Gesture production and speech fluency in competent speakers and language learners

Maria Graziano (maria.graziano@humlab.lu.se)
Humanities Lab, Lund University

Marianne Gullberg (marianne.gullberg@ling.lu.se)
Centre for Languages and Literature & Humanities Lab, Lund University

Abstract

It is often assumed that a main function of gestures is to compensate for expressive difficulties. This predicts that gestures should mainly occur with disfluent speech. However, surprisingly little is known about the relationship between gestures and fluent vs. disfluent speech. This study investigates the putative compensatory role of gesture by examining competent speakers' and language learners' gestural production in fluent vs. non-fluent speech. Results show that both competent and less competent speakers predominantly produce gestures during fluent stretches of speech; ongoing gestures during disfluencies are suspended. In all groups, the few gestures that are completed during disfluencies are both referential and pragmatic. The findings strongly suggest that when speech stops, so do gestures, thus supporting the view of speech and gesture as an integrated system.

Keywords: Gesture; speech production; language development; second language acquisition; crossmodal coordination.

Introduction

Despite their different formats, gesture and speech are temporally and semantically interrelated in discourse production such that both modalities express related information at the same time. The semantic coordination arguably motivates the frequent assumption that speakers in general and language learners in particular (both children and adults) use gesture to overcome expressive difficulties. This notion is implicit in studies of child language acquisition (e.g., Capirci et al., 1996; Goldin-Meadow, 2003; Pine et al., 2007), in which gestures are seen as precursors to language, as markers of transitional knowledge states, etc. The compensatory notion is explicit in many studies of adult second language learners and bilinguals and is based on the observation that adult learners typically produce more gestures than native speakers (e.g., Nicoladis et al., 2007). Furthermore, some theoretical proposals concerning the gesture-speech relationship assume that representational gestures in particular, i.e., gestures conveying meaning related to the content of speech (Kendon, 2004; here also called referential, following Kendon), have a facilitating function, either aiding lexical retrieval (Krauss et al., 2000) or helping conceptualisation, or information packaging (Alibali et al., 2000; Kita, 2000).

Importantly, these assumptions predict that gestures should be more frequent during disfluent stretches of speech. However, there is some evidence that this may not be the case, but that gestures rather stop when speech stops in stuttering, in general disfluency, and in adult second language speakers (Gullberg, 1998; Mayberry et al., 1998; Ragsdale & Silva, 1982; Seyfeddinipur & Kita, 2001). An additional observation is that if adult learners gesture during disfluencies, they often produce pragmatic rather than representational gestures, that is, gestures that comment on the breakdown itself (pragmatic gestures) rather than on the content of the sought words (representational gestures; Gullberg, 1998).

Previous studies are arguably built on small samples. Moreover, very little is known about children's corresponding behaviour. This study therefore aims to test the putative compensatory role of gestures and their functional distribution relative to disfluencies by comparing the gestural behaviour of competent and learning language users, both children and adults. We ask (1) whether speakers predominantly produce gestures with fluent or with disfluent speech; (2) what articulatory features gestures have during disfluent speech; and (3) what functions gestures completed during disfluencies might have. For all questions we further explore whether there are differences across (child and adult) learners and adult competent speakers

Methodology

Corpora

The analyses draw on three multimodal sets of narrative discourse in the form of cartoon retellings in dyadic, interactive settings. The first set consists of narrative production by 33 Italian children divided into three age groups (4-5; 6-7; 8-10 years), each containing 11 subjects. The recordings took place in Naples and Rome in an environment familiar to the children, either home or at school. Children were asked to retell the story of an episode of the series Pingu to an adult they already knew (a friend of the family or their teacher). During the child's retelling, the adult, who had also seen the cartoon, did not interrupt the child but provided feedback. The second set contains

narratives of the same cartoon by 11 Italian adults. Participants were recruited in Naples, at the Università degli Studi di Napoli "L'Orientale". Two participants were involved in each session: one person was asked to watch the cartoon and to retell it to a friend who had not seen it. The listener was instructed to only listen to the story and to avoid interrupting the narrator, or to ask questions at the end of the story. This procedure was adopted in order to make the adult narratives comparable with those produced by the children.

The third data set includes retellings of a printed wordless cartoon produced by 11 Dutch adult learners of French as a second language at low to intermediate levels of proficiency. They had studied French for a minimum of 4 years and had never lived in a French-speaking country. The story was told to a confederate native speaker who was presented as unfamiliar with the cartoon. The native speaker (the listener) was instructed to participate in the interaction by asking clarification questions and providing feedback.

Coding and analysis

All stories were transcribed verbatim by native speakers. All spoken disfluencies were identified and classified into filled and unfilled pauses, interruptions and lengthenings. Disfluencies occurring at clause boundaries or following discourse markers were excluded from analysis, since they may signal an ongoing planning process rather than an expressive difficulty.

