DOI: 10.17469/O2109AISV000008

SIMONA SBRANNA, SIMON WEHRLE, MARTINE GRICE

The use of Backchannels and other Very Short Utterances by Italian Learners of German[1]

Backchannels (BCs) positively contribute to fluency in social interactions. However, their realisation is language-specific, which can cause miscommunication in intercultural contexts. Nevertheless, backchanneling is not formally taught in most classroom settings. To find out whether L2 learners still manage to acquire a target-like BC behaviour, we carried out an exploratory study on Italian learners of L2 German. We recorded Map Task dialogues performed by 6 dyads speaking L1 Italian and L2 German at different proficiency levels and 5 dyads speaking L1 German. We extracted BCs, defined as acknowledgment tokens, and other very short utterances (VSUs) with the same lexical realisation as BCs, but different functions. We analysed their frequency, length and lexical type according to their function. Preliminary results suggest that dyad-specific patterns play a larger role than L2 proficiency when predicting BC frequency and length. As for lexical choice of BC types, L2 learners prefer items shared with their L1 Italian. Specifically German types are only used by advanced learners, indicating a role of proficiency in this aspect of BC production.

Keywords: backchannels, very short utterances, L2 acquisition, individual variability, communicative competence.

1. *Introduction*

One issue in second language acquisition (SLA) research has been the question of how to assess communicative competence in a quantitative and homogeneous way, while taking into account idiosyncratic and contextual factors impacting the L2 learning process as well as L2 oral performance. Fluency has been widely recognised as a central aspect in the assessment of L2 oral performance (De Jong, 2016) and is listed under the abilities for oral interaction in the official European guidelines of language competence, the Common European Framework of Reference for Languages (Council of Europe, 2001; Figueras, North, Takala, Van Avermaet & Verhelst, 2009). However, fluency is a complex phenomenon, as further demonstrated by the lack of a clear-cut definition (e.g. Lennon, 1990, Lennon, 2000, Wood, 2001; Wolf, 2008).

[1] The authors covered different roles in the realisation of this article. Following the CRediT authorship contribution statement, Simona Sbranna contributed with conceptualisation, data curation, visualisation, writing – original draft; Simon Wehrle contributed with conceptualisation, methodology, visualisation, writing – review and editing; Martine Grice contributed with conceptualisation, methodology, supervision.

Most studies on L2 fluency have focussed on individual measures of fluency, such as speech and articulation rate; amount, location and duration of pauses; or repairs and repetitions (for a comprehensive list, see Saito, Ilkan, Magne, Tran & Suzuki, 2018). However, the majority of our real-life oral performances are interactions and much less often monologues. For this reason, 1) training and testing learners in monologic settings does not sufficiently help them to develop L2 fluency and 2) using individual measures in L2 fluency research risks providing only a partial picture of learners' oral competence, ignoring many other contributing factors.

Apart from idiosyncratic and contextual factors, such as a speaker's status, L2 proficiency level, L1 background as well as topic and setting (He, Young, 1998), fluency in dialogue is determined to a great extent by the specific interactional mechanisms that influence a conversation between two interlocutors. Therefore, unique dyad-related factors play a fundamental role, along with strictly individual factors. It is thus clear that the contribution of both interlocutors to the interaction, and the possible accommodation to the conversation partner's speech style, cannot be ignored when studying fluency in a conversation (for an extensive discussion on this topic see Sbranna, Cangemi & Grice, 2020). The specificity of the cooperation between parties shaping a conversation together has been described by Jacoby and Ochs (1995), who view interaction as a form of co-construction and a joint creation of discourse between interlocutors; by Hall (1993, 1995 as reported in He, Young, 1998), who argues that interactional competence emerges in varied interactive practices to which participants contribute with the appropriate linguistic and pragmatic resources; and by McCarthy (2009), who defines interactional flow as a jointly achieved harmonisation of tempo.

The smoothness of a conversation is achieved, among other factors, through the rhythm of turn-taking (Sacks, Schegloff, & Jefferson, 1974). In particular, smooth or disfluent turn transitions can take place at turn-boundaries (also called transition relevance places–TRP–in conversational analysis), and interlocutors have to appropriately foresee the end of the other party's turn and react to it quickly and accordingly (Levinson, 2015; Bögels, Torreira, 2015).

