

## Effects of the restriction of hand gestures on disfluency

Sheena Finlayson, Victoria Forrest, Robin Lickley & Janet Mackenzie Beck

Queen Margaret University College, Edinburgh, Scotland.

### Abstract

This paper describes an experimental pilot study of disfluency and gesture rates in spontaneous speech where speakers perform a communication task in three conditions: hands free, one arm immobilized, both arms immobilized.

Previous work suggests that the restriction of the ability to gesture can have an impact on the fluency of speech. In particular, it has been found that the inability to produce iconic gestures, which depict actions and objects, results in a higher rate of disfluency. Models of speech production account for this by suggesting that gesture and speech production are part of the same integrated system. Such models differ in their interpretation of the location of the gesture planning mechanism in relation to the speech model: some authors suggest that iconic gestures relate closely to lexical access, while others suggest that the link is located around the conceptualization stage.

The findings of this study tentatively confirm that there is a relationship between gesture and fluency – overall, disfluency increases as gesture is restricted. But it remains unclear whether the disfluency is more related to lexical access than to conceptualization. Proposals for a larger study are suggested.

The work is of interest to psycholinguists focusing on the integration of gesture into models of speech production and to Speech and Language Therapists who need to know about the impact that an impaired ability to produce gestures may have on communication.

### 1. Introduction

A growing body of research suggests that many hand and arm gestures stem from the same basic process as the generation of spoken language, resulting in one interactive and co-expressive system. Gestures are assumed to enhance and elaborate on the content of accompanying speech but also form a part of the speech planning process. In some cases, like the description of spatial relationships between objects, gestures may be crucial to conveying the complete message. If this is so, what effect does the restriction of the ability to use gesture have? In this paper we describe preliminary research that compares some of the characteristics of speech produced with and without restrictions on arm movements: in particular, we investigate the relationship between restricted gestures and the production of disfluencies.

While many studies demonstrate that gesture may have a communicative function, conveying various forms of information to a listener [6], it is clear that gestures also serve some function in the speaker's encoding of speech. Some authors contend that gesture has a role in facilitating lexical access [2, 11, 17], while others, following McNeill [15], take the view that gesture is involved at the level of conceptual planning of speech [1, 4]

These differing viewpoints can be described with reference to Levelt's model of speech production [13], incorporating the basic components Conceptualiser, Formulator and Articulator and extending the basic model with some version of a gesture planning module. While Butterworth & Hadar's [2]

explanation of apparent lexical facilitation by gesture would locate the source of iconic gestures within the lexicon itself, more recent accounts suggest that they are generated around or within the conceptualiser. In the model proposed by Krauss, Chen & Gottesman [12], iconic gestures (lexical, in their terminology) derive from non-propositional representations in working memory, just prior to the conceptualiser component of speech production. In their view, the gestures thus produced are able to facilitate lexical access by feeding into the phonological encoder within the formulator. De Ruiter's [4] Sketch model and the Information Packaging Hypothesis of Kita and colleagues [1, 9] locates the source of gestures within the conceptualiser itself. In the Sketch model, the gesture planning module branches out of the conceptualiser, taking input from a sketch generation subcomponent, which uses spatio-temporal information, within the conceptualiser and feeding back a signal to the message generator as well as producing a motor program for the gesture. Unlike Krauss *et al.*'s model, there is no external feed into the lexical selection process: any such interaction must thus take place via the conceptualiser. Outside the conceptualiser, speech and gesture are produced independently and in parallel. While Krauss *et al.* argue that gestures can help to activate lexical items via some kind of cross-modal priming, de Ruiter's model allows some spatial features to be activated and reactivated by gestures via a feedback loop from the gesture planner to the conceptualiser.

All authors agree that more hard data on gesture planning is needed before such models can be much more than speculative.

All of these models suggest that gesture may have a facilitatory role in the production of speech. By implication, it is suggested that the removal of the ability to gesture should therefore result in less efficient speech production. In particular, a lack of gesture could lead to lexical access difficulties or more general planning difficulties, particularly with spatial content phrases, where iconic gestures are very prevalent [11]. Such planning and lexical access difficulties typically induce disfluencies, especially hesitations – silent and filled pauses and stalling repetitions. Studies with restricted gestures have indeed shown that under such conditions, the time spent pausing [5] and the rate of disfluency [17] increase.

