Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Backchannelling across Languages: Rate, Lexical Choice and Intonation in L1 Italian, L1 German and L2 German

Simona Sbranna, Eduardo Möking, Simon Wehrle, Martine Grice

I*f*L-Phonetik, University of Cologne, Germany
{s.sbranna, emoekin1, simon.wehrle, martine.grice@}uni-koeln.de

Abstract

Backchannel (BC) are short verbal and non-verbal signals of acknowledgments, whose realisations differ from language to language and between native (L1) and non-native (L2) speakers. Our study is the first comprehensive, cross-linguistic analysis of backchannels in L1 and L2 speakers.

We recorded 20 dyads of Italian learners of German in both their L1 and L2 (9 beginner and 11 advanced) and 5 dyads of German native speakers performing a Map Task. We analysed backchannel rate, lexical type, function (marking passive recipiency, PR, or incipient speakership, IS) and prosodic realisation.

BC rate was similar across languages in all groups and across proficiencies in L2 speech. Intonation was dependent on the lexical choice of BC and/or the function expressed. "Mm-hm" was mostly rising and marked PR. "Genau" was predominantly falling across functions. German "ja" and Italian "sì" were produced with more falling contours for IS and more rising contours for PR. Learners showed a high degree of variability overall, but clearly preferred German lexical BCs that were shared with their L1 Italian.

Overall, we found a complex, non-arbitrary mapping between lexical type, function and intonation in both languages. For L2 speech, speaker-specific behaviour (across languages) had a stronger effect than level of proficiency.

Index terms: backchannels, intonation of backchannels, second language acquisition, backchannels in second language acquisition.

1. Introduction

Backchannels are very short utterances [1] used by a listener to show understanding and acknowledgement of the interlocutor's speech [2], [3]. By signalling the active engagement of listeners, backchannels have been shown to positively contribute to fluency in social interactions.

However, cross-language differences have been found regarding BC rate and function [4], [5], [6], intonation contours [7], [8] and lexical types [9]. Further perception studies [4], [7], [10], [11], [12] have demonstrated that listeners are highly sensitive to changes relating to backchannel timing or acoustic parameters, which can lead to miscommunication and misunderstanding in cross-cultural settings [11], [13]. Therefore, it is important to investigate how backchannels differ across languages and how they develop in interlanguages with the goal of raising awareness in multicultural communicative contexts and improving L2 speakers' interactional skills.

To date, very few studies have investigated backchanneling in second language acquisition ([6], [8], [9] and [13] among others). Results mostly show a transfer of learners' L1 backchannelling behaviour to the L2. These studies differ in design (i.e. how participants are matched, their proficiency levels in the second language, the setting of the dialogue, the task used for dialogue elicitation and aspects of BCs analysed), making it difficult to compare their findings. Moreover, most studies analysed the speech of subjects with different L1 backgrounds, making it impossible to distinguish between mere transfer phenomena and cross-linguistic, speaker-specific characteristics.

In this contribution, we present an in-depth, cross-linguistic, descriptive analysis of backchannels. A homogeneous design for data collection and analysis allows for a direct comparison of backchannels across languages and nativeness, i.e. of L1 Italian, L1 German and L2 German speakers, the latter with Italian as L1. For the L2 analysis, we use a within-subjects design so as to compare learners' backchannelling behaviour in their L1 and L2 and in order to differentiate transfer phenomena from idiolectal features.

We take into account time-related, functional and intonational aspects of backchannels following [14] and [15], to offer a comprehensive picture of backchanneling behaviour.

2. Methods

In our analysis, we defined BCs as tokens showing the interlocutor's acknowledgement of what the speaker is saying either without claiming the floor (marking 'passive recipiency'), or signalling the listener's intention to start a turn of their own (marking 'incipient speakership'), as in [16].

2.1 Participants

We recorded 40 Italian native speakers, who were learners of German, and 10 German native speakers.

All Italian speakers had grown up in the dialectal area of Naples with parents of the same origin, to rule out variation in their L2 intonation resulting from their native linguistic substratum. Learners were studying L2 German either at university level or at the Goethe Institute in Naples.

