Function and Prosodic Form of Backchannels in L1 and L2 German

Simon Wehrle & Martine Grice

University of Cologne, Germany

simon.wehrle@uni-koeln.de; martine.grice@uni-koeln.de

Backchannels (BC) are used in conversation to signal understanding or agreement. They are a ubiquitous and essential feature of spoken communication. It has been claimed that deviances from language-specific conventions of BC usage have negative consequences for understanding, comprehensibility and character attribution [1][2][3][4]. In a previous study involving mouse-tracking to indicate how (un)friendly, (in)attentive or (dis)interested a speaker sounded, German native listeners rated rising and falling BCs predominantly positively, whereas they rated flat intonation negatively [5]. In light of these perception results, we investigate how this flat intonation is distributed in BC productions, both in native speakers of German as well as in learners with L1 Vietnamese. In Vietnamese, flat (or slightly falling) BCs are most frequent and perceived as polite and acceptable according to [4].

Our corpus consists of Map Task data from 12 age-matched female speakers in three groups: 4 L1 speakers of German (G_{L1}), 4 L2 speakers of German with L1 Vietnamese (G_{L2}) and 4 L1 speakers of Vietnamese (V_{L1}). G_{L2} speakers were students in Germany and had a proficiency level of B2 CEF. In total, we analysed 135 minutes of dialogue containing 812 BCs. For German, all BC types except "*ja* (yes)", "mm(hm)" and "*genau* (exactly)" had previously been excluded, as these were the only BC types investigated in [5]. 73.2% of all BC tokens remained for analysis. The pitch contours of all BCs were manually corrected and smoothed. f0 was measured at 10% and 90% of token duration and the difference between the two points was calculated in semitones (ST) (cf. [4]). Results are presented on a continuous scale and tokens with a pitch movement of less than 1 ST are considered as "flat". Duration did not enter into the analysis as all tokens are very short, with very little variance (mean: 0.4s; SD: 0.11s).

Figure 1 shows ST values for all BC tokens. For G_{L1}, relatively few tokens are flat (16.7%), whereas for G_{L2}, almost twice as many are flat (30.9%) and both the mean and median values fall within the area +/- 1 ST. For V_{L1}, the overwhelming majority of BC tokens are falling, with very little variance, but also some level tokens (10%). This overall pattern is indicative of transfer effects from V_{L1} to G_{L2}.

Next, we consider different lexical types of BCs in German. Our working hypothesis is that prosodic form plays a special role for the non-lexical "mm(hm)" type, as speakers cannot rely on literal meaning in this case. For $G_{L1,}$ we found that all 67 "mm(hm)" BC tokens are rising, except one produced with a flat contour. For G_{L2}, on the other hand, "mm(hm)" tokens seem to be evenly distributed across rising, level and falling contours. We hypothesise that the G_{L1} distribution arises from an impetus to clearly distinguish "mm(hm)" BCs from filled pauses, which have a similar segmental form in German (*ähm*) and which are also non-lexical. Functionally, however, FPs and BCs play opposite roles in dialogue management. While BCs are used by listeners to signal to the interlocutor that the message is understood and that they may continue their turn, FPs are used by speakers to signal that their own turn is not finished and that they intend to hold the floor. Therefore, speakers may aim to distinguish non-lexical BCs and FPs as much as possible in their prosodic realisation.

Figure 2 shows a comparison of FP tokens (n = 194) with "mm(hm)" BC tokens (n = 119) in German. FPs in G_{L1} are firmly within the region of flat pitch contours, mostly with a very slight fall (cf. [6,7]). FPs in G_{L2} have a similar, but more variable distribution. Crucially, in G_{L1}, there is almost no overlap between FPs and non-lexical BCs. G_{L2} does not show this complementary distribution: FPs and non-lexical BCs overlap completely in their prosodic realisation. This overlap in G_{L2} BC and FP productions may result from the speakers' L1, as speakers of Vietnamese do not appear to produce FPs in any comparable sense, instead repeating or lengthening words to achieve the same function [8].

In sum, we have shown that G_{L2} speakers produce more BC tokens with a flat intonation contour than G_{L1} speakers. Furthermore, G_{L2} speakers do not distinguish the prosodic realisation of non-lexical BCs from that of FPs in the clear and precise way of G_{L1} speakers. In line with [5], both of these behaviours are likely to lead to negative character attributions and misunderstandings in conversation with G_{L1} speakers.