All gestures were identified and classified as co-occurring with fluent or disfluent speech. They were further coded for structural properties (that is, complete vs. suspended stroke). The purpose of this coding was to be able to ascertain whether gestures during disfluent speech tend to be completed strokes or to be interrupted along with speech. Finally, following Kendon (2004), completed gestures were functionally classified as referential (also labelled as representational in the literature) or pragmatic gestures. Referential gestures express semantic content through depiction of referential properties (e.g., size, shape, function). Pragmatic gestures, in contrast, do not express referential content but rather function like speech acts, parsing speech or commenting on the speaker's spoken production. In this coding, we excluded those gestures that could not be determined as being either referential or pragmatic (n=20 gestures or 8% of the total number of gestures in disfluencies).

Non-parametric statistical tests were used, specifically Mann-Whitney tests for comparisons of two independent samples and Kruskal-Wallis tests for comparisons of several independent samples. Because the dependent variables are proportions, they were arcsine transformed for statistical analysis (Howell, 2002); however, non-transformed values are reported in tables, figures and text.

Results

Gestures with fluent vs. disfluent speech

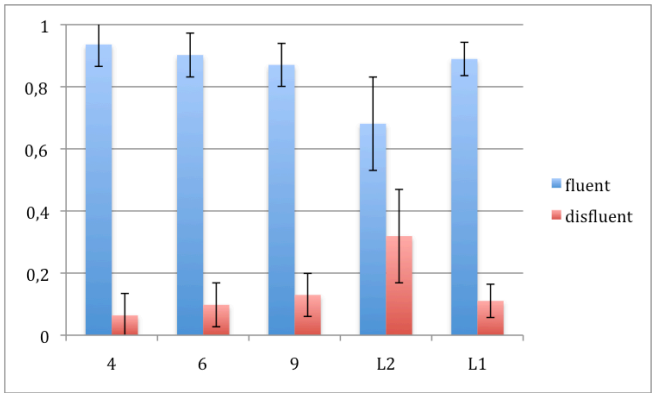

Figure 1: Mean proportion of gestures occurring with fluent vs. disfluent speech across groups, children aged 4, 6, 9 years, adult L2 learners, and adult L1 speakers.

First, we examined the extent to which speakers gesture during fluent vs. disfluent speech across the groups (Figure 1). The vast majority of all gestures occurred with fluent speech in all groups. A Kruskal-Wallis test revealed a significant difference in the use of gestures during fluent stretches between groups (χ^2 (4, N=55) = -22.11, p < 0.0001). Mann-Whitney tests revealed that adult L2 speakers were significantly less likely to align gestures with fluent speech than 4-year-olds (z = -3.68, p < 0.0001), 6-year-olds (z = -3.41, p < 0.0003), 9-year-olds (z = -2.95, p < 0.0016), and native L1 speakers (z = -3.25, p = 0.0006), who did not differ.

Articulatory properties of gestures in disfluency

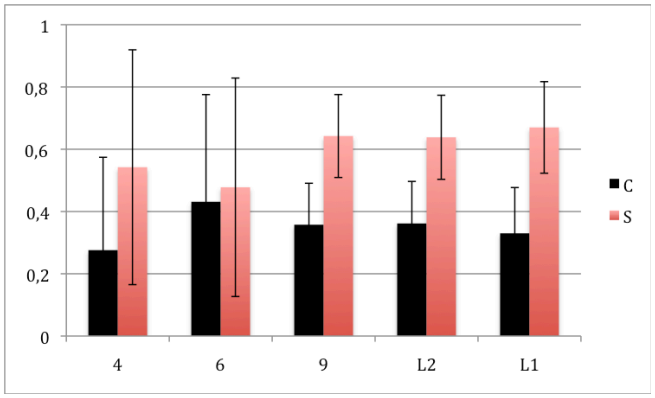

Figure 2: Mean proportion of completed (C) and suspended (S) gestures occurring during disfluency across groups, children aged 4, 6, 9 years, adult L2 learners, and adult L1 speakers.

Next, we examined the tendency for gestural activity during disfluencies to be completed or suspended strokes across the groups (Figure 2). The gestural activity in all groups during disfluencies tended to be suspended gestures. The groups did not differ in this respect (χ^2 (4, N=55) = 1.07, p = 0.899.

Gesture functions in disfluency

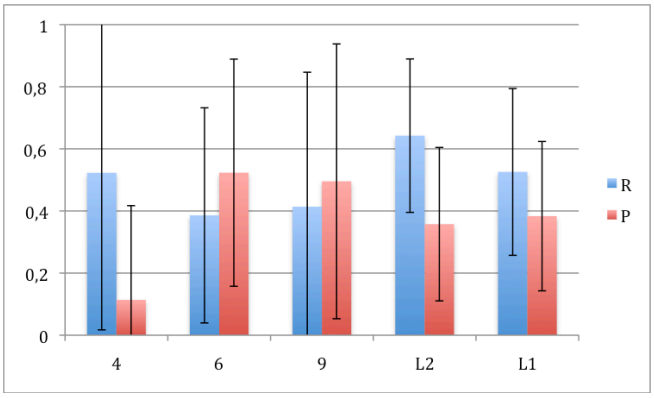

Figure 3: Mean proportion of representational (R) and pragmatic (P) gestures occurring during disfluency across groups, children aged 4, 6, 9 years, adult L2 learners, and adult L1 speakers.