Their proficiency levels were established on the basis of the language courses they were attending at the time of the recordings, corresponding to the levels described by the Common European Framework of Reference for Languages (CEFR) [17]. For the sake of determining potential effects of proficiency, we divided them into two groups: beginners (from A1 to B1 levels) and advanced (from B2 to C2 levels).

10.21437/SpeechProsody.2022-149

L1 German participants had grown up in North Rhine-Westphalia and were students at the University of Cologne.

2.2 Corpus and data collection

Our corpus includes 19 dialogues in L1 Italian (a further file was damaged; total duration = 100 minutes), 20 dialogues in L2 German, 9 of which were performed by beginners (78 min.) and 11 by advanced learners (89 min.), and 5 dialogues in L1 German (52 min.).

Data were collected using a collaborative goal-oriented map task [18]. In this game two participants, an instruction giver and an instruction follower, are each given a map showing several different landmarks. The aim of the task is for the instruction follower to find out about and reproduce a route, which is on the instruction giver's map but missing on their own map. The two maps contain mismatches in the type and position of landmarks. Participants were not informed of the mismatches prior to the task and could not see each other or each other's maps during the task.

Italian learners of L2 German were recorded at the Goethe Institute in Naples and were matched by proficiency level to carry out the task in dyads. They first read instructions and performed the task in their L1, then watched video instructions for the task in German and, finally, performed the task again in their L2. Native German speakers were recorded at the Department of Phonetics of the University of Cologne.

2.3 Procedure

We extracted 1451 backchannels. We analysed their rate per minute, lexical type, function and prosodic realisation.

241 tokens were excluded because they did not display the necessary amount of periodic energy to perform a prosodic analysis, e.g. items produced with creaky voice, or items with a voiced portion that was too short, such as "sì". As a consequence, 1210 BCs (83% of the extracted data) underwent prosodic analysis.

2.3.1. BC type

All BCs produced during the dialogues were annotated using *Praat* [19]. Tokens included lexical and non-lexical items, which were subsumed under the types "mmhm", "okay", "ja", "sì", "va bene" (Engl. "all right"), "genau", "esatto" (German and Italian forms for "exactly"). These token types cover 92% of the whole corpus. We used a category called "other" to label remaining token types that were used on very few occasions.

2.3.2. BC function

Backchannels were further distinguished according to the turn-taking functions of passive recipiency and incipient speakership. Tokens that were produced without the speaker taking the floor and simply served as signals to the primary speaker that they may continue, were labelled as acknowledgement tokens marking passive recipiency (PR). When a speaker used backchannels to acknowledge the interlocutor's turn but then took the floor by continuing to speak, (when a turn transition took place), these backchannels were labelled as marking incipient speakership (IS).

Notice that our corpus includes considerably more instances of acknowledgement tokens marking PR (1164) than IS (290).

2.3.3. BC intonation

After token annotation and extraction, the F0 trajectory of the extracted tokens was pre-processed through smoothing and manual correction of pitch points [20].

The contour was categorized as either rising, flat or falling and measured in semitones (ST; with a reference value of 1 Hz). To do so, pitch points were taken from two time points in the audio files of each token, one at the beginning and one at the end of the signal. The script was programmed to sample pitch points at 10%-90%, 20%-80% or 30%-70% of the token duration depending on where voiced material was first available. This means that if a token began with a voiceless consonant and no pitch points were present at 10% of its total duration, the script tried to sample at 20% and, if necessary, again at 30% of its total duration. Following [15], differences between the first and the last pitch point above +1 ST were defined as *rises* in intonation. Values within the range +/- 1 ST were defined as *level*. Values below -1 ST were categorized as *falling*.

3. Results

3.1 Rate

The rate of BCs was very similar across native languages (L1 German = 6.12 BCs/min, SD = 1.86; L1 Italian = 5.13 BCs/min, SD = 2.33). Learners' BC rate was lower than in the native groups, with beginners (2.87 BCs/min, SD = 4.99) producing fewer BCs per minute than advanced speakers (4.81 BCs/min, SD = 3). This result may lead to the conclusion that with increasing proficiency and overall fluency, advanced learners tend to approximate native backchannelling behaviour. However, Fig. 1 reveals similarities within dyads and across languages for BC rate.

