The effect of gesture on verbal fluency

Jonne Beeks

ANR 962213

Bachelor Thesis

Communication and Information Sciences

Specialization Intercultural Communication

Faculty of Humanities

Tilburg University, Tilburg

Supervisor: M.W. Hoetjes

Second Reader: dr. M. B. Goudbeek

July 2013

Abstract

Spontaneous speech is filled with verbal disfluencies. Spontaneous speech, however, consists of more than just its verbal aspect. Previous research has strengthened the view that speech and gesture production seem to interact. Verbal fluency might thus be related to whether a speaker gestures or not. However, there have not been many studies relating these two aspects. The present study, therefore, attempts to identify a link between gesturing and verbal fluency by comparing non-gesturers' verbal fluency to gesturers' verbal fluency. In the current research, an existing data set containing spontaneous speech was analysed. It was found that the overall disfluency rate of gesturers was lower in comparison to non-gesturers. Also, gesturing individuals produced fewer hedges, and generated longer utterances and a greater number of words compared to their non-gesturing opposites. These results provide evidence that gesturing indeed prevents disfluencies.

Keywords: spontaneous speech, verbal fluency, gesture

Introduction

Individuals differ greatly in the fluency with which they speak. Imagine for instance the professional sports commentator who can talk continuously without even pausing to think about what to say next. At the other extreme, imagine the extremely nervous student who can barely deliver a coherent presentation, who is constantly stuttering, pausing and stammering as he gropes for what to say to his fellow students and teacher and how to say it. Verbal fluency addresses these differences.

Spontaneous speech, which is notoriously disfluent (Clark & Wasow, 1998), presents an interesting type of data for researchers interested in this aspect of human speech. Spontaneous speech consists of more than just its verbal aspect. Consider, for instance, this snippet of conversational speech reproduced from a study where respondents narrated a cartoon story (Beattie & Aboudan, 1994), in which the brackets indicate that a gesture is produced at this point in time and in this manner:

> 'So the hand is now trying to start the car.'
> [hand moves in a winding movement]

This example illustrates the multimodality of spontaneous speech, in which the speaker uses both auditory and visual cues to transmit a message to the listener, and in which the accompanying visual element completes the full interpretation of the utterance. The concurrent production of speech and gestures could define these movements as speech-accompanying, suggesting that they are closely related to the speech they accompany. Even though the earlier example illustrates the multimodality of spontaneous speech, it does not contain any of the verbal disfluencies which human speech is filled with. Now consider the following example which addresses a different speaker narrating the same cartoon (Beattie & Aboudan, 1994):

> '(pause) starting it at the front with the (pause) winder thing'
> [hand moves in a winding movement]

Here, the speech-accompanying gesture starts and finishes before its lexical affiliate is even uttered. Most notably, however, while both narrating the cartoon and producing gestures, the individual's speech contains disfluencies. Earlier research on speech-accompanying gestures (see for instance McNeill & Levy, 1982; McNeill, 1992) emphasizes the importance of gestures in relation to the thoughts of a speaker. Gestures let us observe thoughts as they occur, and are not constrained by systems and rules or standards, but how does verbal fluency relate to this connection? It has been concluded that gestures are more common during unfilled pauses than during actual articulation (Butterworth & Hadar, 1989; Schegloff, 1984; Beattie & Aboudan,1994). Other verbal fluency factors (e.g., elongated segments, fillers, repeated words, self-correction and editing expressions), however, have not been clearly investigated in relation to a speaker's gestures.

Hoetjes, Koolen, Goudbeek, Krahmer and Swerts (2011) conducted a study on repeated references with regard to their verbal and nonverbal nature. In their study, a director-matcher test was set up in which the director had to refer to the same abstract object several times. They mainly addressed the question whether reduction during repeated references to these abstract objects also occurred in speech-accompanying gestures. The present study uses the data set collected in their study. Its focus is on the spontaneous speech people produce in consultation with the simultaneously given visual cues, as compared to the spontaneous speech of people who do not produce gestures at all while communicating. The question is addressed whether people's verbal fluency is affected by the presence or absence of these speech-accompanying gestures. The results provide a closer look at the added value of gestures during the production of spontaneous speech and its overall verbal fluency.

Theoretical framework

Speech

The production of speech is a difficult process, consisting of various stages a speaker has to go through (Levelt, 1996). During the lexical selection stage, a speaker selects words that appropriately express his or her intentions. Subsequently, the speaker activates the lexical concept associated with the intention, and develops a phonological code, which, in turn, becomes encoded as an articulatory gesture (Levelt, 1996). This sequential order of speech production becomes even more difficult when speakers have to design their utterances to suit the needs of their listeners. This issue has been addressed by the development of the Dual Process Hypothesis (Bard, Anderson, Sotillo, Doherty-Sneddon & Newlands, 2000; Bard & Aylett, 2005). This hypothesis proposes a division between fast and automatic processes of speech production, and slower, more costly processes of audience design, which require inference or attention. The latter includes complex reasoning used to construct a model of the listener, while the former includes priming (Balota, Boland & Shields, 1989; Mitchell & Brown, 1988), an effect of the speaker's own recent experience.

While addressing a listener in conversation, speakers have to take the listener's lexicon into account. These lexical choices, referring to the words and phrases employed to transfer meaning to others, are independently constructed by the contribution of three different factors (Brennan & Clark, 1996). Firstly, with regard to a listeners' presence, the informativeness of an expression is an important factor while communicating. Especially while addressing a specific object or event, a speaker should provide enough information to avoid confusion on the listener's side of the conversation. Speakers should, therefore, construct their lexical choices in such a way that listeners are able to distinguish that particular referent from a set of alternatives.(Brown, 1958; Olson, 1970). These expressions are specifically selected by the speaker, in order to enable their addressee to pick out the referent uniquely. A referent should also meet Grice's (1975) maxim of quantity in which two conditions should be met: an expression should be as informative as is required for the current purposes of the informational exchange, and the expression should not be more informative than is required. An example which illustrates the application of this maxim could be a speaker's reference to a "little yellow flower". If this specific flower would grow in a field with different flowers, the referent "little yellow flower", or perhaps even "yellow flower", would suffice. The expression will, however, be more informative than required if the speaker would refer to the flower as "the little yellow flower next to the purple flowers on your left" when there is only one yellow flower present. The speaker hereby provides the listener with unnecessary information, the former, shorter, phrase would have been sufficient. Recent research has indicated that the language processing system is only moderately Gricean with respect to the maxim of quantity. Speakers often over-specify (and employ underspecified-descriptions as well; Ferreira, Slevc & Rogers, 2005) their lexical choices, without it bothering listeners (see for instance Engelhardt, Bailey & Ferreira, 2006; Koolen, Gatt, Goudbeek & Krahmer, 2009).

