Pitch characteristics of filled pauses

Malte Belz\textsuperscript{1}, Uwe D. Reichel\textsuperscript{2}

\textsuperscript{1} Humboldt-Universit\text{"a}t zu Berlin, \textsuperscript{2} Ludwig-Maximilians-Universit\text{"a}t M\text{"u}nchen

Research question

\(\Rightarrow\) Do filled pauses (FP) in German (\text{\^{a}}h/\text{\^{a}}hm) show systematic patterns of pitch characteristics?

\(\Rightarrow\) If yes, what functions might they have?

Floor-holding account and clitization

- FP are used to hold the turn during speaking (Maclay and Osgood 1959; Malandro et al. 1989).
- Turn-holding needs to prevent silences.

\(\Rightarrow\) Hyp. 1: Isolated FP might be more prominent than cliticized FP.

Visual account

- FP occur more frequently in audio-only situations (Kasd and Mahl 1965).
- Speakers in audio-visual dialogues mark turn-holding by non-verbal cues as well (Goodwin 1981).

\(\Rightarrow\) Hyp. 2: FP are more frequent in audio-only communication due to compensation processes.

Intonation patterns

- Gradual downssteps of \(f_1\) (O’Shaughnessy 1992), also within clauses (Shriberg and Lickley 1993).
- FP show flat intonational contours (Tseng 1999).

\(\Rightarrow\) Hyp. 3: Rising \(f_1\) contours on FP serve a different function (e.g. turn management).

Corpus GECO (Schweitzer and Lewandowski 2013)

- 48 German spontaneous mono- and multimodal dialogues (25 min each).
- 13 women aged from 20 to 25 years who are strangers to each other.
- Only the multimodal condition allows subjects to see each other (audio-visual), as opposed to the monomodal condition (audio-only).

CoPaSuL approach (Reichel 2014) for \(f_1\) stylization

- Contour-based (Co), parametric (Pa), superpositional (Su) representation.
- Intonation is stylized automatically.
- Superposition of linear global contours represent the \(f_1\) baseline.
- Third order polynomial local contours represent the local pitch movements related to accent groups.
- Additional kmeans clustering of the stylization coefficient vectors to describe the \(f_1\) shapes also in categorical terms.

Linear regression model of prominence

- Prominence: high \(c_2\) indicates a prominence-lending \(f_1\) deviation from the baseline.
- Global class interacts with modality \((p = 0.01)\) as well as with isolation \((p = 0.01)\).

Results

\(\Rightarrow\) Prominence on cliticized FP is stronger for rising pitch on cliticized than on isolated FP (rejecting Hyp. 1).

- In audio-only communication, FP do not occur more often than in audio-visual communication (rejecting Hyp. 2).
- 88% of FP are uttered with a steady \(f_1\) contour. However, FP are more prominent in the audio-only condition for rising pitch (supports Hyp. 3).

Conclusion

\(\Rightarrow\) Filled pauses differ systematically in their intonational appearances.

- Prominent FP in cliticized positions might hold the floor more easily.
- Compensation processes in audio-only dialogues are rather linked to prominence than to frequencies.
- FP with rising and prominent \(f_1\) contours might be better turn management signals than non-prominent contours.
- FP with non-prominent contours, in turn, might serve as more prototypical hesitation signals.

References