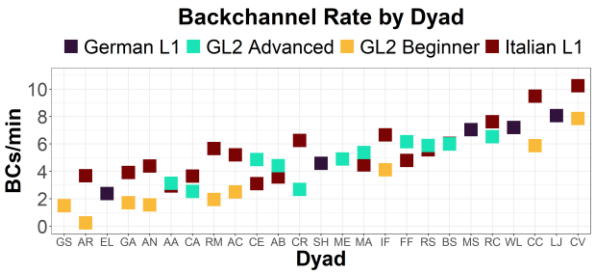

Figure 1: *BC rate by dyad.*

Note that GS did not produce any BCs in L2 (only other VSUs not represented in the plot) and that there are no L1 data for ME (damaged file). Dyads AA, RS and BS display overlapping squares, signalling that values of BC rate are almost identical across L1 and L2.

Advanced speakers show a more similar BC rate across L1 and L2 than beginners, but both learner groups clearly tend to reproduce their own L1 rate when speaking in L2.

Finally, the five L1 German dyads also show high variability, which in turn means that there is no ideal target for BC rate that learners could or should approach. Instead, BC rate seem to be dependent on dyad-specific behaviour across languages and groups.

We will henceforth report on learners independently of their proficiency level, since no proficiency effect could be

observed for type, function and prosodic realisation of BCs, which we will comment on in the following paragraphs.

3.2 BC type by function

The preferred choice of BC type differs according to function, i.e. acknowledgements marking passive recipiency and acknowledgements marking incipient speakership (Fig.2).

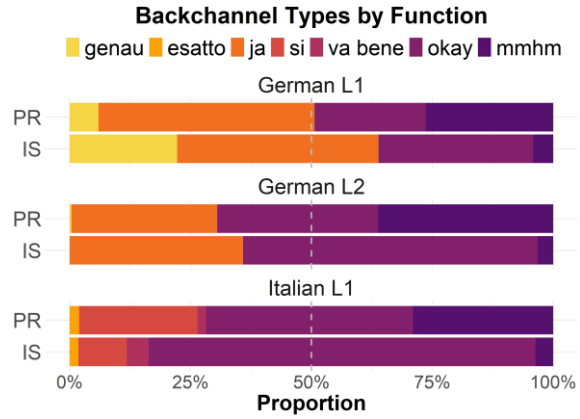

Figure 2: *BC type by function.*

For PR, L1 German and L1 Italian speakers behave similarly, tending to use "okay", "ja/sì" and "mmhm". In particular, the two L1s share approximately the same amount of "mmhm" (25% and 22% respectively), but L1 German speakers use more "ja" (43%) than "okay" (22%), while the converse is observed for Italian, with "okay" (41%) being more frequent than "sì" (23%). In both languages, "genau" (6%) and "esatto" (2%) are seldom used to express this function.

For IS, L1 German and L1 Italian differ more in preferred backchannel type, with "okay" being used in Italian almost exclusively (72%) and "genau" (22%) being more frequently used in German as opposed to Italian "esatto" (2%). Interestingly, in both languages the non-lexical "mmhm", that was largely used to express PR, is only occasionally used to express IS (4% in German and 3% in Italian).

Italian learners of German transfer the types preferred in L1 to their L2 in the case of PR (apart from one single instance of "genau" produced in the advanced group). Surprisingly, in the case of IS learners show a greater use of "ja" as compared to "sì" in their native Italian, but still the preferred type is the same as in their native language, i.e. "okay". Moreover, none of the learners use the most typical German "genau" in combination with this function.

3.3 BC intonation by type and function

In line with previous results on Bari Italian [14], we find that PR and IS functions of backchannels tend to be expressed with rising and falling contours respectively (Fig. 3).

This trend is common across the languages under investigation, but the pattern is clearer in L1 German than in L1 Italian and L2 German (as can be seen in the violins representing distributions), with the latter showing a substantial transfer of native patterns with higher variability than in L1. In particular, Italian learners of German approximate a bimodal distribution for PR in both their L1 and L2 (violins with a narrower section around zero and two wider sections above and below zero), indicating that a small proportion of items with a falling contour is also used to express PR at the expense of level contours, which are used more in German L1.