Informativeness is not the only contributing factor to specific lexical choices in conversation. While addressing a listener, speakers should also take into account the lexical availability of an object. This concept measures the ease with which a word can be generated as a member of a given category, and sometimes overrides Grice's maxim of quantity (Brennan & Clark, 1996). The "little yellow flower" in the previous example could, for instance, be referred to as "flower". According to the maxim of quantity such an expression would suffice if only one flower is present. While addressing the flower, however, the expression could be specified to "buttercup", hereby referring to the specie of the plant which is more informative than required in comparison to simply addressing the flower as "flower", hereby addressing its category "flowers and plants" instead of its specific specie. While addressing the "little yellow flower", "buttercup" proves to be less lexical available in comparison to the more preferable and more lexical available "flower".

Finally, another factor contributing to a specific lexical choice, that could also override the factor of informativeness, is perceptual salience. Speakers can address what is salient about an object instead of the minimally required information a listener requires to understand a reference. The "little yellow flower", for instance, addresses the salient features of an object (e.g., its size and colour) whereas the object could simply be described as "flower".

Even while considering informativeness, lexical availability, and perceptual salience, lexical variability in conversation can still be high. The likelihood, for instance, that people in one conversation would choose similar terms for the same common objects as people in another conversation is only 10% (Brennan & Clark, 1996).

Spontaneous speech and verbal fluency

The construction of lexical choices in conversation is a difficult process for speakers. Formulating an entire utterance at once can create a problem when speakers have not entirely thought out its construction or change their mind about what they are saying. Spontaneous speech, therefore, is filled with disfluencies; unwanted pauses, elongated segments, fillers (e.g., *uh* and *well*), repeated words, self-corrections, and editing expressions or hedges (such as *I mean* and *you know*) (Clark & Wasow, 1998). Verbal fluency is a term which is generally used to describe the rate at which an individual can produce words, including the amount of disfluencies a speaker produces. A fluent speaker could, for instance, be described as a speaker with normal, or even excessive, speed, normal melody and articulation, who produces long and grammatically complex sentences (Goodglass, Quadfasel & Timberlake, 1964; Howes, 1967; Benson, 1967; Goodglass & Kaplan, 1972). Hinder from disfluencies could affect these typically verbal fluency characteristics.

A conservative estimate (which excludes unfilled pauses) for the rate of disfluencies in spontaneous speech is 6 disfluencies per 100 words (Fox Tree, 1995). A corpus of speech displayed a similar rate of 5.97 per 100 words (Bortfeld, Leon, Bloom, Schober & Brennan, 2001). Earlier research on speech disfluencies has associated this typical speech phenomenon with a speaker's planning load. In Oviatt's (1995) study on disfluencies in six types of task-oriented conversations for instance, utterance length ranged from 1 to 26 words. It was concluded that long utterances had higher disfluency rates than short ones. This finding is supported by a later finding which indicates that the longer the sentence in spontaneous speech, the more likely it is to contain fillers (Shriberg, 1996). The association of disfluencies with planning load is consistent with findings that disfluencies are more likely to appear near the beginning of turns or sentences since planning load is presumably higher there (Boomer, 1965; Shriberg, 1996). The topic or domain of conversation appears to affect the planning load of utterances as well. Schachter and colleagues (Schachter, Christenfeld, Ravina & Bilous, 1991; Schachter, Rauscher, Christenfeld & Crone, 1994), for instance suggested that speakers use more fillers when they are able to choose from a larger range of expressive options. In situations in which a speaker is restricted to a certain set of terminology (e.g., hard science lectures), spontaneous speech appears to contain less disfluencies, compared to situations in which a speaker has an extensive range of phrasing (e.g., humanities lectures) (Schachter, Christenfeld, Ravina & Bilous, 1991).

Another possibility for why disfluencies are so common in spontaneous speech, besides the speaker's planning load, is that they might be used to coordinate interaction and, hereby, serve a communicative function (Shriberg, 1996). They may provide information which could enable people in conversation to better coordinate interaction, manage turn-taking, or align their mental states (see for instance, Bortfeld, Leon, Bloom, Schober & Brennan, 2001; Brennan & Schober, 2001; Brennan & Williams, 1995; Clark & Wasow, 1998; Fox Tree, 1995; Shriberg, 1996). Time, for instance, is a resource that people jointly manage in conversations. When a speaker takes a long time to produce an utterance, he or she might risk losing a listener's concentration or attention (Clark & Brennan, 1991). By producing a filler like *um* or *uh*, the speaker might warn the addressee of a delay in producing the correct word or phrase (Clark, 1994; Smith & Clark, 1993), or, on the contrary, direct the attention of the listener towards the upcoming target word (Fox Tree, 1993). Brennan and Williams (1995) showed in their research on paralinguistic cues that listeners are sensitive to filled disfluencies (like *um*, *uh*, clicks, and other noises) versus unfilled pauses in conversation. They indicated that these pauses could be used to estimate other people's knowledge of the topic of conversation, which is a necessity for coordinating individual mental states during communication.

Optimal fluency performance in a naming/verbal fluency task, as described by Troyer (2000), involves generating words within a subcategory, and switching to another subcategory when the previous one has been exhausted. Individuals who, for instance, have to recall as many animals as possible, could revert back to searching for and recalling items from a new subcategory (e.g., animals you could find at a zoo), after naming animals you could find in another setting. Subcategories can be divided into two general categories; a semantic category, such as animals or types of flowers, and a phonemic category which involves the systematic organization of sounds in languages (e.g., words that begin with the letter *p*) (Troyer, 2000).

Researchers, largely beginning with the work of Bousfield and Sedgewick (1944) on restrictive associative responses, have suggested a two-stage cyclical search process model used for long-term memory retrieval in verbal fluency tasks (Gruenewald & Lockhead, 1980; Herrmann & Pearle, 1981; Wixted & Rohrer, 1994). In the first stage of this model, it is assumed that participants search for overall categories. In the second stage, participants search for specific items within these categories. In a typical verbal fluency task, an individual has to say as many words as possible from a category in a given time. Troyer (2000), for instance, administered a fluency task by administering participants to a category in which they had to recall as many animals as possible in a 60 second time-period, or asking them to recall as many supermarket items in the same time-period.

