Title
On what happens in speech and gesture when communication is unsuccessful

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

ISSN
1069-7977

Authors
Hoetjes, Marieke
Krahmer, Emiel
Swerts, Marc

Publication Date
2014

Peer reviewed

On What Happens in Speech and Gesture when Communication is Unsuccessful

Marieke Hoetjes (m.w.hoetjes@tilburguniversity.edu)

Emiel Krahmer (e.j.krahmer@tilburguniversity.edu)

Marc Swerts (m.g.j.swerts@tilburguniversity.edu)

Tilburg center for Cognition and Communication (TiCC), Tilburg University, the Netherlands.

Abstract

Repeated references have been found to be reduced as compared to references that are not repeated, both in speech and in gesture. In the present study we wanted to see whether certain factors can inhibit this reduction in repeated references. In a production experiment, speakers were confronted with negative feedback after an initial description of an object, indicating that the communication was unsuccessful. We found that after initial negative feedback, second references were reduced with regard to all speech variables. When the speakers were confronted with additional negative feedback, the ensuing third references were increased in the number of words and the duration, as compared to the second references, but further reduced in their speech rate. Gesture rate increased in third references as compared to initial references. After (repeated) instances of unsuccessful communication, speakers speak slower and increase their gesture rate, thereby making their repeated references clearer for the addressee.

Keywords: Gesture; repeated references; negative feedback.

Introduction

Repeated references in speech and gesture

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.

Research has found that repeated references in speech are often reduced, as compared to initial references. This reduction can be acoustic (Aylett & Turk, 2004; Bard, et al., 2000; Fowler, 1988; Lindblom, 1990), lexical (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986), or both (Galati & Brennan, 2010). For example, a person might first be described as ‘that tall guy with the funny hat’, followed by a lexically reduced reference to ‘the hat guy’.

This reduction process in repeated references extends to gesture production. It has been found that repeated references contain fewer gestures (de Ruiter, Bangerter, & Dings, 2012; Hoetjes, Koolen, Goudbeek, Krahmer, & Swerts, 2011, under revision; Jacobs & Garnham, 2007) and that gestures in repeated references are smaller (Vajrabhaya & Pederson, 2013) and less precise (Galati & Brennan, 2014; Gerwing & Bavelas, 2004; Hoetjes, et al., 2011, under revision) than gestures in initial references.

We can relate these reduced repeated references to the idea that during conversation, speakers and addressees try to minimize their collaborative effort (Clark & Wilkes-Gibbs, 1986), while being as communicatively efficient as possible (Jaeger & Tily, 2011; Zipf, 1936). This is also in line with Grice’s (1975) Maxim of Quantity which states that speakers should make their contribution to the discourse “as informative as required” but “not more informative than is required”.

However, speakers may also repeat words in a markedly different context, for example in a situation in which a problem occurred during interaction and it is unsure whether the reference to an object is shared by the speaker and the addressee. The question is whether in cases like this, an opposite process could also take place, in which speakers do *not* reduce, or might even increase, their (speech and) gesture production of repeated references, especially if doing so is communicatively necessary or efficient. In this paper we investigate a likely scenario in which a repeated reference may be expanded rather than reduced, namely when a speaker receives (negative) feedback.

Effect of feedback on gesture production

It can be argued that in cases of successful repeated references the addressee (implicitly) informs the speaker that less (gestural) information suffices. However, what about situations of miscommunication, or situations in which the addressee needs *more* information? Lindblom (1990) claims, in his theory of hyper & hypo-speech, that speakers reduce when this is possible, but hyper-articulate when this is necessary. In fact, Oviatt, Bernard and Levow (1998) found that when users of a simulated automatic speech recognition system corrected recognition errors, they tended to hyper-articulate, for example by lengthening speech segments.