An analysis of intonation contours by lexical type and function reveals a more complex picture behind this apparent relationship between contour and pragmatic function, showing that intonation is also dependent on word choice. We present BC contours in both a continuous (Fig. 4) and a categorical fashion (Fig. 5), so that proportions can be related to the amount of data points for each type and function.

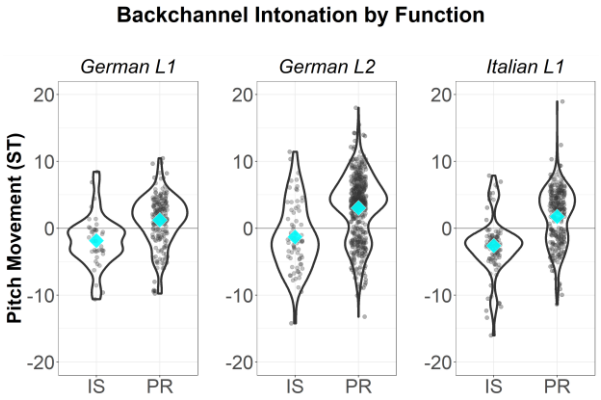

Figure 3: *BC contours by function. Values above zero represent items with a rising contour; values below zero represent items with a falling contour. Cyan diamonds represent mean values.*

Fig. 4 shows that the distribution of rising or falling contours by function can be highly variable for some BC types, e.g. in the case of "okay", and that the previously identified function-contour relation does not apply to all types, e.g. in the case of "mmhm" and "genau".

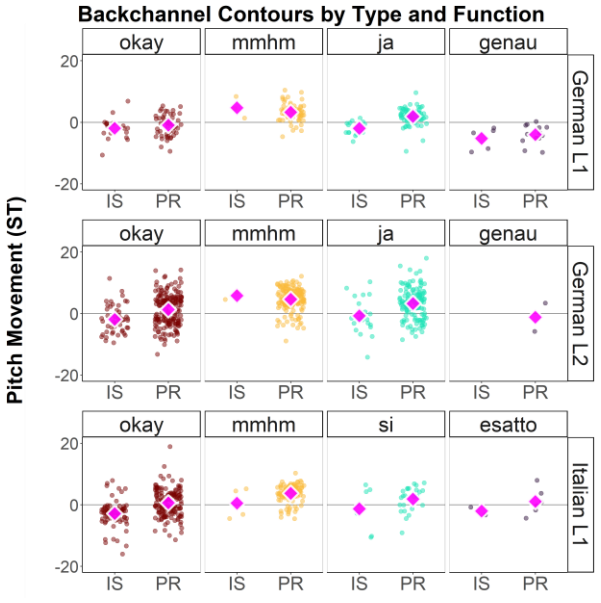

Figure 4: *BC contours by type and function. Values above zero represent items with a rising contour; values below zero represent items with a falling contour. Magenta diamonds represent mean values.*

In both native languages, "okay" displays a similar amount of rising and falling contours for PR (Fig.4), which is even clearer from the categorical analysis (Fig.5). We find 38% falling, 12% level and 50% rising contours in L1 Italian, and 46% falling, 20% level and 34% rising contours in L1 German for this type, suggesting that the variability of contours associated with "okay" with a PR function explains the bimodal distribution observed for PR in L1 Italian and L2 German in fig.3. Differently, when expressing IS, "okay" also follows the general trend in fig.3 with L1 Italian presenting 83% falling contours and L1 German 61% falling and 27% level contours.

Two types present a preferred contour independently of their function, i.e. "mmhm" and "genau". "Mmhm" is typically rising across the two languages and L1/L2, while "genau" is predominantly falling in L1 German. By contrast, the corresponding Italian "esatto" presents equal proportions of rises, falls and levels in the PR condition. Learners transfer this variability to the corresponding word in their L2, using for "genau" both rising and falling contours. Notice that these observations are based on only few data points for the types "esatto" in L1 Italian (eight tokens) and "genau" in L2 German (two tokens produced by the advanced group).