Even though traditional verbal fluency tasks are good predictors of an individual's word fluency, researchers have described other factors that contribute to verbal fluency, and which help determine the analysis of spontaneous speech and dialogue. Speech rate, which addresses the number of words per minute, is a good example of a verbal fluency contributor (Howes & Geschwind, 1964; Benson, 1967; Kerschensteiner, Poeck & Brunner, 1972). Utterance length is also described to be an important factor in analysing verbal fluency (Goodglass & Kaplan, 1972; Wagenaar, Snow & Prins, 1975). Kreindler, Mihailescu and

Fradis (1980) added a factor they used in their research on speech fluency: the duration (in seconds) of an individual speaking. They state that utterance length alone is not a good predictor of verbal fluency. An individual could, for instance, produce a large number of utterances. This individual would in most cases be considered fluent. However, if this individual produces these utterances within a much longer time, he or she might rather be rated as non-fluent.

Analysing verbal fluency is a process which involves multiple techniques and distinctions. As indicated before, a method that has been used before is the verbal fluency task. These tasks require a targeted data set fit for these means of analyses. Spontaneous speech production is, therefore, not a valid type of data fit for the use of a verbal fluency task. To analyse this type of information, it is important to consider specific aspects for the analysis of spontaneous speech production (Wagenaar, Snow & Prins, 1975). The scoring of the items, for example, should take place in an objective and quantified manner, which facilitates the replication or application of these scales to other samples.

Gesture

Gestures are not just random movements that an individual produces. Arms waving in the air, for instance, serve not only a purely kinetic function by their movement, but are also symbols that exhibit meanings in their own right, for example: ‘I am standing here’ or ‘goodbye’ (McNeill, 1992).

There are different types of gestures that can be found in everyday conversation; iconic gestures, metaphoric gestures, emblems, beat gestures and deictic gestures. These will be described below.

Iconic gestures (McNeill, 1992) are especially relevant to the semantic content of speech because of their speech-accompanying function, which suggests that these movements are closely related to the speech they accompany. They are created by the speaker at the moment of speaking, and coexist with the words and sentences of speech, but sometimes differ from those words and sentences regarding their exact meaning. These gestures, as defined in Melinger and Levelt (2004), share a transparent relationship with some semantic aspect of the concurrent speech, often representing concrete entities, traits or activities. Beattie and Shovelton (1999b) illustrate this definition in their research by presenting informants with cartoon pictures that contained some words that represented the speech of the main characters, and asked them to narrate the stories so that any gestures they made would be recorded.

Consider the following example:

‘By squeezing his nose’
[Left hand opens
then closes]

Beattie and Shovelton (1999b) claim that this sentence conveys the action involved, in this case squeezing. It does, however, not tell how it is accomplished. The iconic gesture is critical to the communication here because it shows the method of squeezing.

Gestures that present imagery, but present an image of an abstract concept (i.e., ‘knowledge’, ‘language’), are called “metaphoric” gestures. Although this type of gesture appears to be very similar to iconics (which also present imagery), they are intrinsically more complex due to their representation of an abstract concept (McNeill, 1992). Imagine, for instance, a situation in which a speaker uses a metaphoric gesture (making a rolling motion with the hands) while talking to a listener about the length of a meeting. This specific gesture could imply that the meeting was too long, hereby addressing an abstract concept like time while making this specific motion.

A widely known type of gesture that shows cultural specificity is the “emblem”, or Italianate gesture (Ekman & Friesen, 1969; McNeill, 1992). Emblems are part of a social code but are not fully structured as a language. McNeill (1992) exemplifies this specific type of gesture by the so-called Hand Purse, in which the fingers and thumb are pressed together at the tips and pointed upwards. This emblem, which has no use at all by North Americans, symbolizes six or seven different meanings in Europe. Italians, for instance, interpret this gesture as “what do you really want? What do you mean?”, while in Belgium and France, it represents “fear”. Kendon’s (2002) study on the ‘head shake’ in multiple situations and locations illustrated a similar result. It was concluded that this well-known head gesture appeared to be an expression in its own right. It is, therefore, concluded that not all gestures are speech-accompanying. Emblems, for instance, implicate meaning on their own, as do so-called “beats”.

Beat gestures can be recognized in terms of their prototypical movement characteristics (McNeill, 1992). A beat is, in fact, a staccato strike that creates emphasis and grabs attention (Beattie, 2003). A short and single beat, for instance, can mark an important point in a conversation, while repeated beats can continuously hammer on a critical concept.

A final gesture type is the “deictic”, or pointing, gesture. Pointing gestures, which are typically performed with the pointing finger, could also be performed using extensible objects or body parts, including the head, nose, and chin. These gestures most obviously indicate objects and events in the concrete world. Most of these deictic gestures in conversation and spontaneous speech, however, play a part when there is nothing concrete to point at (McNeill, 1992; Levy & McNeill, 1992; Cassel, McNeill & McCullough, 1998). McNeill (1992) illustrates this type of gesture in the following example in which a student is asked the following question: “where did you come from before?”. The student subsequently uses a deictic gesture while answering the question, hereby aiming the gesture at the abstract concept of where he has been, instead of the actual existing physical place.

Relationship between speech and gesture

Almost all speakers spontaneously produce gestures while they speak. These gestures support and elaborate on the information that is carried out verbally (see for instance McNeill, 1985; 1992; Kendon, 2004). Gestures, as shown by Church and Goldin-Meadow (1986), also identify certain underlying reasoning processes that were not or could not initially be articulated by its speaker. The gestural channel, therefore, can provide insight into the mental representations of the speaker.

Willems, Özyürek and Hagoort (2007) address the relationship between speech and gesture as systematic and consisting of three different levels. The first level they describe expresses the semantic overlap between the representation in gestures and the meaning expressed in the concurrent speech. Speech and gesture typically convey similar or related information (see for instance McNeill, 1992). This level could be explained by an example in which a speaker makes a steering movement with his hands while simultaneously saying “He was steering the car”. The gesture in this example depicts the event as a whole by describing its manner (“steering”). The second level of the relationship between speech and gestures describes its temporal alignment to each other. Typically, the onset of the gesture (i.e., preparation) precedes the onset of the relevant speech segment by less than a second. More importantly however, in most speech and concurrent gesturing, the stroke of a gesture (which is the semantically most meaningful part of the gesture) coincides with the relevant speech segment (McNeill, 1992). Finally, it has been shown that spontaneous gesturing has a function similar to speech, namely to communicate the intended message to the listener (see for instance Kendon, 2004; Melinger & Levelt, 2004).