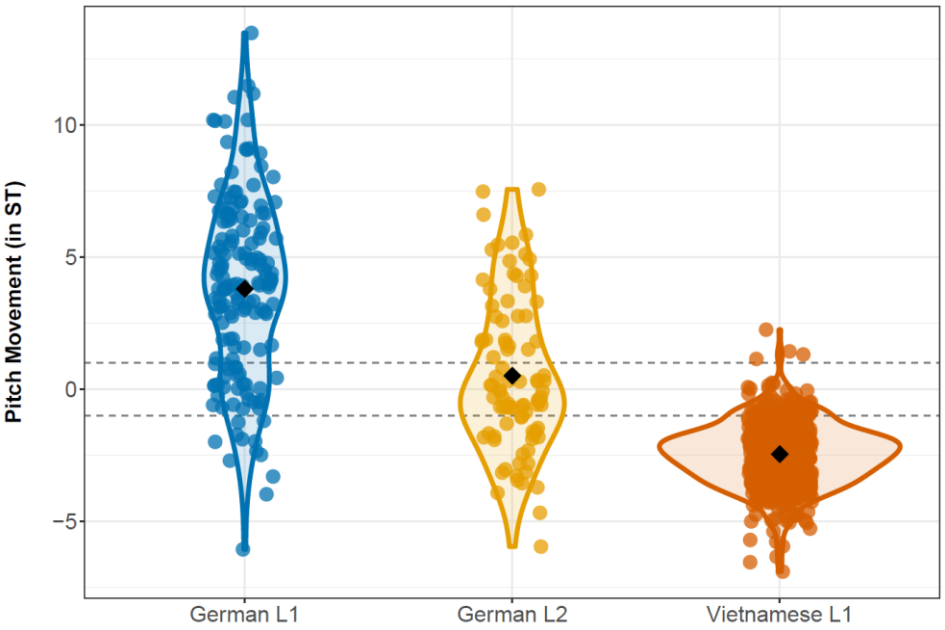

Fig.1 *Prosodic realisation of BCs in G_{L1}, G_{L2} and V_{L1}. Pitch movement in semitones on the y-axis. The dashed lines indicate the "flat" region of +/- 1 semitone. Black diamonds indicate mean values per group. Dots represent individual BC tokens.*

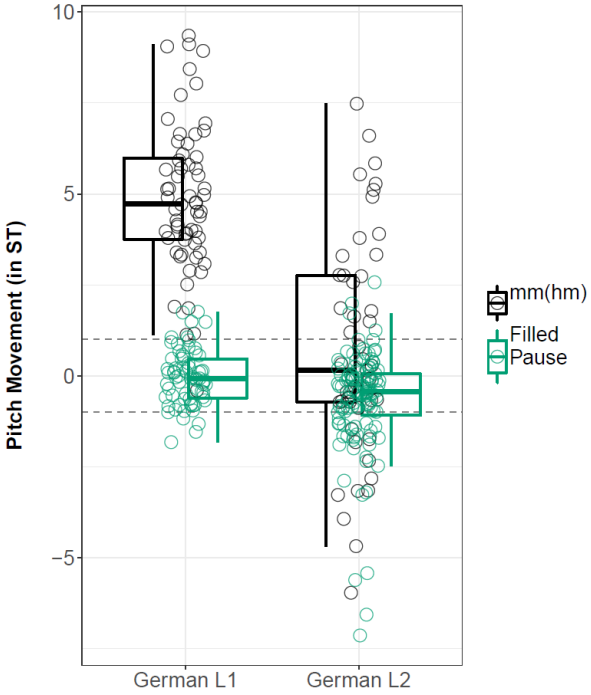

Fig.2 *Prosodic realisation of non-lexical BCs and filled pauses in G_{L1} and G_{L2}. Dashed lines indicate the flat region of +/- 1 semitone. Circles represent individual tokens.*

References

[1] Cutrone, P. (2011). Politeness and face theory: Implications for the backchannel style of Japanese L1/L2 speakers. *Language Studies Working Papers, 3*, 51-57.

[2] Li, H.Z. (2006). Backchannel responses as misleading feedback in intercultural discourse. *Journal of Intercultural Communication Research, 35*(2), 99-116.

[3] Shelley, L., & Gonzalez, F. (2013). Back Channeling: Function of Back Channeling and L1 Effects on Back Channeling in L2. *Linguistic Portfolios, 2*(1), 97-108.

[4] Ha, K.P., Ebner, S. & Grice, M. (2016). Speech prosody and possible misunderstandings in intercultural talk - A study of listener behaviour in Vietnamese and German dialogues. *Proc. Speech Prosody 8*, Boston, 801-805.

[5] Wehrle, S., Roettger, T.B. & Grice, M. (2018, November). *Exploring the dynamics of backchannel interpretation: The meandering mouse paradigm*. Paper presented at ProsLang: Workshop on the Processing of Prosody across Languages and Varieties, Wellington, New Zealand.

[6] Shriberg, E. (2001). To 'errrr'is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association, 31*(1), 153-169.

[7] Adell, J., Bonafonte, A., & Escudero-Mancebo, D. (2010). Modelling filled pauses prosody to synthesise disfluent speech. In *Speech Prosody 2010-Fifth International Conference*.

[8] Marc Brunelle, personal communication, March 28, 2019.