2000; McNeill, 2000). The act of gesturing may have even laid the foundation for the development of language over biological and cultural evolution (Corballis, 1999). Several models propose that language emerged from communicative hand movements, and that vocalized communication is a more-recent development in hominid evolution (Corballis). While this theory is still debated, the physical act of gesturing has been undeniably linked with the overarching language system (Goldin Meadow, 2003; Goldin Meadow, Kim, & Singer, 1999; McNeill, Cassell, & McCullough, 1994). Gestures are ubiquitous in face-to-face interaction, and represent a form of movement that is yoked to verbal utterances (McNeill, 1992; Williems & Hagoort, 2007). In fact, it is possible that a true gesture cannot be independent of speech (McNeill). This highly-correlated system of gesture and language is often referred to as co-speech (Hostetter, 2011; Marstaller & Burianova, 2014; William & Hagoort). The study of co-

speech has resulted in several hypotheses in which the construction of language and gesture falls under a single, cohesive system (McNeill). These models imply that gesture affects speech comprehension and production at all stages of processing (Kelly, Barr, Church, & Lynch, 1999; Kelly, Kravitz, & Hopkins, 2003; McNeill). However, distinct archetypes exist to describe the formation of representational gestures versus beat gestures.

Representational gestures often function as co-speech acts (Austin & Sweller, 2014). It has been hypothesized that the pairing of a representational gesture and a vocal utterance within a co-speech act is demonstrative of the larger communication hierarchy (McNeill, 1992). Proposed by Hostetter and Alibali (2008), Gesture as a Simulated Action (GSA) framework postulates that cognitive processing of representational co-speech acts simultaneously arouses the motor and visuo-perceptual systems in speakers and in listeners. As follows, the listener receives a direct mental image of the speaker's message contained within the vocalization and the accompanying representational gesture, contributing to increased listener comprehension (Alibali & Hostetter, 2010; Hostetter & Alibali, 2008).

While the GSA framework provides a comprehensive model of representational gestures, it does not account for the role of beat gestures within cognition and communication (Austin & Sweller, 2014). Unlike representational gestures, the majority of beat gestures are co-speech acts; they are thought to facilitate listener-speaker communication through emphasis of pertinent information encoded in verbal utterances, rather than through the stimulation of the GSA framework (Hostetter & Potthoff, 2012). According to McNeill (1992), "beats reveal the speaker's conception of the narrative discourse as a whole. The semiotic value of a beat lies in the fact that it indexes the word or phrase it accompanies as being significant, not for its own semantic content, but for its discourse-pragmatic content" (p. 15). Despite differences, these

dichotomous classes of gesture are both considered directly communicative, where the speaker's message is directly fixed within the co-speech movement (Grice, 1957).

The Effect of Gesture on Speech Production and Comprehension

The relationship between gesture and language is inherently complex, and not fully understood. It is necessary to note that there is a body of evidence which contradicts the "gesture as direct communication" model (Kelly, Kravitz, & Hopkins, 2003; Krauss, Morrel-Samuels, & Colasante, 1991). That is, some researchers propose that gestures are produced to facilitate the production of language, and that listener comprehension is not affected by a speaker's gestures (Krauss et al.). Thus, some forms of gesture may be regarded as indirectly communicative within the academic community (Kelly et al.).

Still, most scholars agree that the meaning of an utterance is not limited solely to the speaker's message and is encoded within the communicative act of co-speech (Kelly et al., 1999, Willems & Hagoort, 2007). As an example, consider the ambiguity of figurative language. The intended meaning of a metaphorical utterance cannot be captured by the speaker's content alone; rather, the listener depends on the information provided by different communicative modalities in order to construct the speaker's intent. For instance, consider the phrase "I saw a man on a hill with a telescope". In isolation, this sentence is linguistically-vague. The listener would have to decide if the speaker saw a man holding a telescope, or if the speaker used a telescope to look a man standing on a hill. In order for the speaker to effectively convey intended meaning to his audience, the semantic and morphologic data provided by the utterance must be augmented by additional information. Gesture is an energy-and time-efficient way to provide clarification. Face-to-face communication is a cooperative, fluid exchange in which speaker and listener try to deduce the significance behind each interaction. Enrici, Cappa, Bara, and Tettamanti (2011)

affirm that "people use communication as a social action to affect and modify others' mental states using both linguistic and gestural modalities" (p.2415). Consequently, it is suggested that the information encoded within gestures is pertinent to the comprehension of figurative language. Cornejo et al., (2009) examined the processing of metaphors and found that verbal and gestural information need to be integrated simultaneously in order to successfully discern metaphorical meaning.

Several bodies of research indicate that co-speech plays a significant function in speech production (Kelly et al., 1999; McNeil, 1992) and that speaker communication directly benefits as a result of gesture use (Maricchiolo, Gnisci, Bonaiuto, & Ricca, 2009). In fact, recent studies have revealed that gestures enhance cognition and expressive and receptive language across a multitude of contexts, including: learning and memory (Hostetter, 2011; McNeill), problem-solving (Goldin-Meadow, Wein, & Chang, 1992), narrative discrimination (Driskell & Radtke, 2003; McNeill, Cassell, & McCullough, 1994), and lexical processing (Thompson & Massaro, 1994). Gesturing also decreases the cognitive constraints on verbal working memory during language production, resulting in more effective communication (Gillespie, James, Federmeier, & Watson, 2014; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).

Various studies have found that gestures locked into a co-speech framework enhance listener attention and comprehension (Valenzo, Albali, & Klazky, 2003). According to von-Raffler-Engel, "eliminating the visual modality creates an unnatural condition which strains the auditory receptors to capacity" (1980, p. 235). Co-speech augments listener knowledge of a speaker's message in order to reduce strain on the auditory comprehension system (von-Raffler Engel, 1980). It has been demonstrated that listeners will retain information presented only through the modality of gesture, resulting in better narrative comprehension (Alibali, Flevares, &

Goldin-Meadow, 1997). Austin and Sweller (2014) argue that gesturing indirectly supplements the speaker's verbalized meaning, resulting in better listener attention and rapport. This is accomplished through co-speech gestural imitation, where increased frequency of gesture by the speaker is proportional to increased frequency of gesture by the listener in response (Holler & Wilkin, 2011; Austin & Sweller).

While the overt benefit of co-speech during face-to-face communication is well documented, there is a body of evidence that suggests additional discrete communicative benefits of co-speech which occur in the absence of a communicative partner. Alballi, Health, and Meyers (2001) found that speakers will use gestures in the absence of an audience. These findings imply that gestures do not simply enhance the speaker's communicative content for the listener. Rather, it would seem that they play a role in speech production itself. Additionally, speakers who are blind from birth also gesture when they speak (Iverson & Goldin Meadow, 2001). Interestingly, results were similar when the blind speaker interacted with other blind conversational partners and with sighted speaking partners. This finding is significant because it implies that gestures are not exclusively social, learned behaviors, and indicates that they may play a direct role in speech production. Finally, gestures occur at the same rate when the speaker is talking over a telephone as opposed to a speaking partner (Rimé, 1982). It can be concluded, therefore, that gestures do not solely facilitate direct communication, but rather serve an additional, indirect communicative function.

Gesture and Fluency

Given the intricacy of human communication, it is unsurprising that speakers make occasional, unplanned errors when the talk; these speech disfluencies, which include false starts, repetitions, revisions, unfilled pauses, and filled pauses (e.g. "um", "uh"), occur in up to 10% of

all spoken words (Fox Tree, 1995; Shriberg, 2001). Most experts agree that speech disfluency that falls between one and three percent of all spoken words is within normal limits for adult speakers (Lutz & Mallard, 1986; Webster, 1980). Disfluency that exceeds ten percent is considered severe (Webster). Among the most common types of disfluencies for non-stuttering adults are interjections of word fillers (e.g. "I'll meet you at, *um*, at five"), and revisions (e.g. "*I shaw*- I saw Suzie") (Johnson, 1961; Lutz & Mallard). Pauses, both filled and unfilled, are also a universal characteristic of naturalistic speech (Krauss, 1998). Speech disfluencies appear to be more prevalent during spontaneous speech (Shriberg) and stressful speaking situations (Lass, 2014). Considering the grammatical and semantic links within co-speech, gestures likely contribute to the overall fluency of speech (McNeill, 1992). The Lexical Retrieval Hypothesis postulates that gestures assist the construction of verbal utterances (Alibali, Kita, & Young, 2000; Hostetter et al., 2007; Krauss and Hadar, 2001). According to this model, speakers use representational gestures to articulate their mental constructs (Hostetter et al.; Kita, 2000; Krauss, Chen, & Chawla, 1996). Finlayson, Forrest, Lickley, and Beck (2003) found that an inability to gesture due to hand restriction increases the number of disfluencies in elicited speech. Additionally, Pine, Lufkin, Kirk, and Messer (2007) found that speakers were more likely to retrieve words during tip-of-the-tongue situations when they were able to freely gestures as compared to speaking situations in which hands were bound. Hostetter et al. suggest that speaker's vocabularies become less prolific when they are unable to use their hands to gesture. Finally, Krauss (1998) found a statistically-significant correlation between filled pauses and hand restraint, suggesting a relationship between disfluency and lexical retrieval. However, empirical evidence on the correlations between gesture and fluency is inconclusive.

Gesture, Fluency, and Political Speech

Nowhere are gesture and fluency more important than in political speech, where small details of one's behavior have important effects on listener perception, and can be used to elicit reactions from an audience (Bull, 1986). Atkinson (1984) demonstrated that several rhetorical devices are highly effective in evoking applause during political discourse. Specifically, he found that three-part lists and opposites/contrasts were highly correlated with an increase in audience applause during political debate. The effects of speaker fluency on audience perception in political debate are also well-documented. Speakers with higher rates of speech disfluencies, generally above three percent, are rated as more-negative by audiences, while speakers with comparatively lower rates of disfluencies are rated as more-positive (Panico, Healey, Brouwer, & Susca, 2005).

According to some accounts, a speaker's gestures and non-verbal behaviors may directly influence audience reaction to the speaker nearly as much as the content of the verbal message itself, affecting overall communicative success (Pennebaker, Mehl, & Niederhoffer, 2003; Willems & Hagoort, 2007). For instance, hand-gestures can be used to control audience applause, either to sustain or to quell it. (Bull, 1986, Atkinson 1984).Similarly, non-verbal cues (e.g. facial expression, eye gaze, gross body movements), non-linguistic auditory cues (e.g. prosody, tone, rate of speech), and speech disfluencies such as "um" and "uh" can be indicative of dishonesty (Arciuli, Mallard, & Villar, 2010; Villar, Arciuli, & Mallard, 2012). Repeated self-soothing or grooming movements, such as patting one's hair or scratching one's face, can also cause a speaker to appear untrustworthy (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). Audiences are able to detect sub-contextual information about speaker personality based upon speech-related cues (Rosenberg, & Hirschberg, 2009). One of the most

noteworthy examples of this phenomenon is the Kennedy-Nixon Presidential Debate. A retrospective study found that audience members were more-likely to attribute positive qualities to Nixon when listening to debate audio only, while positive qualities were more-likely to be attributed to Kennedy when watching the debates without audio (Druckman, 2003). The Kennedy-Nixon Debate also demonstrates the influence of image and body language on political debate. Unlike facial expression, gesture can be seen from a great distance; therefore, public speakers convey the majority of their non-verbal message to a live audience through the modality of gesture (Bull, 1986).

Surprisingly, the relationship between gesture and speech fluency in political debate has not been a focus of research. Zhang, Sherwin, Dmochowski, Sajda, and Kender (2014) examined gestures in the 2012 Presidential Debates, but focused on correlations between gesture and audience engagement. Likewise, an analysis of the 2004 and 2008 Presidential Debates (Casanto & Jasmin, 2010) found relationships between positive messages conveyed using dominant-hand gestures, as well as negative messages with non-dominant hand gestures, but did not draw correlations between speaker gesture and fluency. Some direct relationships have been established between non-verbal movements and speech disfluencies; for instance, Dittman and Llewellyn (1968, 1969) found that body movements occurred with greater frequency during and directly after speech disfluencies. However, these studies have examined only non-verbal movements across a subjective continuum, and have not been restricted to the analysis of gestures defined as co-speech. Numerous media outlets have published stories describing the differences in candidate gestures during the political debates, but eschew an examination of speech fluency. The lack of information on the relationship between gestures and fluency within

10

the political sphere is unexpected, given the social, economic, political, and cultural effects of political debate upon the world stage.

The 2012 Presidential Debates

The 2012 Presidential Debates were selected for study due to their distinctive linguistic, political, and communicative significance. The podium-style debate, composed of two speakers, a moderator, and a live and televised audience, represents a largely-unstudied environment for the examination of gesture and speech disfluency, especially from a linguistic standpoint. Public speech takes place in front of an audience by definition; therefore, the speaker is less likely to be affected by the presence of a camera. In the past, the presence of a camera has made it difficult to study gesture and speech within conversation, given the pressures that the camera places on the speaker (Bull, 1986). Currently, most co-speech studies utilize staged or elicited conversations, resulting in narrative-like dialogue with concrete topics and referents (McNeill, 1992; Mendoza-Denton & Jannedy, 2011). However, political discourse involves topics that are both more-nebulous and more-complex than expected with elicited narrative topics (Bohman, 1996; Mendoza-Denton & Jannedy). Additionally, there are very specific sociolinguistic and communicative constraints placed upon the speaker within a debate format (Mendoza-Denton & Jannedy). The speaker is given a restricted response period. At the same time, it is likely that the opposing political candidate or the moderator may interrupt the speaker without warning, causing the audience to shift their attention away from the speaker. Therefore, the "sociolinguistic pressure [is] great for the speaker to pack as much information as possible" into each utterance (Mendoza-Denton & Jannedy). These factors result in a novel speaking atmosphere that greatly differs from the speech conditions observed with naturalistic conversation or elicited narrative, both of which are common methods of elicitation used to study

11

gesture and fluency. Politically, the Presidential Debates are a natural choice for analysis of

gesture and speech. According to Stewart (2015, p.362):

> "These events often attain large viewing audiences hoping to not only catch glimpses of insight into their presumptive leaders' policy positions, but also to compare these competitors in terms of intelligence, personality, and values… These events can play a key role in defining who the major contenders for a party's nomination will be, leading to change in opinions toward, and support of, candidates among the undecided…As a result, self-presentation and connection with the audience, both live at the debate venue and watching via mass media (Peifer and Holbert 2013), is important for conveying personality and character—of which nonverbal behavior plays an important role".

The 2012 Presidential Debates were of particular interest when considering non-verbal

behaviors. Barack Obama's precision-grip gesture (in which the thumb lies against the tip of the

index finger) is considered a trademark of his oration-speech style, and has been previously

studied (Lempert, 2011). However, no correlations between Obama's precision-grip gesture and

speech disfluencies have been explored. In contrast, Mitt Romney's gesture repertoire is largely

un-documented, creating an intriguing juxtaposition between the two candidates.