As indicated by McNeill (1992), gestures coexist with the words and sentences of speech, but are qualitatively different from those words and sentences. They are separate symbolic vehicles that could convey meaning independent of their concurrent speech, but can also enhance its meaning and interpretation. Consider for instance the following oft-cited example of an iconic gesture (McNeill, 1992):

> 'she [chases him out again]'
> [hand appears to swing an object through the air] (p. 13)

The speech in this example gives information about the two characters that are involved in the action where one of them is chasing the other, and that this is a recurring event. The concurrent gesture reveals that the chasing is done using an object. This example illustrates the simultaneous use of speech and gestures, where the gesture carries information that is not contained in the speech. This phenomenon lends support to the notion that an utterance is formed by both speech and gesture, and that gesture is as much a part of language as speech (McNeill, 1985; 1992; Kendon, 2004). Gestures, in this way, contribute to communication.

The earlier example illustrates the notion that gestures can convey rich semantic information, some of which is redundant with speech and some of which is supplementary. In their study on gestures and the communicative intention of the speaker, Melinger and Levelt (2004) used a picture description elicitation task in a face-to-face conversation to collect descriptions in which speakers could freely choose whether or not to simultaneously produce gestures and speech. They found that deciding to gesture while speaking influences decisions about what is explicitly mentioned in a person's speech. Some speakers divided the informational load between gesturing and actual speech. This is, according to Melinger and Levelt (2004), an obvious sign that the gestures were intended to be informative.

Graham and Argyle (1975) investigated the relationship between speech and gestures by asking six speakers to describe abstract line drawings to a small audience of listeners who subsequently tried to reproduce them. On half of the trials, speakers were prevented from gesturing by being instructed to keep their arms folded throughout the experiment. The results indicated that listeners who listened to speakers that were allowed to gesture reproduced the figures more accurately. The design of this experiment did not control for the possibility that speakers who were gesturing also produced better verbal descriptions of the stimuli, which could, in turn, enable the listeners to reproduce the figures more accurately.

Another aspect of the relationship between speech and gesture can be illustrated by the trade-off-hypothesis. This hypothesis claims that there is a relationship between speech and gesture in terms of their cognitive load (de Ruiter, Bangerter & Dings, 2012). It is assumed that when speech becomes more difficult to produce (i.e., requires more effort to verbally express intended meaning), the likelihood of gestural occurrence, to “take over” or enhance meaning of the communicative load, is higher. The trade-off-hypothesis also works the other way around, in which it suggests that when gesturing becomes harder, speakers will rely relatively more on speech.

An alternative to the trade-off-hypothesis also exists. So, Kita and Goldin-Meadow (2009) found that speakers can use gestures to identify referents across a stretch of discourse. These gestures, however, paralleled rather than compensated for underspecifications in speech when describing scenes to an experimenter. They propose that referent identification in gesture appears to be redundant with, or goes hand-in-hand with lexical specificity in speech, a theory that de Ruiter, Bangerter and Dings (2012) identified as the hand-in-hand hypothesis.

The question of whether and how gestures make a contribution to communication is complex and has resulted in a great deal of contradicting research outcomes. While some research has shown that listeners incorporate gesturally expressed information into their broader understanding of a narrative (see for instance Beattie & Shovelton, 1999a; Beattie & Shovelton, 1999b; McNeill, 1985), other research has shown that listeners’ comprehension of speech is not influenced or supplemented by the presence of gestures (see for instance Krauss, Morrels-Samuels, & Colasante, 1991; Krauss, Dushay, Chen, & Rausher, 1995). McNeill (1985), for instance, argues that gestures and speech are parts of the same psychological structure, due to their close temporal, semantic, pragmatic, pathological, and developmental parallels. Kraus, Dushay, Chen and Rauscher (1995), on the other hand, argue that visual access to gestures that are accompanied by a spoken message do not improve that message’s communicative effectiveness.

Present study

Even though the relationship between speech and gesture production has a considerable body of evidence to prove its existence, verbal fluency relating to these two aspects has not been studied as much. Several researchers have investigated disfluency rates in spontaneous speech with the intention to further extend knowledge about the different aspects of verbal fluency (see for instance Bortfeld, Leon, Bloom, Schober & Brennan, 2001; Brennan & Schober, 2001; Corley & Stewart, 2008; Oviatt, 1995; Shriberg, 1996; Unsworth, Spillers & Brewer, 2011; Seyfeddinipur, 2006; Barr & Seyfeddinipur, 2010). The specific effects of verbal fluency in relation to gesturing have not been clearly demonstrated. However, as indicated above, some research on gestures occurring during disfluencies has been done. It is, therefore, interesting to place verbal fluency in a context which has not been researched as much, and in which no evident research outcomes have been presented yet to demonstrate a possible relationship between verbal fluency and gesturing. The present study attempts to identify a link between the use of gesture and verbal fluency.

Method

The present study used the data set collected in Hoetjes, Koolen, Goudbeek, Krahmer and Swerts' (2011) research on the verbal and nonverbal nature of repeated references. In their study, a director-matcher test was set up in which the director had to refer to the same abstract object several times. For their experiment, two picture grids containing 16 abstract (fairly similar) figures were created. The figures' similarity forced the director to give a detailed description, often using many gestures in the process (see Figure 1 for an example abstract object as was used in Hoetjes et al., 2011).

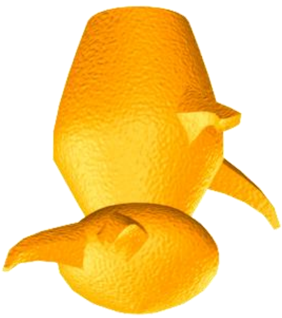

Figure 1. Example of abstract object.

Both picture grids were used for 15 trials, which made a total of 30 trials per participant. During each trial, the director had to refer to a target object (marked by a red square around the image), which was surrounded by 15 distractor objects. Subsequently, the matcher had to pick the corresponding card from a box in front of him/her which showed the same images the director was able to see. As soon as the matcher had selected a card, the next trial was started. During the director's descriptions, the matcher was allowed to ask for additional information about the object.