Despite usually going unnoticed in conversation (Shelley, Gonzalez, 2013), one important linguistic means that can facilitate the flow of conversation is the use of so-called backchannels. Backchannels are very short lexical and non-lexical utterances, like 'okay' or 'mm-hm', which have traditionally been described as non-intrusive tokens—that is, as not claiming a floor transfer—used to signal the listeners' active engagement by showing acknowledgement and understanding (Yngve, 1970, Schegloff, 1982). By supporting the ongoing turn of the interlocutor, backchannels positively contribute to fluency in social interactions (Amador-Moreno, McCarthy & O'Keeffe, 2013) as they maintain flow and contribute to creating a shared structure in dyadic conversation (Sacks, Schegloff, & Jefferson, 1974; Schegloff, 1982; Kraut, Lewis & Swezy, 1982).

On the other hand, backchannels can be potentially misleading in cross-cultural contexts, where different culturally-shaped communicative conventions come into

contact (Cutrone, 2005, 2014; Ha, Ebner & Grice, 2016; Li 2006; among others). Research has indeed provided evidence for language- or variety-specific backchannel characteristics concerning length, duration, frequency, location, intonation and function, and these can potentially have negative social implications in a communicational setting where the interlocutors' linguistic backgrounds diverge.

For these reasons, the importance of backchannels in L2 learning becomes clear. The CEFR (Figueras et al., 2009) already lists the use of feedback expressions under passive competence at the A2 level, an early stage in the learning process. However, backchannels are not explicitly taught or thematised in most L2 classroom settings and it cannot be taken for granted that learners will acquire appropriate backchannelling behaviour simply from exposure to the target language. Moreover, teachers are not always native speakers, and input on this particular interactional feature might be completely absent from many classroom settings.

Against this background, two possible outputs in the learners' interlanguage can be expected. On the one hand, it is possible to assume that backchannels go unnoticed in conversation, resulting in a transfer of features from the L1 to the L2. On the other hand, backchannels might be perceived by learners as salient features of foreign speech and receive an appropriate level of attention, which would favour an adaptation to target language patterns. In the latter case, more target-like backchannel behaviour should be observed, especially at an advanced level, i.e. with more experience and exposure to the target language.

With these two scenarios in mind, we will explore the use of backchannels in second-language learning. The paper is structured as follows: in section two, we offer a brief overview of the literature about the phenomenon of backchannels, their differences across cultures and languages and their acquisition in L2 learning; in section three, we present the methods used, including information on participants, data collection, corpus and measures; in section four, we present our results; and in section five we conclude the paper by summarising the findings and discussing their implications.

2. *Background research on backchannels*

In the literature, there is little agreement on the definition of backchannels (as noticed by Lennon, 1990, 2000; Rühlemann, 2007; Wolf, 2008 among others), resulting in a variety of names and categorisations that are often imprecise and overlapping.

In his analysis of telephone conversations, Fries (1952) was probably the first to recognise 'signals of attention' that do not interrupt the speaker's talk. Since then, other terms have been used to define this phenomenon, such as 'accompaniment signals' (Kendon, 1967), 'receipt tokens' (Heritage, 1984), 'minimal responses' (Fellegy, 1995), 'reactive tokens' (Clancy, Thompson, Suzuki & Tao, 1996), 'response tokens' (Gardner, 2001), 'engaged listenership' (Lambertz, 2011) and 'active listening responses' (Simon, 2018).

The term 'backchannel communication' was first coined by Yngve (1970) to define the channel of communication used by the listener and recipient to give

useful information to their interlocutor without claiming a turn, in opposition to the main channel used by the speaker holding the floor.

Early research on backchannels was mostly conducted on American English (Fries, 1952; Yngve, 1970; Duncan, 1974; Duncan, Fiske, 1977; Schegloff, 1982; Jefferson, 1984; Goodwin, 1986). These studies aimed at defining the phenomenon and tried to offer a categorisation of backchannel types, generally based on their pragmatic function or formal realisation.

