# Which gesture types make a difference?

de Beer, Carola; Carragher, Marcella; van Nispen, Karin; de Ruiter, Jan Peter; Hogrefe, Katharina ; Rose, Miranda

[Link to publication](Link to publication)

# Which gesture types make a difference?

# Interpretation of semantic content communicated by PWA via different gesture types

*Carola de Beer[1], Marcella Carragher[2], Karin van Nispen[3], Jan P. de Ruiter[1], Katharina Hogrefe[4] & Miranda L. Rose[2, 5]*

[1] Faculty of Linguistics and Literature Science, Bielefeld University, Bielefeld, Germany
[2] Department of Community and Clinical Allied Health, La Trobe University, Melbourne, Australia
[3] Tilburg Center for Cognition and Communication, Tilburg University, Tilburg, The Netherlands
[4] Clinical Neuropsychology Research Group (EKN), Clinic of Neuropsychology, Bogenhausen Hospital, Munich, Germany
[5] Centre for Clinical Research Excellence in Aphasia Rehabilitation (CCRE), Australia

carola.de_beer@uni-bielefeld.de, M.Carragher@latrobe.edu.au, K.vanNispen@uvt.nl, jan.deruiter@uni-bielefeld.de, katharina.hogrefe@ekn-muenchen.de, M.Rose@latrobe.edu.au

## Abstract

People with aphasia (PWA) spontaneously use various gesture types. Such gestures can potentially express semantic content that complements speech.

We investigated whether production of different gesture types adds crucial semantic content to the spoken output produced by PWA. In a perception experiment using multiple choice questions, naïve judges reported their information uptake from messages communicated by PWA in a speech-only vs. gesture+speech condition. The results show that the choice of response-options differed between conditions for all tested gesture types. We conclude that gestures in PWA disambiguate the interpretation of communicated messages and therefore markedly influence the expression of semantic content.

**Index Terms**: gesture, aphasia, spontaneous communication, semantic content

## 1. Introduction

The relationship between gesture and speech is assumed to vary between different gesture types. Kendon [1] distinguishes between gesticulation, pantomimes, emblems and sign language. These gesture types show different characteristics in terms of their relationship to speech, their degree of conventionalization and their linguistic properties. Gesticulations are not conventionalized, only appear with speech and have no linguistic properties. In contrast, emblems and pantomimes are conventionalized to a certain degree and hold some linguistic properties. Therefore, the latter two gesture types hold the potential to be understood without accompanying speech, whilst the interpretation of gesticulations is closely related to the accompanying speech.

The role of gestures in the expression of semantic content has been investigated in a number of studies. One line of enquiry relates to whether the content expressed via gesture is redundant to the accompanying speech or complementary. Some researchers argue that iconic gestures do not play an important role in the communication of relevant information [e.g. 2]. This assumption is based on the finding that participants' interpretation of semantic content was not improved with the accessibility of visual information compared to only audio information. In contrast, Bangerter [3] as well as Melinger and Levelt [4] report that spatial information is completely omitted from spoken output in the presence of deictic or iconic gestures in target-identification tasks. Furthermore in narratives, [5] parts of the informational content expressed via gesture was not inferable from the content of the spoken output.

The coordination and link between gesture and speech can be conceptualised by the planning and production processes underlying each. Non-parallel expression of content in gesture and speech can be accounted for by models of gesture production that assume a shared origin of gesture and speech and tightly coordinated but separate production processes of the two channels, for example the *Sketch Model* [6, 7]. Parts of a speaker's communicative intention can be conveyed via gesture and do not necessarily have to be specified in speech as well. This is especially evident in people with impaired spoken output, as is the case in PWA [8, 7]. However there is evidence against this compensatory or *trade-off relationship* of gesture and speech in non-impaired speakers [9]. Regarding people with aphasia some researchers were able to demonstrate a spontaneous and compensatory use of gestures that is especially true for those individuals presenting with severe aphasia [10, 11]. But this potential compensatory role of gesture for PWA has been debated, with evidence against an effective compensatory use of gestures [12]. Furthermore, there is evidence that both gesture and speech are vulnerable to simultaneous break down in PWA [13]. These findings clearly call into question the view that gesture plays a compensatory role in the case of aphasia.