Data selection

As mentioned above, the current research used the data collected in Hoetjes et al. (2011). From this data set, four trials (two trials per picture grid) consisting of participants' first descriptions of the target objects were selected.

The entire set of participants consisted of 106 undergraduate students (31 male, 75 female, age range 18-29 years old, M = 21 years and 7 months), who took part in the experiment in pairs as a fulfilment of course credits. From the four trials, data of 40 participants (20 directors and 20 matchers) was selected. Because of this research's focus on verbal fluency, only the data originating from the directors (who produced the descriptions) was analysed. For each of the four trials, non-gesturing directors were selected. These participants added up to a total of ten non-gesturers. Of the ten non-gesturers, two participants produced no gestures in all four trials, one participant produced no gestures in three trials, another participant produced no gestures in two trials, and the remaining six participants produced no gestures in one trial. This, therefore, provided for a data set of 19 non-gesturing data points. Subsequently, the non-gesturing directors were matched to gesturing directors. To distinguish maximal differences between gesturers and non-gesturers, gesturers were selected based on their total amount of produced gestures in one or multiple trials. The two participants who produced the most gestures in all trials combined were matched to the two participants who produced no gestures in all trials. The participant who produced the most gestures in three trials was matched to the non-gesturing participant in the same three trials, and so on. These gesturing directors (ten in total) provided us with a data set of 19 gesturing data points. Combined, the data set consisted of 38 data points. See Table 1 for a progressive scheme of the current study's data selection.

Table 1. Progressive scheme of data selection.

Goal	Action (in sequence)
Selection entire data set	• Discriminating first descriptions • Selecting trials containing first descriptions • Discriminating directors
Selection non-gesturers	• Discriminating non-gesturers from gesturers • Sorting non-gesturers by trial
Selection gesturers	• Note gestures produced per trial (per gesturer) • Sort trials according to non-gesturers • Add up gestures produced in all trials (per gesturer) • Discriminate gesturers who produced most gestures according to amount of non-gesturers in same trials • Add up gestures produced in three trials (per gesturer) • Discriminate gesturers who produced most gestures according to amount of non-gesturers in same trials • Add up gestures produced in two trials (per gesturer) • Discriminate gesturers who produced most gestures according to amount of non-gesturers in same trials • Note gestures produced in one trial (per gesturer) • Discriminate gesturers who produced most gestures according to amount of non-gesturers in same trial

Data analysis

Speech variables

Duration of trial

To measure the total duration of a trial, the beginning and end point of a description were taken, and this total speaking time was measured in seconds. The original data set provided this information. The total duration of a trial provided a reference point for the analysis of a director's speech rate and utterance production.

Word production

The present study analysed the word production of an individual by counting the total number of words produced per trial. Filled pauses (e.g, "uh,", "um,") were not considered part of the word production due to their nonlexical function. Other disfluencies (self-corrections, hedges) were considered part of the word production. The word production provided a reference point for the analysis of participants' speech rate and disfluencies.

Verbal fluency variables

To analyse the verbal fluency of an individual's spontaneous speech, multiple aspects of this verbal feature have been shown to be of importance (see for instance Goodglass, Quadfasel & Timberlake, 1964; Howes, 1967; Benson, 1967; Goodglass & Kaplan, 1972; Fox Tree, 1995; Oviatt, 1995; Shriberg, 1996) . The present study analysed the speech rate, utterance production, total number and rate of several disfluencies, and the overall disfluency rate of gesturing and non-gesturing directors.

Speech rate

The speech rate per trial was analysed by looking at the total number of words per second.

Utterance production

The present study determined the beginning and end of an utterance by analysing different cues: (1) pauses, often followed by a filled pause (e.g., "uh,", "um,"), provided cues to utterance boundaries because of the temporal stop in speaking, (2) sentence final intonation, analysed by a lower vocal intonation at the end of an utterance, (3) conjunctions (e.g., "and", "or", "at least") provided cues to the continuation of an utterance, (4) sentence continuation intonation, analysed by a higher vocal intonation, hereby indicating a following continuation of the utterance. For each data point, the total number of utterances were analysed and used to compute the mean number of words produced per utterance.

Disfluencies

Disfluencies were classified into the following types: (1) filled pauses – spontaneous nonlexical sounds that fill pause in running speech (e.g., "uh," and "um,"), often signalling the start of a new utterance or self-correction, (2) self-corrections – spontaneously corrected misspoken words or utterances (e.g., “they are, uh, it is”), (3) hedges – words or phrases that make statements less forceful or assertive (e.g., “I think”, “maybe”, “somewhat”). For the purpose of this study, no attempt was made to code relatively minor mispronunciations such as elongated or slurred sounds, which are also more difficult to identify reliably during scoring. The total number of disfluencies for each type were converted to a rate per 100 words.

Disfluency rate

The disfluency rate addresses the total amount of disfluencies (filled pauses, self-corrections and hedges) for each data point, converted to a rate of number of disfluencies per100 words.

Statistical analysis

To test the differences on verbal fluency between gesturing and non-gesturing participants, a one-way ANOVA was used with gesturing as a between-subjects factor.

Results

As can be seen in Table 2, a one-way ANOVA test showed that gesturers speak longer (M = 53.9, SD = 23.4), compared to non-gesturing participants (M = 42.5, SD = 22.8). These results, however, were found to be non-significant ($F_{(1,36)} = 2.306$, $p = .14$). However, the mean number of words uttered was significantly higher for gesturing participants (M = 111, SD = 49), than for non-gesturing participants (M = 78.5, SD = 46.6) ($F_{(1,36)} = 4.391$, $p < .05$).

Table 2. Overview of overall mean results for dependent speech variables, for non-gesturers and gesturers.