"Ja" and "sì" tend to follow the trend displayed in fig. 3. When expressing PR, they tend to be used with a rising contour in both L1s (73% in German and 63% in Italian). When expressing IS, L1 German shows a preference towards falling and level contours (50% and 43% respectively). Differently in L1 Italian both contours are equally possible, but also in this case the number of instances is limited (six tokens) and more tokens may provide a different result. As stated in 2.3, a prosodic analysis of many "sì" tokens was not possible due to the shortness of the voiced portion. In L2 German, however, the amount of "ja" is higher (129 tokens for PR and 22 tokens for IS) and confirms that this type tends to be produced with a rising contour when expressing PR (70% rises and 7% levels) and a falling contour when expressing IS (59% falls and 13% levels).

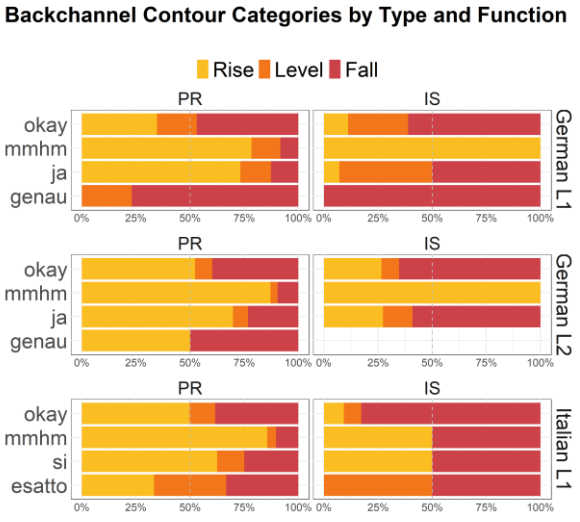

Figure 5: *Categorical classification of BC contours by type and function.*

4. Conclusion

With this contribution, we aimed at providing an in-depth analysis of BCs across native Italian and German and in L2 German spoken by Italian learners. We took into account rate, lexical and non-lexical types, function and prosodic realisation. We found a complex relationship between intonation, BC type and function across languages.

We did not find major differences in BC rate across L1s, with both groups presenting dyad-specific variability. Taking into account dyad-specific rate of backchannels revealed that learners do not approximate an ideal L2 target and, instead, tend to reproduce their own L1 rate. Advanced learners appear to approach their L1 BC rate better than beginners do, possibly owing to overall improved fluency in their L2.

We found that choice of BC type relates to the function they express, with a preference for certain type-function relations across languages. L2 German speakers prefer BC types that are shared with Italian over specifically German ones.

An even more complex, non-arbitrary mapping between lexical type, function and intonation is found in both languages. Overall, there is a preference for producing PR acknowledgements with a rising contour and IS acknowledgements with a falling contour across languages and groups of speakers. However, when BC type is taken into account, it becomes clear that it is not a one-to-one relation. A possible example of this function-contour relation is the case of "ja" and "sì" types. The limited amount of data in L1 Italian did not allow us to reliably report this trend. Hypothesising that learners are transferring their L1 patterns to the L2, we could assume that this tendency may also be valid in L1 Italian, confirming that the prosodic realisation of some types is influenced by the function they express. This hypothesis could be addressed by future research.

On the other hand, we found that the prosodic contour can also be highly variable for certain types, as in the case of "okay", which is used for PR with both rising and falling contours, or that it can even be type-specific, as in the case of "mmhm" and "genau", where the prosodic outcome is mostly unidirectional and independent of the function expressed.

Our results suggest that future studies should not address single aspects of BCs, but rather the relation among them to provide a complete picture of backchanneling behaviour across different languages. Moreover, since some language-specific aspects of BC use emerged from our study (e.g., the use of more level contours and of the type "genau" in L1 German), L2 language pedagogy would benefit from cross-linguistic studies to improve learners' communicative skills in cross-cultural settings.

5. Acknowledgements

This research was funded by a.r.t.e.s. Graduate School for the Humanities Cologne and the German Research Foundation (DFG), grant number 281511265, SFB 1252 Prominence in Language. We thank the Goethe Institute of Naples for making rooms available for the recordings and Prof. Silvia Palermo for recruitment of participants at the University of Naples.