Purpose

The primary purpose of this study is to investigate the relationship between hand gestures

and speech disfluencies in the First 2012 Presidential Debate between President Barack Obama

and Governor Mitt Romney. The study examined associations between the frequency and type of

both speech disfluencies and hand gestures. The rate of disfluency in political speech has been

shown to directly influence audience perception of success (Panico et al., 2005). Gestures have

also been linked to success in speech comprehension and production (e.g. Kelly et al., 1999).

Furthermore, several studies indicate that a speaker's gestures directly influence audience

reaction (Pennebaker, Mehl, & Niederhoffer, 2003; Willems & Hagoort, 2007). Yet, no literature

is known which examines the interactions between gestures and fluency, and no studies have

been conducted which investigate a nationally-televised Presidential debate. The following aims were addressed:

1. This study determined the extent to which gestures coincide with speech disfluencies. Analysis of the First 2012 Presidential Debate was conducted to measure relationships between speech disfluencies and the five taxonomies of gestures (iconic, metaphoric, deictic, cohesive, and beat).

2. This study compared the frequency of representational, beat, and cohesive gestures employed between the two Presidential candidates. These data were compared to the frequency of speech disfluencies within the First 2012 Presidential Debate. This analysis was performed at the turn level.

3. This study compared differences in gestures and speech disfluencies between the two Presidential candidates. These data were compared to reported audience outcomes for success.

CHAPTER 2

METHODS AND PROCEDURES

Methods

Stimulus

The First Presidential Debate between President Obama and Governor Romney took

place on October 3, 2012 at The University of Denver in Denver, Colorado. The podium-style

debate included no opening statements, as agreed upon in a prior memorandum of understanding

(Halperin, 2012). The debate was divided into six, 15-minute segments, during which specific

topics were presented by moderator Jim Leher of PBS. Each candidate was given two minutes to

respond to the question, with a coin toss determining the order of respondents for the opening

segment. Speaker order then alternated between the two candidates for the remainder of the

debate. Topics included the economy, jobs, Social Security, the Affordable Care Act, the federal

deficit, and the role of the federal government (The CNN Political Unit, 2012). Although the

candidates were limited to two-minute response times, both Romney and Obama frequently

exceeded the set limit, resulting in increasingly-shorter speaking turns towards the end of the

debate. The audience did not participate, and applause was limited to the beginning and the end

of the program.

Video footage of the First 2012 Presidential Debate, lasting one hour and thirty-two

minutes, was obtained from ABC News (Obama vs. Romney: The First Debate, 2012). The

footage was chosen for its use of multiple camera angles and tight focus on the debate speaker,

providing a comprehensive view of the speaker's upper body. Although the camera angle did

shift to a split view of both candidates several times over the course of the debate, ABC News footage imparted the best comprehensive focus on the speaking candidate, as well as the candidate's hands as compared to other news outlets. Several gestures for both speakers were out of view due to restricted camera angles; candidate's hands were also obscured by the podium on occasion. These gestures were omitted from study. In total, 3,082 hand gestures were deemed visible and subsequently coded.

A speech transcription was obtained from The Commission on Presidential Debates (October 3, 2012 Debate Transcript). The speech was transcribed in ordinary English orthography. The transcription was divided into speaking turns lasting approximately two minutes during which each candidate replied to a question voiced by moderator Leher. Any interjections by candidates speaking out of turn were omitted from the transcript, as were all interjections. Dialogue from Jim Leher was also removed. In sum, 1,021 utterances comprised of 12,478 words were collectively documented between President Obama and Governor Romney. Eighteen responses were recorded for each candidate including closing statements, resulting in thirty-six total speaking "turns", or allocated response segments.

Measures

Disfluency Annotation Measures. The orthographic transcription of the First 2012 Presidential Debate was sourced from The Commission on Presidential Debates (October 3, 2012 Debate Transcript) and examined for speech disfluencies using InqScribe (Inquirium, LLC) software. All utterances, divided into thirty-six total speaking turns, were uploaded into the program alongside video footage sourced from ABC News. A time-stamp was added to the beginning of each candidate's utterance to link the orthographic transcription with the

corresponding video (Figure A.2). Thus, coders were able to note speech disfluencies while watching and listening to the candidates in real time.

Researchers utilized The University of Alabama Department of Communicative Disorders Transcription Procedures Manual, Parts 2 and 3 (Table A.3), to identify and code speech disfluencies, including syllable repetitions, prolongations, syllable breaks, filled pauses, unfilled pauses, revisions, and disfluency clusters. Table 1 below illustrates the notations which were used for annotation.

Table 1. Disfluency annotations used for speech analysis

Code	Definition
[S#]	Syllable repetition: repetitions of parts of syllables.
[P#]	Syllable Prolongation: sound prolongation identified within a syllable.
[B#]	Syllable Break: a block or break in intonation at the beginning or within a syllable.
[W#]	Word Repetition: repetitions of whole, single-syllable words.
[M#]	Multi-Syllable Repetition: repetition of multiple words that constitute a phrase.
[F#]	Filled Pause: pauses associated with an utterance such as "um", "uh", and "er". Starter words such as "like" and "well" are not included in this category. Interjections such as "hmmm" and "ahhh" are not included in this category.
[R]:	Revision: instances in which an utterance is modified. A revision may be a correction of an error, or a reformation of an intention.

Gestural Annotation Measures. Gestural analysis was accomplished using video footage from ABC News uploaded into InqScribe software (Inquirium, LLC). All video was analyzed at 0.2-0.5 of normal video speed, allowing coders to observe the distinct phases of subtle hand, arm, and head movements.

Two stages of gestural analysis took place. First, Kendon's gestural segmentation scheme (1980, p. 207) was executed to partition every gesture into gesture phases (Table A.1). All movements of the upper extremities and head were recognized except for self-soothing or self-grooming movements (e.g. brushing back hair) and object manipulation (e.g. picking up a pen). The preparation phase of each hand gesture was then marked, and a time-stamp was coded to

16

indicate the point in time at which the hands or limb moved from the rest position into the gesture space. Each timestamp contained a hyperlink that navigated the user to the onset of each gesture in the video feed, accurate to a one-hundredth of a second. That is, clicking on the hyperlink aligned hand gestures with corresponding speech, providing a precise measure of gesture between researchers. The timestamp feature was especially important given the high frequency of gestures; a single minute of video feed could contain upwards of eighty distinct gesticulations. Finally, the stroke and retraction phases were noted. If a pause in motion occurred preceding the retraction phase, or the point at which the hands returned to resting position, the researcher used hand shape to establish the division between gestures. If the hand shape changed after an arrest in motion, a separate gesture was coded, even if the hands did not return to resting position.

A gestural annotation scheme adapted from McNeill (1992, p.375-387) was applied to classify each gesture as beat, representational, or cohesive. See Table 3.2 for the adapted gestural annotation scheme. Representational gestures were annotated as hand positions that reference semantic content, and include metaphorical, iconic, and deictic gesticulations (Hostetter & Alibali, 2008; McNeill). Iconic gestures refer to actions or objects, metaphoric gestures portray abstract concepts; and deictic gestures involve pointing to a referent, either tangible or intangible (McNeill; Terken, 1991). Representational gestures include all three obligatory gesture phases (preparation, stroke, and retraction), making them distinguishable from beat gestures. Beat gestures appear as rhythmic flicks of the hands that make no reference to the content or function of an utterance. That is, beat gestures align with the prosody of an utterance to provide stress. They only contain stroke and retraction phases. Finally, cohesive gestures were annotated if a series of motions occurred which 1) utilized the same gesture form and 2), linked thematically-

related statements. For example, a speaker may use the same closed fist gesture when introducing each point of a three-point argument; cohesive gestures may include any combination of representational and/or beat gesticulations.

Although the requirements for gesture classification are unambiguous, the complexity of gestures and the speed and frequency with which they occur can lead to difficulty during annotation. Therefore, a checklist was developed based the works of McNeill (1992), McNeill and Levy (1982), and Kendon (1980), designed to formally classify all abstruse hand movements. See Table A.4 for the Transcription Checklist for Gesture Annotation. Table 2 below illustrates the notations which were utilized in transcription:

Table 2. Gesture annotations used for gesture analysis

Code	Definition
B	Beat gesture
I	Iconic gesture
M	Metaphorical gesture
D	Deictic gesture
C	Cohesive gesture
U	Unknown gesture type

Measures of Debate Outcome. Outcome measures for audience-perceived success were established using CNN/ORC International polling (CNN, 2012) and Gallup polling (Gallup, 2012). A CNN/ORC International post-debate poll was performed on October 3, 2012 via telephone (CNN, 2012). Survey respondents were interviewed previously in an ORC International random national sample, conducted from September 28 – October 2, 2012. The sample contained responses from 430 adult Americans who watched the First 2012 Presidential Debate; 362 interviews were accomplished via landline, and 68 via cell phone. Of the sampled respondents, 33 percent identified as Republican, 37 percent as Democrat, and 29 percent as Independent. Eight questions were asked of each respondent in both the pre-debate questionnaire

(September 28 – October 2, 2012) and the post-debate questionnaire (October 3, 2012). The margin of sampling error for results based on the total sample was determined to be ±4.5 percentage points.

Results of the CNN poll revealed that 67 percent of American viewers thought that Governor Romney won the debate, while 25 percent of viewers thought that Obama was the winner (CNN, 2012). Additionally, 57 percent of respondents deemed Romney to be the "stronger leader", as compared to 37 percent for Obama. Thirty-five percent of respondents reported that watching the debate made them more-likely to vote for Romney, as compared to 18 percent for Obama; forty-seven percent reported that the debate did not affect their feelings towards either candidate. Sixty-one percent reported that Obama did a "worse job" than expected, while 82 percent reported that Romney did a "better job" than expected.

A Gallup telephone poll was conducted from October 4-October 6, 2012 via the Gallop Daily tracking survey (Gallup, 2012). The random sample population contained registered American voters living in the United States, ages 18 and older. A total of 1,387 respondents were surveyed via landline and cell phone. Sampled landline and cell phone numbers were chosen arbitrarily by random-digit dial. Interviews were conducted in both English and Spanish. The margin of sampling error for results based on the total sample as determined to be ±3 percentage points.

Results of pre-debate Gallup tracking survey found that President Obama retained a 5-percentage-point lead in the three days prior to the debate (Gallup, 2012). Post-debate polling conducted from October 4-October 6, 2012 found no clear respondent preference for either candidate (47 percent for both Obama and Romney). Among respondents who watched the First 2012 debate, 72 percent reported that Governor Romney won the debate, while 20 percent

thought that Obama was the winner. This 52-percentage-point win reflects the largest debate victory margin ever recorded by Gallup. Ninety-seven percent of Republicans and forty-nine percent of Democrats rated Romney as "doing a better job" than Obama.

Technical Measures. Selection of measures, methods, and procedures was largely based upon the collective published works of McNeill (1992), McNeill and Levy (1982), and Kendon (1980). Similar measures for speech and gesture analysis are currently implemented as standard practice at the McNeill Center for Gesture and Speech Research at The University of Chicago. The McNeill annotation method is based on observation and analytic heuristics from the McNeill Lab, and is designed for "novice gesture/speech/discourse analysts" (McNeill Lab Gesture Topic: Analysis). This method aims to "observe speech-gesture synchrony to a degree of accuracy that permits assessment of how meaningful gesture movements co-occur with speech, syllable by syllable" (McNeill Lab Gesture Topic: Intro to Annotation). This model was deemed appropriate as it is an established, vetted system that can be easily understood by a novice coder. Results of the McNeill Lab approach have been used in a multitude of published works, including *Hand and Mind* (McNeill, 1992). This study departs from the elicitation protocol utilized for data collection, given that the debate was pre-recorded (McNeill Lab Gesture Topic: Elicitation Protocol). However, all other phases of fluency and gesture analysis were put into practice using McNeil Lab standard measures.

Procedure

Setting and Materials

All annotation took place in The University of Alabama Department of Communicative Disorders Fluency Laboratory. Researchers watched stimulus material with InqScribe (Inquirium, LLC) software on a 21.5 inch high-resolution monitor. Video was analyzed at 0.2-

0.5 normal video speed to allow detection of rapid, minute hand movements. Audio volume remained at an audible level during annotation to ensure accurate comprehension of all utterances.

Annotation

Fluency Annotation. Two stages of annotation, one for disfluency and one for gesture, were executed over the course of this study. Disfluency annotation took place first, spanning a period of twelve months. Annotation was performed by five undergraduate student researchers, with supervision and training provided by an experienced professor with a Ph.D. in speech science. The debate stimulus was divided into five, 18-minute segments that were randomly assigned to a student researcher for analysis. Audio and video files were aligned with an orthographic English transcription of the First 2012 Presidential Debate, sourced from The Commission on Presidential Debates (October 3, 2012 Debate Transcript). InqScribe (Inquirium, LLC.) software was used for transcription analysis. Speech disfluencies were annotated according to measures established by The University of Alabama Department of Communicative Disorders Transcription Procedures Manual, Parts 2 and 3 (Table A.3). The number of disfluencies per 100 words was calculated for each speaker (e.g. Oomen & Postma, 2001).

Following the completion of disfluency annotation, the transcription was added to a spreadsheet for analysis (Figure A.5). All utterances were ordered chronologically, divided by speaker and speaker's turn, and marked by corresponding timestamp (Figure A.3). Turns for both Presidential candidates were categorized as having a high, medium, or low level of speech disfluency based on the average number of disfluencies per 100 words per speaking turn. Speaking turns containing ≤3% disfluency were categorized as low-level. Turns containing 4-5% disfluency were categorized as medium-level, and turns containing disfluency ≥6% were

classified as high-level. Classification was based on speech fluency norms for adults (Fox Tree, 1995; Lutz & Mallard, 1986; Shriberg, 2001; Webster, 1980).

Gesture Annotation. Two undergraduate student assistants, in addition to the lead researcher, served as coders for the gesture-phase of the study. Coders participated in six, one-hour training sessions guided by the lead researcher directed towards classification of gestures according to schemes based on the work of McNeill (1992) and Kendon (1980).

Transcription was initially completed by the lead researcher, who identified every gesture contained in the debate footage according to Kendon's gestural analysis scheme. In total, 3,082 gestures were identified and classified between the two speakers. A timestamp was added at the onset of each gesticulation, during the preparation phase. Gesture type was coded according to measures established by the Gesture Annotation Classification Checklist (Table 2), and transcribed according to the previously-defined gesture notation shorthand (Table 3). Two identical InqScribe files were created containing video footage of the debate; both files contained 3,978 timestamps and functioned as a blank annotation template (Figure A.4). The lead researcher completed gesture transcription for the first file, and the two undergraduate student assistants each coded fifty percent of the second file.