	Non-gesturers (*SD*)	Gesturers (*SD*)
Duration (sec)	42.5 (*22.8*)	53.9 (*23.4*)
Words*	78.5 (*46.6*)	111 (*49*)

* $p < .05$

As can be seen in Table 3, compared to gesturers, non-gesturers showed a different verbal fluency in several ways. It was found that the participants showed a higher speech rate (in words per second) when they were gesturing (M = 2.11, SD = .41), compared to when they were not gesturing (*M* = 1.87, *SD* = .44). These results, however, were not significant ($F_{(1,36)}$ = 2.813, *p* = .10). Also, the mean utterance length was longer for gesturers (*M* = 19.2, *SD* = 7.7), compared to non-gesturers (*M* = 15.2, *SD* = 3.8) ($F_{(1,36)}$ = 4.112, $p < .05$). Participants who did gesture had less filled pauses (per 100 words) in their spontaneous speech (*M* = 5.53, *SD* = 3.78) than non-gesturers (*M* = 6.66, *SD* = 2.79). These results, however, did not reach significance ($F_{(1,36)}$ = 1.090, *p* = .30). Also, gesturers used less self-corrections (per 100 words) to describe the abstract objects (*M* = 3.72, *SD* = 2.44), compared to non-gesturing participants (*M* = 4.10, *SD* = 3.34). The results did, however, also not reach significance ($F_{(1,36)}$ = .161, *p* = .69). However, gesturing participants used fewer hedges (per 100 words) (*M* = 3.01, *SD* = 1.42), than non-gesturers (*M* = 5.32, *SD* = 2.76) ($F_{(1,36)}$ = 10.601, $p < .01$). Finally, gesturers also expressed a smaller total disfluency rate (per 100 words) (*M* = 12.3, *SD* = 5.6), in comparison to non-gesturing participants (*M* = 16.1, *SD* = 5.0) ($F_{(1,36)}$ = 4.910, $p < .05$). While duration, speech rate, filled pauses and self-corrections were found to be not significant, the participants' number of words uttered, mean utterance length, number of hedges (per 100 words) and total disfluency rate (per 100 words) were considered significantly different between the gesturers and the non-gestures.

Table 3. Overview of overall mean results for dependent verbal fluency variables, for non-gesturers and gesturers.

	Non-gesturers (*SD*)	Gesturers (*SD*)
Speech rate (w/sec)	1.87 (*.44*)	2.11 (*.41*)
Mean utterance length (words)*	15.2 (*3.8*)	19.2 (*7.7*)
Filled pauses (per 100 words)	6.66 (*2.79*)	5.53 (*3.78*)
Self-corrections (per 100 words)	4.10 (*3.34*)	3.72 (*2.44*)
Hedges (per 100 words)**	5.32 (*2.76*)	3.01 (*1.42*)
Disfluency rate (per 100 words)*	16.1 (*5.0*)	12.3 (*5.6*)

** $p < .01$, * $p < .05$

Discussion

The present study addressed the question whether people's verbal fluency is affected by the presence or absence of simultaneous gesturing. The research presented in this paper has shown that people who gesture produce more words in longer utterances, with a longer duration and a higher speech rate, but produce fewer disfluencies than people who do not gesture. However, not all of these findings reached statistical significance.

As discussed earlier, verbal fluency is a term which is often used to describe the rate at which an individual can produce words. In the current research there was no significant difference between the speech rate of gesturing and non-gesturing individuals. The overall mean production of words between these two groups did show a significant difference: gesturers produced considerably more words in comparison to non-gesturing individuals. It could therefore be concluded that gesturers do significantly produce more words while uttering spontaneous speech, but are not necessarily faster speakers than their non-gesturing opposites. The findings on the overall mean production of words and duration of speaking time could also provide insight into the non-significant effect on speech rate between gesturing and non-gesturing participants. Even though speech rate was found to be non-significant, the result show a difference between the gesturing- and non-gesturing-group. This difference could result from the difference on duration of speaking time between gesturers and non-gesturers, which, although not significant, was clearly present.

With regard to the mean utterance length, there was a significant difference between gesturing and non-gesturing participants. The results expressed a lower mean utterance length for non-gesturers in comparison to gesturing individuals. These results seem to contradict earlier research outcomes on utterance length in relationship to disfluencies. It has been indicated that longer utterances had higher disfluency rates than short utterances, mainly addressing filled pauses like *uh* and *um*. (Oviatt, 1995; Shriberg, 1996). In the present study, however, although the utterance length was longer for gesturers, gesturers did not produce significantly more filled pauses in comparison to non-gesturers.

Although not included in the research analysis, non-filled pauses may contribute to the overall speech rate of both gesturing and non-gesturing individuals. As indicated earlier, there was no significant difference between filled pauses by gesturers and by non-gesturers. Fillers like *um*, *uh*, or *let me think* serve as warnings to the addressee of a delay in producing the correct word or phrase (Clark, 1994; Smith & Clark, 1993). The pause following these filled pauses could have accounted for the equal results in speech rate. Speakers who had uttered a filled pause, most likely delayed in continuing their description because of the warning already provided to the addressee of the following delay. Consider, for instance, the following exemplifying sentences of a non-gesturer and a gesturer who produced a similar speech rate while providing a description. The numbers between the brackets stand for the total pause in seconds produced after uttering a filled pause:

(1) 'Eh, deze figuur eh [3.056] heeft boven een soort van eh [2.296] ja een rondje waar je 'm op zou kunnen neerzetten, lijkt het.'
'Eh, this figure eh [3.056] has on top a sort of eh [2.296] yes round shape on which you could put it, it seems.'

(2) 'Eh, deze heeft een eh [2.707] een beetje een langere neus, aan die aan die onderkant.'
'Eh, this one has a eh [2.707] rather longer nose, on the on the bottom.'

Both the non-gesturing participant in example 1, and the gesturing participant in example 2 appeared to have difficulty with providing a correct description. As indicated before, the utterance of a filled pause could provide the listener with information about the upcoming pause in speaking. Both participants continued their descriptive utterance after a short pause. Even though these utterances do not display the entire description of the participants, the small unfilled pauses within the utterances could have provided for their similar speech rate.

The results showed no significant difference between self-corrections in gesturing individuals and non-gesturing individuals. Interestingly, however, there appeared to be a significant difference in hedges between these two opposites, with gesturers producing fewer hedges than non-gesturers. As indicated earlier, hedges serve a function in which they make words or phrases less forceful or assertive (Clark & Wasow, 1998). It could be assumed that this function of transferring information between two speakers is already embedded in the gestures they simultaneously produce. Consider, for instance, the following example from the analysed data set, in which the brackets indicate the onset and ending of the gesture:

(3) 'Eh, hier zit onderaan weer zo'n bol en dan loopt 'ie weer zo uit naar beneden.'
[makes diamond shaped movement with hands]
'Eh, here at the bottom again a ball and it runs out below like this.'