Providing feedback is the simplest way in which the addressee can explicitly tell the speaker that more (or less) information is required. Previous research on the effect of feedback in spoken conversation found that more details were given when the addressee was attentive and provided feedback (Kuhlen & Brennan, 2010). Previous research on the effect of feedback on gesture showed that gesture rate was reduced when addressees were less attentive and did not provide positive feedback (Jacobs & Garnham, 2007) or when they did not provide any feedback at all (Beattie & Aboudan, 1994). Holler and Wilkin (2011) found that when speakers provided confirmatory responses (i.e., after positive feedback by the addressee), speakers' gesture rate was reduced, but when addressees provided feedback that "encouraged clarification, elaboration or correction" (Holler & Wilkin, 2011, p. 3522), gesture rate was not reduced but remained constant. These results suggest that the type of feedback that is provided matters for gesture production.

In our previous study on reduction in speech and gesture in repeated references (Hoetjes, et al., 2011, under revision), the speaker was always given implicit positive feedback (since he or she could only go to the next trial once the matcher had found the correct object). In that study, we found that the number of gestures was reduced in repeated references. However, it is reasonable to assume that when it is communicatively necessary or efficient for the speaker to be clearer, for example in the case of negative feedback, speakers do not reduce the gestures in their repeated references, but keep their number, rate and execution constant or even increase them. Increased gesture production due to negative feedback would fit the general idea that speakers tend to be communicatively efficient. After all, if a reference is not merely repeated in the discourse because it happens to re-occur, but because the negative feedback suggests that the reference is not shared (as is the case in repeated references in unproblematic types of discourse), then the speaker could decide to try and put more effort in a following reference, for example by making this reference clearer, not just in speech but also in gesture. We aim to find out whether this is the case by looking at the evolution of repeated references after negative feedback.

Present study

We conducted a production experiment where, in some trials, speakers were confronted with negative feedback from an addressee after producing a referring expression, which lead to repeated reference production. We studied several aspects of speech and gesture to see whether these repeated references were increased with regard to speech and gesture production. To facilitate comparison, the method that was used was largely identical to our previous study on reduction in speech and gesture, as reported in the 2011 CogSci proceedings (Hoetjes, et al., 2011).

Considering the fact that speakers tend to take the addressee into account and tend to be communicatively efficient, we expect that after negative feedback speakers will try to make their repeated references clearer than their initial references, in speech especially by speaking with a lower speech rate, and in gesture by producing more gestures that are also increased in their execution (i.e., gestures that take longer, are larger and more often two-handed).

Method

In order to study unsuccessful communication, participants were videotaped while taking part in a director-matcher task. In this task, the director had to describe an abstract object to the matcher in such a way that the matcher could identify the correct object from a range of similar looking objects. In some trials, the matcher purposely and repeatedly chose the wrong object, causing the director to produce a repeated reference.

Participants

Participants were 38 Dutch undergraduate students (9 male, 29 female, age range 18-30 years old, M = 21 years and 7 months), who took part as partial fulfilment of course credits. The participants took part in the experiment in the role of director. A confederate took part in the role of matcher. This confederate was the same person (female, 20 years old) for all participants.

Stimuli

Following Hoetjes, et al. (2011, under revision), stimuli consisted of picture grids of so-called Greebles[1] (see Gauthier & Tarr (1997) for a detailed description of the Greebles and their properties). We used Greebles because they caused the director to provide detailed descriptions, accompanied by many (mainly iconic) gestures (see Hoetjes, et al., 2011, for more details on why the Greebles were used). We created two picture grids, each containing 16 Greebles. In each trial, one object in the picture grid was surrounded by a red square, indicating that this was the target object of that trial. An example of a picture grid can be seen in Figure 1. The order in which the two picture grids were presented was counterbalanced over participants.

The experimental manipulation was that several objects had to be described repeatedly due to apparent communication problems: in each of the picture grids, seven objects had to be described once (these were fillers), one object had to be described twice, and two objects had to be described three times. The objects that had to be described repeatedly were always preceded and followed by an object that had to be described only once. The repeated references to the same object had to be given one straight after the other, when negative audio feedback provided by the matcher made clear to the participant that an incorrect object had been chosen. For the current paper we analysed all three descriptions of the objects that had to be described three times.