Other studies which examine the relationship between gesture and disfluency demonstrate that the timing of gesture and speech overlaps considerably – gesture does not have the function of filling a pause while a speaker plans, self-corrects or searches for a word. Seyfeddinipur & Kita [19] found that for disfluent stretches of speech, gestures are suspended just before speech stops and resume just before speech restarts. Similarly, in the speech of people who stutter, Mayberry & Jaques [14] found that iconic gestures did not occur during episodes of blocking or repetition, but only coincided with stretches of fluent speech. If, as suggested by the studies reported above, gesture has a role in the planning of speech or in accessing lexical items, its timing seems to be very closely linked to the relevant speech events.

In the present study, we aim to add to our understanding of the relationship between gesture and fluency, by partially replicating previous work, while extending the scope of the research to include partial (one-handed) gesture restriction.

Studies of the effects of partial immobilization on speech production are hard to find. Rimé, Schiarature, Hupet & Ghysseleux [18] experimented with the immobilisation of various body parts (head, legs and arms) during spontaneous conversation. They found increased levels of movement in the body zones that were left free, but found no effect on the speech rate nor on fluency. Their study did not focus specifically on gestures nor did it examine closely the relationship between partial gesture restriction and fluency. Several studies suggest that gestures are most reliant on the speaker's dominant hand [7, 8, 15, 20]. Given the findings reported above on complete restriction of hand movement, this poses the question of whether the restriction of the dominant hand only will produce similar effects on fluency, or whether the ability to use the non-dominant hand will compensate.

In this study, we compare the performance of subjects in 3 conditions: hands-free, both hands restricted, dominant hand restricted. The task was based on one used in previous studies (e.g. [16]): the narration of the story of a children's animated cartoon. The story was useful in eliciting gestures, since subjects were required to describe a lot of movement and the cartoon contained barely any dialogue.

Given previous findings [5, 17], it was hypothesised that, relative to the hands-free condition, there would be a higher rate of disfluency in the condition where both hands were constrained.

In the condition where the dominant hand is constrained, current models do not suggest clear hypotheses. If the dominant hand is the more important in performing iconic gestures, and the other hand does not easily compensate, we would expect similar problems with formulating sentences or with lexical access, resulting in increased disfluency compared to the hands-free condition. However, if the non-dominant hand is able to compensate, then no effects on disfluency should be found (of course, this is indistinguishable from the null hypothesis, that there is no relationship between the ability to gesture and the ability to speak fluently).

The design of the task also allows us to look at relationships between disfluencies and another frequent gesture type, beat gestures, which, according to previous work, are less reliant on use of the dominant hand.

## 2. Method

### 2.1. Subjects

Six subjects took part as speakers in the study. All were female, aged between 17–25 and from various social and educational backgrounds. None reported a history of mobility, auditory or communication problems and none had more than a minimal knowledge of sign language. Other participants took part as passive listeners: these people were within the same age range as the subjects. No participants were paid for taking part and all were free to withdraw from the experiment at any stage.

### 2.2. Materials

The experiment took place in a sound-proofed room, measuring about 4m × 4.5m. Two straight-backed armchairs were placed in the room, facing each other, about 2m apart. The subjects' armchair was fitted with strips of Velcro on the arms, to allow arm movements to be restricted when required. A digital video camera was placed behind the listeners' chair,

facing the subject, so that the subjects' head, trunk and arms were in focus.

### 2.3. Recording Procedure

Subjects were informed in advance that the study aimed to examine communicative behaviour in story telling. No reference to gesture was made in the instructions.

Each subject was required to watch a cartoon on video, while sitting alone in a quiet room. When the cartoon was finished, the subjects were asked to retell the story of the cartoon to a listener in the same room. Listeners took no part in any dialogue, but offered appropriate backchannels. This procedure was performed three times for each subject, each time with the same cartoon, but with a different listener. In the first session, subjects retold the story with no restrictions on arm movement. In the second session, three of the subjects had their dominant arm fastened to the arm of the chair with a Velcro strip, while the other three had both arms immobilised. In the third session, the arm-binding conditions were reversed, so that all subjects took part in all three conditions. There was a break of five minutes between each session.

### 2.4. Analysis

Orthographic transcriptions were made from the video recordings, and subsequently checked using digital sound files and speech waveforms on a PC, when disfluencies were annotated on the transcriptions.