Finally, we examined whether gestures completed during disfluencies were predominantly referential or pragmatic across the groups (Figure 3). All groups produced referential gestures during disfluencies. Despite the numerical differences the groups did not differ in this regard given the small sample size and the substantial individual variation (χ^2 (4, N=55) = 3.52, p = 0.4748). However, all groups, especially form age 6 onwards, also produced a non-negligable number of pragmatic gestures during disfluencies.

Conclusion

This preliminary study of speakers with different degrees of expertise in a given language shows that speakers' gestural production differs in fluent vs. disfluent stretches of speech such that gestures overwhelmingly occur with fluent speech. This is true for child and adult learners and competent speakers alike. Notably, adult L2 learners are more likely to be performing gestures in disfluencies than any of the other groups, presumably because they occasionally use representational gestures to elicit lexical help from their interlocutors (cf. Gullberg, 1998; 2011). Moreover, ongoing gestural activity during the onset of disfluencies is mostly suspended in all groups, providing further evidence that gesture production and speech production are suspended in tandem. Finally, although the majority of gestures actually completed in disfluencies are referential, speakers also produce pragmatic gestures, relating not to lexical content but to aspects of difficult interaction that ensues from the disfluencies.

Clearly, a range of more fine-grained questions will need to be examined based on these findings. For example, a pressing issue is to categorise the different types of speech disfluencies to probe whether different types of disfluencies are accompanied by different types of gesture behaviour, and whether child and adult learners and native speakers differ in these regards.

Although preliminary, the findings constitute an important challenge both for gesture theories assuming a mainly (lexical) compensatory role for (referential/representational) gestures and for theories of language acquisition. The data strongly suggest that when speech stops, so does gesture in both competent and less competent language users. Moreover, the observation that gestures that do accompany disfluencies often are pragmatic gestures raises further important challenges for theories concerning the speech-gesture relationship which have hitherto been based on subsets of gestures (referential/representational) and solely on adult, competent, fluent speakers. Together, the findings provide strong support for the notion that speech and gestures form a tightly integrated system in both competent speakers and developing language users, both children or adults.

Acknowledgments

We gratefully acknowledge financial support from the European Science Foundation (Short Term Scientific Mission - COST Action 2102, n. 17), the Swedish Research Council (Vetenskapsrådet; A0667401), the Max Planck Institute for Psycholinguistics, and the Erik Philip-Sörensen Foundation.

References

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore, we gesture. *Language and Cognitive Processes*, 15, 593-613.

Capirci, O., Iverson, J.M., Pizzuto, E. & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language*, 23, 645- 673.

Colletta, J.-M. (2004). *Le développement de la parole chez l'enfant âgé de 6 à 11 ans. Corps, langage et cognition*. Hayen: Mardaga.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us thinking*. Cambridge, MA: The Belknap Press.

Graziano, M. (2009). *Rapporto fra lo sviluppo della competenza verbale e gestuale nella costruzione di un testo narrativo in bambini dai 4 ai 10 anni.* SESA – Scuola Europea di Studi Avanzati - Università degli Studi "Suor Orsola Benincasa", Napoli, Italy & Université Stendhal – Grenoble 3, Grenoble, France. Unpublished doctoral dissertation.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund: Lund University Press.

Gullberg, M. (2011). Multilingual multimodality: Communicative difficulties and their solutions in second language use. In J. Streeck, C. Goodwin & C. LeBaron (Eds.), *Embodied interaction: Language and Body in the material world* (pp. 137-151). Cambridge: Cambridge University Press.

Howell, D. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury.

Iverson, J.M., Capirci O. & Caselli, M.C. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23-43.

Kendon, A. (1980). Gesture and speech: two aspects of the process of utterance. In M.R. Key, (Ed.), *Nonverbal Communication and Language* (pp. 207- 227). The Hague: Mouton.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162-185). Cambridge, UK: Cambridge University Press.

Krauss, R. K., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: a process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). Cambridge: Cambridge University Press.

Mayberry, R., Jaques, J. and DeDe, G. (1998). What stuttering reveals about the development of the gesture-speech relationship. *New Directions for Child Development*, 79, 77-87.

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

McNeill, D. (Ed.) (2000). *Language and gesture*. Cambridge: Cambridge University Press.

Nicoladis, E. Pika, S, Yin, H. & Marentette, P. (2007). Gesture use in story recall by Chinese–English bilinguals. *Applied Psycholinguistics*, 28, 721–735.

Pine, K.J, Bird, H & Kirk, E. (2007). The effects of prohibiting gestures on children's lexical retrieval ability. *Developmental Science*, 10(6): 747–754.

Ragsdale, J. D., & Silvia, C. F. (1982). Distribution of kinesic hesitation phenomena in spontaneous speech. *Language and Speech, 25*(2), 185-190.

Seyfeddinipur, M. & Kita, S. (2001). Gestures and disfluencies in speech. In S. Santi, I. Guaïtella, C. Cavé et G. Konopczynski (Eds.), *Oralité et gestualité. Communication multimodale, interaction* (pp. 266-270). Paris: L'Harmattan.