[1] Images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology Carnegie Mellon University. URL: http://www.tarrlab.org/

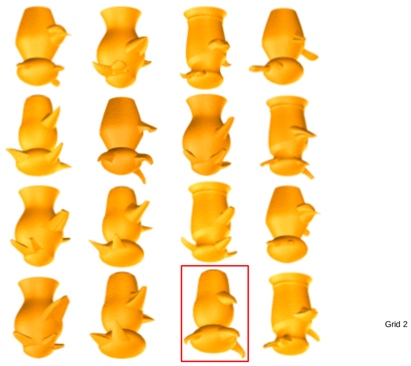

Figure 1: Example of one of the picture grids. The target object of this particular trial is marked by a red square.

Procedure

The experiment consisted of a director-matcher task that was performed in a lab, where the director and the matcher were seated at a table opposite each other (see figure 2 for an example of the setup). The director was presented with the trials on a computer screen, and the task was to provide a description of the target object in such a way that it could be distinguished, by the matcher, from the 15 distractor objects. The director was told that the matcher was shown the same objects on her screen as on the director's screen, but that these objects were ordered differently for the director and the matcher (this meant that the director could not use the object's location in the grid for the target descriptions).

Figure 2: Example of experimental setup. The director is seen from the back, viewing one of the picture grids.

The director was told that, on the basis of his/her target description, the matcher picked the object that she thought was being described. After the matcher had picked one of the objects, a sound would tell whether the matcher had chosen the correct object or not (a low buzzing sound was played for an incorrect object identification and a high ping sound was played for a correct object identification). In reality however, and unknown to the director, the director and the matcher both viewed the same picture grid and all the matcher had to do was play one of the sounds after the director had given a description of the target object of that particular trial.

When the sound indicating incorrect object identification was played, the director had to continue describing the same target picture until the sound indicating correct object identification was played. Once finished successfully, the director went on to the next trial. The feedback given by the matcher only consisted of the sounds that were played after each trial and there was no free conversation between the director and the matcher. This was done in order to ensure that the same unspecified negative feedback was provided across participants and (target) items (cf. Healey, Mills, & Eshgi, 2013, for a similar approach) .

After 10 trials (and a total of 15 descriptions), the director was shown a second picture grid containing 16 new objects, and continued for another 10 trials (i.e., 15 descriptions). The participants were debriefed at the end of the experiment, and none of the participants expressed any suspicions concerning the experimental set-up.

For half of the participants, there was a screen between the director and the matcher, causing a lack of mutual visibility. However, due to lack of space we will not describe the effect of (lack of) visibility in detail here. Overall, the lack of visibility caused references to be reduced, in speech and gesture, as compared to when there was mutual visibility. Important for the present paper, is that we found no interaction between visibility and repetition. We will discuss the effect of repetition in detail below.

Data analysis

The experiment was filmed, the video recordings were digitized, and the recordings showing the director were annotated using the multimodal annotation program ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). The subsequent speech and gesture annotation and data analysis were largely based on our previous research on repeated references, as reported in Hoetjes, et al. (2011, under revision).

For the *speech analyses* we analysed the duration (in seconds), the number of words, and the speech rate (in number of words per second), all per trial. For the *gesture analyses* all gesture stroke phases (main expressive part of the gesture, as proposed by Kendon (2004) and McNeill (1992) were selected. We then ran analyses on the quantity of gestures (in numbers and gesture rate, i.e. number of gestures per 100 words, per trial), and also on several aspects of the execution, or quality, of the gestures. The analyses on the quality of the gestures were based on a smaller dataset (of 29 participants) in which all descriptions were accompanied by at least one iconic gesture (the focus on iconic gestures was due to the fact that, unsurprisingly given the affordances of the stimuli, iconic gestures were the type of gesture produced most often). The following gesture aspects were taken into account. We measured the duration of the gesture stroke, the size of each gesture (by coding whether the stroke was produced using a finger (1), the hand (2), the forearm (3) or the entire arm (4)), annotated whether the gesture was produced using one hand or two hands (resulting in a range from 1-2, with e.g. 1.3 indicating that 30% of gestures were two-handed), and we counted the number of repeated strokes per trial (if any). A stroke was considered to be repeated when identical strokes followed each other without a retraction phase in between.

Speech and gesture analysis were conducted for all three reference descriptions of the objects that had to be described three times. The statistical procedure consisted of two repeated measures ANOVAs, one by participants (F_1) and one by items (F_2). On the basis of these, *minF'* was computed (Barr, Levy, Scheepers, & Tily, 2013; Clark, 1973), so that the results can be generalised over participants and items simultaneously, while keeping the experimentwise error rate low (Barr, et al., 2013, p. 268). Bonferroni tests were used for post hoc multiple comparisons.