Whilst acknowledging the lack of consensus regarding the role of gesture in communication, it is widely accepted that PWA make use of various gesture types in spontaneous communication [e.g. 14, 15]. Amongst many other gesture types, Sekine and colleagues [15] identified emblems, pantomimes and referential gestures as frequently used by PWA in spontaneous communication. Whilst we know that PWA with different aphasic types and severities make spontaneous use of a variety of gestures in communication, previous studies have not investigated the content expressed via gesture. Furthermore, it cannot be inferred from previously reported evidence what information listeners were able to comprehend when gesture, speech or both channels were accessible.

Hogrefe et al. [11] investigated the comprehensibility of cartoon-narratives produced by PWA based on the responses of naïve judges. The PWA recalled the cartoon narratives in two conditions: 1) they were asked to retell the cartoons they watched without any specific instructions (speech+gesture condition) and 2) they were explicitly asked to retell the cartoons only by the use of gestures (gesture only condition). Judges' information uptake from the first condition was compared between gesture and speech. The reactions to the audio stimuli were more accurate for 8 (out of 16) PWA. For 2 of the 16 PWA, judges' reactions to the gesture stimuli were more accurate. Judges' reactions to the gesture stimuli from the first condition (speech+gesture) were also compared to the gesture stimuli from the second condition (gesture only). The judges' responses were more accurate for 8 PWA in the second

condition. In summary, speech was more informative than speech+gestures in most PWA. However, for some PWA their speech-replacing gestures (gesture only) were more informative than their speech-accompanying gestures (speech+gesture).

In an additional analysis, Hogrefe et al. [16] evaluated the information content that six judges identified from the speech vs. gesture (speech+gesture condition) stimuli used by PWA. The judges were presented with choices from a list of predefined content-related propositions and asked to identify which propositions they were able to recognize from the stimuli. For 5 of the 16 PWA, more propositions were correctly detected from the gestures. Similarly, for 5 of the 16 PWA, more propositions were correctly detected from the speech by the judges. A subsequent analysis per proposition was carried out to investigate if there were a) any cases in which no information was understood from either of the communication channels, b) propositions were recognized from both modalities (redundant), c) propositions were solely recognized from gesture, and d) propositions were solely recognized from speech. The redundant score did not significantly differ from the gesture-only score for the whole group. For individuals presenting with severe aphasia, more propositions were shown to be conveyed solely by gesture. These results suggest that individuals with severe aphasia produce gestures to compensate for their reduced verbal output. However, whilst Hogrefe et al. [16] considered the effects of all gestures used in the narrative, they did not distinguish between different gesture types and their respective influence on the judges' perception.

Rose and colleagues [17] tested the comprehensibility of pantomimes produced by PWA. The data were extracted from spontaneous conversations and presented in a) audio+video b) audio only and c) video only. Seventy-four student participants answered open-ended questions (OQ) and multiple-choice questions (MCQ). The combined audio+video stimuli led to the most accurate responses to both the OQ and MCQ.

In a follow-up study by De Beer et al. [18], the impact of gestures on the communicative effectiveness in PWA was investigated. The accuracy of information uptake from messages communicated by PWA was studied for three different gesture types; referential gestures, emblems and pantomimes. Clips from conversation samples of PWA were presented in a gesture+speech condition or a speech-only condition. Participants answered OQ and MCQ and their responses were scored. Participants' responses were more accurate in the gesture+speech condition for all tested gesture types for both OQ and MCQ. The choice of the MCQ options was compared between conditions: analysis indicated that participants' responses differed significantly between the two conditions. In other words, the participants' perception of information content differed between the gesture+speech and speech-only conditions. However, the choice of response options was not tested for each of the specific gesture types. Hence it is not possible to infer from the data if all three different gesture types (pantomimes, emblems and referential gestures) express information that differs from verbal speech to a different extent.

The present study represents a follow-up analysis of participants' choice of response-options from the multiple-choice questionnaire for pantomimes (as defined by Kendon [1]), emblems (as defined by Kendon [1]) and referential gestures (reflecting what Kendon [1] named gesticulations and subsuming McNeill's [19] deictic and iconic gestures). We compared participants' responses between two different presentation conditions 1) gesture+speech (G+S) and 2) speech-only (S-O). The analysis aimed to further differentiate various gesture types and their respective effects on listeners' uptake of messages produced by PWA.

# 2.    Method

A subsequent analysis was conducted using data collected in a perception experiment. In the original study, we tested participants' reactions to 30 stimulus clips taken from spontaneous conversation samples of PWA [18].

## 2.1. Participants

10 participants with aphasia were chosen from the AphasiaBank Database (http://www.talkbank.org/ Aphasia Bank). They presented with primarily productive deficits and varying degrees of severity of aphasia (for details on the participants, see De Beer et al. [18]).

