Breathing and speech planning in turn-taking

Francisco Torreira, Sara Bögels & Stephen C. Levinson
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

In conversation, turn transitions between speakers often occur smoothly, usually within a time window of a few hundred milliseconds (Stivers et al., 2009). Since planning and producing a simple word takes around 600 ms (Levelt et al., 1999), it has been argued that conversational participants must start planning their utterance in overlap with their interlocutor's turn to achieve smooth turn transitions (Levinson 2013). In order to avoid overlaps and long gaps, speakers must also time their utterances accurately with respect to the end of their interlocutor's turn. However, direct evidence of early speech planning and accurate turn-end identification in conversation still remains scarce. In this talk, we present preliminary data from a project aimed at investigating whether the breathing behavior of conversational participants can provide such evidence. While previous studies on read speech have identified a relationship between breathing behavior and utterance duration (Whalen & Kinsella-Shaw, 1996; Fuchs et al., 2013), it remains to be investigated whether breathing behavior can be informative about speech planning in conversational speech as well.

Six dyadic unscripted conversations between Dutch male friends were recorded with head-mounted microphones and an Inductotrace^{TM} system of inductive plethysmography for around 40 minutes each. Each participant wore an Inductotrace band attached around his chest at the level of the axilla and a head mounted-microphone coupled to an amplifier. The speech and breathing signals were recorded simultaneously via an A/D converter connected to a computer.

We extracted and segmented 144 question and answer sequences from our data, and annotated all answerer's inbreaths that occurred between the start of the question and the start of the answer. We then examined a) if a relationship exists between the answerer's breathing behavior and the length of the answer, as has been found for read speech, and b) the timing of the answerer's inbreath relative to the end of the question (i.e. the moment when an answer is expected).

We found that 37% of the answers were not preceded by an inbreath, and, interestingly, that the presence vs. absence of an inbreath before the answer was related to the length of the answer. Figure 1 shows boxplots of answer duration for answers preceded and not preceded by an inbreath. It can be seen in this figure that answers preceded by an inbreath tended to be significantly longer than answers not preceded by an inbreath. This difference was statistically significant in a regression model with answer length as response and the presence of an inbreath as a predictor (β = 1087, t = 3.88, $p < .001$). It indicates that, as in controlled read speech, speakers' breathing behavior can be informative about the scope of speech planning in conversational speech as well.

Regarding the timing of the answerer's inbreaths relative to the questioner's turn end, we observed that the most frequent timing (i.e. the mode of the distribution) was located at the end of the question. This is illustrated in Figure 2a, which shows a density plot of this measure. It can also be seen in this figure that the distribution of inbreath timings is skewed to the left, with more inbreaths starting in overlap with the question (with negative values) than in the gap following the question (with positive values). Subsequent inspection of early vs. late inbreaths revealed that inbreaths occurring before long answers, for which inbreaths are presumably required, tended to cluster closely around the end of the question, whereas inbreaths preceding short answers displayed a wider spread and earlier timings in general. This is illustrated in Figure 2b, which shows the timing of inbreaths for answers shorter and longer than 2.5 s. These observations suggest that many of the early inbreaths could be vital or partly-vital inbreaths not primarily intended for speech, and that it is the late inbreaths that could be more specifically designed for speech. In our question and answer sequences, therefore, the timing of speech inbreaths before the longer answer appears to be sensitive to where precisely the question ends.

Additionally, given that preparing an inbreath requires at least a few hundred ms (i.e. activation of internal intercostals alone requires 140 to 320 ms, Draper et al. 1960) and that speech inbreaths are contingent on the planned length of the answer, the fact that inbreaths mostly occur close to or before the end of the question provides evidence that participants in a conversation often start planning their speech in overlap with their interlocutors' turn.

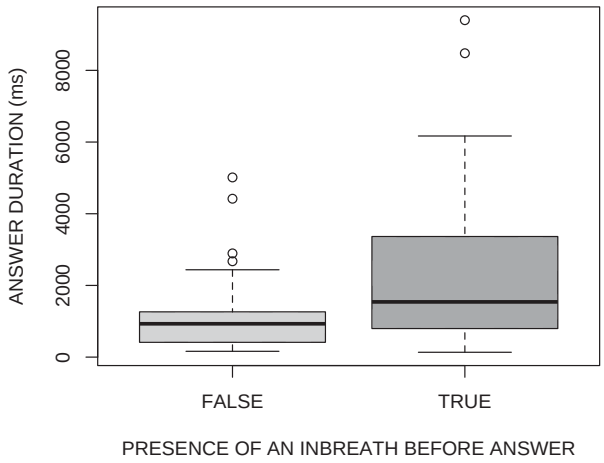

Figure 1 Answer duration as a function of the presence or absence of an inbreath before the answer.

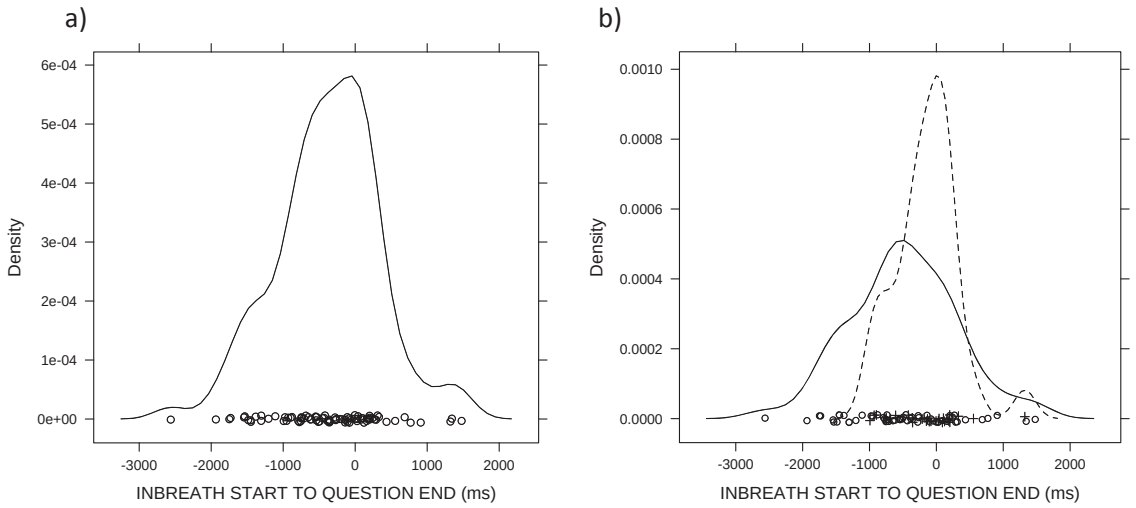

Figure 2 Answerer's inbreath start relative to question end (ms), divided in two groups of answer duration in the right panel (dashed: > 2500 ms; solid: < 2500 ms).

References

Draper, M. H., Ladefoged, P., and Whitteridge, D. (1960) Expiratory pressures and airflow during speech. *British Medical Journal*, 1(5189): 1837–1842.

Fuchs, S., Petrone, C., Krivokapic, J., and Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics*, 41(1):29–47.

Levelt, W., Roelofs, A., and Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1):1–37.

Levinson, S. C. (2013). Action formation and ascription. In Stivers, T. and Sidnell, J., editors, *Handbook of Conversation Analysis*, pages 103–130. Wiley-Blackwell.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., and Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS*, 106(26):10587–10592.

Whalen, D. H. and Kinsella-Shaw, J. M. (1997). Exploring the relationship of inspiration duration to utterance duration. *Phonetica*, 54:138–152.