Available online at www.sciencedirect.com

ScienceDirect

Speech Communication 72 (2015) 160–175

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

On what happens in gesture when communication is unsuccessful

Marieke Hoetjes *, Emiel Krahmer, Marc Swerts

Tilburg Center for Cognition and Communication (TiCC), Tilburg University, The Netherlands

Received 9 July 2014; received in revised form 16 April 2015; accepted 2 June 2015
Available online 6 June 2015

Abstract

Previous studies found that repeated references in successful communication are often reduced, not only at the acoustic level, but also in terms of words and manual co-speech gestures. In the present study, we investigated whether repeated references are still reduced in a situation when reduction would not be beneficial for the communicative situation, namely after the speaker receives negative feedback from the addressee. In a director–matcher task (experiment I), we studied gesture rate, as well as the general form of the gestures produced in initial and repeated references. In a separate experiment (experiment II) we studied whether there might (also) be more gradual differences in gesture form between gestures in initial and repeated references, by asking human judges which of two gestures (one from an initial and one from a repeated reference following negative feedback) they considered more precise. In both experiments, mutual visibility was added as a between subjects factor. Results showed that after negative feedback, gesture rate increased in a marginally significant way. With regard to gesture form, we found little evidence for changes in gesture form after negative feedback, except for a marginally significant increase of the number of repeated strokes within a gesture. Lack of mutual visibility only had a significant reducing effect on gesture size, and did not interact with repetition in any way. However, we did find gradual differences in gesture form, with gestures produced after negative feedback being judged as marginally more precise than initial gestures. The results from the present study suggest that in the production of unsuccessful repeated references, a process different from the reduction process as found in previous studies in repeated references takes place, with speakers appearing to put more effort into their gestures after negative feedback, as suggested by the data trending towards an increased gesture rate and towards gestures being judged as more precise after feedback.

Keywords: Gesture; Speech; Repeated references; Negative feedback

1. Introduction

People often refer to objects and persons during a communicative exchange. In many cases, the same target is referred to repeatedly in the discourse, and these references may be multimodal, using both speech and manual co-speech gesture. It is well established that repeated references in successful communication tend to be reduced variants of initial references, consisting of less words and less gestures. For example, a speaker who wants to point out a particular person for an addressee might produce an initial description such as "that tall girl with the long blond hair", accompanied by two gestures, first one indicating the height of the girl, followed by another one indicating the length of the girl's hair. Later on in the conversation, the speaker might refer back to the same girl by saying "the tall girl from before", accompanied by only one gesture, say, indicating the girl's height.

These reduction effects have been explained in terms of increased common ground (e.g., Clark and Wilkes-Gibbs, 1986; Galati and Brennan, 2014; Gerwing and Bavelas, 2004; Holler and Stevens, 2007; Jacobs and Garnham,

* Corresponding author at: Room D 404, PO Box 90153, 5000 LE Tilburg, The Netherlands. Tel.: +31 13 466 2918.
E-mail addresses: m.w.hoetjes@tilburguniversity.edu (M. Hoetjes), e.j.krahmer@tilburguniversity.edu (E. Krahmer), m.g.j.swerts@tilburguniversity.edu (M. Swerts).

http://dx.doi.org/10.1016/j.specom.2015.06.004
0167-6393/

2007). The initial description introduces an entity in common ground, after which a reduced reference can be sufficient. The emergence of common ground is the result of a process often referred to as information grounding (Clark and Schaeffer, 1989; Traum, 1994), and generally understood as involving two phases: a presentation phase, in which a speaker sends a message to the addressee, and an acceptance phase, in which the addressee signals whether the message came across in good order or not. If our addressee knows which tall, long-haired girl the speaker is referring to, he[1] can signal this using a positive "go on" signal (using the terminology of Krahmer et al., 2002). This can, for example, be an explicit backchannel cue such as "OK", but it may also be a more implicit signal, because the addressee correctly identifies the target girl, e.g., by looking at her.

