Individual and Systematic Pitch and Intensity Variation in Turn-End Construction

Kathrin Feindt

Kiel University

Turn-taking —as the basis of any conversation—is an intricate process with a strict set of norms that guide speaker change. Throughout the conversation, there are so-called Transition Relevance Places (TRPs) [1] at which a speaker change is possible, though not obligatory. In cases where a change follows, speakers tend to avoid gaps and overlapping speech. Instead, turns follow immediately after each other, on average within 225-250 ms [2]. In order to achieve this fast pace, listeners are sensitive to certain linguistic cues that enable them to project the upcoming TRP. More specifically, though, the current turn has to contain information on whether the speaker intends to keep or cede the floor and thus whether the listener is expected to launch speech. To achieve this, speakers use a cluster of different cues, amongst which pitch and intensity are seen as important factors.

To examine pitch and intensity patterns towards TRPs and their influence on the timing of the next conversational action, 488 turn ends of two-party natural German and Swedish conversation were analysed. Comparing these two languages will show whether there are cross-linguistic differences in the preference for certain cues. This will be especially interesting in the case of F0, as it is assumed that the Swedish pitch accent restricts the use of pitch for communicative purposes. After annotating the corpora with labels referring to sentence type, syntactic completeness, next action (speaker change, keep or backchannel) and the time elapsed till the next action (gap, overlap or no-gap-no-overlap), a Praat script was used to extract pitch and intensity values [3, 4]. Specifically, F0 values (in semitones above 1 Hz) and intensity (in dB) were measured at three points (i) P1 at the turn end, (ii) P2 at 200 ms before the TRP, (iii) P3 at 500 ms before the TRP. The F0 values were normalised with the speaker's baseline pitch [3, 4].

A two-way Manova was conducted to determine whether there is a difference between language groups compared to individual speakers and the occurrence of certain pitch or intensity patterns. This would thus answer the question of whether the deployment of pitch and intensity as a turn-taking cue is language specific or rather driven by individual preferences. Results show that there is neither a significant difference in intensity ($F(3, 959) = 0.168$; Pillai-Spur = 0.918) based on the speaker's language, nor in F0 ($F(3, 681) = 3.595$; Pillai-Spur = 0.013). There is, however, a significant effect of individual speakers on the realisation of F0 ($F(33, 2049) = 18.885$; Pillai-Spur = 0.000) as well as intensity ($F(33, 2883) = 16.578$; Pillai-Spur = 0.000). Thus, there are no language specific patterns found that lead to either gaps, overlaps or no-gap-no-overlap transitions. An even greater influence on the prosody at turn ends is the individual preference of the speaker. Great intra-speaker variability has also been observed for example in backchannel inviting cues by [5]. Thus, the phonology of conversation is composed of manifold unique realisations that are perceived and interpreted by the listener to achieve the goal of a smooth speaker transition in a short amount of time. Just as phones are subject to abstraction processes to match them to certain phonemes, unique realisations of turn-taking cues are abstracted to match the phonology of conversation.

doi:10.11576/pundp2022-1049

References

[1] H. Sacks, E. A. Schegloff, and G. Jefferson. A simplest systematics for the organisation of turn-taking for conversation. *Language*, 50(4):696–735, 1974.

[2] T. Stivers, N. J. Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, F. Rossano, J. P. de Ruiter, K. E. Yoon, and S.C. Levinson. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26):10587–10592, 2009.

[3] K. Feindt, M. Rossi, and M. Zellers. The time course of pitch variation towards possible places of speaker transition in German and Swedish. In *1st International Conference on Tone and Intonation (TAI)*, pages 112–116, 2021.

[4] M. Rossi, K. Feindt, and M. Zellers. Individual variation in F0 marking of turn-taking in natural conversation in German and Swedish. In *Speech Prosody 2022*, pages 185–189, 2022.

[5] A. Gravano and J. Hirschberg. Backchannel-inviting cues in task-oriented dialogue. In *Tenth Annual Conference of the International Speech Communication Association*, pages 1019–1022, 2009.

doi:10.11576/pundp2022-1049