

# The temporal relationship between feedback and pauses: a pilot study

Kristina Lundholm Fors

Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, Sweden

kristina.lundholm@gu.se

## Abstract

In this pilot study we investigated the temporal relationship between pauses and feedback. We found that the majority of feedback items occur in the proximity of trp-pauses (pauses that occur at a trp, within a speaker's turn), but also that most intraturn pauses do not coincide with feedback units. This suggests that when modeling feedback in human-computer interaction, a method to identify trp-pauses will also provide suitable places for feedback.

**Index Terms:** feedback, back-channels, pauses

## 1. Introduction

Both feedback and pauses are closely linked to turntaking. Pauses can occur when a speaker is trying to determine whether the other speaker wants the turn, or if she/he may continue speaking, and feedback can be used to signal that the other speaker is free to continue.

In this study we will explore the temporal relationship between pauses and feedback items. Pauses in general do not seem to be sufficient predictors for feedback [1], but our hypothesis is that some subtypes of pauses are more closely related feedback than others. Since we are using a rather small sample of dialogue, we will not be able to draw general conclusions, but rather we will use this pilot study to examine whether the relationship between feedback and different types of pauses might yield results that could be useful in feedback modeling.

### 1.1. Feedback

Feedback and related phenomena have also been referred to as, for example, backchannels and continuers. One of the defining characteristics of feedback is that it is not uttered in an attempt to claim the turn. Further, feedback can be produced at the same time as someone else is speaking, but it is not perceived as interruption. All feedback items do not have the same function. Allwood et al. suggest a model based on four basic functions of feedback: to signal willingness and ability to continue the conversation, to perceive the message, to understand the message and to convey attitude, specifically acceptance or rejection, towards the message [2].

Feedback tends to be preceded by certain cues, such as differences in intonation, duration and voice quality,

but cues vary between individual speakers [3]. Regions of low pitch may be good predictors of upcoming feedback [1]. The amount and type of feedback given is also dependent on cultural background of the speakers [4].

Feedback does not have to be verbal; smiles and nodding are common feedback signals. However, in this study we will focus on verbal feedback behavior.

### 1.2. Pauses

Sacks et al. drew a distinction between *pauses* and *gaps*, where pauses are silent intervals that occur within a speaker's turn, and gaps occur when a speaker has stopped speaking and no one else has been nominated or has taken the turn [5]. Previous work has shown that speakers tend to vary their pause lengths in synchrony, which means that when one speaker is lengthening his/her pauses, so will the other speaker [6]. In this study we concentrated on pauses, and subdivided pauses into two groups: pauses that occurred at a transition relevance place (TRP) and pauses that occurred elsewhere within a speaker's turn. The pauses are referred to as trp-pauses and ntrp-pauses respectively. When categorizing the pauses, only the activity of the current speaker was taken into account. A trp-pause was defined as a silent interval within a speaker's turn, where the speaker could have finished and yielded the turn. Pauses that occurred where the speaker did not seem finished were categorized as ntrp-pauses. Hjalmarsson used a similar method when judging semantic completeness, and found that interrater agreement for this measure was high [7].

### 1.3. Feedback in human-computer interaction

When modeling feedback in human-computer interaction, timing is highly important. Numerous different models have been developed to identify feedback places, based on for example prosody, POS-tagging and pause duration [8], pitch variations [1] and multimodal output features [9].

## 2. Method and material

The material used was an approximately 10 minute long, spontaneous dialogue with two Swedish female speakers. The speakers were recorded in a recording studio

and they were free to discuss any topic.

The recordings were transcribed orthographically and analyzed in Praat [10]. Pauses were identified manually, based on the acoustic signal. After identification, pauses were categorized as trp-pauses and ntrp-pauses respectively, as outlined in section 1.2. Feedback items were operationalized as a short, isolated utterance that could, but did not have to, be produced by one speaker while the other speaker had the turn.

### 3. Results

The dialogue contained 75 feedback items (the distribution is shown in table 1), and "mm" was the most common feedback unit. The mean length of feedback units was 312 ms (SD 124 ms).

| Type of backchannel | Frequency |
|---------------------|-----------|
| mm                  | 51        |
| aa                  | 8         |
| ja                  | 5         |
| jaa                 | 4         |

Table 1: Feedback items that occurred more than once in the material

The distribution of the pauses found in the material is shown in table 2.

|            | Feedback present | No feedback |
|------------|------------------|-------------|
| Trp-pause  | 33               | 30          |
| Ntrp-pause | 7                | 98          |
| Sum        | 40               | 128         |

Table 2: Trp-pauses and ntrp-pauses

76% of pauses did not overlap with feedback. Of the 75 feedback items, 40 occurred completely or partially during a pause. The majority of the pauses that coincided with feedback were trp-pauses: 33/40 feedback units occurred at at trp-pause.