Gestures were analysed using the video recordings viewed frame by frame. They were classified using McNeill's definitions: *iconics* (depicting actions and objects), *metaphorics* (relating to abstract aspects of the topic of speech), *deictics* (pointing to an area of the speaker's gestural space) and *beats* (movements reflecting rhythmic aspects of the speech) [15].

Disfluencies were identified by careful auditory examination of the digitised audio recordings by two of the authors. These included pauses (silent and filled), repetitions and reformulations. They were subclassified as non-juncture disfluencies, where they occurred mid-clause, and juncture disfluencies, where they occurred between clauses. Most of the latter were filled pauses.

Spatial content phrases, defined as phrases containing spatial prepositions, were identified by inspection of the transcriptions.

## 3. Results

Table 1 shows raw results by condition for word counts and rate per 100 words of spatial content phrases (SCP), iconic and beat hand gestures, and non-juncture, juncture and total disfluencies. Other gesture types are disregarded in the rest of this study, as their number was too low. Because of the small number of subjects and the large amount of variability between subjects for most of these factors, we restrict the analyses to descriptive and non-parametric statistics. The most important examples of inter-subject variability for this study are in rate of iconic gestures (e.g., range in hands-free condition: 3.5–7.9 per 100 words) and in disfluency rates (e.g. range for all disfluencies in hand-free condition: 2.0–12.4 per 100 words).

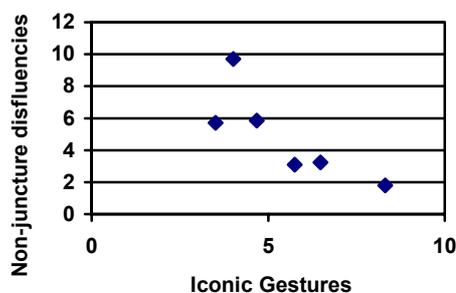
Word counts were greater in the two restricted-gesture conditions than in the hands-free condition, but this may be because speakers were able to recall more detail on their later attempts at the task rather than because of gesture restriction.

**Table 1:** Mean number of words and rate per 100 words of spatial content phrases, iconic and beat gestures and non-juncture and juncture disfluencies, by condition.

	Hands Free	One hand	No hands	Overall	N
Mean number of words	269.3	345	316.67	310.3	5586
Spatial content phrases	24.5	30	30.3	28.3	509
Iconic Gestures	5.82	2.32	N/A	3.85	142
Beat Gestures	2.41	4.73	N/A		
Non-juncture disfluencies	4.64	4.35	5.16	4.71	263
Juncture disfluencies	1.61	2.31	2.05	2.02	113
Total disfluencies	6.25	6.67	7.21	6.73	376

The rate per 100 words of SCPs also increased from the hands-free condition to the restricted gesture conditions. There was no difference between the one-hand and the both-hands restricted conditions for the rate of SCPs. Iconic hand gestures reduced dramatically between the hands-free and one-hand condition – with their dominant hand immobilised, subjects did not compensate by using their non-dominant hand to produce iconics.

In the hands-free condition, which had the highest rate of iconic gestures, we found evidence of a relationship between use of iconic gesture and fluency – speakers who used more iconic gestures also produced fewer non-juncture disfluencies ( $\rho = -.807$ ,  $N=6$ ,  $p < .05$ ) (Figure 1). In the one-handed condition, where iconic gestures were much rarer, no such relationship was found. Conversely, for beat gestures, a numerically higher rate was found in the one handed condition, though this failed to reach significance.

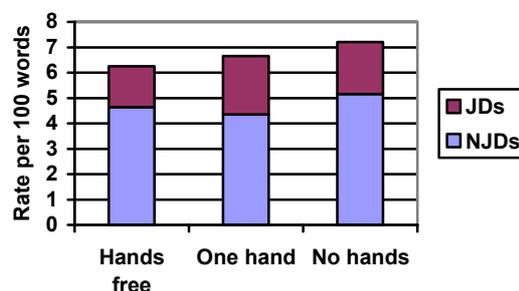


**Figure 1:** Rates per hundred words of Iconic Gestures and Non-Juncture Disfluencies for all six speakers in the hands-free condition.

Interestingly, perhaps, the difference between conditions for beat gestures coincided with a slight change in disfluency rates between the two-handed and one-handed conditions: a lower rate of non-juncture disfluency was found in the one-handed condition than in the two-handed condition, though, again, the difference was not significant.