Schegloff (1982) noted that these short utterances were mainly used by the listener not only to acknowledge the interlocutor's turn, but also to invite the primary speaker to carry on with his turn. For this reason, he defined the minimal utterances used in the specific contexts of an ongoing turn by the interlocutor as "continuers". Jefferson (1984) introduced the term "acknowledgement tokens". Indeed, in its narrow use the term backchannel refers to tokens used to signal acknowledgement and understanding of what the interlocutor is saying, while inviting the main speaker to continue (also used in this sense by Beňuš, Gravano & Hirschberg, 2007, Hasegawa, 2014, and others).

In its broader use, the term backchannel has also been matched to numerous other functions, and some attempts at establishing a function-based categorisation have been made. For example, Drummond and Hopper (1993) further distinguish acknowledgement tokens marking 'passive recipiency', as in the case of continuers, from those marking 'incipient speakership', signalling a listener's intention to start a turn of their own. Maynard (1997) categorises backchannels according to the functions of continuer, understanding, agreement, support, strong emotional answer and minor additions. Kjellmer (2009) recognises five functions of backchannels: regulative, supportive, confirmatory, attention-showing and empathetic. Tolins and Fox Tree (2014) distinguish context-generic backchannels, used as continuers and promoting the production of new information, and context-specific backchannels, also called assessments in previous studies (Goodwin, 1986), such as 'really' or 'wow', eliciting further elaboration of what has just been said.

As far as their formal realisation is concerned, backchannels present a high degree of lexical variability, but they can also be realised through vocal noises (Wong, Peters, 2007), visual modalities such as facial expressions, head movements, gestures (Tolins, Fox Tree, 2014) and responsive laughter (Hasegawa, 2014). Some structurally motivated proposals of classification have been advanced to categorise backchannel utterances. Tottie (1991) classifies them into simple, double and complex types. Simple backchannels are composed of one single utterance, e.g. 'yes', double backchannels are repeated simple types, e.g. 'okay okay', and complex backchannels are a combination of different simple types, such as 'okay yes right'. Wong and Peters (2007) differentiate between minimal, lexical and grammatical types. Minimal types are defined as non-lexical items that are semantically empty and items expressing polarity, e.g. 'mmhm', 'yes', and 'no'. Lexical types are considered to be all single words that are codified in dictionaries and show an increase in semantic weight, such as 'really', 'right', and 'good'. Finally, by grammatical types they mean

predications in the form of short codified phrases, such as 'I see', brief questions, repetitions, sentence completions and commentaries.

Later studies shifted their attention towards languages other than English and revealed that there are cross-cultural and cross-linguistic differences in the use of backchannels (e.g. Tao, Thompson, 1991; Tottie, 1991; Berry, 1994; Clancy et al., 1996; Ward, Tsukahara, 2000; Heinz, 2003; Cutrone, 2005, 2014; Li, 2006; Nurjaleka, 2019; Kraaz, Bernaisch, 2022).

2.1 Backchannel use across languages and cultures

Since then, one main question in the field of backchannel research has been variation across languages and cultures, and differences in backchannel use have been identified regarding frequency, duration, location, lexical type, function and intonation.

Being bound to culture, backchannel behaviour has been found to diverge even across varieties of the same language. Tottie (1991) reports differences with regard to frequency and type across American and British English, showing that in American conversations there was an average of sixteen backchannels per minute, compared with just five backchannels per minute in British conversations. Similarly, differences were observed across Sri Lankan and Indian English in type, frequency and function (Kraaz, Bernaisch, 2022).

Some studies report the impact of different backchanneling behaviour on the turn-taking system. For instance, in a cross-linguistic study on Spanish and North-American English, Berry (1994) found that backchannels were more frequent and longer among Spanish speakers, resulting also in longer stretches of overlapping speech. In turn, American English speakers were shown to use more overlapping backchannels than Germans, as reported in a comparative study by Heinz (2003).