Now, consider what would happen if the initial reference is somehow not successful, which our addressee would indicate during the acceptance phase using a negative, "go back" signal (e.g., "Sorry, which girl?"). Then, how would our speaker realise her second, repeated reference to said girl? We know from other studies that speakers tend not to reduce their utterances (in terms of number of words or articulatory effort) in response to negative feedback, but we know remarkably little about whether, and if so, how, speakers' gestures would change. To the best of our knowledge only a handful of earlier studies asked this question, of which Holler and Wilkin (2011) is arguably the most detailed. However, these authors present their work as "a first glimpse of speakers' gestural behaviour in response to addressee feedback" (Holler and Wilkin, 2011, p. 3534), and point out that more work is "urgently needed" (ibid.).

In the present study we address the above questions by comparing gestures produced in initial references with those in repeated references following negative feedback. The experiments that were conducted for this purpose are based on the experimental paradigm of our previous work on successful repeated references (Hoetjes et al., 2011, 2015). As in this previous work (as well as in various other studies, including the aforementioned Holler and Wilkin, 2011), we concentrate on two aspects: the gesture rate and the qualitative form of the gestures. Before describing our current study in detail, we provide an overview of relevant background literature.

2. Background

2.1. Reduction in successful repeated references

Repeated references occur in discourse whenever a particular person or object is mentioned or described more than once. These references are never exactly the same. The differences in the ways in which references are realised are not only due to naturally occurring variability in speech, but are also influenced by the mere fact that the information status of the referent changes when it gets repeated. For instance, when an object is mentioned a second time, it already belongs to the discourse model of speaker and addressee, and can be assumed to be common ground (that is, when communication was successful). Research has found that when information is given or predictable, such as is the case in repeated references and increased common ground, speech is often reduced.

For example, Lieberman (1963) found that words produced in contexts in which they were predictable, had a shorter duration, and a lower pitch peak (F0). In addition, they were less intelligible when they were taken out of context. In a similar vein, references to given information have been found to be less intelligible when taken out of context and presented in isolation (e.g., Bard et al., 2000; Fowler and Housum, 1987), and to have a shorter duration and a lower pitch peak (e.g., Aylett and Turk, 2004; Brown, 1983; Fowler and Housum, 1987; Lam and Watson, 2010), than references to information that is new in the discourse.

Reduction in repeated references at the lexical level has also been well established. For example, Clark and Wilkes-Gibbs (1986) showed that when speakers repeatedly (and successfully) refer to the same object, they lexically reduce their references (e.g. from an initial description such as "a person who's ice skating, except they're sticking two arms out in front", to a sixth description of the same figure as "the ice skater", Clark and Wilkes-Gibbs, 1986, p. 12). This robust finding has often been explained in terms of the creation of a conceptual pact (Brennan and Clark, 1996), which occurs as more common ground emerges between speakers.

These findings relate to spoken language, but human speakers are known to produce speech in tandem with a variety of visual cues, of which manual gestures are our main focus of attention in this study. Such manual speech-accompanying or co-speech gestures (which we will call gestures for short) can generally be defined as symbolic movements of the arms and hands that people produce when they speak (Kendon, 1980, 2004; McNeill, 1992). Most researchers agree that there is a close, co-expressive relationship between speech and gesture (Kendon, 1972, 1980, 2000, 2004; McNeill, 1985, 1992; McNeill and Duncan, 2000), with speech and gesture arguably going "hand-in-hand" (e.g., Kita and Özyürek, 2003; So et al., 2009). To take one, more or less arbitrary, example, consider the study reported by So et al. (2009), who asked English speakers to retell stories to an experimenter. So and colleagues found that speakers often used gestures to identify a referent in the story, by producing it in the same location used for the previous gesture for this referent. However, importantly, they did this most often when the referent was also uniquely specified in the accompanying speech. This led these authors to conclude that for

[1] Throughout this paper, 'she' will be used to indicate the speaker, and 'he' to indicate the addressee.