60 student participants were recruited as naïve judges for the study. The participants were blinded to the aims of the study.

## 2.2. Material
### a) Video and Audio Stimuli
The clips for the experiment were chosen from conversational samples of the AphasiaBank Database. These clips are recordings of PWA reporting their stroke story and also an important event of their lives. For each PWA, one clip per gesture type was chosen (i.e., pantomimes, emblems and referential gestures). An exception to this was Subject 2, who did not produce any pantomimes in the samples. To ensure an equal number of clips per gesture type, two clips with pantomime gestures were chosen from the conversation sample of Subject 4. This yielded a total of 30 clips containing the gestures of interest. For each of the 30 clips, an audio and a video version were created. The chosen clips were of varying lengths (2 to 10 seconds) due to differing complexities of the communicated messages. Gesture classification was conducted by the first author. The classification for the 30 gestures was checked by a second blinded rater who was familiar with the categorisation system used. Agreement between the two raters was reached for 83.3% of all cases. Cohen's kappa for inter-rater reliability was acceptable at .75.

### b) Multiple Choice Questions
MCQ were constructed to identify the information that the judges understood from the clips. The four multiple choice options included:

1) gesture+speech (G+S) message, i.e., the target message based on the information from the video and the audio versions of the clips;

2) G+S distractor which was semantically related to the G+S message;

3) speech-only (S-O) message, i.e., a message solely based on the information from the audio versions of the clips;

4) S-O distractor which was semantically and phonetically related to the S-O message.

The transcript of one of the stimulus clips (clip 20) is presented below. Table 1 displays the four constructed response options for clip 20.

The four response options were generated by two of the authors. For the construction of the S-O messages, one rater listened to the audio versions of the clips without knowing the video versions.

Example for one stimulus clip: Transcript of the target gesture and the accompanying speech for Clip 20.

S: and one le uh left

H: *left hand in front of the body, palm turned upwards (preparation)*

[/1.5/]

H: *pantomime: left hand and arm on chest height, hand is oriented downwards, circular movement above the table, imitates sprinkling something on top of a round object (target gesture)*

S: [and decorate] cakes an'

| S | spoken output |
| H | hand movements (in italics) |
| / | silent pause (duration in seconds reported in brackets) |
| [] | stroke of gesture |

**Table 1.** Overview of the messages and distractors for Clip 20

| 1) G+S message | I was decorating cakes left-handed |
|---|---|
| 2) G+S distractor | I was baking cakes |
| 3) S-O-message | When they left I was decorating cakes |
| 4) S-O-distractor | I was decorating the house and baking a cake after they left |

### 2.3. Procedure

Participants were randomly assigned to one of the two experimental groups. In group 1 (n=30) clips 1 - 15 represented the audio or S-O version and clips 16 - 30 represented the video or G+S version. For group 2 (n=30) the presentation modes were reversed. In the experimental sessions all participants started with the S-O condition to avoid any unwanted effects of order of condition. Each clip was presented twice before participants were asked to report what they understood from the clips by answering to one OQ per clip and the subsequent MCQ (for more information about the OQ see De Beer et al. [18]).
Participants recorded their responses in a response booklet in written form. For the MCQ, participants were asked to choose the option they felt best matched the message the PWA in the respective clip was trying to communicate. Gestures were not mentioned in the instructions or any of the written forms. The number of choices of each option was counted per clip and per condition.

### Analysis

Clip number 4 was removed from the analysis because of poor sound quality. The gesture type presented in clip 4 was an emblem. Thus for the category of emblems only 9 clips were included in the final data analysis.
Two-tailed Wilcoxon Sign Ranked Test for related samples was used for the statistical analysis.