These differences lead to the hypothesis of a potentially negative effect on communication in intercultural conversations. In a study on responsive tokens in English, Mandarin and Japanese, Clancy et al. (1996) observed that Japanese people produced the most frequent reactive tokens, placing them in the middle of the interlocutor's speech. Mandarin speakers, in contrast, produced the fewest backchannels and mostly at TRPs, i.e. at the end of interlocutors' turns. American English lay between the two other languages with regard to frequency, and reactive tokens were placed within interlocutors' turns and at TRPs, but preferably at grammatical competition points. The authors speculate that, in Japanese, backchannels are used as a form of emotional support and cooperation, whereas, on the opposite pole, Mandarin speakers might perceive Japanese backchannels as intrusive in comparison to their tendency to not interrupt the other speaker out of respect. American English speakers, likewise, might find Japanese speakers disruptive while the scarce reactions of Mandarin speakers might leave them wondering what their listeners are thinking (Clancy et al., 1996: 383). Similar hypotheses were tested in a study on backchannel intonation, in which Ha et al. (2016) found differences across Vietnamese and German. While Vietnamese continuers were consistently level or falling, German equivalents were tendentially rising. Based on the results

of a previous perception experiment (Ha, 2012), the authors hypothesise probable misunderstandings in intercultural dialogues. In Vietnamese, rising pitch as used by Germans might be interpreted as impolite. Conversely, for German natives, the level/falling pitch used by Vietnamese speakers might cause irritation (Stocksmeier, Kopp & Gibbon, 2007) and could be interpreted as showing disinterest or as an attempt to end the interlocutor's turn.

Given the observed differences across languages, the immediate next step in research was to put the consequences of this variation in intercultural conversations to the test and find out whether and to what extent differences in backchannel use can lead to miscommunication and/or have negative social implications. Li (2006) conducted a study on Canadian and Chinese speakers in intra- and intercultural conversations and showed that backchannels facilitated communication among speakers of the same language. However, when Canadian speakers were paired with Chinese speakers, the opposite effect was observed, leading to the claim that backchannel responses can be misleading in intercultural conversations and cause miscommunication. It was also found that the Chinese speakers produced the most backchannels and the Canadians the fewest, but when crossed, speakers tended to produce a number of backchannels in between. In a follow-up study providing an analysis of backchannel types (Li, Cui & Wang, 2010), it was found that, in intercultural conversations, both Canadian and Chinese speakers used other backchannels than in their respective native languages, showing some degree of speech convergence for both frequency and lexical type.

However, accommodation in intercultural conversations does not always take place automatically, as knowledge of language- and culture-specific conventions is likely to be essential. For example, White (1989) reports that Japanese speakers did not adapt their active listening style in conversations with Americans, while Americans did, because "they clearly have the linguistic ability to do so" (White, 1989:74), suggesting that language proficiency might be a prerequisite for accommodation. A high level of L2 proficiency can, indeed, provide the speaker with diverse linguistic means to select verbal tokens according to the respective context, and with the flexibility to recognise and switch among language conventions.

2.2 Backchannel productions by L2 speakers

To date, only relatively few studies have investigated backchannels in L2 speech. The relevant findings reinforce the assumptions made on the basis of intercultural studies in showing that L1 backchannel behaviour is generally carried over to the L2, frequently causing miscommunication and misperceptions.

For example, Cutrone (2005) examined the use of backchannels in dyadic interactions between Japanese EFL (English as a Foreign Language) and British speakers. Differences were found in frequency, type and location, and these negatively affected intercultural communication. The frequent backchannels used by the Japanese participants were interpreted as interruptions by the British speakers, and their interlocutors were perceived as impatient. In a follow-up study, Cutrone

(2014) reports that Japanese EFL speakers used a greater number of backchannels because it helped them to feel comfortable as listeners, showing a behaviour similar to the one reported for L1 Japanese (Clancy et al., 1996).

Wehrle and Grice (2019) also report on the negative effect of transfer on intercultural communication. In a pilot experiment they compared the intonation of backchannels in L2 German spoken by Vietnamese and observed that Vietnamese learners produced twice as many non-lexical backchannels (e.g. 'mmhm') with a flat intonation contour as German native speakers, showing a transfer from their L1. As previously mentioned, and corroborated in Wehrle and Grice (2019) through a mouse-tracking experiment, a flat backchannel contour in German might be interpreted as a signal of disinterest and cause irritation (Ha et al. 2016).

Another study that hypothesises a transfer of backchannel features from the L1 to the L2 was conducted by Castello and Gesuato (2019). They investigated the frequency and lexical types of 'expressions of convergence' in Chinese, Indian and Italian learners of English in a language examination setting. They found that Chinese learners used the most backchannels and Indian learners the least, with Italian learners lying between these two groups. They also observed differences in the choice of backchannel types between groups, suggesting an effect of L1 background.