Reliability

Fluency Annotation Reliability. Inter-rater reliability was measured via percent agreement between four student research assistants and a principal coder, an experienced professor with a Ph.D. in speech science. Fidelity training was conducted during one-on-one sessions led by the principal coder. Measures were established according to The University of Alabama Department of Communicative Disorders Transcription Procedures Manual 3 (Table A.3), Parts 2 and 3. A copy was provided to each student, and an additional copy was stationed

in the fluency lab for reference. Coders were responsible for determining 1) presence versus absence of speech disfluency, and 2) type of speech disfluency. Video and audio analysis was performed at full video speed using InqScribe (Inquirium, LLC.) software. Percent agreement between the student assistants and the principal coder was determined to be greater than 80% prior to the conclusion of fidelity training.

Following transcription of the First 2012 Presidential Debate, four undergraduate student researchers were each assigned a four-minute debate segment to be annotated and compared to determine inter-rater reliability (Figure A.8). Segment start times coincided with the beginning of candidate speaking turns. Half of the selected stimulus contained speaker President Obama, while the other half covered Governor Romney. Inter-rater reliablity was conducted across all three debates. For each, four exchanges were chosen and coded for speech disfluencies a second time to produce measures of number of disfluencies per 100 words. A Pearson correlation across the 12 pairs of measures revaled a correlation coefficient of 0.98.

Gestural Annotation Reliability. Inter-rater reliability was measured at multiple levels for both Kendon's gestural scheme (1980) and McNeill's classification of gestures (1992). Fidelity training was completed over six, one-hour meetings by all three coders. The stimulus was a five-minute video sample taken from the beginning of the First 2012 Presidential Debate. Hand-written reliability sheets (Figure A.2) were individually completed by coders for the assigned segment, which featured President Obama and which contained a low level of disfluency. Coders were responsible for determining 1) the time of gesture onset, 2) gesture classification, 3) gesture handedness, 4) gesture form, and 5) gesture meaning. If the gesture was verified as beat, the number of beats within in the beat segment was recorded as well. Finally, coders rated each gesture according to a confidence scale ranging from 1-5, with 1 being "very

23

unsure" and 5 being "very sure" of gesture classification. A reference packet containing

Kendon's Gestural Segmentation Scheme: Adapted (Table A.1), McNeill's Gestural Annotation

Scheme: Adapted (Table A.2), McNeill's Division of the Gesture Space (Figure A.1), and the

Gesture Annotation Classification Checklist (Table 2) were issued to all coders as a reference for

specific annotation guidelines. Annotators watched and coded the training segment using

InqScribe (Inquirium, LLC.) software at 0.2-0.5 normal video speed. Then, all three coders met

to discuss results. Any gesture which did not achieve 100% inter-rater reliability was reviewed

(Figure A.7). Gestures that received a low-level confidence rating scale were also reviewed.

Inter-rater reliability greater than 80% was established between each undergraduate student

assistant and the lead researcher prior to the conclusion of the training period.

Next, the lead researcher completed gesture transcription for the whole debate, while the

two undergraduate student assistants each coded half of the debate; undergraduate student coders

alternated between response turns so that each student assistant annotated an equal number of

utterances from President Obama and Governor Romney. Six additional meetings were held to

discuss findings, with impromptu meetings occurring upon undergraduate assistant request. A

Pearson correlation was conducted, yielding a correlation coefficient of 0.86.

Following the completion of gesture annotation, the lead researcher identified 411

gestures which constituted 13.4% disagreement between coders. An additional meeting was held

to review all contested gesticulations and reach consensus regarding classification. After group

discussion, a Pearson correlation a corrrelatoin coefficient of 1.0 was established.

Data Reduction

Prior to statistical analysis, data for fluency and gesture were transferred to an Excel

spreadsheet (Microsoft, 2013). Speaker was denoted with a dichotomous nominal value

(Obama=0, Romney=1). Raw data for all speaking turns were condensed and systematized according to level of speech disfluency (Table A.7). Then, ordinal scaling was applied (low-level=0, medium-level=1, high-level=2).

Data were also reduced for all gestures according to classification of gesture. First, gesture classification was converted into a nominal scale (beat (B)=0, iconic (I)=1, metaphoric (M)=2, deictic (D)=3, cohesive (C)=4). Finally, gestures were assigned a dichotomous nominal value (beat=0, non-beat-1).

Variables

Independent variables are factors which produce or influence research outcomes (Creswell, 2003). The independent variables in this study are speaker (Obama=0, Romney=1), a categorical variable, and level of speech disfluency (low=0, medium=1, high=2), an ordinal variable. A numerical value was assigned to each independent variable. Dependent variables are assumed to arise from the independent variables (Creswell). The dependent variable in this study was the type of gesture, measured as either beat or non-beat gestures. Beat gestures are equal to 1, and are 0 otherwise, representing a dichotomous categorical variable. Since the dependent variable is discrete, the ordinary least squares regression can be manipulated to fit a linear probability model. However, non-beat gestures were further classified according to type of representational gesture (beat=0, iconic=1, metaphoric=2, deictic=3, cohesive=4), representing a nominal categorical variable. Therefore, a logistic regression model was chosen for analysis.

CHAPTER 3

RESULTS

To characterize the relationship between speech disfluencies and the five taxonomies of gesture (iconic, metaphoric, deictic, cohesive, and beat) and to predict future behaviors, this study utilized a linear regression model augmented by a count and percent comparison for variables of interest. Domains of interest included frequency of disfluency and gesture, speaker, and audience-reported outcomes. Aims were addressed across four topics. First, general trends were observed in the speech and language of each candidate. Second, fluency was evaluated for speakers at two different levels of analysis, utterance level and speaking turn level. Third, investigation of gesture was accomplished. Finally, associations between gesture and fluency were made at utterance and turn level for both President Obama and Governor Romney.

Speech

Utterance Level. A study of the relationship between gesture and speech disfluency was completed between President Obama and Governor Romney across two conditions, utterance and speaking turn. Excel (Microsoft, 2013) and R statistical software (R Development Core Team, 2016 were used for data analysis and extrapolation. Data summary at the utterance level revealed that President Obama had fewer total utterances than Governor Romney, but that Obama utilized a greater number of words per utterance. Romney demonstrated a faster average rate of speech, yet he spoke for four minutes less than Obama over the course of the debate. Romney was also more-likely to speak out of turn than Obama. Tables 3 and 4 below depict these findings:

26

Table 3. Summary of total utterances across candidates

	Number of Utterances	Words/Utterance	seconds/Utterance	Words/second
Obama	440	17.04	5.84	2.92
Romney	581	13.43	3.97	3.38
Total	1204			

Table 4. Summary of total utterance time across candidates

	Time (s)	Time (min)	Overlap Speaking Time (min)
Obama	2571	42:50.5	00:03.9
Romney	2306	38:25.7	00:19.1
Total	88.967		

Turn Level. A summation of collected data at the turn level suggests that Obama used a greater number of words per speaking turn than Romney; on average, Obama's turn length was 307.56 words, compared to 306.16 words for Romney, reflecting a difference of 64.4 words. Romney's turn length was greatest towards the middle of the debate, while Obama was most-verbose at the beginning of the debate. Figure 1 below shows the frequency of words per turn for both speakers.
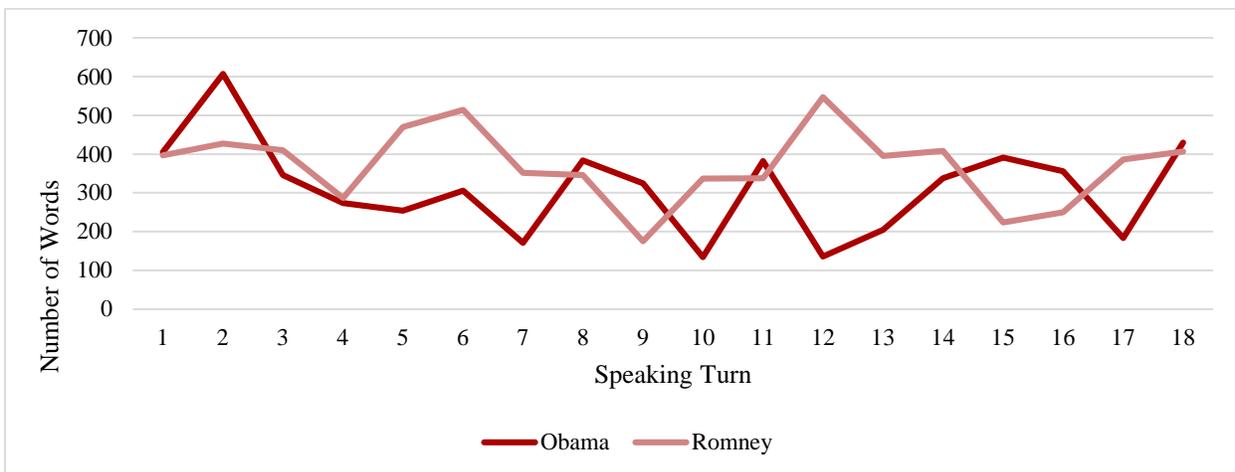

Figure 1. Words per Speaking Turn

Fluency

      Utterance Level. Speech disfluencies were catalogued by type and frequency for both speakers at the utterance level. President Obama exhibited a greater number of total speech disfluencies when compared to Governor Romney, as reflected in Table 5 below.

Table 5. Total disfluencies according to type across candidates

Disfluency Type	FP	R	M1	M2	W1	W2	W3	S1	S2	S3	Sum
Obama	221	21	7	4	35	5	1	19	3	1	317
Romney	29	30	27	6	68	6	1	35	7	1	210
Total	268	63	37	11	113	12	3	57	10	2	527

President Obama also spoke with a greater disfluency rate at the utterance level; his percentage of speech disfluency per 100 words was almost double that of Governor Romney. According to fluency norms for adult speakers, Obama's rate of disfluency is outside of the expected range and can be characterized as clinically-disfluent (Webster, 1980; Lutz & Mallard, 1986; Shriberg, 2001; Fox Tree, 1995). Expected speech disfluency for adult males is between one and three percent. In contrast, Romney's rate of disfluency at the utterance level is within functional limits for his age and gender. Figure 2 below illustraits these findings. Note, all repetitions were collapsed for the measure depicted.

Figure 2. Type and Rate of Speech Disfluency at Utterance Level

Turn Level. Level of speech fluency (low, medium, high) was evaluated across the condition of speaking turn for Romney and Obama. A simple count and percent comarison was peformed; these measuers were selected as a result of the categorical classification of the variables. President Obama evidenced a greater percentage of total speaking turns containing medium (4-5%) or high levels (≥6%) of speech disfluency as compared to Governor Romney. Table 6 below reflects these findings.

Table 6. Level of speech disfluency across turn

Low, Medium, and High Speaking Turns			
	Obama	**Romney**	**Sum (Percent)**
Low	5 (33.33%)	10 (66.67%)	15 (100%)
Medium	7 (53.85%)	6 (46.15%)	13 (100%)
High	6 (75.00%)	2 (25.00%)	8 (100%)

Calculation of disfluency rate at the turn level revealed a peak at the midpoint of the debate for both candidates. See Tables A.5. Distribution of disfluency frequency per speaking turn was more variable. That is, Obama's number of disfluencies per turn spiked towards the middle of the debate, while Romney's was greatest towards the end of the debate.

Figure 3. Rate of Speech Disfluency across Speaking Turns

29

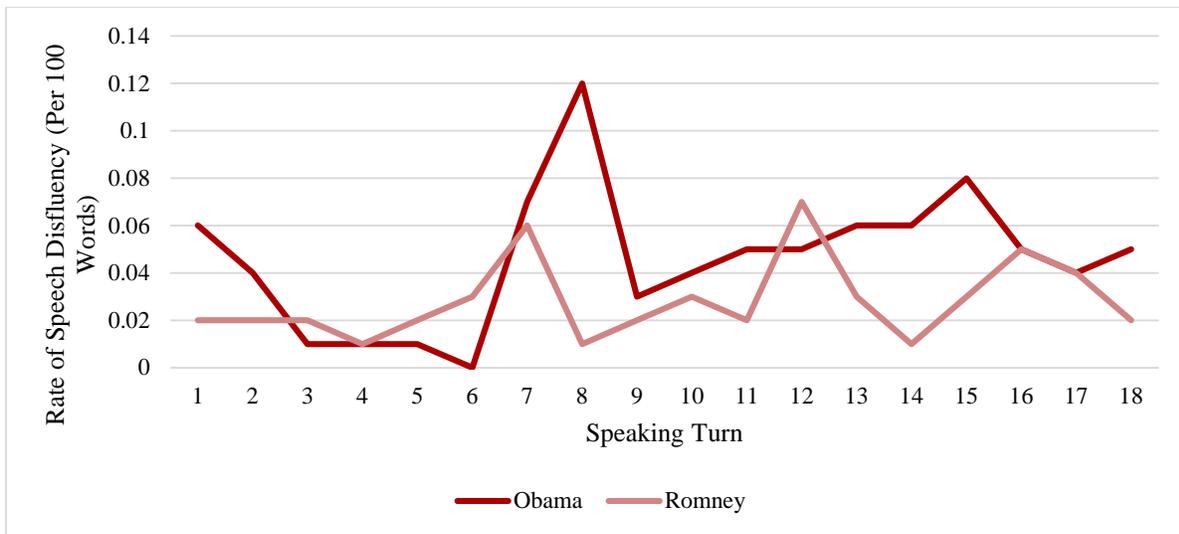Figure 3 illustraits the rate of speech disfluency per 100 words across 18 speaking turns for both candidates, reflecting changes in disfluency rates over time. Levels of speech disfluency ≤3% are low, between 4-5% are medium, ≥6% and greater are high.

A Pearson chi-squared test (X^2) was performed to assess the relation between speaker and speech disfluency at the turn level. Specifically, the frequency of gesturing was analyzed across the three levels of speech disfluency (high, medium, low) by speaker (Obama, Romney). The relation between these variables was significant, $X^2 = 304.85$ (df = 2, p < .001).

Subjective Observation. A subjective analysis of fluency disparity between speakers was conducted. Researchers made informal observations that were reported during a roundtable discussion. Coders found President Obama to be more-disfluent than Governor Romney. Moreover, researchers felt that Obama's disfluencies would be audible to an unfamiliar, untrained listener. Filled pauses were especially marked, as they appeared to coincide with a perceptible facial expression, an open-mouthed posture that approximated groping. It was noted that Obama's prosodic features and oration style did positively impact researcher-perceived success. Governor Romney's speech disfluencies were deemed imperceptible to an unfamiliar, untrained listener. Coders felt his speech was "mostly clear". The group also reported that Romney's discernible errors appeared to be repetitions or revisions made during impassioned speech. These results should be interpreted cautiously, given the subjective nature of undocumented observation and the potential for researcher bias.

