Turn-timing and the body:
Gestures play a core role in coordinating conversation

Judith Holler, Kobin H. Kendrick, & Stephen C. Levinson

Conversation is the core niche of human language use and it is based on a turn-taking system. How we coordinate who says what and when is a significant pragmatic and psycholinguistic challenge. This becomes particularly evident when we consider that conversational turn-taking is remarkably fast, with gaps between speaking turns averaging around just 200 ms (Stivers et al., 2009). Considering that the production of single word utterances takes a minimum of 600 ms alone (Indefrey & Levelt, 2004), language production and comprehension must largely run in parallel; that is, while listening to an on-going turn, a next speaker has to predict the upcoming content, understand the speech act, and start preparing their own turn to be able to launch it on time (Levinson, 2013, 2016).

Considering that the primordial site of conversation is face-to-face social interaction where participants do not just speak but make use of a host of visual signals to communicate, a fundamental question arises: what is the role of the body in the coordination of speaking turns in conversation? In order to investigate this question, we carried out two studies, one quantitative analysis of multimodal conversational corpus data, and one based on spontaneous conversation combined with experimentally manipulating the availability of bodily signals.

For study 1, we analyzed a corpus of 7 casual face-to-face conversations between English speakers by identifying all question-response sequences (N=281), as well as the gestures that accompanied the identified set of questions, and the timing of these gestures with respect to the speaking turns they accompanied. Moreover, we measured the length of all inter-turn gaps in our set. To gain a first insight into whether gestures contribute to conversational coordination we asked whether the length of the gap between turns varied systematically as a consequence of questions being accompanied by gesture. Our results revealed that this is indeed the case: Questions with a gestural component were responded to significantly faster than questions without a gestural component. This finding holds when we consider head and hand gestures separately, when we control for points of possible turn completion in the verbal utterance prior to turn end, and when we control for the complexity associated with question type. Furthermore, our findings revealed a second, independent effect; namely, even within the group of questions accompanied by gestures, those questions whose gestures retracted prior to turn end were responded to faster than questions whose gestures retracted following turn end.

Study 2 is based on conversations that involved a within-participants manipulation: 20 dyads talked while they were able to see one another as well as while they were not. As for study 1, we measured the gaps between turns and compared the face-to-face with the no vision condition. The findings are in line with those from study 1 in that gaps between turns are shorter when interlocutors have bodily signals at their disposal, thus suggesting that bodily signals play important coordinative functions. A further experimental study is currently underway testing which types of gestures and other bodily signals facilitate early responding to speaking turns and what

mechanisms lie beneath this effect. Results are expected in time for the conference and will further elucidate the issue at hand.

In sum, the two studies suggest that the body plays an important role in the coordination of face-to-face conversation. Rather than burdening our cognitive system, gestures i) facilitate language processing, even in the rich and cognitively challenging context of conversational interaction, and ii) they seem to play a role also in the prediction of upcoming turn ends. Both of these contributions appear to contribute to interlocutors being able to respond fast in face-to-face conversation. The findings suggest an urgent need for adapting existing turn-taking models that focus primarily on the verbal modality (Sacks et al., 1974).

References

Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92,* 101–144.

Levinson, S. C. (2013). Action formation and ascription. Ins J. Sidnell and T. Stivers (Eds.), *The Handbook of Conversation Analysis*, pp. 101–130. Malden, MA:Wiley-Blackwell.

Levinson, S. C. (2016). Turn-taking in human communication, origins, and implications for language processing. *Trends in Cognitive Sciences, 20,* 6-14.

Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50,* 696–735.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. U.S.A., 106,* 10587.