A similar conclusion is reached by Shelley and Gonzales (2013), who analysed backchannel functions in informal interviews in four ESL (English as a Second Language) speakers with different L1 backgrounds as well as one American native speaker of English. They identify four backchannel functions: continuers (the listener is paying attention and does not hold the floor), acknowledgements (the listener agrees or understands), newsmakers (the listener communicates an emotional reaction) and change of activity (the listener signals to move toward a new topic). They report an effect of the L1 as differences in the preferred backchannel functions across the four speakers were found.

Finally, there are studies showing that higher proficiency in the L2 implies a better ability to use backchannels. Galaczi (2014) compared the frequency of backchannels and expressions of confirmation among learners of English with different proficiency levels. Results show that intermediate learners provided less feedback than highly proficient learners, among which the "ability to act as supportive listeners through backchanneling and confirmations of comprehension was found to be more fully developed" (Galaczi, 2014: 570).

To summarise, previous research on various languages provides converging evidence 1) for miscomprehension and misperception of the interlocutor's intentions due to a use of backchannels that diverges from the native conventions, 2) for a transfer of the L1 backchanneling behaviour to the L2, and 3) for proficiency as a positive factor in the improvement of learners' L2 backchannelling ability.

At the same time, these studies have some limitations. Their results are not easily comparable as they differ considerably in design and methodology: how participants in the dialogue were matched, their status, their proficiency level in the language of the conversation, the setting of the dialogue, the method used for dialogue elicitation

and aspects of backchannels analysed. Moreover, most studies have focussed on subjects with different L1 backgrounds, which is useful for detecting cultural-specific differences among groups of learners, but does not permit differentiation between transfer phenomena and cross-linguistic, speaker-specific characteristics.

Still, these findings have significant implications for the relevance of backchannels in language teaching environments. In order to better understand the mechanism behind cross-cultural backchannel behaviour, it is important to shed light on how the backchannelling ability develops in interlanguages, with the goals of raising awareness in multicultural communicative contexts and improving L2 speakers' interactional skills.

Therefore, in the present paper we try to overcome some of the limitations mentioned and relate the results to language pedagogy. In order to assess transfer phenomena and/or the acquisition of target-like backchannel features, we carried out an exploratory study using a within-subjects design. In particular, we investigate backchannel use across Italian learners' L1 and L2 German and compare their realisation to a German native group. We pay particular attention to dyad-specific behaviour in order to differentiate idiosyncratic factors from actual transfer or the acquisition of patterns. Finally, we take into account several aspects of BCs to offer a more comprehensive view of the phenomenon, i.e. backchannel frequency, length, type and function.

3. *Method*

The definition of the term 'backchannel' varies considerably in previous literature. For the purpose of this exploratory study, we will adopt the term 'very short utterances' (VSUs), proposed by Edlund, Heldner and Pelcé (2010) as a loose definition for the wide variety of interactional dialogue phenomena providing feedback to the interlocutors. According to this definition, backchannels are to be considered as a specific sub-category of VSUs with an acknowledging function. Words such as 'yes', which are used both as BCs and as VSUs with a different function – in our corpus as positive answers to yes-no and tag questions – will also be analysed to perform a comparison. The difference between BCs and positive replies is motivated by the fact that BCs are 'unsolicited', whereas in the case of replies the primary speaker gives up their turn by asking a question. For this reason, we will refer to 'backchannels' and 'acknowledgments' interchangeably, while we will refer to tokens that fulfil a function other than acknowledgment as 'other VSUs'.

3.1 Participants

For exploration purposes, we selected 22 speakers from a larger corpus: 12 speakers of L1 Italian and L2 German at different proficiency levels (6 beginner and 6 advanced), as well as 10 speakers of native German as a control group.