Gesture

Frequency. Gesture frequency was examined across speaker at the turn level, as well as across the entire debate. A simple count revealed that President Obama used a slightly-greater

number of gestures as compared to Governor Romney. However, the margin was very small,

reflecting a difference of only 0.01%.

Table 7. Gesture frequency across Debate 1

	Obama	Romney	Sum
Total Gestures	1543	1539	3082
Percentage	50.06	49.94	100.00

Gesture frequency was further compared across speaking turns. Candidates evidenced a

high level of variance of gesture frequency. For President Obama, Min=50, Max=83,

Range=133, Avg.= 121.72. For Romney, Min=43, Max=75, Range=132, Avg.=90.06.

Figure 4. Frequency of Gesture across Speaking Turns

Type. Type of gesture was also analyzed, both within and across speakers. A simple

count and percent comarison was performed to contrast gesticulations across speaker, the results

of which are shown in Table 8. Although Obama used a greater number of total hand gestures,

Romney evidenced a far greater number of representational gestures across the debate (945

compared to 623).

Table 8. Gesture type across speakers

	Obama	Romney	Sum (Percent)
Beat	920 (60.77%)	594 (39.23%)	1514 (100%)
Iconic	355 (35.97%)	632 (64.03%)	987 (100%)
Metaphoric	182 (48.28%)	195 (51.72%)	377 (100%)
Deictic	75 (58.59)	53 (41.41%)	128 (100%)
Cohesive	11 (15.07%)	65 (84.93%)	76 (100%)

A Pearson chi-squared test (X^2) was then performed to investigate the associations between the classification of gesture (beat, iconic, metaphoric, deictic, and cohesive) and speaker (Obama, Romney). As a result, $X^2 = 79.467$ (df = 4, p < .001).

Within speaker, the most frequently occurring type of gesture was beat, which fit with expected results (McNeill, 1992). However, there was a marked difference in distribution between the two candidates. Beats accounted for fifty-nine percent of all of Obama's gestures but only thirty-nine percent of Romney's. Differentiations between the candidates were also reflected in the total percent of iconic gestures. There were comparable percentages of metaphoric, deictic, and cohesive gestures between the two candidates. Figures 5 and 6 illustrate the results of gesture classification within speaker:

Figure 5. Distribution of Obama's Gestures According to Type

Figure 6. Distribution of Romney's Gestures According to Type

Subjective Observation. A subjective study of gesture was made during annotation that further describes variations among speakers. Observations were reported and discussed collectively by researchers during inter-rater reliability annotation sessions. Researchers agreed that President Obama's gesticulations were easier to annotate due to high predictability. He frequently repeated a limited set of hand shapes across diverse speaking contexts. Moreover, Obama this small set of hand shapes were noted across different classifications of gesture. As an example, President Obama utilized the same hand shape, motion, and gesture space when talking about "the tax system" (metaphoric gesture) and "closing the deficit" (iconic gesture). His hands were cupped within the central space, one stacked over the other, as if holding a ball. The retention of form across many speaking conditions made identification of Obama's gesture phases and boundaries straightforward. Obama's beat gestures were typically bi-lateral within central gesture space, having a loosely-curled hand shape. When he did expand his gesture space to the right, left, or bilateral periphery, hand shape often remained the same. Obama habitually returned his hands to the classical rest position in the central gesture space, signaling clear pauses between gesture sequences.

Researchers found that annotation of Governor Romney's gestures required more attention to detail, given that hand shape regularly changed. Romney favored a larger gesture space than Obama, with one or both hands extending into the periphery. He less-frequently returned to classical resting phase between gesture sequences. Also, the angle of wrist-flexion during beat gestures was greater than Obama's, resulting in a "palms-out" beat style. Finger splay was also greater during beat and representational gesticulations. Interestingly, Romney sometimes spoke for periods of time without moving his hands, which appeared "unnatural" to researchers.

Gesture and Fluency

First, a simple sum and percent comarison of frequency of gestures across all three domains of speech disfluency (low, medium, high) was executed in Excel (Microsoft, 2013). Mesures were chosen to accommodate the categorical classification of both variables. Outcomes revealed that the fewest number of gestures arose under the condition of high disfluency for both speakers. Also, Romney had a greater number of total gestures under low disfluency conditions. Results should be interpreted cautiously; number of gestures under each fluency level condition is affected by the number of speaking turns categorized as either low, medium, or high. That is, each gesture reflects a data point contained within a speaking turn dataset; level of fluency is calculated as the average rate of disfluency for the entire speaking turn and ordered according to a fluency norm continuum, rather than a simultaneous gesture-disfluent event at a single point in time. Therefore, it was expected that Romney would have a smaller number of gestures occurring under the condition of high disfluency, given that only two of his total speaking turns were above 6 percent disfluency. Table 9 below reflects these results.

Table 9. Frequency of gesture across disfluency speaking turns

Number of Low, Medium, and High Speaking Turns			
	Obama	**Romney**	**Sum**
Low	560	826	1386
Medium	565	310	875
High	418	90	508
Percentage of Low, Medium, and High Speaking Turns			
	Obama	**Romney**	**Sum**
Low	40.40	56.57	100.00
Medium	64.57	35.43	100.00
High	82.28	17.72	100.00

A logistic regression analysis was then performed to predict gesture type according to level of disfluency across speaker turn. Our dataset retained a binary response variable, beat vs. non-beat (0,1). There were two predictor variables; level of disfluency, low, medium, and high, (0,1,2), and speaker, Obama versus Romney (0,1). The data were screened for missingness and violation of logistical regression assumptions prior to analysis. A statistically-significant negative correlation was found between beat gestures and increasing levels of speech disfluency. Expected coefficient change of gestures from beat to non-beat given an increase of disfluency level from low to medium was equal to -1.34830 (SE= 0.12879, z= -10.469, p<.001). Expected coefficient change of gestures from beat to non-beat given an increase of disfluency level from medium to high was equal to -2.40992 (SE= 0.12879, z= -12.401, p<.001). See table A. 6 for further analysis.

A second logistic regression model was conducted to further predict the type of gesture according to the level of speech disfluency per speaker turn. This statistical analysis model was selected due to the categorical classification of the dependent variable (beat=0, iconic=1, metaphoric=2, deictic=3, cohesive=4). There were two predictor variables: disfluency level (high, medium, low) and speaker (Obama, Romney). A statistically-significant positive

interaction were observed between iconic gestures and both medium (p<.001) and high (p<.001) disfluency levels. No other significant results were observed. See Table A.7 for further analysis.

Subjective Observation. Ongoing subjective assessment of gesture type during discrete disfluent events was conducted during annotation. Researchers compared observations during inter-rater reliability annotation sessions. Gesture space appeared to be inversely proportional to fluency. Additionally, researchers felt that there was a negative relationship between beat gestures and disfluencies, especially when speaking topic was highly abstract (e.g. interest rates). Researchers agreed that no prevailing pattern emerged from representational gesticulations during episodes of speech disfluency. Subjective observation should be considered carefully; it is likely that gesture patterns are highly variable between individual speakers, which could account for differences.

CHAPTER 4

DISCUSSION

The purpose of this study was to investigate relationships between gesture and speech disfluency in the First 2012 Presidential debate between President Barack Obama and Governor Mitt Romney. Specifically, interactions between gesture type and level of speech disfluency for both candidates were examined at the turn level. Measures were derived from the following sources: 1) Fluency: The University of Alabama Department of Communicative Disorders Transcription Procedures Manual, Parts 2 and 3 (Table A.3); 2) Gesture: Kendon's Gestural Segmentation Scheme: Adapted (Table A.1), McNeill's Gestural Annotation Scheme: Adapted (Table A.2), Gesture Annotation Classification Checklist (Table 2); and 3) Debate Success: CNN/ORC International polling (CNN, 2012), Gallup polling (Gallup, 2012).

Fluency Differences

When measuring differences in speech disfluency levels (low, medium, high) between the two Presidential candidates at the turn level, a statistically-significant relation between the two variables was found. This relationship was illustrated by President Obama's greater rate and frequency of disfluency as compared to Governor Romney. Thirteen out of eighteen speaking turns were classified as having medium (4-5%) or high (≥6%) disfluency; in contrast, Romney only had eight speaking turns above 3% disfluency. A significant portion of Obama's disfluencies involved filled pauses. Filled pauses are characteristic of hesitation phenomena, and entail delayed transfer of speaker message (Krauss,1998). Additionally, the filled pause, specifically pauses at the juncture between utterances, has been previously correlated with a

37

breakdown of lexical access (Krauss). In contrast, the greater part of Romney's disfluencies were repetitions.

These findings suggest that the preponderance of Obama's utterances, contained disfluency elevated above expected norms for adult males, had speech errors which would be noticeable to the average, untrained listener. Speaking turns which were highly disfluent ($\geq 6\%$) are especially marked. In contrast, Romney's average speech fluency remained within normal limits for adult males, resulting in speaking turns which would most-likely be unremarkable to the average, untrained listener.

Gesture Differences

Gesture comparison between the two debate speakers revealed a statistically-significant relationship between gesture type and speaker. The most common type of gesture observed during the First 2012 debate was beat, which fits with previously reported outcomes (McNeill, 1992). Iconic gestures had the second-highest frequency, followed by metaphoric, deictic, and cohesive. President Obama and Governor Romney had a nearly-identical number of gestures over the debate. Both candidates evidenced a comparable number of representational gestures. Differences in gesture frequency are therefore attributable to an increase in Obama's use of beat gesticulations. Previous studies have found a positive correlation between beat gestures and filled pauses (Krauss, 1998). Given that Obama demonstrated a larger number of beat gestures and speech disfluencies (most of which were classified as filled pauses) as compared to Romney, this finding is unsurprising. Unexpectedly, Romney utilized a greater number of cohesive gestures; Obama rarely employed his signature precision-grip gesture, and never as a part of a cohesive gesture sequence.

Gesture, Fluency and Perceived Success

Multiple statistically-significant interactions were found between level of speech disfluency and gesture type. A strong negative relationship was found between beat gestures and increasing levels of speech disfluency. That is, speaking turns containing a level of speech disfluency above expected norms for age and gender were correlated with a higher frequency of representational gestures. As disfluency level increased from medium to high, the expected change in gesture type from beat to non-beat became increasingly negative. Additionally, statistically-significant positive interaction between iconic gestures and medium and high disfluency levels. These findings were in accordance with subjective observation made by researchers. There are many possible explanations. Abstract topics are associated with increased speech disfluency (Bortfield, Leon, Bloom, Schober, & Brennan, 2001). There is also evidence, although far from definitive, that suggests that gestures augment communication during difficulty speaking situations (e.g. von-Raffler Engel, 1980; Krauss, 1998). Therefore, a shift from beat to iconic gestures during disfluent turns could explain a subconscious effort to clarify arguments containing nebulous topics for the audience. Second, it is known that beat gestures contain no semantic content (McNeill, 1992). Rather, beats follow the prosodic line of the utterance and provide emphasis. It is therefore logical that a breakdown of the syntax and prosody of an utterance during a disfluent event would impact beat gestures. These results pose future research questions regarding relationships between gesture types during discrete disfluency events.

As previously mentioned, Governor Mitt Romney was declared the winner of the First 2012 Debate by a large margin, according to snap polling. Interestingly, Romney was a more-fluent speaker when compared to President Obama. Therefore, results are in accordance with

other studies that proposed correlations between fluency and audience-perceived success (e.g. Villar, Arciuli, & Mallard, 2012). Additionally, positive, statistically-significant interactions established in this study between increasing levels of speech disfluency and representational gestures indicate that there may also be a relationship between gesture type and audience-measured debate success. However, further study is warranted.

Study Limitations

Although every effort was made to conduct an effective, efficacious study, there were constraints which undoubtedly affected presented findings. The overall design of the study was largely based upon the measures, methods, and procedures established by McNeill (1992) and Kendon (1998). However, design modifications were made to accommodate an additional variable of interest, speech fluency. Methods and procedures were also altered as a result of the selected stimulus, the First 2012 Presidential Debate. Intrinsic factors which affected the study included stimulus constitution and size of dataset. Extrinsic factors which affected the study included coder limitations, time constraints, and poverty of resources.

The selected stimulus presented many intrinsic challenges that contributed to annotation difficulty. Video of the debate was sourced with hand view in mind. Researchers determined that ABC News footage provided the broadest view of the candidate's upper body and hands. Still, many gestures were not visible due to camera angle or network logos, resulting in lost data. Attempts were made to gain access to raw debate footage from multiple camera angles through organizations such as the Associated Press, but video was not immediately available.

The designated stimulus also contained less narrative content as compared to samples used by McNeill (1992), who elicited speech samples with film or cartoon retell. Researchers established gestural fidelity using McNeill's annotation protocol, which was intended for speech

samples containing a high number of embedded spatial-temporal referents. The lower number of tangible referents within political speech likely contributed to the complexity of gesture annotation, given intangible referents (e.g. "the deficit", "cutting Medicare") often depict complex topics, especially when compared to concrete referents (e.g. "he climbed a tree"). There were a number of disputed gestures reviewed during inter-rater reliability annotation that reflected this uncertainty; highly contested gestures were predominately iconic or metaphoric. Although McNeill maintains that metaphoric gestures "are de facto iconic gestures, given that metaphor entails iconicity… [and] therefore, should be considered as a continuum --with a given gesture having more or less iconicity, metaphoricity, etc.", a decrease in coder confidence likely contributed to a reduction in annotation fidelity (McNeill Lab Gesture Topic: Intro to Annotation).

Finally, researcher partiality must be considered. Given the inherent political nature of this study, preconceived political beliefs and opinions constitute a source of bias. Both President Obama and Governor Romney are well-known politicians. Additionally, this study was ongoing during the 2016 Presidential election cycle. Therefore, both candidates were frequently in the public eye. Study design aimed to eliminate this source of bias through the selection of apolitical variables.

Future Application

While the rate of fluency has been much studied resulting in derived norms for adult speakers, gesture studies are fewer. Additionally, gesture is highly individualized. All persons have a unique and nuanced gesture style. While a rich dataset was extracted from the First 2012 Presidential Debate, it reflects a short speaking sample between two individuals speaking under unnatural conditions. Candidates also have an immense amount of public speaking experience.

Finally, it is likely that the candidates received some level of speech or debate coaching which would affect oration style as well as gesture and fluency. Both Obama and Romney are middle-aged American males with a high level of education and socio-economic status. These factors all contribute to a highly distinctive speaking and gesturing style. Consequently, the boarder application of these findings is limited.