We also examined the distance from the beginning of each feedback unit to the beginning of the nearest pause (either preceding or succeeding the feedback). 55 feedback items (79%) were closest to a trp-pause, and the mean difference between the beginning of the feedback unit and the beginning of the trp-pause was 15 ms (SD 669ms). 15 feedback items were closer to a ntrp-pause, with a mean distance of -490ms (SD 1252 ms).

As can be seen in figure 1, the majority of feedback units began  $\pm 1$  second of the beginning of a trp-pause.

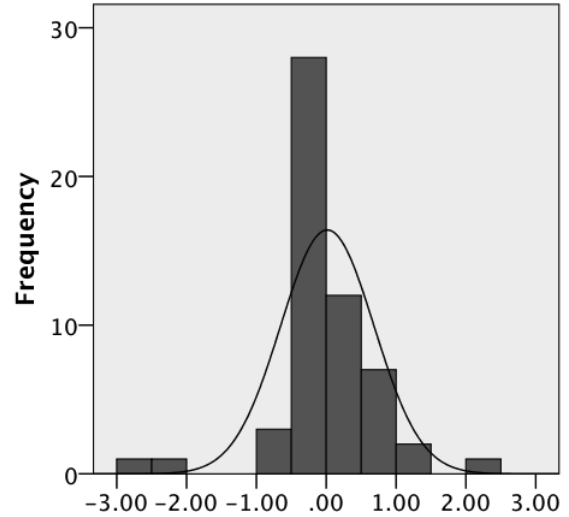


Figure 1: Distance between the start of feedback items and the start of the closest trp-pause

### 4. Discussion

Feedback may coincide with pauses, but pauses in general are not sufficient indicators of possible feedback locations. In our sample, 76% of pauses within a speaker's turn did not contain any feedback from the other speaker. However, the majority of feedback occurred at or in the proximity of trp-pauses, which was in line with our hypothesis. This suggests that the distance to the nearest trp-pause is a more useful indicator of a suitable feedback place than, for example, pause duration.

Feedback items are likely to occur at the same time as a trp-pause, or slightly before such a pause. This raises the question whether the feedback has an effect on pause length, that is, if trp-pauses preceded by feedback are shorter than other trp-pauses. It could be argued that if the feedback is produced shortly before the trp, the interlocutor would take this as a signal to continue and might shorten or even eliminate the upcoming trp-pause.

This pilot study of a short dialogue has shown clear tendencies for feedback to occur close to trp-pauses. We plan to investigate this further to see if these findings will hold in other dialogues as well.

### 5. References

- [1] N. Ward and W. Tsukahara, "Prosodic features which cue back-channel responses in English and Japanese," *Journal of Pragmatics*, vol. 32, no. 8, pp. 1177–1207, 2000.
- [2] J. Allwood, J. Nivre, and E. Ahlsen, "On the semantics and pragmatics of linguistic feedback," *Journal of Semantics*, vol. 9, no. 1, pp. 1–26, 1992.
- [3] A. Gravano and J. Hirschberg, "Backchannel-inviting cues in task-oriented dialogue," in *Tenth Annual Conference of the International Speech Communication Association*, 2009.
- [4] M. Stubbe, "Are you listening? Cultural influences on the use of

supportive verbal feedback in conversation,” *Journal of Pragmatics*, vol. 29, no. 3, pp. 257–289, 1998.

- [5] H. Sacks, E. Schegloff, and G. Jefferson, “A simplest systematics for the organization of turn-taking for conversation,” *Language*, vol. 50, no. 4, pp. 696–735, 1974.
- [6] K. Lundholm Fors, “Pause length variations within and between speakers over time,” in *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue, Los Angeles, USA*, 2011.
- [7] A. Hjalmarsson, “Human interaction as a model for spoken dialogue system behaviour,” Ph.D. dissertation, Royal Institute of Technology (KTH), Sweden, 2010.
- [8] N. Cathcart, J. Carletta, and E. Klein, “A shallow model of backchannel continuers in spoken dialogue,” in *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1*. Association for Computational Linguistics, 2003, pp. 51–58.
- [9] L. Morency, I. de Kok, and J. Gratch, “A probabilistic multimodal approach for predicting listener backchannels,” *Autonomous Agents and Multi-Agent Systems*, vol. 20, no. 1, pp. 70–84, 2010.
- [10] P. Boersma and D. Weenink, “Praat: doing phonetics by computer [computer program],” 2012. [Online]. Available: <http://www.praat.org/>