Results

Effects on speech

In table 1, the means and standard errors of the dependent variables in speech for all three object descriptions can be found.

Table 1. Overview of means and standard errors (SE) for duration (in seconds), number of words, speech rate (in number of words per second)), as a function of Repetition (three levels). Star indicates significant *minF'*.

	Initial (SE)	Second (SE)	Third (SE)
Duration *	39.7 (2.5)	28.9 (1.6)	33.2 (1.8)
Words*	85.0 (6.0)	55.4 (3.4)	58.7 (3.9)
Speech rate*	2.1 (.05)	1.9 (.05)	1.7 (.05)

Firstly, we found that the second descriptions were shorter than the initial descriptions and the third descriptions were longer than the second descriptions, but shorter than the initial ones (see table 1). This effect of repetition was significant, $F_1\,(2,72) = 29.22$, $p < .001$, $\eta_p^2 = .448$; $F_2\,(2,9) = 7.20$, $p < .05$, $\eta_p^2 = .616$; *minF'*(2,14) =5.78, $p < .05$. Post hoc Bonferroni analyses showed that all three descriptions differed from each other ($p < .05$).

Secondly, we analysed the number of words used in the object descriptions. The second descriptions contained fewer words than the initial descriptions, and the third descriptions contained more words than the second descriptions, but fewer words than the first descriptions (see table 1). This effect of repetition was significant, $F_1\,(2,72) = 29.22$, $p < .001$, $\eta_p^2 = .448$; $F_2\,(2,9) = 15.91$, $p < .01$, $\eta_p^2 = .780$; *minF'* (2,21) = 10.29, $p < .001$. Post hoc Bonferroni analyses showed that the initial descriptions differed from the second descriptions and from the third descriptions (both $p < .01$). The second and third descriptions did not differ significantly from each other.

Thirdly, we analysed the speech rate. Speech rate was reduced for each repeated reference (see table 1). This effect of repetition was significant, $F_1\,(2,72) = 30.61$, $p < .001$, $\eta_p^2 = .460$; $F_2\,(2,9) = 18.19$, $p < .01$, $\eta_p^2 = .802$; *minF'*(2,22) = 11.40, $p < .001$. Post hoc Bonferroni analyses showed that all descriptions differed from each other ($p < .01$).

Effects on gesture

In table 2, the means and standard errors of the dependent variables in gesture for all three object descriptions can be found.

Table 2. Overview of means and standard errors for number of gestures (per trial), gesture rate per trial (in number of gestures per 100 words), gesture duration (in seconds), gesture size (range 1-4), number of hands (where 1.4 indicates that 40% of the gestures were produced using two hands), and gesture repetition (number of repeated strokes), as a function of Repetition (three levels). Star indicates marginally significant *minF'*.

	Initial (SE)	Second (SE)	Third (SE)
Number of gestures	3.3 (.49)	2.6 (.38)	3.3 (.52)
Gesture rate *	4.1 (.67)	4.8 (.79)	5.3 (.74)
Gesture duration	1.1 (.07)	1.2 (.09)	1.1 (.06)
Gesture size	2.9 (.10)	2.9 (.09)	2.9 (.09)
Number of hands	1.5 (.06)	1.4 (.06)	1.3 (.05)
Gesture repetition *	.33 (.06)	.50 (.10)	.55 (.09)

The analyses for the gesture rate showed a higher gesture rate for each repetition (see table 2). This effect of repetition was significant over participants and items, $F_1\,(2,72) = 7.1$, $p < .01$, $\eta_p^2 = .165$; $F_2\,(2,9) = 4.8$, $p < .05$, $\eta_p^2 = .516$, but only marginally significant for *minF'*, *minF'*(2,24) = 2.86, $p = .07$. Post hoc Bonferroni analyses showed that the gesture rate of the initial descriptions differed from gesture rate of the third descriptions ($p < .01$).