Additional study is warranted to further explore the relationship between gesture and fluency in the 2012 Presidential Debate. A simultaneous analysis of gesture type during a discrete event of speech disfluency would provide detailed insight into the relationship between these two factors at the word level. Furthermore, data used in this study were sampled from a single debate in a field of three possible debates. Therefore, a logical progression of this study would involve examination of the Second and Third 2012 Presidential Debates. These debates include unaddressed variables which would like have an effect on speech and gesture, such as debate format (podium style versus town hall, hand-held microphone versus clip-on microphone). President Obama was also widely considered the victor of both the Second and Third Presidential Debates, which would allow researchers to examine variables under different conditions of audience-perceived success. Finally, speaking samples sourced from more-naturalistic speaking conditions, such as interviews or home videos, could provide a speech and gesture baseline for both candidates. These analyses could offer valuable information regarding the effects of high-stress, debate-style oration on fluency and gesture.

Conclusions

Overall, the results of this study suggest a positive correlation between elevated speech disfluency at the turn level and frequency of representational hand gestures. That is, when the semanticity of hand movements is increased, fluency decreases. Statistical findings mirror

subjective observations made by researchers while observing individualized fluency and gesture patterns between the two candidates, President Barack Obama and Governor Mitt Romney. As previously stated, Romney was acknowledged as the victor of the First 2012 Debate according to snap polling measures. Romney was also a more-fluent speaker when compared to President Obama. These findings were in accordance with subjective observation made by researchers. Given the strong predictive value of interaction between increased fluency and frequency of representational gestures, it is postulated that this model would hold true across the Second and Third 2012 Presidential Debates.

REFERENCES

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, *5*(4), 307-331.

Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology*, *89*(1), 183.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*(2), 169-188.

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

Arciuli, J., Mallard, D., & Villar, G. (2010). "Um, I can tell you're lying": Linguistic markers of deception versus truth-telling in speech. *Applied Psycholinguistics*, 31(03), 397-411.

Atkinson, M. (1984). Our masters' voices: The language and body language of politics. *Psychology Press*.

Austin, E. E., & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of experimental child psychology*, *122*, 92-103.

Bohman, J., & Rehg, W. (1996). Discourse and democracy: the formal and informal bases of democratic legitimacy.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and speech*, *44*(2), 123-147.

Bull, P. (1986). The use of hand gesture in political speeches: Some case studies. *Journal of Language and Social Psychology*, *5*(2), 103-118.

Cable News Network- CNN. (2012). ORC International Poll, October 3, 2012. ORC International. Retrieved from http://i2.cdn.turner.com/cnn/2012/images/10/03/top12.pdf.

Casasanto, D., & Jasmin, K. (2010). Good and bad in the hands of politicians: Spontaneous gestures during positive and negative speech. *PLoS One*, 5(7), e11805.

Corballis, M. (1999). The Gestural Origins of Language Human language may have evolved from manual gestures, which survive today as a" behavioral fossil" coupled to speech. *American Scientist*, *87*(2), 138-146.

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73-111.

Cornejo, C., Simonetta, F., Ibáñez, A., Aldunate, N., Ceric, F., López, V., & Núñez, R. E. (2009). Gesture and metaphor comprehension: electrophysiological evidence of cross-modal coordination by audiovisual stimulation. *Brain and cognition*, *70*(1), 42-52.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin*, 129(1), 74.

Dittmann, A. T., & Llewellyn, L. G. (1968). Relationship between vocalizations and head nods as listener responses. *Journal of personality and social psychology*, 9(1), 79.

Dittmann, A. T., & Llewellyn, L. G. (1969). Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology*, 11(2), 98.

Driskell, J. E., & Radtke, P. H. (2003). The effect of gesture on speech production and comprehension. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 445-454.

Druckman, J. N. (2003). The Power of Television Images: The First Kennedy-Nixon Debate Revisited. *Journal of Politics*, 65(2), 559-571.

Duncan, Susan and McNeill. *ANNOTATIVE PRACTICE (Under Perpetual Revision) — Due To Susan Duncan (Revision Of D. McNeill (2005) Gesture & Thought, Appendix.)*. 1st ed. Chicago: N.p. Web. 6 Feb. 2016.

Efron, D. (1941). Gesture and environment.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and annotation. *Semiotica*, *1*(1), 49-98.

Enrici, I., Adenzato, M., Cappa, S., Bara, B. G., & Tettamanti, M. (2011). Intention Processing in Communication: A Common Brain Network for Language and Gestures. *Journal of Cognitive Neuroscience*, 23(9), 2415-2431.

Feyereisen, P., & De Lannoy, J. D. (1991). *Gestures and speech: Psychological investigations*. Cambridge University Press.

Finlayson, S., Forrest, V., Lickley, R., & Beck, J. M. (2003). Effects of the restriction of hand gestures on disfluency. Proceedings of Diss, Gothenburg Papers in Theoretical Linguistics, 90, 21-24.

Gallup, I. (2012, October 08). Romney Narrows Vote Gap After Historic Debate Win. Retrieved February 04, 2016, from http://www.gallup.com/poll/157907/romney-narrows-vote-gap-historic-debate-win.aspx

Gillespie, M., James, A. N., Federmeier, K. D., & Watson, D. G. (2014). Verbal working memory predicts co-speech gesture: Evidence from individual differences. *Cognition*, *132*(2), 174-180.

Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, *71*(1), 231-239.

Goldin-Meadow, S. (2003). The resilience of language.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology*, *91*(4), 720.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*(6), 516-522.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction*, *9*(3), 201-219.

Grice, H. P. (1957). Meaning. *The philosophical review*, 377-388.

Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, *35*(2), 133-153.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological bulletin*, *137*(2), 297.

Hostetter, A.B., & Alibali, M.W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture*. (7)1, 73-95. Doi:10.1075/gest.7.1.

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

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis.*Psychological bulletin*, *137*(2), 297.

Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! A test of the Gesture as Simulated Action framework. *Journal of Memory and Language*, *63*(2), 245-257.

Hostetter, A. B., & Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture*, *12*(1), 62-83.

Iverson, J. M., & Goldin‐Meadow, S. (2001). The resilience of gesture in talk: Gesture in blind speakers and listeners. *Developmental Science*, *4*(4), 416-422.

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of memory and Language*, *40*(4), 577-592.

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and language*, *89*(1), 253-260.

Kendon, A. (1972). Some relationships between body motion and speech. *Studies in dyadic communication*, *7*(177), 90.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. *The relationship of verbal and nonverbal communication*, *25*, 207-227.

Kita, S. (2000). How representational gestures help speaking. *Language and gesture*, 162-185.

Krauss, R. M. (1998). Why do we gesture when we speak?. *Current Directions in Psychological Science*, *7*(2), 54-54.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?. *Advances in experimental social psychology*, 28, 389-450.

Krauss, R. M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. *Gesture, speech, and sign*, 93.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate?. *Journal of personality and social psychology*, *61*(5), 743.

Lass, N. J. (Ed.). (2014). *Speech and language: advances in basic research and practice* (Vol. 5). Academic Press.

Lempert, M. (2011). Barack Obama, being sharp: Indexical order in the pragmatics of precision-grip gesture. *Gesture*, 11(3), 241-270.

Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes*, *24*(2), 239-266.

Lutz, K. C., & Mallard, A. R. (1986). Disfluencies and rate of speech in young adult nonstutterers. Journal of Fluency Disorders, 11(4), 307-316.

Marstaller, L., & Burianová, H. (2014). The multisensory perception of co-speech gestures–A review and meta-analysis of neuroimaging studies. *Journal of Neurolinguistics*, *30*, 69-77.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

McNeill, D. (2000). *Language and gesture* (Vol. 2). Cambridge University Press.

McNeill, D., Cassell, J., & McCullough, K. E. (1994). Communicative effects of speech-mismatched gestures. *Research on language and social interaction*, *27*(3), 223-237.

McNeill Lab Gesture Topic: Analysis. (n.d.). Retrieved February 01, 201, from http://mcneilllab.uchicago.edu/analyzing-gesture/elicitation-protocol.html

McNeill Lab Gesture Topic: Elicitation Protocol. (n.d.). Retrieved February 01, 201, from http://mcneilllab.uchicago.edu/analyzing-gesture/elicitation-protocol.html

McNeill Lab Gesture Topic: Intro to Annotation. (n.d.). Retrieved February 01, 201, from http://mcneilllab.uchicago.edu/analyzing-gesture/elicitation-protocol.html

Mendoza-Denton, N., & Jannedy, S. (2011). Semiotic layering through gesture and intonation: a case study of complementary and supplementary multimodality in political speech. *Journal of English Linguistics*, 0075424211405941.

Obama vs. Romney: The First Debate. (2012, October 04). Retrieved August 18, 2015, from http://abcnews.go.com/Nightline/video/obama-romney-debate-17393182.

October 3, 2012 Debate Transcript. (n.d.). The Commission on Presidential Debates. Retrieved January 18, 2017, from http://www.debates.org/index.php?page=october-3-2012-debate-transcript

Panico, J., Healey, E. C., Brouwer, K., & Susca, M. (2005). Listener perceptions of stuttering across two presentation modes: A quantitative and qualitative approach. *Journal of fluency disorders*, 30(1), 65-85.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1), 547-577.

Peifer, J. T., & Holbert, R. L. (2013). Developing a systematic assessment of humor in the context of the 2012 US general election debates. *Argumentation and Advocacy*, 49(4), 286-301.

Pine, K. J., Lufkin, N., Kirk, E., & Messer, D. (2007). A microgenetic analysis of the relationship between speech and gesture in children: Evidence for semantic and temporal asynchrony. *Language and Cognitive Processes*, 22(2), 234-246.

Ping, R. M., Goldin-Meadow, S., & Beilock, S. L. (2014). Understanding gesture: Is the listener's motor system involved?. *Journal of Experimental Psychology: General*, 143(1), 195-204. doi:10.1037/a0032246

Reston, Maeve. (2012). When Romney trounced Obama. Retrieved February 04, 2016, from http://www.cnn.com/2016/09/25/politics/obama-debate-election-2012/

Rimé, B. (1982). The elimination of visible behaviour from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European journal of social psychology*, *12*(2), 113-129.

Rosenberg, A., & Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Communication*, 51(7), 640-655.

Shriberg, E. (2001). Toerrrr'is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31(1), 153-169.

Stewart, P. A. (2015). Do the presidential primary debates matter? Measuring candidate speaking time and audience response during the 2012 primaries. *Presidential Studies Quarterly*. (2). 361.Terken, J. (1991). Fundamental frequency and perceived prominence of accented syllables. *The Journal of the Acoustical Society of America*, *89*(4), 1768-1776.

Thompson, L. A., & Massaro, D. W. (1994). Children′s Integration of Speech and Pointing Gestures in Comprehension. *Journal of Experimental Child Psychology*, *57*(3), 327-354.

Tree, J. E. F. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of memory and language*, 34(6), 709-738.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*(2), 187-204.

Villar, G., Arciuli, J., & Mallard, D. (2012). Use of "um" in the deceptive speech of a convicted murderer. *Applied Psycholinguistics*, 33(01), 83-95.

von Raffler-Engel, W. (1980). Kinesics and Paralinguistics: A Neglected Factor in Second-Language Research and Teaching. *Canadian Modern Language Review*, *36*(2), 225-37.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In Proceedings of the ACL 2012 System Demonstrations (pp. 115-120). Association for Computational Linguistics.

Webster, R. L. (1980). Evolution of a target-based behavioral therapy for stuttering. Journal of Fluency Disorders, 5(3), 303-320.

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex*, *17*(10), 2322-2333.

Zhang, J. R., Sherwin, J., Dmochowski, J., Sajda, P., & Kender, J. R. (2014, November). Correlating Speaker Gestures in Political Debates with Audience Engagement Measured via EEG. In Proceedings of the ACM International Conference on Multimedia (pp. 387-396). ACM.

APPENDIX

Table A.1. Kendon's Gestural Segmentation Scheme

Kendon's Gestural Segmentation Scheme: Adapted		
Phase Classification	Description	Optional/Obligatory
Gesture-Unit (G-Unit)	The period of time in which the limbs move from the position of rest and return to the position of rest.	Obligatory
Gesture-Phrase (G-Phrase)	The G-Phrase occurs within a G-Unit. Several G-Phrases may occur within a single G-Unit. A G-Phrase consists of one or more stroke phases. A G-Phrase must contain a stroke phase; other phases are optional.	Obligatory
Preparation Phase	The limb moves away from the rest position into the gesture space in order to initiate the stroke phase.	Optional
Pre-Stroke Hold Phase	A hand position reached at the end of the stroke phase prior to the onset of the stroke. This phase occurs when the stroke phase is delayed.	Optional
Stroke Phase	The meaning of the gesture is expressed during this phase. The stroke phase usually occurs within the central gesture space. It is semantically and temporally linked to the verbal utterance it accompanies.	Obligatory
Post-Stroke Hold Phase	The hand position which occurs following the stroke phase and preceding the retraction phase. This phase occurs when the retraction is delayed.	Optional
Retraction Phase	The retraction of the hand to the rest position.	Optional
Adapted from Kendon (1980), p. 207		

Table A.2. McNeil's Gesture Annotation

McNeill's Gestural Annotation Scheme: Adapted
I. Identify the movements that are gestures.
II. Determine the onset and the offset of the gesture.
III. Determine the phases of the gesture (See Table 3.1).
IV. Identify the boundaries of the gesture phases in the corresponding transcription of speech.
V. Identify the boundaries of the gesture phases within the physical gesture space.
VI. Analysis and Classification of Gestures
a. Code the gesture type.
i. Representational: represents attributes, actions, and relationships
1. Iconic
a. Hand: describe the dimensions of the hand movements
i. Form: code the form the hand takes
ii. Meaning: code what the hand represents
b. Motion: describe the dimensions of the hands in space
i. Hand: code the motion of the hand in space
ii. Form: code the form of the movement in space
iii. Meaning: code what the motion represents
2. Metaphoric
a. Hand: describe the dimensions of the hand movements
i. Form: code the form the hand takes
ii. Meaning: code what the hand represents
b. Motion: describe the dimensions of the hands in space
i. Hand: code the motion of the hand in space
ii. Form: code the form of the movement in space
iii. Meaning: code what the motion represents
3. Deictic
a. Hand: describe the dimensions of the hand movements
i. Form: code the form the hand takes
ii. Meaning: code what the hand represents
b. Motion: describe the dimensions of the hands in space
i. Hand: code the motion of the hand in space
ii. Form: code the form of the movement in space
iii. Meaning: code what the motion represents
ii. Beats: small flicks of the hand which convey no semantic information, but move in time with the prosodic features of speech
1. Beat Filter: apply the beat filter to any gesture that takes place between two rest positions. Add 1 point to the score for any yes answer. A score of 0 equates to a formal beat gesture, while a higher score indicates that the gesture is likely representational.
a. Does the gesture have more than two movement phases (yes for ± 2).

	b. How many times does wrist or finger movement appear in any movement phase which does not end in a rest phase? (Add this number to the score)
	c. If the first movement takes place outside of the central gesture space, does any other movement take place inside the central gesture space? (Yes or no?)
	d. If there are only 2 movement phases, is the initial phase in the same place as the second phase? (Yes or no?)
iii.	Confidence Scale: rate coder confidence on each gesture ranging from 1 (not confident) to 5 (very confident)
Adapted from McNeill, 1992, p. 375-387	