## 3.      Results

### a) Referential Gestures

For the category of referential gestures, the G+S message was chosen significantly more often ($Z = -2.549$, $p = .011$) in the G+S condition (*mean* = 21.6, *SD* = 6.931) compared to the S-O condition (*mean* = 10.6, *SD* = 8.249). The G+S distractor was chosen more often in the G+S condition (*mean* = 3.6, *SD* = 4.671) compared to the S-O condition (*mean* = 2.8, *SD* = 4.686), but this difference did not reach statistical significance ($Z = -.06$, $p = .952$). The S-O message was chosen more often in the S-O condition (*mean* = 8.90 , *SD* = 5.744) than in the G+S condition (*mean* = 2.2, *SD* = 3.155). This difference was significant ($Z = -2.553$, $p = .011$). Also the S-O distractor was picked significantly more often ($Z = -2.492$, p = .013) in the S-O condition (*mean* = 7.7, *SD* = 7.273) compared to the G+S condition (*mean* = 2.6, *SD* = 2.989). See Figure 1.
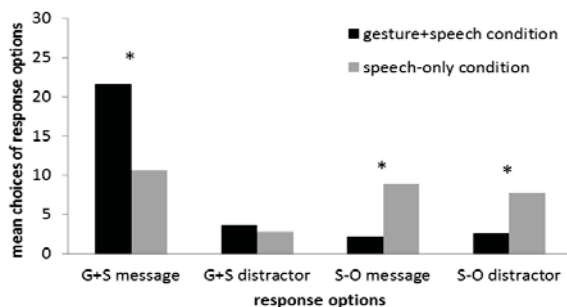


Figure 1: *Frequencies (means) of the four different choices of response options for referential gestures compared between the gesture + speech condition (black) and the speech-only condition (grey). Significant differences are indicated by asterisks.*

### b) Emblems

For the category of emblems, participants chose the G+S message more often in the G+S condition (*mean* = 16.56, *SD* = 10.43). This difference to the S-O condition (*mean* = 9.11, *SD* = 7.132) reached statistical significance ($Z = -2.556$, p = .011). The difference for the G+S distractor between the G+S condition (*mean* = 3.22, *SD* = 5.426) and the S-O condition (*mean* = 3.56, *SD* = 5.615) was not significant ($Z = -.632$, $p = .527$). Participants' choices of the S-O message differed significantly between conditions ($Z = -2,075$, $p = .038$) and it was more often chosen in the S-O condition (*mean* = 11.22, *SD* = 7.513) compared to the G+S condition (*mean* = 6.33, *SD* = 8.602). Participants chose the S-O distractor significantly more often ($Z = -2.2$, $p = .028$) in the S-O condition (*mean* = 6, *SD* = 7.826) compared to the G+S condition (*mean* = 4, *SD* = 7.632). See Figure 2.
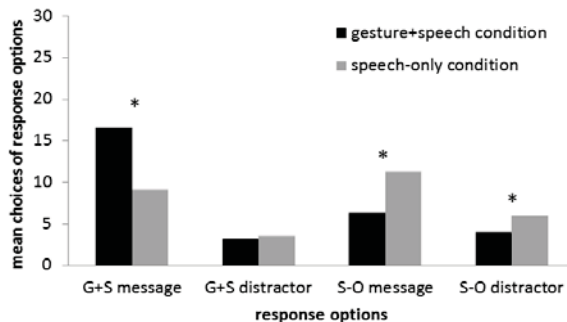


Figure 2: *Frequencies (means) of the four different choices of response options for emblems compared between the gesture + speech condition (black) and the speech-only condition (grey). Significant differences are indicated by asterisks.*

### c) Pantomimes

For the category of pantomime gestures, the G+S message was chosen more often in the G+S condition (*mean* = 20, *SD* = 8) compared to the S-O condition (*mean* = 11.7, *SD* = 9.638). This difference was statistically significant ($Z = -2.67$, $p = .008$). No significant difference ($Z = -.768$, $p = .443$) was found for the choice of the G+S distractor between the G+S condition (*mean* = 2.2, *SD* = 3.736) and the S-O condition (*mean* = 3.5, *SD* = 5.642). The S-O message was chosen more often in the S-O condition (*mean* = 10.6, *SD* = 8.884) compared to the G+S condition (*mean* = 6.5, *SD* = 5.421). This difference did not reach statistical significance ($Z = -1.899$, $p = .058$). Participants' choices of the S-O distractor differed significantly between conditions ($Z = -2.536$, $p = .011$). It was chosen more often in the S-O condition (*mean* = 4.3, *SD* = 3.622) compared to the G+S condition (*mean* = 1.2, *SD* = 1.135). See Figure 3.
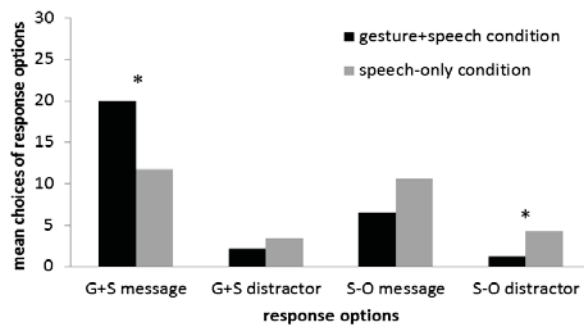