All Italian speakers had grown up in the province of Naples with parents of the same origin, ruling out variation in their L2 resulting from the native linguistic substratum. Learners were studying L2 German either at university level at the faculty of foreign languages and literatures or at the Goethe Institute in Naples. Their proficiency levels

ranged from A2 to C1 and 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 CEFR (A1-A2: beginner; B1-B2: intermediate; C1-C2: advanced). For the sake of determining potential effects of proficiency by using two balanced groups, we categorised them into beginners (from A1 to B1 levels) and advanced (from B2 to C2 levels). We acknowledge that this not a very precise way of identifying the proficiency of individual learners, as there is likely to be a high degree of variability with regards to different language skills both within and across language courses. Nevertheless, the classification used here can serve as a valid starting point, especially as we have a stated interest in how the definitions and demarcations outlined in the CEFR relate to language production in naturalistic conversational interactions.

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

3.2 Data collection and Corpus

Recordings were performed using headset microphones (AKG C 544 L) connected through an audio interface (Alesis iO2 Express) to a computer running Praat (Boersma, Weenink 2022). All participants were recorded in pairs, with L2 learners being matched by their proficiency level.

To collect data, we used the Map Task (Anderson et al., 1991; Grice, Savino, 2003 for set up, map layout and instructions), which matches the goal-oriented cooperation task described in the CEFR. For the task, participants sit opposite one another but have no eye contact. They are given two maps showing several landmarks, but only one map has a path drawn on it. The objective is to co-operate so that the participant without the path can reproduce it on their own map with the help of instructions given by the partner. The task is made more difficult as some landmarks are intentionally not identical across the two maps. The participants are not informed of these mismatches, as the purpose is to create situations requiring collaborative problem solving. This task is particularly useful when dealing with a mixed group of learners including beginners, as it can be performed at every proficiency level. Indeed, learners should have acquired the grammar and vocabulary necessary for giving directions at the beginner level, according to the CEFR.

Italian learners were recorded at the Goethe Institute in Naples. Learners first read the game instructions and carried out the task in Italian. Afterwards, before performing the task in their L2, they watched a video with a German native speaker (S.W.) explaining the instructions again in German to help them get into the language mode and reduce L1 bias. German native speakers were recorded at the University of Cologne. They watched the same German language video instructions and then played the game.

The resulting corpus for the 22 selected speakers includes 6 dialogues in L1 Italian (30 minutes in total); 6 dialogues in L2 German, with 3 performed by beginners (39 minutes in total) and 3 by advanced learners (22 minutes in total); and 5 dialogues in L1 German (52 minutes in total). We extracted and analysed a total of 924 VSUs, of which 646 were BCs.

3.3 Procedure

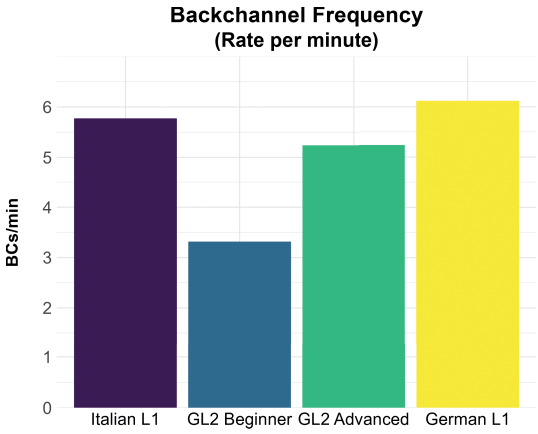
In our analysis, we took into account different aspects of backchannels, i.e. frequency, length, type and function. Specifically, frequency is operationalised as backchannel rate per minute. Length is their duration in milliseconds. Type refers to the lexical and non-lexical realisations. In our corpus, we found that the most frequently used lexical types were 'ja' and 'sì' ('yes' in German and Italian, respectively), 'genau' and 'esatto' ('exactly'), and 'okay'. The most common non-lexical type was 'mmhm'. These token types cover 92% of the whole corpus. We used a category called 'other' for the remaining tokens. Finally, we distinguished three possible functions: acknowledgement, categorised as 'BC', as well as positive replies to tag questions and to yes–no questions, both categorised as 'other VSUs'.

4. *Results*

4.1 Backchannel frequency

Figure 1 shows Backchannel frequency across groups, i.e. the rate of BCs per minute of dialogue. The rate of BCs is very similar across native languages (5.7 BCs per minute for L1 Italian and 6.11 BCs per minute for L1 German). The learners' BC rate is lower than that of both native groups. Beginners produced the fewest BCs, at a rate that was almost half of their output in the target language (3.31 BCs per minute). Advanced learners showed a rate of BCs more similar to the German L1 control group (5.32 BCs per minute). This result might lead to the appealing, but simplistic conclusion that learners acquire a native-like backchanneling behaviour with increasing proficiency. However, this characterisation is incomplete.