Table A.3. Transcription Manual for Fluency Annotation

<table>
<tr><td colspan="2">Department of Communicative Disorders
University of Alabama
Transcription Procedures Manual
Updated as of 04-04-2012</td></tr>
<tr><td colspan="2">OUTLINE:
</td></tr>
<tr><td colspan="2">Part 1. Basic Transcription
I. Bound Morpheme Conventions:
    1. Always use the root word followed by a slash "/" and the bound morpheme</td></tr>
<tr><td colspan="2">Examples:

- *Taking* →
  - TAKE/ING
- *There are no ghosts in the graveyard* →
  - C THERE ARE NO GHOST/S IN THE GRAVEYARD.
</td></tr>
<tr><td colspan="2">   2. Do **not** mark bound morpheme if tense (e.g., past) cannot be determined from context.</td></tr>
<tr><td colspan="2">Examples:

- *The dog bark* →
  - C THE DOG BARK.
- *He eat grass* →
  - C HE EAT GRASS.
</td></tr>
<tr><td colspan="2">   3. Only mark ***grammatical bound morphemes***:</td></tr>
<tr><td colspan="2">Examples:

a. possessive	(*dad's* → DAD/Z*; yours* → YOUR/Z*; hers* → HER/Z)
b. plural	(*babies* → BABY/S*; houses* → HOUSE/S)
c. plural/possessive	(*babies'* → BABY/S/Z
d. 3rd person singular	(*goes* → GO/3S)
e. regular past tense	(*died* → DIE/ED*; loved* → LOVE/ED)
f. contractions	(*it's* → IT/'S*, I'm* → I/'M*; they've* → THEY/'VE)
g. negative contractions	(*didn't* → DID/N'T*, can't* → CAN/'T)
h. present progressive	(*having* → HAVE/ING)
</td></tr>
</table>

4. Do **not** mark the following as grammatical bound morphemes:
   - Irregular past tense (*went; saw*)
   - Overgeneralizations (*goed; Playdohs*)
   - Singular forms (*pants; lots; scissors; glasses*)
   - Past participle (*used to*; *supposed to*; *scrambled eggs*; *am tired*)
   - Present participle (*swimming hour*; *looking glass*)
   - Present/past perfect (*has finished*; *had looked*)
   - Present/past progressive (*am running*; *was looking*)
   - Concatenitives (*gonna*; *hafta*; *wanna*)
   - Gerands (*swimming is fun*)
   - Frozen forms (*downstairs; o'clock*)
   - Derivational morphemes (*reddish*; *quickly*; *bigger*)
- Finally, these contractions: *don't*; *won't*; *ain't*; *let's*

II. Combining Words and Titles:

Generally, if a child **under five years** of age uses a word independently elsewhere in the transcript, it can be transcribed separately. Otherwise, transcribe as one unit.

1. *Names or titles*:

Examples:
   - *Little Red Riding Hood* →
     - C LITTLEREDRIDINGHOOD
   - *Ninja Turtles* →
     - C NINJATURTLES
   - *Grandma Jones* →
     - C GRANDMAJONES

2. *Frozen forms*:

Examples:
   - *Macaroni and cheese is good* →
     - C MACARONIANDCHEESE IS GOOD.
   - *Here kitty cat* →
     - C HERE KITTYCAT.
   - *Thank you* →
     - C THANKYOU.

3. If a *speech code* occurs within a proper name or frozen form, separate them.

Example:
   - *Little red r___riding hood* →
     - C LITTLERED [P]RIDINGHOOD
   - *I read three little p p pigs* →
     - C I READ THREELITTLE [S2]PIGS.

III. Rote Language Production:

1. **Counting** items or objects: transcribe the first two items and put the rest in a maze.

Examples:
   - *I see one two candles* →
     - C I SEE ONE TWO CANDLE/S.

• *I can see one to three four five of them* →
○ C I CAN SEE ONE TWO (THREE FOUR FIVE) OF THEM.

2. **Listing** items or objects; place in parentheses after the second word of the list.

Example:
- *I want the dog the cat the sheep and the horse* →
  - O C I WANT THE DOG THE CAT (THE SHEEP AND THE HORSE).
- *I see a fork, spoon, knife, and spatula* →
  - O C I SEE A FORK SPOON (KNIFE AND SPATULA).

3. **Sequential** information may include pseudo-stuttering, repetition for emphasis, or listed information. Place in parentheses after the second item.

Examples:
- *I used to say my name is Hai-Hai-Hailey* →
  - O C I USED TO SAY MY NAME IS (HAI HAI) HAILEY.
- *No no no go* →
  - O C (NO) NO NO GO!
- *Whale shark eel fish stingray dolphin* →
  - O C WHALE SHARK (EEL FISH STINGRAY DOLPHIN).

4. Recitation of songs, poems, etc. from memory should be placed in parentheses

Example:
- *I pledge allegiance to the flag...* →
  - O C (I PLEDGE ALLEGIENCE TO THE FLAG…)
- *Mary had a little lamb, little lamb...* →
  - o C (MARY HAD A LITTLE LAMB…)

IV. Sound Effects and Discourse Markers:

1. *Sounds effects*: transcribe if they function as a subject or object

Examples:
- *Give me the vroom vroom.* →
  - O C GIVE ME THE VROOMVROOM.
- *Meow wants some more food.* →
  - O C MEOW WANT/3S SOME MORE FOOD.

2. *Whispering*: any speech that is whispered should be placed in brackets.

3. *Singing:* any speech that is sung should be placed in brackets.

4. *Interjections* should be placed in parentheses and coded with an "I" in brackets.

Examples:
- *Ouch, that hurts!* →
  - O (OUCH [I]) THAT HURT/3S!
- *Wow, look at that!* →
  - o (WOW [I]) LOOK AT THAT!

5. *Discourse markers* should be transcribed according to the following categories, placed in parentheses and coded with a "D" in brackets. Indicate positive or negative with **accent mark**: `
   - *Positive*: OKAY, YES, YEP, MHM`, UHUH`
   - *Negative*: NO, NOPE, M`HM, U`HUH

Examples:
- *Mhm, give it to me* →
  - (MHM` [D]) I WANT IT.
- *Uhuh, I don't want to go!* →
  - (U`HUH [D]) I DON'T WANT TO GO!

V. Spelling Conventions:

1. A diminutive form should always use the suffix "y" (as opposed to "ie"):

Examples:
- *Looky, there's a kitty!* →
  - LOOKY THERE/'S A KITTY!
- *Oopsy, I knocked it off.* →
  - OOPSY I KNOCK/ED IT OFF.

2. For less than full productions, change to *intended* form, and in parentheses, use asterisk to indicate missing syllable.

Examples:
- *Give me nothe oner* →
  - C GIVE ME ANOTHER (*NOTHER) ONE.
- *I don't member* →
  - C I DON'T REMEMBER (*REMEMBER).
- *We have a puter* →
  - C WE HAVE A COMPUTER (*PUTER).

VI. Utterance Segmentation:

1. A phrase counts as one utterance when it is preceded and followed by adult utterances. This rule supersedes all other rules

Example:

E WHAT WERE YOU DOING IN CLASS TODAY?
C PLAY/ING WITH CAR/S
E WHICH CAR/S?
C THE BIG ONE/S WITH THE SIREN/S.

2. **Use intonation** to make decisions about utterance segmentation. A falling intonation contour defines comments and statements. A rising intonation contour defines questions.

3. **Independent clauses** joined by a coordinating conjunction constitute one utterance. Segment any additional independent clause if joined by a conjunction.
   - AND, BUT, OR, ALTHOUGH, SO

Examples:
- *I went to my aunt's house and my brother and parents did too* →
  - C I WENT TO MY AUNT/Z HOUSE AND MY BROTHER AND PARENT/S DID TOO.
- *I went to my aunt's house and my brother went to my aunt's house and my dad and my mom went to my aunt's house* →
  - C I WENT TO MY AUNT/Z HOUSE AND MY BROTHER WENT TO MY AUNT/Z HOUSE.

	o C AND MY DAD AND MY MOM WENT TO MY AUNT/Z HOUSE.
4.	**Dependent clauses** joined by a conjunction should be segmented with the *second* occurrence.

Examples:	• *I want the doggy and the kitty.* →
	  o C I WANT THE DOGGY AND THE KITTY.
	• *I want to go the park and then to go shopping.*
	  o C I WANT TO GO TO THE PARK AND THEN TO GO SHOPPING.
	• *The baby woke up and then put on her clothes and then went to play* →
	  o C THE BABY WOKE UP AND THEN PUT ON HER CLOTHES.
	  o C AND THEN WENT OUT TO PLAY.

5.	If the child **pauses** for longer than two seconds count the phrase as two utterances. This rule supersedes the conjunction rule.

Example:	• *The bird saw the beehive (pause) and the bees came out and said bye bye.* →
	  o C THE BIRD SAW THE BEEHIVE.
	  o C AND THE BEES CAME OUT AND SAID BYEBYE.

6.	**Tag questions** count as one utterance and should not be segmented.
Example:	• *You like school, don't you* → o E YOU LIKE SCHOOL DON'T YOU.
7.	**Discourse markers** functioning as an independent clause (i.e., "self talk") should be segmented as a separate utterance.

Example:	• *Yesterday they made colors I mean not colors they made pictures with what are they called sponges* →
	  o C YESTERDAY THEY MADE COLOR/S.
	  o C I MEAN NOT COLOR/S.
	  o C THEY MADE PICTURES WITH WHAT ARE THEY CALLED SPONGES.

8.	**Interruptions**. Instances in which a speaker is interrupted by another speaker, marked with <

Example:	• *When are we going to- Plese give that to me.* →
	  o C WHEN ARE WE GOING TO<
	  o E PLEASE GIVE THAT TO ME.

9.	**Abandoned utterances**. Instances in which an utterance is not completed, mark with >

Example:	• *I like the other the horsie fell down.* →
	  o C I LIKE THE OTHER>
	  o C THE HORSY FELL DOWN.

- *Yesterday I went no wait what's that over there?*
  - C YESTERDAY I WENT>
  - C (NO [D]) WAIT.
  - C WHAT/'S THAT OVER THERE?

VII. Unintelligible Utterances:

1. For language analyses, listen to a segment *3 times* before marking it as unintelligible, but for stuttering, take as much time as you need.
2. If there is a single unintelligible word, go ahead and make an *educated guess* if possible.

Example:

- *I was in Mr. XX's room last year* →
  - C I WAS IN MISTER XX/Z ROOM LAST YEAR.

3. Mark unintelligible syllables with and "X, with one "X" for each unintelligible syllable.

Part 2: Fluency Annotation

I. Syllable Level Events

Speech does should be attached to the beginning of the word with no intervening space. In general, only code if an event would also be perceived by another listener.

1. **[S#] Syllable Repetition**. Repetitions of parts of syllables. Examples: "a-and," "f-five," "ba-baby," "mo-mo-mommy." Code number of repetition units, not including the final production. If a syllable repetition occurs on a syllable *other than the first*, indicate the omitted segment in parentheses. In addition, if a syllable repetition occurs *with another word*, indicate the omitted segment in parentheses. Finally, if a syllable repetition occurs within revised material, include a code only with parentheses.

Examples:

- a-a-I →
  - [S2]I
- *mo-mo-mommy* →
  - [S2]MOMMY
- *becau-because* →
  - [S1]BECAUSE (BECAU*)
- I want to play with the f- the f- the farm. →
  - C I WANT TO PLAY WITH (THE F* THE F*) THE [S2]FARM.
- *I want to eat some s-s-soup cereal* →
  - I WANT TO EAT SOME ([S2]SOUP [R]) CEREAL

2. **[P#] Syllable Prolongation**. Sound prolongation identified within a syllable. If possible, include the number of seconds the prolongation lasts, with 1 being the minimum. Look for:

- atypical length
- atypical intonation pattern
- quaver or tremor-like quality

• respiratory irregularity

Examples:
- It's *ffffffunny* →
  - IT/'S [P1]FUNNY.
- *aaaaai* don't know→
  - [P1]I DON'T KNOW.

3. **[B#] Syllable Break**. A block or break in intonation to being or within a syllable, characterized as a brief cessation of voicing. If possible, include the number of seconds the block lasts, with 1 being the minimum. Look for:
- sudden, tense release of position, audible burst of air
- audible or visible struggle/tension during abnormally long pause
- insertion of a glottal stop in the middle of a word, not as phonemic substitution (broken word)

Example:
- *st_ore* →
  - [B1]STORE
- __give it to me
  - [B2]GIVE IT TO ME!

4. **[W#] Word Repetition**. Repetitions of whole, single-syllable words. For example, "but-but," "and-and." This does not include words repeated for emphasis, as in "big, big". If possible, include the number of seconds the prolongation lasts, with 1 being the minimum. A word repetition separated by a filled pause such as "um" or "uh" should be set apart in parentheses.

Examples:
- *And and and and then put it there*→
  - C [W3]AND THEN PUT IT THERE.
- I want um want to go now.
  - C I (WANT [F1]UM) [W]WANT TO GO NOW

Finally, if a syllable-level event *might not* be perceived as such by someone else, don't mark it. In general, avoid making assumptions about what is going on in the mind of the speaker.

II. Multi-Syllable Annotation
Speech events associated with multiple words should set aside in parentheses, with code placed at the end of the parentheses.
1. **[M#] Multi-Syllable Repetition.** Repetition of multiple words that constitutes a phrase. If the repetition is not exactly the same in length as the final production, write out the repetitions in a maze, with an asterisk indicating what was left off.

Examples:

- *The green the green the green one is yucky.* →
  - C (THE GREEN [M2]) THE GREEN ONE IS YUCKY.
- Once up- once upon a time →
  - (ONCE UP* [M1]) ONCE UPON A TIME.
- *Because because I don't like the red one.* →
  - C [M1]BECAUSE I DON'T LIKE THE RED ONE.
- *Look into into the house* →
  - C LOOK [M1]INTO THE HOUSE.

2. **[F#] Filled Pause.** Pauses associated with an utterance such as "um," "uh," and "er". Starter words such as "like" and "well" are not included in this category. Interjections such as "hmmm" and "ahhh" are not included in this category. All filled pauses should be placed in parentheses as they are not actually words. If different types of filled pauses occur in sequence, transcribe them within the maze in the order in which they occur.

- Um, um, um, that's the pool. →
  - C ([F3]UM) THAT/'S THE POOL.
- I um like the blue one. →
  - C I ([F1]UM) LIKE THE BLUE ONE.
- Uh, um, uh, don't! →
  - C ([F3]UH UM UH) DON/T!

3. **[R] Revision.** Instances in which an utterance is modified. A revision may be the correction of an error, or a reformulation of an intention.