Models discussed in the introduction, as well as previous findings, would predict that disfluency rates would be higher when gesture was restricted. Numerically, *overall* disfluency rates *were* higher in both restricted gesture conditions (One hand: 6.67. No hands: 7.21/100 words) than in the hands-free condition (6.25). This was not a statistically significant finding. For non-juncture disfluencies, however, the lowest rate of disfluency was in the *one-hand* condition and this was

significantly lower than the rate for the no hands condition (Wilcoxon signed ranks test:  $W=0$ ,  $N=6$ ,  $p < .05$ , two-tailed). Finally, for juncture disfluencies (mostly clause-initial filled pauses), the numerical increase from the hands-free condition to the restricted condition was also not significant (Figure 2).



**Figure 2:** Disfluency rates per 100 words by condition. (NJD = Non-juncture disfluency; JD = Juncture disfluency).

While spatial content phrases were more frequent in the two gesture-restricted conditions than with hands free, no relationship between SCP rates and disfluency rates could be found.

#### 4. Discussion

This study can be seen as a pilot for a larger and more carefully designed study to be undertaken shortly. Various methodological issues and large variability in disfluency and gesture rates in this small group of subjects mean that any conclusions that we draw have to be seen as tentative. Accordingly, we present the following outcomes as possible indicators for further research.

- In the hands-free condition, speakers who used more iconic gestures, also produced significantly fewer non-juncture disfluencies.
- With the ability to gesture restricted, the overall disfluency rate is higher.
- In the one-handed condition, with the dominant hand restricted, subjects used fewer iconic gestures, not compensating by using the non-dominant hand. However, no effect on disfluency rates was seen as a result of the reduction in iconic gestures.
- Beat gestures occurred at a higher rate in the one-handed condition than in the hands-free condition, coinciding with a lower rate of non-juncture disfluencies.

The first two points support previous findings [17] which demonstrated an increase in disfluency with restricted gesture. This is compatible with models of speech production which incorporate gesture as part of the same system. From this study, it is hard to argue in favour of a model which links gesture planning to lexical access rather than the conceptual level. The data in Table 1 suggest that the increase in disfluency rate is not restricted to non-juncture disfluencies, as a lexical access account might predict: there is an increase in disfluency rates for clause-onset disfluencies (mostly filled pauses), as well as for non-juncture disfluencies. This may indicate that more general utterance planning is affected (i.e., at the conceptual level) rather than, or at least as much as, lexical access, thus supporting a model which relates gesture-planning more closely to the conceptualiser than to lexical access.

The third point, that a drop in rate of iconic gestures in the one-handed condition with respect to the hands-free condition

did not coincide with an increase in disfluency, is problematic for models that assume a link between use of iconic gesture and fluency. We speculate that the increase in the use of beat gestures may have a confounding role here. Evidence from experiments which impose external timing on spontaneous speech by use of a metronome suggests that such rhythmical support dramatically reduces the rate of filled pauses without affecting the overall speech rate [3]. The same method is commonly reported to reduce disfluency rates in people who stutter (e.g., [10]). If the use of beat gestures has a similar effect on disfluency as an artificial timing device, then we would expect to find more fluent speech as the rate of beat gesture increases.

A more mundane explanation for an increase in disfluency in the restricted-gesture conditions might be that the restrictions simply constituted a minor distraction from the speaking task, which had an impact on the speakers' ability to focus on speech planning alone. The study by Rimé *et al.* [18], reported above, provides some evidence against this explanation, but in future work, control for this possible confound should be implemented.

There are clear methodological problems in this relatively small pilot project. Firstly, given the high rate of variation in disfluency and gesture rates by subject, a larger sample is needed. Secondly, the method of restricting arm movements allowed a certain amount of gesture leakage – despite the subjects' arms being restricted, it was observed that some subjects still attempted to perform gestures with their hands and fingers: the amount of such gesturing was too small to be reported here, but future studies should ensure that this is not possible. Thirdly, all speakers performed the same communication task three times, introducing the possibility of rehearsal effects. We might hypothesise that this would *decrease* the disfluency rates in the gesture-restricted conditions, since the planning load on retelling the story would be reduced. In fact, we still found an *increase* in disfluency rates overall. If a rehearsal effect is there, then this suggests that the design fault in our experiment may have reduced the observable effect of gesture restriction on disfluency.

In summary, we have some evidence to support the view that gesture is helpful to fluent speech production, but we can not yet explicitly support either a lexicon-linked or a conceptualiser-linked model of integrated speech and gesture production. While interesting for psycholinguists working in speech production, the work is also relevant to Speech and Language Therapists dealing both with clients with fluency disorders and mobility problems.