6. References

[1] J. Edlund, M. Heldner, and A. Pelcé, "Prosodic features of very short utterances in dialogue," in *Nordic Prosody: Proceedings of the Xth Conference*, pp. 57–68, 2009. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.174.658&rep=rep1&type=pdf. [Accessed Nov. 18, 2021].

[2] V. H. Yngve, "On getting a word in edgewise," in *Papers from the sixth regional meeting*, M. A. Campbell et al, Eds. Chicago Linguistic Society, 1970, pp. 567–578.

[3] E.A. Schegloff, "Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences," in *Analyzing Discourse: Text and Talk*, D. Tannen, Ed., Washington, DC: Georgetown University Press, 1982.

[4] G., Tottie, "Conversational style in British and American English: The case of backchannels," in *English Corpus Linguistics*. Longman, pp. 254–271, 1991.

[5] A. Berry, "Spanish and American turn-taking styles: A comparative study," in *Pragmatics and Language Learning Monograph Series*, L.F. Bouton, ed., vol. 5, pp. 180-190, Urbana: University of Illinois, 1994.

[6] E. D. Galaczi, "Interactional Competence across Proficiency Levels: How do Learners Manage Interaction in Paired Speaking Tests?" *Applied Linguistics*, vol. 35, no. 5, pp. 553–574, 2014.

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

[8] S. Wehrle, and M. Grice, "Function and Prosodic Form of Backchannels in L1and L2 German," Poster at *Hanyang International Symposium on Phonetics and Cognitive Sciences of Language*, Seoul, South Korea, 2019.

[9] E. Castello, and S. Gesuato, "Holding up one's end of the conversation in spoken English Lexical backchannels in L2 examination discourse," *International Journal of Learner Corpus Research*, vol. 5, no. 2, pp. 231-251, 2019.

[10] P. Cutrone, "A cross-cultural examination of the backchannel behavior of Japanese and Americans: Considerations for Japanese EFL learners," *Intercultural Pragmatics*, vol. 11, no. 1, pp. 83-120, 2014.

[11] H. Z. Li, "Backchannel responses as misleading feedback in intercultural discourse," *Journal of Intercultural Communication Research*, vol. 35, no. 2, pp. 99-116, 2006.

[12] S. Wehrle, T. B. Roettger, and M. Grice, „Exploring the Dynamics of Backchannel Interpretation: The Meandering Mouse Paradigm," presented at ProsLang: Workshop on the processing of prosody across languages and varieties, Wellington, New Zealand, 2018.

[13] L. Shelley, and F. Gonzalez, "Back Channeling: Function of Back Channeling and L1 Effects on Back Channeling in L2," *Linguistic Portfolios*, vol. 2, no. 1, p. 9, 2013.

[14] M. Savino, "The Intonation of Backchannel Tokens in Italian Collaborative Dialogues," in *Human Language Technology Challenges for Computer Science and Linguistics. LTC 2011. Lecture Notes in Computer Science*, Z. Vetulani, J. Mariani, Eds., vol. 8387, Springer, Cham, 2014.

[15] S. Wehrle, "A Multi-Dimensional Analysis of Conversation and Intonation in Autism Spectrum Disorder", Ph.D. dissertation, University of Cologne, 2022.

[16] K. Drummond, and R. Hopper, "Back channels revisited: Acknowledgement tokens and speakership incipiency," *Research on Language and Social Interaction*, vol. 26, no. 2, pp. 157-177, 1993.

[17] Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge, U.K.: Press Syndicate of the University of Cambridge, 2001.

[18] A. H. Anderson, M. Bader, E. G. Bard, E. H. Boyle, G. M. Doherty, S. C. Garrod, and R. Weinert, "The HCRC map task corpus," *Language and speech*, vol. 34, no. 4, pp. 351-366, 1991.

[19] P. Boersma & D. Weenink, *Praat: doing phonetics by computer* (Version 6.2) [https://www.praat.org], 2021.

[20] F. Cangemi, *mausmooth*, [http://phonetik.philfak.uni-koeln.de/fcangemi.htlm], 2015.