Examples:

- *I went we went to the store* →
  - C (I WENT [R]) WE WENT TO THE STORE.
- *I want the ball the red ball.* →
  - C I WANT (THE BALL [R]) THE RED BALL.
- *The the girl goed went.* →
  - C [W1]THE GIRL (GOED [R]) WENT.
- *A girl g went* →
  - C A GIRL (G* [R]) WENT.
- *The we put the rest on.* →
  - C (THE [R]) WE PUT THE REST ON.

III. Disfluency Clusters.

If one more than one types of disfluency occur on the *same* syllable, this is called a within-unit cluster, and are coded in the order in which they occur, separated with a dash. If disfluencies occur on adjacent syllables, a code for between-unit clusters [BUC] is placed at after the final disfluency within the cluster.

Examples:

- *m-m-m____onkey* →
  - C [S2-P]MONKEY
- s-s-s-_____sand-sand →
  - C [S3-P-W]SAND
- *u-u-u-um um* →
  - C [S3-F2]UM
- *um h____he wwwwwants to go* →
  - C ([F]UM) [B]HE [P3]WANT/3S [BUC] TO GO.

Part 3: Error Annotation

Some possibilities:

Speech sound errors use {} brackets with the error type coded within the brackets.

Fronting = F

Backing = B

Syllable Deletion = D

Gliding = G

Cluster Reduction = R

Table A.4. Transcription Checklist for Gesture Annotation

Gesture Annotation Classification Checklist		
Is there a movement present?	Y/N	If yes, continue.
Do the hand leave the rest position?	Y/N	If yes, continue.
Pauses		
If the handshape does not change during the pause, the pause is considered part of the gesture. Does the handshape change?	Y/N	If yes, then they are two different gestures. If no, then the pause usually marks the beginning of a new gesture.
Multiple or Ambiguous Gestures		
A local rest position occurs when the hands begin and end their movement from the same place in space (usually not the rest position). Does the series of gestures begin and end at a local rest position?	Y/N	If yes, then the gesture is usually a beat.
A canonical rest position occurs when the hands are static and relaxed. Does the gesture begin and end at a canonical rest position?	Y/N	If yes, then the gesture is usually a beat.
A series of gestures can begin with a representational and have beats layered on top of it, but a representational cannot follow a beat in a series. Does a representational gesture follow a beat gesture in a series?	Y/N	If yes, then the gesture is usually the beginning of a new gesture, and does not belong to that series.
Beat Gestures		
Do the hands remain in the central gesture space while they move from resting position back to resting position?	Y/N	If yes, then the gesture is usually beat.
If the first movement is in a non-central space, is any other movement performed in the central space?	Y/N	If yes, then the gesture is usually NOT beat.
Does the gesture incorporate small, low-energy flicks of the wrist?	Y/N	If yes, then the gesture is usually beat.
Does the gesture contain two phases (preparation and retraction?)	Y/N	If yes, then the gesture is usually beat.
How many times does the wrist or finger movement or tensed stasis appear in any movement phase not ending in a rest position?		If this number is greater than zero, the gesture is usually NOT a beat.
Does the gesture include movements of the wrists?	Y/N	If yes, then the gesture is usually beat. If no, it is usually representational.
Consider the verbal message. Is the gesture tied directly to the utterance?	Y/N	If yes, then the gesture is usually NOT beat.
Does the gesture coincide with the prosodic features of speech?	Y/N	If yes, then the gesture is usually beat.

Is the gesture repeated several times throughout the utterance/turn as a part of a list?	Y/N	If yes, then the gesture is usually part of a cohesive.
Representational Gestures		
Does the hands leave the central gesture space in front of the body?	Y/N	If yes, then the gesture is usually representational.
Does the gesture have more than two phases (preparation and retraction) to include the stroke phase? (The effort of the hands was concentrated on reaching the end-point of the phase).	Y/N	If yes, then the gesture is usually representational.
Does the gesture have ties to the semantic content of the utterance?	Y/N	If yes, then the gesture is usually representational.
Does the gesture include movements of the wrists after the gesture has reached the furthest point in the stroke?	Y/N	If yes, then the gesture is most likely NOT representational.
Does the gesture include two or move changes in feature (shape change, movement change, or space change?)	Y/N	If yes, then the gesture is most likely representational.
Iconic Gestures		
Does the gesture refer to a concrete object or action?	Y/N	If yes, then the gesture is usually iconic.
Metaphoric Gestures		
Does the gesture refer to a metaphorical concept, or aspects of time or space?	Y/N	If yes, then the gesture is usually metaphorical.
Deictic Gestures		
Do the hand or fingers point to something concrete or abstract?	Y/N	If yes, then the gesture is usually deictic.
Cohesive Gestures		
Does the gesture appear as part of a list or collection of topics?	Y/N	If yes, then the gesture is usually a cohesive.
A cohesive may be composed of beats or representational, but the form must remain the same. Does the form remain the same? Cohesive gestures may be separated by other gestures or gesture sequences.	Y/N	If yes, then the gesture is usually a cohesive.

Table A. 5. Disfluency at the Turn Level

Summary of Disfluency at the Turn Level: Obama*																	
Turn	Words	FP	D	I	R	M1	M2	W1	W2	W3	S1	S2	S3	P	B	Sum	Rate
1	397	22	0	0	1	0	0	0	0	0	0	0	0	0	0	23	0.06
2	427	16	0	0	0	1	0	0	0	0	0	0	0	0	0	17	0.04
3	410	3	0	0	0	0	0	2	0	0	1	0	0	0	0	6	0.01
4	287	1	0	0	0	0	1	1	0	0	0	0	0	0	0	3	0.01
5	470	3	0	0	0	2	0	2	0	0	0	0	0	0	0	7	0.01
6	514	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2	0.00
7	352	17	0	0	1	1	0	1	0	0	1	2	0	0	0	23	0.07
8	**346**	**27**	**0**	**0**	**3**	**2**	**0**	**4**	**0**	**0**	**2**	**2**	**0**	**0**	**0**	**40**	**0.12**
9	175	3	0	0	0	1	0	0	0	0	1	0	0	0	0	5	0.03
10	337	9	0	0	3	0	0	3	0	0	0	0	0	0	0	15	0.04
11	338	14	0	0	2	0	0	0	0	0	0	0	0	0	0	16	0.05
12	547	19	0	0	0	0	0	2	0	0	5	0	1	0	0	27	0.05
13	395	17	0	0	0	0	0	2	1	0	3	1	0	0	0	24	0.06
14	408	22	0	0	1	0	0	1	0	0	0	0	0	0	0	24	0.06
15	224	14	0	0	1	0	0	2	0	0	0	0	0	0	0	17	0.08
16	250	7	0	0	1	0	2	2	0	0	1	0	0	0	0	13	0.05
17	386	9	0	0	2	0	0	3	1	0	0	0	0	0	0	15	0.04
18	407	12	0	0	0	0	0	3	0	1	5	0	0	0	0	21	0.05
Summary of Disfluency at the Turn Level: Romney*																	
Turn	Words	FP	D	I	R	M1	M2	W1	W2	W3	S1	S2	S3	P	B	Sum	Rate
1	405	5	0	0	0	0	0	1	0	0	3	0	0	0	0	9	0.02
2	607	2	0	0	3	1	0	7	0	0	1	0	0	0	0	14	0.02
3	346	0	0	0	0	0	0	4	0	0	2	1	0	0	0	7	0.02
4	274	0	0	0	1	0	0	2	1	0	0	0	0	0	0	4	0.01
5	254	0	0	0	0	3	0	2	0	0	0	0	0	0	0	5	0.02
6	306	0	0	0	0	2	0	4	0	1	1	0	0	0	0	8	0.03
7	171	1	0	0	0	2	0	2	0	0	5	0	0	0	0	10	0.06
8	384	1	0	0	1	2	0	0	0	0	0	1	0	0	0	5	0.01
9	325	1	0	0	1	1	0	3	0	0	0	0	0	0	0	6	0.02
10	134	0	0	0	1	1	0	1	0	0	1	0	0	0	0	4	0.03
11	382	3	0	0	1	1	0	2	0	0	0	1	0	0	0	8	0.02
12	**136**	**5**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**3**	**0**	**0**	**0**	**0**	**10**	**0.07**
13	205	2	0	0	1	0	0	0	0	0	2	1	0	0	0	6	0.03
14	338	1	0	0	1	0	0	2	0	0	0	0	0	0	0	4	0.01
15	391	0	0	0	1	0	3	3	0	0	2	0	1	0	0	10	0.03
16	356	2	0	0	2	0	2	8	2	0	1	1	0	0	0	18	0.05
17	183	0	0	0	4	0	1	1	0	0	1	0	0	0	0	7	0.04
18	430	4	0	0	3	0	0	2	0	0	1	0	0	0	0	10	0.02

*Highest rate of disfluency is highlighted in red; highest frequency is in bold.

Table A.6. Correlation Coefficients: Beat and Non-Beat

Deviance Residuals				
Min	1Q	Median	3Q	Max
-1.3670	-1.0923	-0.4895	1.0334	2.0886

Coefficients				
	Estimate	**St. Error**	**Z Score**	**Pr(>|z|)**
(Intercept)	0.34850	0.07454	4.675	2.93e-06 ***
Medium Disfluency1	-1.34830	0.12879	-10.469	< 2e-16 ***
High Disfluency2	-2.40992	0.19433	-12.401	< 2e-16 ***
Representational1	0.08682	0.10985	0.790	0.42931
Medium Disfluency1: Representational1	0.70938	0.18127	3.913	9.10e-05 ***
High Disfluency2: Representational1	1.01909	0.26353	3.867	0.00011 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
(Dispersion parameter for binomial family taken to be 1)				
Null deviance: 3802.3 on 2768 degrees of freedom Residual deviance: 3428.4 on 2763 degrees of freedom				
AIC: 3440.4				
Number of Fisher Scoring iterations: 4				

Table A.7. Correlation Coefficients According to Gesture Type

Deviance Residuals				
Min	1Q	Median	3Q	Max
-2.8042	-0.4895	-0.4895	1.0348	2.1899

Coefficients				
	Estimate	St. Error	Z Score	Pr(>|z|)
(Intercept)	0.34525	0.07447	4.636	3.55e-06 ***
Medium Disfluency1	0.12875	-2.40667	-10.447	< 2e-16 ***
High Disfluency2	-2.40667	0.19430	-12.386	< 2e-16 ***
Iconic1	-0.23720	0.13521	-1.754	0.079366 .
Metaphoric2	0.27695	0.18127	1.698	0.089475
Deictic3	-0.04898	0.25626	-0.191	0.848420
Cohesive4	3.56678	1.01262	3.522	0.000428 ***
Medium Disfluency1: Iconic1	1.04030	0.21249	4.896	9.79e-07 ***
High Disfluency2: Iconic1	1.74656	0.30324	5.760	8.43e-09 ***
Medium Disfluency1: Metaphoric2	0.52033	0.27332	1.904	0.056945 .
High Disfluency2: Metaphoric2	0.19524	0.42334	0.461	0.644662
Medium Disfluency1: Deictic3	0.27559	0.44553	0.619	0.536198
High Disfluency2: Deictic3	-0.19218	0.80491	-0.239	0.811290
Medium Disfluency1: Cohesive4	-0.95754	1.27923	-0.749	0.454143
High Disfluency2: Cohesive4	-1.97536	1.17584	-1.680	0.092967
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
(Dispersion parameter for binomial family taken to be 1)				
Null deviance: 3802.3 on 2768 degrees of freedom Residual deviance: 3352.7 on 2754 degrees of freedom				
AIC: 3382.7				
Number of Fisher Scoring iterations: 6				

Figure A.1. McNeill's Division of the Gesture Space

EXTREME PERIPHERY

upper

upper right

PERIPHERY

upper left

CENTER

right

CENTER-CENTER

left

lower right

lower left

lower

Figure A.1. Division of the gesture space for transcription purposes. From Pedelty (1987).

McNeill, 1992, p. 378

Figure A.2. Hand-Written Reliability Sheets Used For Gesture Code Training

Figure A.3. InqScribe (LLC.) Sample of Fluency Annotation

Figure A.4. InqScribe (LLC.) Sample of Gesture Annotation

Figure A.5. Sample of Fluency Annotation using Excel (Microsoft, 2013)