The participant in this example used a gesture to represent the seemingly diamond shaped abstract figure she had to describe. A non-gesturer could have addressed this figure by verbally indicating its diamond shaped figure, or, most likely, by adding a less forceful expression like ‘sort of’, to indicate its non-exact shape. Consider, for instance, earlier discussed example 1 of a non-gesturer describing the same abstract figure. This participant does not gesture, he also verbally tries to address the diamond shaped figure on top of ‘*the round shape*’. Without the use of gestures, he appears to have difficulty with describing the abstract figure. Adding hedges (‘*sort of*’, ‘*it seems*’) to his description helps in providing a message to the listener about his uncertainty on the shape of the object.

Overall, the different disfluency types displayed mixed results in their statistical significance. The mean results of non-gesturers, however, were constantly lower in comparison to their gesturing opposites, hereby indicating a specific difference between the two groups. The total disfluency rate showed a significant difference between gesturing individuals and non-gesturing individuals. The results display a lower disfluency rate for gesturers in comparison to non-gesturing participants, and could therefore signify a specific relationship between verbal fluency and gesturing, which could not be indicated by merely analysing the separate disfluency types (which do show clear differences but do not all reach statistical significance).

The results of this study showed a connection between an individual’s verbal fluency and their concurrent use or absence of gestures. Even though not all variables expressed significant outcomes, it is shown that the mean values of disfluencies of non-gesturers are higher compared to the mean values of gesturing individuals. Also, the mean values of the total number of words, speech rate, duration and utterance length of gesturers are lower in comparison to the mean values of non-gesturers. This means that the participants who gestured produced more speech, but that speech contained fewer disfluencies. It could, therefore, be concluded that the results show a connection between fluent speech and gesturing, in which the use of gesturing causes a speaker to be more fluent in their concurrent spontaneous speech.

Even though the results show a connection between gesturing and verbal fluency, not all outcomes were considered significant (duration, speech rate, filled pauses, self-corrections). These findings could be explained by specific details considering the data set. The participants were, for instance, asked to describe abstract figures. As indicated earlier, the topic or domain of conversation can affect the planning load of utterances and the disfluencies coupled within this spontaneous speech production (Schachter, Christenfeld, Ravina & Bilous, 1991). By describing a specific abstract image, the total range of expressive options appeared to be fairly large. Participants, therefore, had the possibility to express themselves differently while addressing the (much alike) abstract figures. Confusion at the listener's side could have caused them to be in a search of words, hereby filling their speech with more disfluencies. Consider, for example, the following description provided by a participant who displayed a high speech rate in comparison to other participants:

(4) 'Eh, dit is een eh, ja die heeft zo'n groef aan de aan de bovenkant en eh, het is een eh. Ja, die steeksels zijn heel dun en smal en eh aan de, aan het ronde gedeelte onder en er zit een zo'n eh, ja driehoekig eh uitsteeksel.'
'Eh, this is a eh, yes this has a kind of cut on the on the top side and eh, it is a eh. Yes, those protrusions are really thin and narrow and eh on the, on the round section below and there is this eh, yes triangular eh protrusion.'

The participant in example 4 is clearly in search of words. Also, the participant's speech is filled with disfluencies (e.g., '*eh*', '*…een eh, ja die heeft…*', '*…aan de aan de…*'), which could indicate her uncertainty about the specific choice of words on the hard to describe abstract objects. To rule out that the overall high number of disfluencies was caused by the content of this specific dataset, future work could consider a different approach for collecting the spontaneous speech data. A method could be used in which participants are stimulated to talk about a topic they have an affinity with. The high knowledge level of the participants on the topic of their choosing would cause them to have a select lexicon in which they could express themselves. The small lexicon (in comparison to a larger lexicon on topics they are less knowledgeable of) could cause the participants to be more fluent in their speech. This method could enable us to test whether spontaneous speech production, in which a speaker is more restricted in their set of terminology, translates into a clearer relationship between an individual's verbal fluency and use of gesturing.

Earlier research has indicated a conservative estimate for the rate of disfluencies in spontaneous speech (Fox tree, 1995; Bortfeld, Leon, Bloom, Schober & Brennan, 2001). This estimate is situated around 6 disfluencies per 100 words, and is considerably lower in comparison to the disfluency rates (M = 16.1 for non-gesturers, M = 12.3 for gesturers) expressed in the current research. It could therefore be assumed that the data collected in this research differs at least in this respect from that in earlier researches with regard to the spontaneous speech production. As indicated before, the large set of terminology the directors were able to use could have accounted for this difference. Employing the method described above could intercept the dissimilarity between the current research and earlier researches on the rate of disfluencies.

Since participants were not specifically motivated or restricted in their use of gestures, the data set was restricted to those individuals who did and did not *spontaneously* gesture. These participants showed a more natural way of communicating by not being restricted or forced to gesture (as e.g. in Graham and Argyle, 1975), however, because of this naturalness (in which most people do gesture), the entire data set was quite small. It could be assumed that, due to its small size, the data set employed was not entirely representative for the current research. Even though the current study provided for significant research outcomes, which indicated a relation between gesturing and verbal fluency, a follow-up study, in which a larger data set would be utilised, would possibly show even more significant results. The larger data set would enable us to compare larger groups of gesturers and non-gesturers, hereby examining a more representative view of the target groups.

Conclusion

In conclusion, in this study on the relationship between verbal fluency and gestures, it was found that people who gesture display a better verbal fluency in comparison to non-gesturers. In particular, the total number of words produced during an individual’s expression of spontaneous speech, their mean utterance length, the total amount of hedges employed, and their overall disfluency rate differ between gesturing and non-gesturing individuals. Gesturing individuals generated fewer hedges, hereby indicating a better verbal fluency than their non-gesturing opposites. Gesturers also exhibited longer utterances and used more words while uttering spontaneous speech in comparison to non-gesturers. These results indicate an important characteristic of the ‘fluent speaker’, who produces long sentences. Also, when addressing several disfluency types (filled pauses, self-corrections, hedges) together, research outcomes showed a significant difference between gesturing and non-gesturing individuals. Overall, gesturers generated fewer disfluencies, and therefore demonstrated a lower disfluency rate, in comparison to their non-gesturing opposites, hereby indicating a difference between gesturers and non-gesturers with regard to their verbal fluency. Gesturing, therefore, causes an individual to be more fluent, in comparison to the absence of gesturing.

In previous research, the effect of gesturing on an individual’s verbal fluency has not yet been clearly demonstrated. The results in the current study, however, provide a beginning for investigating these interesting aspects of human conversation.