For the number of stroke repetitions within each gesture we found an effect of repetition over items and a marginal effect over participants and for *minF*', $F_1\,(2, 54) = 3.24$, $p = .06$, $\eta_p^2 = .107$; $F_2\,(2,9) = 13.64$, $p < .05$, $\eta_p^2 = .752$; *minF'*(2,62) = 2.61, $p = .08$, with an increase for each following description. Post hoc Bonferroni analyses showed a difference between gestures from initial descriptions and from third descriptions ($p < .05$).

The other gesture aspects that were analysed (number of gestures, gesture duration, gesture size, number of hands), showed no significant effect of repetition in *minF'*, although in some cases there was an effect of repetition over participants (for number of gestures), or over items (for number of hands).

Discussion and Conclusion

For speech duration and the number of words we found that there was reduction in the second description compared to the initial description, and an increase in the third description as compared to the second description. This may be explained by different strategies that participants used for the second and third descriptions. When an object had to be described for a second time, many participants provided some additional information to their initial description but did not repeat most of the information from the initial description, making the second description shorter than the first. When the same object had to be described for a third time, many participants decided to take time to repeat information from the initial description, and sometimes also added additional information (see example 1). This increase, or reset, in the third descriptions as compared to the second descriptions is in line with Oviatt et al.'s (1998) findings on repeated descriptions after recognition errors.

Example 1. Initial, second and third descriptions of the same object, translated into English (original in Dutch). Gestures occurred during the speech that is placed between square brackets.

Initial description (30.2 seconds, 86 words, 3 gestures):
"This is more like a vase with a kind of shell as as yes arm or something, how should I know what it is. Let's say at the widest part of the vase there is a kind kind ***[****of shell that you can listen to****]*** *coming out. To the right of it there is* ***[****a point going to the right****]****. And then yes it is quite smooth going down, you can see a* ***[****clear ball****]*** *as a kind of head let's say, that is only one."*

Second description (20.9 seconds, 50 words, 4 gestures):
"And damn. Uh ***[****the shell****]*** *of the arm by the way is pointing its opening* ***[****downward****]****. Uh and then yes just above the bottom of that shape there are, sticking out* ***[****a point to the left****]*** *and a kind of blunt* ***[****round****]*** *nose there seems to be in the middle."*

Third description (26.5 seconds, 60 words, 5 gestures):
"Uh oh wait. It's the only shape that has let's say if you look ***[****at****]*** *the the* ***[****middle****]*** ***[****of the shape that****]*** *goes down, there seems to be a kind of elongated ellips, a kind of nose-like something and to the left of that nose there is* ***[****a point****]****, quite a large point like a toucan's beak* ***[****to the left****]****."*

The reduction we found in speech rate means that speakers tended to produce fewer words per second for each subsequent reference. This could be because they slowed down their speech, because they produced more empty pauses in repeated references, or both. Future, more detailed analysis could clarify this. Although this is a reduction in numbers, this is consistent with a hyper-articulation interpretation, with speakers trying to make their references clearer, in the sense that a reduced speech rate arguably makes it easier for the matcher to determine which object was referred to.

We found a marginally significant increase (in *minF'*) for gesture rate and gesture repetition. Participants produced more gestures per 100 words and more repeated strokes in third descriptions as compared to first descriptions. The increase in gesture rate was caused by the fact that in repeated references, the number of words was reduced, but the number of gestures stayed constant. This means that the relative reliance on gesture increased in repeated references.

The other gesture variables did not show significant differences across references. Although this is not an increase as such, it can also be seen as an absence of reduction. We may interpret this absence of reduction as some sort of reset back to the level of the initial description.

The results from this study add to those from previous studies on repeated references, in which it was found that repeated references are reduced, in speech (e.g. Galati & Brennan, 2010) and in gesture (e.g. Hoetjes, et al., 2011; under revision). For example, in Hoetjes et al. (2011), we found that, after implicit positive feedback, gestures in repeated references were reduced with regard to their number, whereas in the present study, after receiving negative feedback, the number of gestures stayed constant (and their rate increased).