	Debate Number	Time 1	Speaker	Utterance	Time 2	Duration	Overlap	Words	FP
2	1	1	00:00.0 L	GOOD EVENING FROM THE MAGNESS ARENA AT THE UNIVERSITY OF DENVER IN DENVER COLORADO.		00:07.0		14	
3	1	2	00:07.0 L	I'M L OF THE PBS NEWSHOUR.		00:02.2		6	
4	1	3	00:09.3 L	AND I WELCOME YOU TO THE FIRST OF THE 2012 PRESIDENTIAL DEBATES BETWEEN PRESIDENT BARACK OBAMA THE DEMOCRATIC NOMINEE		00:08.9		19	
5	1	4	00:18.1 L	AND FORMER MASSACHUSETTS GOVERNOR MITT ROMNEY THE REPUBLICAN NOMINEE.		00:05.1		8	
6	1	5	00:23.2 L	THIS DEBATE AND THE NEXT THREE TWO PRESIDENTIAL ONE VICE PRESIDENTIAL ARE SPONSORED BY THE COMMISSION ON PRESIDENTIAL [		00:08.9		18	
7	1	6	00:32.1 L	TONIGHT'S NINETY MINUTES WILL BE ABOUT DOMESTIC ISSUES.		00:04.0		8	
8	1	7	00:36.1 L	AND WILL FOLLOW A FORMAT DESIGNED BY THE COMMISSION.		00:04.0		9	
9	1	8	00:40.1 L	THERE WILL BE SIX ROUGHLY FIFTEEN MINUTE SEGMENTS WITH TWO MINUTE ANSWERS FOR THE FIRST QUESTION.		00:06.9		16	
10	1	9	00:47.0 L	THEN OPEN DISCUSSION FOR THE REMAINDER OF EACH SEGMENT.		00:04.2		8	
11	1	10	00:51.2 L	THOUSANDS OF PEOPLE OFFERED SUGGESTIONS ON SEGMENT SUBJECTS OR QUESTIONS VIA THE INTERNET AND OTHER MEANS.		00:06.8		16	
12	1	11	00:58.0 L	BUT I MADE THE FINAL SELECTIONS.		00:02.1		6	
13	1	12	01:00.1 L	AND FOR THE RECORD.		00:01.1		4	
14	1	13	01:01.3 L	THEY WERE NOT SUBMITTED FOR APPROVAL TO THE COMMISSION OR THE CANDIDATES.		00:05.8		11	
15	1	14	01:07.0 L	THE SEGMENTS AS I ANNOUNCED IN ADVANCE WILL BE THREE ON THE ECONOMY AND ONE EACH ON HEALTH CARE THE ROLE OF GOVERNME		00:10.0		25	
16	1	15	01:17.0 L	WITH AN EMPHASIS THROUGHOUT ON DIFFERENCES SPECIFICS AND CHOICES.		00:06.2		9	
17	1	16	01:23.3 L	BOTH CANDIDATES WILL ALSO HAVE TWO MINUTE CLOSING STATEMENTS.		00:04.8		8	
18	1	17	01:28.0 L	THE AUDIENCE HERE IN THE HALL HAS PROMISED TO REMAIN SILENT.		00:04.0		11	
19	1	18	01:32.0 L	NO CHEERS APPLAUSE BOOS HISSES AMONG OTHER NOISY DISTRACTING THINGS SO WE MAY ALL CONCENTRATE ON WHAT THE CANDIDATE		00:10.2		22	
20	1	19	01:42.3 L	THERE IS A NOISE EXCEPTION RIGHT NOW THOUGH AS WE WELCOME PRESIDENT OBAMA AND GOVERNOR ROMNEY.		00:29.8		16	
21	1	20	02:12.0 L	GENTLEMEN.		00:06.2		1	
22	1	21	02:18.2 L	GENTLEMEN WELCOME TO YOU BOTH.		00:02.9		4	
23	1	22	02:21.1 L	LET'S START THE ECONOMY SEGMENT ONE.		00:04.0		6	
24	1	23	02:25.0 L	AND LET'S BEGIN WITH JOBS.		00:02.2		5	
25	1	24	02:27.2 L	WHAT ARE THE MAJOR DIFFERENCES BETWEEN THE TWO OF YOU ABOUT HOW YOU WOULD GO ABOUT CREATING NEW JOBS?		00:11.0		19	
26	1	25	02:38.3 L	YOU HAVE TWO MINUTES EACH OF YOU HAVE TWO MINUTES TO START.		00:02.9		12	
27	1	26	02:41.2 L	THE COIN TOSS HAS DETERMINED MR PRESIDENT YOU GO FIRST.		00:03.8		9	
28	1	26	02:45.0 O	WELL THANK YOU VERY MUCH JIM FOR THIS OPPORTUNITY.		00:02.0		9	
29	1	28	02:47.1 O	I WANT TO THANK GOVERNOR ROMNEY AND THE UNIVERSITY OF DENVER FOR YOUR HOSPITALITY.		00:03.2		13	
30	1	29	02:50.3 O	([F]UH) THERE ARE A LOT OF POINTS THAT I WANT TO MAKE TONIGHT.		00:02.9		13	
31	1	30	02:53.2 O	BUT ([F]UH) THE MOST IMPORTANT ONE IS THAT ([F]UH) TWENTY YEARS AGO I BECAME THE LUCKIEST MAN ON EARTH BECAUSE MICHELLE O		00:07.9		26	
32	1	31	03:01.1 O	AND SO ([F]UH) I JUST WANT TO WISH ([F]UH) SWEETIE ([F]UH) YOU HAPPY ANNIVERSARY		00:05.9		14	
33	1	32	03:07.0 O	AND LET YOU KNOW THAT A YEAR FROM NOW WE WILL NOT BE CELEBRATING IT IN FRONT OF FORTY MILLION PEOPLE.		00:07.1		20	
34	1	33	03:14.1 O	([F]UH) YOU KNOW FOUR YEARS AGO WE WENT THROUGH ([F]UH) THE WORST FINANCIAL CRISIS SINCE THE GREAT DEPRESSION.		00:05.9		18	
35	1	34	03:20.1 O	MILLIONS OF JOBS WERE LOST.		00:01.0		5	
36	1	35	03:21.1 O	THE AUTO INDUSTRY WAS ON ([F]UH) THE BRINK OF COLLAPSE ([F]UH).		00:03.2		11	
37	1	36	03:24.2 O	THE FINANCIAL SYSTEM HAD FROZEN UP.		00:02.9		6	
38	1	37	03:27.2 O	AND BECAUSE OF THE RESILIENCE AND THE DETERMINATION OF THE AMERICAN PEOPLE ([F]UH) WE'VE BEGUN TO FIGHT OUR WAY BACK.		00:06.0		19	
39	1	38	03:33.2 O	([F]UH) OVER THE LAST THIRTY MONTHS WE'VE SEEN 5 MILLION JOBS IN THE PRIVATE SECTOR CREATED.		00:05.8		16	
40	1	39	03:39.0 O	([F]UH) THE AUTO INDUSTRY HAS COME ROARING BACK AND HOUSING ([F]UH) HAS BEGUN TO RISE.		00:06.1		15	
41	1	40	03:45.1 O	BUT WE ALL KNOW THAT WE'VE STILL GOT A LOT OF WORK TO DO.		00:02.9		14	
42	1	41	03:48.1 O	AND SO THE QUESTION HERE TONIGHT IS NOT WHERE WE'VE BEEN BUT WHERE WE'RE GOING.		00:05.9		15	
43	1	42	03:54.0 O	([F]UH) GOVERNOR ROMNEY ([F]UH) HAS A PERSPECTIVE THAT SAYS ([F]UH) IF WE CUT TAXES SKEWED TOWARDS THE WEALTHY AND ROLL E		00:10.2		26	
44	1	43	04:04.2 O	I'VE GOT A DIFFERENT VIEW.		00:01.9		5	
45	1	44	04:06.2 O	I THINK WE'VE GOT TO INVEST IN EDUCATION AND TRAINING.		00:02.0		10	
46	1	45	04:08.2 O	I THINK IT'S IMPORTANT FOR US TO DEVELOP NEW SOURCES OF ENERGY HERE IN AMERICA.		00:04.1		15	
47	1	46	04:12.2 O	THAT WE CHANGE OUR TAX CODE TO MAKE SURE THAT WE'RE HELPING SMALL BUSINESSES AND COMPANIES THAT ARE INVESTING HERE IN		00:06.8		24	
48	1	47	04:19.1 O	THAT ([F]UH) WE TAKE SOME OF THE MONEY THAT WE'RE SAVING AS WE WIND DOWN ([F]UH) TWO WARS ([F]UH) TO REBUILD AMERICA		00:06.1		22	
49	1	48	04:25.2 O	AND THAT WE REDUCE OUR DEFICIT IN A BALANCED WAY THAT ALLOWS US TO MAKE THESE CRITICAL INVESTMENTS.		00:06.8		17	
50	1	49	04:32.0 O	NOW IT ULTIMATELY IS GOING TO BE UP TO THE VOTERS TO YOU WHICH PATH WE SHOULD TAKE.		00:04.2		18	
51	1	50	04:36.2 O	([F]UH) ARE WE GOING TO DOUBLE DOWN ON THE TOP DOWN ECONOMIC POLICIES THAT (G* [R])HELPED TO GET US INTO THIS MESS OR D		00:13.0		40	
52	1	51	04:49.2 O	AND I'M LOOKING FORWARD TO HAVING THAT DEBATE.		00:01.0		6	

DEBATE 1 | Debate 1 copy

69

Figure A.6. Sample of Fluency Analysis using Excel (Microsoft, 2013)

Speaker	Debate	Turn	Time 1	Speaker	Words	FP	D	I	R	M1	M2	W1	W2	W3	S1	S2	S3	P	B	sum	rate
O	1	8	0:43:31		346	27	0	0	3	2	0	4	0	0	2	2	0	0	0	40	0.12
O	1	15	1:17:17		224	14	0	0	1	0	0	2	0	0	0	0	0	0	0	17	0.08
O	1	7	0:39:01		352	17	0	0	1	1	0	1	0	0	1	2	0	0	0	23	0.07
O	1	13	1:06:56		395	17	0	0	0	0	0	2	1	0	3	1	0	0	0	24	0.06
O	1	14	1:11:38		408	22	0	0	1	0	0	1	0	0	0	0	0	0	0	24	0.06
O	1	1	0:02:45		397	22	0	0	1	0	0	0	0	0	0	0	0	0	0	23	0.06
O	1	16	1:18:58		250	7	0	0	1	0	2	2	0	0	1	0	0	0	0	13	0.05
O	1	18	1:26:26		407	12	0	0	0	0	0	3	0	1	5	0	0	0	0	21	0.05
O	1	12	1:01:32		547	19	0	0	0	0	0	2	0	0	5	0	1	0	0	27	0.05
O	1	11	0:56:42		338	14	0	0	2	0	0	0	0	0	0	0	0	0	0	16	0.05
O	1	10	0:51:02		337	9	0	0	3	0	0	3	0	0	0	0	0	0	0	15	0.04
O	1	2	0:07:11		427	16	0	0	0	1	0	0	0	0	0	0	0	0	0	17	0.04
O	1	17	1:24:17		386	9	0	0	2	0	0	3	1	0	0	0	0	0	0	15	0.04
O	1	9	0:47:25		175	3	0	0	0	1	0	0	0	0	1	0	0	0	0	5	0.03
O	1	5	0:26:58		470	3	0	0	0	2	0	2	0	0	0	0	0	0	0	7	0.01
O	1	3	0:16:54		410	3	0	0	0	0	0	2	0	0	1	0	0	0	0	6	0.01
O	1	4	0:21:19		287	1	0	0	0	0	1	1	0	0	0	0	0	0	0	3	0.01
O	1	6	0:32:37		514	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2	0

Speaker	Debate	Turn	Time 1	Speaker	Words	FP	D	I	R	M1	M2	W1	W2	W3	S1	S2	S3	P	B	sum	rate
R	1	13	1:04:54		136	5	0	0	0	0	0	1	1	0	3	0	0	0	0	10	0.07
R	1	7	0:36:22		171	1	0	0	0	2	0	2	0	0	5	0	0	0	0	10	0.06
R	1	17	1:20:25		356	2	0	0	2	0	2	8	2	0	1	1	0	0	0	18	0.05
R	1	18	1:23:20		183	0	0	0	4	0	1	1	0	0	1	0	0	0	0	7	0.04
R	1	10	0:49:55		134	0	0	0	1	1	0	1	0	0	1	0	0	0	0	4	0.03
R	1	14	1:05:29		205	2	0	0	1	0	0	0	0	0	2	1	0	0	0	6	0.03
R	1	6	0:30:13		306	0	0	0	0	2	0	4	0	1	1	0	0	0	0	8	0.03
R	1	16	1:14:04		391	0	0	0	1	0	3	3	0	0	2	0	1	0	0	10	0.03
R	1	19	1:28:39		430	4	0	0	3	0	0	2	0	0	1	0	0	0	0	10	0.02
R	1	2	0:09:37		607	2	0	0	3	1	0	7	0	0	1	0	0	0	0	14	0.02
R	1	1	0:04:53		405	5	0	0	0	0	0	1	0	0	3	0	0	0	0	9	0.02
R	1	11	0:54:41		382	3	0	0	1	1	0	2	0	0	0	1	0	0	0	8	0.02
R	1	3	0:15:06		346	0	0	0	0	0	0	4	0	0	2	1	0	0	0	7	0.02
R	1	5	0:24:53		254	0	0	0	0	3	0	2	0	0	0	0	0	0	0	5	0.02
R	1	9	0:46:01		325	1	0	0	1	1	0	3	0	0	0	0	0	0	0	6	0.02
R	1	4	0:19:50		274	0	0	0	1	0	0	2	1	0	0	0	0	0	0	4	0.01
R	1	8	0:41:02		384	1	0	0	1	2	0	2	0	0	0	1	0	0	0	5	0.01
R	1	15	1:09:28		338	1	0	0	1	0	0	2	0	0	0	0	0	0	0	4	0.01

Sheet1

Figure A.7. Sample of Gesture Reliability using Excel (Microsoft, 2013)

	A	B	C	D	E
1	TIME	GESTURE		TIME	GESTURE
2	PRINCIPLE CODER			ASSISTANT CODERS	
3	[00:02:45.24] B			[00:02:45.24]B	
4	[00:02:47.07] D			[00:02:47.07]D	
5	[00:02:48.22] D			[00:02:48.22]D	
6	[00:02:49.15] B			[00:02:49.15]B	
7	[00:02:51.28] C			[00:02:51.28]C	
8	[00:02:52.10] B			[00:02:52.10]B	
9	[00:02:54.09] C			[00:02:54.09] I	
10	[00:02:54.24] B			[00:02:54.24]B	
11	[00:02:55.16] B			[00:02:55.16]B	
12	[00:02:57.19] B			[00:02:57.19]B	
13	[00:02:58.17] B			[00:02:58.17]B	
14	[00:02:59.21] D			[00:02:59.21]D	
15	[00:03:00.12] B			[00:03:00.12]B	
16	[00:02:59.17] D			[00:02:59.17]D	
17	[00:03:00.06] B			[00:03:00.06]B	
18	[00:03:03.20] D			[00:03:03.20]D	
19	[00:03:04.06] B			[00:03:04.06]B	
20	[00:03:07.26] I			[00:03:07.26] M	
21	[00:03:08.17] B			[00:03:08.17]B	
22	[00:03:08.27] B			[00:03:08.27]B	
23	[00:03:09.13] B			[00:03:09.13]B	
24	[00:03:10.03] B			[00:03:10.03]B	
25	[00:03:14.16] I			[00:03:14.16]I	
26	[00:03:15.08] B			[00:03:15.08]B	
27	[00:03:17.24] B			[00:03:17.24]B	
28	[00:03:18.05] B			[00:03:18.05]B	
29	[00:03:20.09] B			[00:03:20.09]B	
30	[00:03:21.10] B			[00:03:21.10]B	
31	[00:03:22.14] B			[00:03:22.14]B	
32	[00:03:25.05] I			[00:03:25.05] M	
33	[00:03:28.07] B			[00:03:28.07]B	
34	[00:03:28.23] B			[00:03:28.23]B	
35	[00:03:30.07] M			[00:03:30.07] M	
36	[00:03:30.26] I			[00:03:30.26]I	
37	[00:03:32.14] I			[00:03:32.14]I	

All codes which differed were marked in red.

Figure A.8 Sample of Fluency Reliability using Excel (Microsoft, 2013)

Debate 1	Time	Debate 2	Debate 3
Annika	0:00	Sarah	Hayley
Hayley	0:18	Stephen	Katie
Katie	0:36	Annika	Sarah
Sarah	0:54	Hayley	Stephen
Stephen	1:12	Katie	Annika
Remove commas - do this for the next 9 minutes			
	Debate 1	Debate 2	Debate 3
Katie	0:00	36:00:00	1:12
Hayley	9:00	45:00:00	1:21:00
Sarah	36:00:00	1:12:00	18:00
Stephen	45:00:00	1:25:00	27:00:00
Reliability coding - code for the next 4 minutes			
	Debate 1	Debate 2	Debate 3
Katie	39:04:00	25:44:00	7:05:00
Hayley	59:37:00	44:40:00	22:08:00
Sarah	2:45:00	56:55:00	38:19:00
Stephen	23:15:00	10:18:00	1:13:00