Figure 3: *Frequencies (means) of the four different choices of response options for pantomimes compared between the gesture + speech condition (black) and the speech-only condition (grey). Significant differences are indicated by asterisks.*

# 4. Discussion

In summary, the participants' choices of response options in the MCQ differed between conditions for all three gesture types. The G+S message and the S-O message were chosen more often in their respective conditions. These effects were significant, apart from the number of choices of the S-O message for pantomime gestures. For the G+S distractors no remarkable effects of condition were found for either of the three gesture types. The S-O distractor was chosen significantly more often in the S-O condition for all three gesture types.

The number of choices of the response options indicates overall that the participants did pay attention to the type of gesture that the PWA produced in the clips and that the information expressed via all gestures was used for the interpretation of the messages. This supports earlier findings by De Beer et al. [18].

In the G+S condition, participants demonstrated a clear preference for the G+S message (the target message); this was true for all three gesture types. However, in the S-O condition, participants did not choose the S-O message with a similar frequency. Participants' choices of the response options were less stable in the S-O condition; here, the target message was chosen with a similar frequency as the S-O message for all three gesture types. A remarkable number of participants in the S-O condition still chose the target message which is not surprising, because for many clips most of the semantic content was expressed in speech. The presentation of the MCQ options might have influenced participants' interpretation of the messages. Particularly in the S-O condition, when participants did not have access to the complete informational content (i.e., information conveyed via gesture), the presentation of the target message might have led to reinterpretation of the audio-stimuli. Combining these assumptions together with the effects of condition, it can be inferred that the accessibility of the information from the gesture channel decreased the ambiguity of the communicated messages in the stimuli. Therefore in the G+S condition when participants had access to the information from both modalities, they were able to identify the target message with higher accuracy.

Strikingly the G+S distractors were rarely chosen in both conditions across gesture types. There were no clear effects of condition found for this distractor. This finding may be due to the construction of the distractors, because the G+S distractor was only semantically related to the G+S message and not always phonetically related to the information presented in verbal speech. Hence the G+S distractors may not have been sufficiently closely related to the target messages.

The effects of condition were shown for all three gesture types. This indicates that all tested gesture types did influence the participants' information uptake. By their nature, pantomimes and emblems hold the potential to convey content that complements or even replaces spoken output. Referential gestures are assumed to be more tightly related to spoken output and only completely interpretable in the context of the accompanying speech. Surprisingly, within this study, the effect of gesture on participants' interpretation of semantic content was not limited to pantomimes and emblems; participants showed similar effects for all three gesture types on information uptake, though one would expect stronger effects of gestures that can replace speech in the case of impaired production of speech. For at least some PWA, gesture might necessarily be used to replace speech in the event of severely compromised spoken output. It is crucial to mention that some content is still expressed in speech by PWA in most cases. One-word utterances as well as sentences interrupted by unsuccessful word retrieval still serve as a source for semantic content for listeners. Gestures produced in spontaneous conversation can be interpreted in the context of even very reduced speech production. Within this study, all tested gesture types played a significant role in the expression of semantic content. This semantic content can complement spoken output, but it is still interpreted in the context of spoken production. The findings of the current study support our earlier conclusions [18] and serve to further our understanding of the impact of different gesture types on the expression of semantic content in PWA. Therefore, we were able to contribute to the evidence suggesting a compensatory use of gestures in PWA, i.e. argue against the assumption that gesture and speech break down in parallel in PWA.

We acknowledge that the choice of stimulus clips might have influenced the results of the study. This would be true if only sequences were chosen in which gestures were used in a speech-replacing way. However we included clips of sequences in which gestures were complementary but also redundant to the spoken output. Thus the stimuli were chosen to reflect varying degrees of complement or redundancy. Future studies might wish to consider constructing the target messages and distractors on the basis of independent judges' interpretation of the audio and video stimuli to improve validity. We also acknowledge the use of short messages in a perception study has been criticised by Beattie & Shovelton [5], who argued that the information expressed via gestures is often inferable from the wider context of a narrative. In the present study we used parts out of spontaneous conversation samples. Whilst it is plausible that contextual information influenced judges' perception of messages, we took care not to choose any clips that could only be interpreted with context knowledge of the whole conversation. Finally, the work of Hogrefe et al. [16], who investigated the information uptake from narrations produced by PWA, also suggests that in some individuals with aphasia gestures are more informative than speech.

