

Gestural holds and turn holding

On the coordination of speech, gesture and eye gaze in interaction management

When people engage in conversation, they make (mostly) implicit arrangements about who can speak at which moment. This negotiation of speakership is regulated by techniques in different modalities (Sacks et al. 1974). Traditionally, the verbal modality received much attention, but research in various research traditions shows that turn transitions are also regulated by gesture and eye gaze (e.g. Bavelas et al. 2002, Mondada 2007, Navarretta & Paggio 2013).

In this paper we are studying one part of multimodal turn management organization, namely turn holding. When speakers want to keep a turn in a conversation, they have a variety of techniques to choose from. Starting from a multimodal approach, we study turn holding techniques on three levels: verbal elements, eye gaze, and gesture. We aim to get an overview on how these different modalities interact in turn holding. For each modality, we selected one phenomenon that has typically been described as a turn holding technique. For the verbal level, we focus on fillers, the Dutch ‘*eah*’ and ‘*eahm*’, as speakers use fillers to keep the turn in moments of doubt or hesitation (Maclay & Osgood 1959, Clark & Fox Tree 2002). On the gestural level, we focus on the gestural hold, an element that, just like the fillers, has been described as a disfluency (Seyfeddinipur 2006). A speaker’s ongoing gesturing signals to the interlocutor that the turn is also still ongoing, and holding a gesture can thus be used to hold the turn (Duncan 1973, Gullberg & Kita 2009). As a third turn holding cue, we study gaze aversion, as speakers typically gaze away to focus on the planning of their utterance and prevent their interlocutors to take the turn (Kendon 1967, Jokinen et al. 2013, Weiß & Auer 2016).

To explore the interaction of these three semiotic levels, we performed a multimodal analysis on informal dyadic and triadic interactions in Dutch. We studied 6 dyadic and 3 triadic brainstorm sessions, resulting in a corpus of 60 minutes. For these brainstorm sessions, all participants, who were friends, were equipped with mobile eye tracking glasses, to enable us to capture simultaneously the participants’ eye gaze and their behavior in other semiotic channels, such as speech and gesture. The interactions were transcribed in ELAN and all fillers and gestural holds (gestures that were not moving for longer than 200 msec, cfr. Quek et al. 2000) were annotated. The speakers’ gaze behavior during and 500 msec before the filler or gestural hold was also coded, to study the co-occurrence.

The analysis shows that the interaction between these turn holding techniques is not very clear cut. Whereas verbal fillers fairly often occur with gaze aversion, the co-occurrence of a gestural hold and gaze withdrawal does not occur that often. A qualitative analysis of the data shows that eye gaze patterns seem to reflect functional variety on the verbal and gestural level. When the filler or gestural hold is used in situations of speech planning, it often co-occurs with a gaze aversion. In these cases, a close coordination of speech, gesture and eye gaze is found. When the verbal and gestural elements occur in other contexts, however, the elements are mostly accompanied by a sustained gaze at the interlocutor. This is, for example, the case when a filler is used in a turn eliciting context, or when a gestural hold is not a sign of disfluency, but rather an intrinsic part of a gesture, for example a deictic. These results show that functional differences in verbal or gestural elements are reflected in speakers’ eye gaze, and thus indicate the importance of incorporating a study of eye gaze to analyze multimodal turn management.

References

- Bavelas, J., Coates, L. & Johnson, T. 2002. Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, 566-580.
- Clark, H. H. & Fox Tree, J. E. 2002. Using 'uh' and 'um' in spontaneous speaking. *Cognition* 84, 73-111.
- Duncan, S. J. 1973. Toward a grammar for dyadic conversation. *Semiotica*, 9, 29–47.
- Gullberg M. & Kita S. 2009. Attention to speech-accompanying gestures: Eye movements and information uptake. *Journal of Nonverbal Behaviour* 33, 251-277.
- Jokinen, K. Furukawa, H., Nishida, M., Yamamoto, S. 2013. Gaze and Turn-taking behavior in casual conversational interactions. *ACM Transactions on Interactive Intelligent Systems* 3.2, article 12.
- Kendon, A. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26, 22-63.
- Maclay, H. & Osgood, C. E. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15, 19-44.
- Mondada, L. 2007. Multimodal resources for turn-taking: pointing and emergence of possible next speakers. *Discourse Studies* 9.2, 195-226.
- Navarretta, C. & Paggio, P. 2013. Multimodal turn management in Danish dyadic first encounters. In Allwood et al. (Eds.), *NEALT Proceedings. Northern European Association for Language and Technology, 4th Nordic Symposium on Multimodal Communication* 93, 5-12.
- Quek, F., McNeill, D., Bryll, R., Kirbas, C. Arslan, H., McCullough, K.E., Furuyama, N. & Ansari, R. 2000. Gesture, speech and gaze cues for discourse segmentation. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2, 247-254.
- Sacks, H., Schegloff E.A. & Jefferson, G. 1974. A simplest systematics for the organisation of turn-taking for conversation. *Language* 50, 696-735.
- Seyfeddinipur, M. 2006. Disfluency: Interrupting speech and gesture (MPI Series in Psycholinguistics 39). Nijmegen: University of Nijmegen.
- Weiß, C. & Auer, P. 2016. Das Blickverhalten des Rezipienten bei Sprecherhäsitationen: eine explorative Studie. *Gesprächsforschung* 17, 132-167.