Previous research on feedback showed that both lack of feedback and providing positive feedback can cause gesture rate to be reduced (e.g. Holler & Wilkin, 2011; Jacobs & Garnham, 2007). The present study shows that negative feedback can cause gesture rate to increase. However, since research showed that attentiveness matters (Kuhlen & Brennan, 2010), it might be the case that, had the negative feedback been more elaborate and produced by the matcher herself instead of through a pre-recorded sound, the gesture rate would have increased (even) more. Future research could study whether this might be the case.

It can be argued that the fact that the descriptions were unsuccessful caused the director to doubt (especially after the second reference) whether there was any common ground between the director and the matcher. Previous studies (Gerwing & Bavelas, 2004; Holler & Stevens, 2007) found reduction in gesture in cases of common ground. It might be the case that in the present study, the descriptions were not reduced, but produced with more effort, in an attempt by the director to (re)create common ground with the matcher (example 1 suggests such a process).

The current results support our hypothesis that repeated references are not always reduced but can become more effortful following communication problems, and can also be related to the idea that speakers try to be as communicatively efficient as possible (Jaeger & Tily, 2011). In the context of our experiment, the director was given negative feedback indicating that more information was necessary. The ensuing descriptions given by the director provided the matcher with the necessary extra (visual) information (in the form of a lower speech rate and a higher gesture rate), thereby making the overall discourse as communicatively efficient as possible.

Acknowledgments

We would like to thank Elsa Jonkers and Kristel Bartels for their help in collecting, transcribing and annotating the data, Nick Wood for technical support, and Adriana Baltaretu, Martijn Goudbeek, Ruud Koolen, Fons Maes, Ingrid Masson-Carro, Jorrig Vogels, Hans Westerbeek and two anonymous reviewers for useful comments and suggestions. We received financial support from The Netherlands Organization for Scientific Research, via a Vici grant (NWO grant 27770007), which is gratefully acknowledged.

References

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech, 47*(1), 31-56.

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.

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

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*(3-4), 239-272.

Brennan, S., & Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology, 22*(6), 1482-1493.

Clark, H. (1973). The Language-as-Fixed-Effect fallacy: A critique of language statistics in psychological research. *Journal of verbal learning and verbal behavior, 12*, 335-359.

Clark, H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

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.

Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech, 31*(4), 307-319.

Galati, A., & Brennan, S. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language, 62*, 35–51.

Galati, A., & Brennan, S. (2014). Speakers adapt gestures to addressees' knowledge: implications for models of co-speech gesture. *Language, Cognition and Neuroscience, 29*(4), 435-451.

Gauthier, I., & Tarr, M. (1997). Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision Research, 37*, 1673-1682.

Gerwing, J., & Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture, 4*, 157-195.

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

Healey, P., Mills, G. J., & Eshgi, A. (2013). *Making things worse to make them better: The role of negative evidence in the coordination of referring expressions.* Paper presented at the Pre-Cogsci 2013, Berlin.

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. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3250-3255). Boston: Cognitive Science Society.

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., & Swerts, M. (under revision). Reduction in speech and gesture during the production of repeated references.

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 Social Psychology, 26*(1), 4-27.

Holler, J., & Wilkin, K. (2011). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics, 43*, 3522–3536.

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language, 56*, 291-303.

Jaeger, T. F., & Tily, H. (2011). Language Processing Complexity and Communicative Efficiency. *WIREs: Cognitive Science, 2*(3), 323-335.

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

Kuhlen, A., & Brennan, S. (2010). Anticipating Distracted Addressees: How Speakers' Expectations and Addressees' Feedback Influence Storytelling. *Discourse Processes, 47*(7), 567- 587.

Lindblom, B. (1990). Explaning variation: a sketch of the H and H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403-439). Dordrecht: Kluwer Academic Publishers.

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

Oviatt, S., Bernard, J., & Levow, G. A. (1998). Linguistic Adaptations During Spoken and Multimodal Error Resolution. *Language and Speech, 41*(3-4), 419-449.

Vajrabhaya, P., & Pederson, E. (2013). *Repetition vs. listener accommodation: A case study of co-speech gesture in retellings.* Paper presented at the NWAV42, Pittsburgh.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). *ELAN: a Professional Framework for Multimodality Research.* Paper presented at the LREC.

Zipf, G. K. (1936). *The psychobiology of language*: Routledge.