## 5. Acknowledgements

This work is based on final-year undergraduate dissertations by Sheena Finlayson and Victoria Forrest, supervised by Janet Beck, on the BSc course in Speech Pathology and Therapy at Queen Margaret University College, Edinburgh. Their participation at DiSS '03 was partially funded by profits from previous workshops in the series at Berkeley and Edinburgh.

## 6. References

[1] Alibali, M. W., S. Kita & A. J. Young. 2000. Gesture and the Process of Speech Production: We Think, therefore we Gesture. *Language and Cognitive Processes*, vol. 15(6), pp. 593–613.

[2] Butterworth, B. & U. Hadar. 1989. Gesture, Speech and Computational Stages: A reply to McNeill. *Psychological Review*, vol. 96, pp. 168–174.

[3] Christenfeld, N. 1996. Effects of a Metronome on the Filled Pauses of Fluent Speakers. *Journal of Speech and Hearing Research*, vol. 39, pp. 1232–1238.

[4] De Ruiter, J. P. 2000. The Production of Gesture and Speech. In: D. McNeill (ed), *Language and gesture* (284–311). New York: Cambridge University Press.

[5] Graham, J. A. & S. Heywood. 1976. The Effects of Elimination of Hand Gesture and of Verbal Codability on Speech Performance. *European Journal of Social Psychology*, vol. 5, pp. 189–195.

[6] Kendon, A., 1994. Do Gestures Communicate? A Review. *Research on Language and Social Interaction*, vol. 27, pp. 175–200.

[7] Kimura, D. 1973a. Manual Activity during Speaking – I. Right-handers. *Neuropsychologia*, vol. 11, pp. 34–50.

[8] Kimura, D. 1973b. Manual Activity during Speaking – II. Left-handers. *Neuropsychologia*, vol. 11, pp. 51–55.

[9] Kita, S. 2000. How Representational Gestures Help Speaking. In: D. McNeill (ed.), *Language and gesture*. New York: Cambridge University Press, pp. 162–185.

[10] Klingbeil, G. M. 1939. The Historical Background of the Modern Speech Clinic. *Journal of Speech Disfluencies*, vol. 4, pp. 115–132.

[11] Krauss, R. M. 1998. Why do we Gesture when we Speak? *Current Directions in Psychological Science*, vol. 7, pp. 54–59

[12] Krauss, R. M., Y. Chen & R. F. Gottesman. 2000. Lexical Gestures and Lexical Access: a Process Model. In McNeill, D. (ed), *Language and gesture*. New York: Cambridge University Press, pp. 261–283.

[13] Levelt, Willem J. M. 1989. *Speaking. From Intention to Articulation*. Cambridge, Massachusetts: MIT Press.

[14] Mayberry, R. I. & J. Jaques. 2000. Gesture Production during Stuttered Speech: Insights into the Nature of Gesture–Speech Integration. In: D. McNeill (ed), *Language and gesture*. New York: Cambridge University Press, pp. 199–214.

[15] McNeill, D. 1992 *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.

[16] McNeill, D. & E. T. Levy. 1982. Conceptual Representations in Language Activity and Gesture. In: R. Jarvella & W. Klein (eds.) *Speech Place and Action*. Chichester, UK: Wiley and Sons, pp. 271–295.

[17] Rauscher, F. B., R. M. Krauss & Y. Chen. 1996. Gesture, Speech and Lexical Access: The Role of Lexical Movements in Speech Production. *Psychological Science*, vol. 7, pp. 226–230.

[18] Rimé, B., L. Schiarature, M. Hupet & A. Ghyssele. 1984. Effects of Relative Immobilization on the Speaker's Nonverbal Behavior and on Dialog Imagery Level. *Motivation and Emotion*, vol. 8(4), pp. 311–325.

[19] Seyfeddinipur, M. & S. Kita. 2001. Gesture as an Indicator of Early Error Detection in Self-Monitoring of Speech. *Proceedings of DiSS '01, ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech*, Edinburgh, Scotland, pp. 29–32.

[20] Sousa-Poza, J. F., R. Rohrberb & A. Mercure. 1979. Effect of Type of Information (abstract–concrete) and Field dependence on Asymmetry of Hand Movements during Speech. *Perceptual and Motor Skills*, vol. 48, pp. 1323–1330.