# 5. Conclusion

All three gesture types under investigation (pantomimes, emblems and referential gestures) influence the interpretation of the messages communicated by PWA. Gestures produced by PWA are used by listeners to disambiguate messages from spoken output. Gestures do not necessarily have to be used in a speech-replacing way by PWA to play a role in the expression of semantic content. Therefore, communication in PWA has to be viewed as a multi-modal process. Gesture types which differ in the degrees of conventionalisation and relation to speech have been demonstrated to hold the potential of expressing semantic content. This was true even for gestures that are closely related to spoken output (referential gestures). Our results clearly suggest a compensatory use of different gesture types and broaden the knowledge about their role for communication for PWA.

# 6. Acknowledgements

# 7. References

[1]  Kendon, A., "Gesture: Visible Action as Utterance", Cambridge: University Press, 2004.
[2]  Krauss, R., Dushay, R.A., Chen, Y., & Rauscher, F., "The communicative value of conversational hand gestures", J Exp Soc Psychol, 31(6):533–552, 1995.
[3]  Bangerter, A., "Using pointing and describing to achieve joint focus of attention in dialogue", Psychol Sci, 15(6):415–419, 2004.
[4]  Melinger, A., & Levelt, W. J. M. (2004), "Gesture and the communicative intention of the speaker", Gesture, 4(2):119–141, 2004.
[5]  Beattie, G., & Shovelton, H., "An exploration of the other side of semantic communication: How the spontaneous movements

of the human hand add crucial meaning to narrative", Semiotica, 184(1-4):33–51, 2011.

[6]  De Ruiter, J. P., "The production of gesture and speech", In D. McNeill [Ed], Language and Gesture, 284–311, Cambridge: Cambridge University Press, 2000.

[7]  De Ruiter, J. P., & De Beer, C., "A critical evaluation of models of gesture and speech production for understanding gesture in aphasia", Aphasiology, 27(9):1015–1030, 2013.

[8]  De Ruiter, J. P., "Can gesticulation help aphasic people speak, or rather, communicate?", Int J Speech Lang Pathol, 8(2):124–127, 2006.

[9]  De Ruiter, J. P., Bangerter, A., & Dings, P., "The interplay between gesture and speech in the production of referring expression: Investigating the Tradeoff Hypothesis", Top Cogn Sci, 4(2):232–248, 2012.

[10] Goodwin, C., "Gesture, aphasia and interaction", in D. McNeill [Ed], Language and Gesture, 84–98, Cambridge: Cambridge University Press, 2000.

[11] Hogrefe, K., Ziegler, W., Wiesmayer, S., Weidinger, N., & Goldenberg, G., "The actual and potential use of gestures for communication in aphasia", Aphasiology, 27(9):1070–1089, 2013.

[12] Cicone, M., Wapner, W., Foldi, N., Zurif, E., & Gardner, H., "The relationship between language and gesture in aphasic communication." Brain Lang, 8(3):324–349, 1979.

[13] Duffy, R. J., & Duffy, J. R., "Three studies of deficits in pantomimic expression and pantomimic recognition in aphasia." J Speech Lang Hear Res, 24(1):70–84, 1981.

[14] Carlomagno, S., & Cristilli, C., "Semantic attributes of iconic gestures in fluent and non-fluent aphasic adults", Brain Lang, 99(1-2):102–103, 2006.

[15] Sekine, K., Rose, M. L., Foster, A. M., Attard, M. C., & Lanyon, L. E., "Gesture production patterns in aphasic discourse: In-depth description and preliminary predictions", Aphasiology, 27(9):1031–1049, 2013.

[16] Hogrefe, K., Ziegler, W., Weidinger, N., & Goldenberg, G., "Gestural expression in narrations of aphasic speakers: redundant or complementary to the spoken expression?", Proceedings of the Tilburg Gesture Research Meeting (TIGER), Netherlands, 2013.

[17] Rose, M.L., Mok, Z., Katthagen, S., & Sekine, K., "The communicative effectiveness of pantomime gesture in people with aphasia", Aphasiology, in prep.

[18] De Beer, C., Carragher, M., Van Nispen, K., De Ruiter, J.P, Hogrefe, K., & Rose, M. L., "How much information do people with aphasia convey via gesture?", Am J of Speech Lang Pathol, under revision.

[19] McNeill, D., "Hand and Mind", Chicago: University Press, 1992.