References

Balota, D. A., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: beyond pattern-recognition and onset latency. *Journal of Memory and Language, 28,* 14-36.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language, 42,* 1-22.

Bard, E. G., & Aylett, M. (2005). Referential form, word duration, and modeling the listener in spoken dialogue. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: bridging the language-as-product and language-as-action traditions* (pp. 173-191). Cambridge: MIT Press.

Barr, D. J., & Seyfeddinipur, M. (2010). The role of fillers in listener attributions for speaker disfluency. *Language and Cognitive Processes, 25*(4), 441-455.

Beattie, G. (2003). *Visible thought: the new psychology of body language*. New York: Routledge.

Beattie, G., & Aboudan, R. (1994). Gestures, pauses and speech: an experimental investigation of the effects of changing social context on their precise temporal relationships. *Semiotica*, *99,* 239-272.

Beattie, G., & Shovelton, H. (1999a). Do iconic hand gestures really contribute anything to the semantic information conveyed in speech? An experimental investigation. *Semiotica*, *123,* 1-30.

Beattie, G., & Shovelton, H. (1999b). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, *18,* 438-462.

Benson, D. F. (1967). Fluency in aphasia: correlation with radioactive scan localization. *Cortex, 3,* 373-392.

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech*, *8,* 148-158.

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.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of restricted associative responses. *Journal of General Psychology, 30*, 149-165.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1482-1493.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*, 274-296.

Brown, R. (1958). *Words and things*. Glencoe, IL: The Free Press.

Cassel, J., McNeill, D., & McCullough, K. (1998). Speech-gesture mismatches: evidence for one underlying representation of linguistic & nonlinguistic information. *Pragmatics & Cognition, 6*(2), 1-24.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43-71.

Clark, H. H. (1994). Managing problems in speaking. *Speech Communication, 15,* 243-250.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine & S. D. Teasly (Eds.), *Perspectives on socially shared cognition* (pp. 222-233). Washington, DC: Morgan Kaufman Publishers, Inc. (Reprinted from: Groupware and computer-supported cooperative work: assisting human-human collaboration).

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*, 201-242.

Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: the meaning of *um*. *Language and Linguistics Compass, 2*, 589-602.

de Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: investigating the trade-off hypothesis. *Topics in Cognitive Science, 4*(2), 232-248.

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

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language*, *54*, 554-573.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, *96*, 263-284.

Fox Tree, J. E. (1993). *Comprehension after speech disfluencies*. Unpublished doctoral dissertation. Stanford University. Standford, CA.

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

Goodglass, H., & Kaplan, E. (1972). *The Assessment of Aphasia and Related Disorders*.

Philadelphia: Lea & Febiger.

Goodglass, H., Quadfasel, F. A., & Timberlake, W. H. (1964). Phrase length and the type and severity of aphasia. *Cortex, 1*, 133-153.

Graham, J. A., & Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology, 10,* 57-67.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: speech acts* (pp. 41-58). New York: Academic Press.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 225-240.

Herrmann, D. J., & Pearle, P. M. (1981). The proper role of clusters in mathematical models of continuous recall. *Journal of Mathematical Psychology, 24,* 139-162.

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., & Swerts, M. (2011). GREEBLES Greeble greeb. On reduction in speech and gesture in repeated references. In L. Carlson, C. Hoelscher & T. F. Schipley (Eds.), *33rd Annual Conference of the Cognitive Science Society* (pp. 3250-3255). Boston: Cognitive Science Society.

Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Psychology, 26*(1), 4-27.

Howes, D. (1967). Some experimental investigations of language in aphasia. In K. Salzinger & S. Salzinger (Eds.), *Research in Verbal Behavior and Some Neurophysiological Implications* (pp. 181-197). New York: Academic Press.

Howes, D., & Geschwind, N. (1964). Quantitative studies of aphasic language. *Research Publications Association for Research in Nervous Mental Disorders, 42*, 229-244.

Kendon, A. (2002). Some uses of the head shake. *Gesture*, *2*(2), 147-182.

Kendon, A. (2004). *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.

Kerschensteiner, M., Poeck, K., & Brunner, E. (1972). The fluency-nonfluency dimension in the classification of aphasic speech. *Cortex*, *8*, 233-247.

Koolen, R., Gatt, A., Goudbeek, M., & Krahmer, E. (2009). Need I say more? On factors causing referential overspecification. In *Proceedings of the CogSci workshop on the Production of Referring Expressions (PRE-CogSci 2009),* Amsterdam, The Netherlands

Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology, 31*, 533-552.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*, 743-754.

Kreindler, A., Mihailescu, L., & Fradis, A. (1980). Speech fluency in aphasics. *Brain and Language, 9*, 199-205.

Levelt, W. J. M. (1996). *A theory of lexical access in speech production.* Paper presented at the COLING '96 Proceedings of the 16th Conference on Computational Linguistics Stroudsburg, PA.

Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes, 15*, 277-301.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92*, 350-371.

McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago/London: The University of Chicago Press.

McNeill, D., & Levy, E. (1982). Conceptual representations in language activity and gesture. In R. J. Jarvella & W. Klein (Eds.), *Speech, place and action*. New York: John Wiley and Sons.

Melinger, A., & Levelt, W. J. M. (2004). Gesture and the communicative intention of the speaker. *Gesture, 4*(2), 119-141.

Mitchell, D. B., & Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. *JEP:LMC, 14*, 213-222.

Olson, D. R. (1970). Language and thought: aspects of a cognitive theory of semantics. *Psychological Review, 77*, 257-273.

Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language, 9*, 19-35.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychological Science, 7*(4), 226-231.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 60*, 362-367.

Schachter, S., Rauscher, F., Christenfeld, N., & Crone, K. T. (1994). The vocabularies of academia. *Psychological Science, 5*, 37-41.

Seyfeddinipur, M. (2006). *Disfluency: interrupting speech and gesture.* Unpublished doctoral dissertation, Radboud University, Nijmegen.

Shriberg, E. (1996). *Disfluencies in Switchboard.* Paper presented at the International Conference on Spoken Language Processing (ICSLP '96), Philadelphia.

Smith, V., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language, 32*, 25-38.

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

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 370-378.

Wagenaar, E., Snow, C., & Prins, R. (1975). Spontaneous speech of aphasic patients: a psycholinguistic analysis. *Brain and Language, 2*, 281-303.

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.

Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: an integrative review of the empirical literature. *Psychonomic Bulletin & Review, 1*, 89-106.