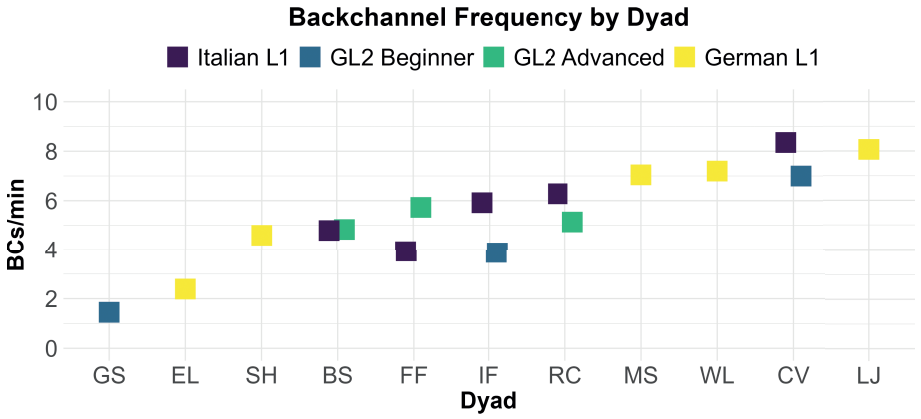
Figure 1 - *Backchannel frequency operationalised as rate per minute of dialogue. The number of BCs per minute is displayed on the y-axis. Language groups are shown on the x-axis and are colour-coded: violet for Italian learners' native speech; blue for beginner learners in L2 German; green for advanced learners in L2 German and yellow for the native German control group*



A closer look at BC rates by dyad [Fig. 2] reveals that dyad-specific behaviour plays a crucial role, in all groups. This becomes obvious when considering the fact that the BC rates across learners' L1 and L2 are strikingly similar. Consider, for example, dyad BS, with almost identical values across the two languages (visible in the overlapping squares). Moreover, the low BC rate in the beginner group as a whole is partly due to the peculiar behaviour of beginner dyad GS, who produced no backchannels whatsoever in their L1 Italian conversation—and only a few other VSUs, which are not displayed in the graph—and also only very few BCs in their L2. This L2 output is most likely not a consequence of low proficiency in German, given that this dyad produced no BCs at all in their L1. Therefore, this peculiarity can be ascribed to dyad-specific behaviour. On the other extreme end, we also observed that another beginner dyad, CV, produced the highest BC rate across all groups, which would not be predicted from group-level results.

A high degree of by-dyad variability can also be seen in the German L1 group. Importantly, one dyad, EL, produced a very similar rate to the beginner dyad GS, showing that a very low BC frequency can also be present among German native speakers, which calls into question the idea of a specific target behaviour to be reached by learners.

Figure 2 - *Backchannel frequency by dyad operationalised as rate per minute of dialogue. The number of BCs per minute is displayed on the y-axis. Dyads are shown on the x-axis and language is colour-coded: violet for Italian learners' native speech, blue for beginner learners in L2 German, green for advanced learners in L2 German, and yellow for the native German control group. Two values are shown for learners, corresponding to speech in their L1 and L2 and distinguished by the colour of the square*



4.2 Backchannel length

Figure 3 shows BC length, i.e. duration in milliseconds (ms). Here, the two native-language groups differ from one another, with L1 Italian speakers producing longer BCs (455 ms) than L1 German speakers (372 ms). A closer inspection of the tokens in the dataset revealed that this difference is mostly due to the fact that Italian

speakers tended to use more complex or repeated BCs, such as 'sì sì okay' or 'okay okay okay'.

Akin to what we already observed for frequency, length values for the two L2 groups seem to suggest an effect of proficiency. There is, indeed, a gradual decrease in BC duration across proficiency levels, from values more similar to the native Italian baseline in the beginner group (416 ms) to values approximating the target in advanced learners (392 ms). This hypothesis would be more appropriate than in the case of BC frequency previously discussed, due to the relatively clear difference between the native and the target languages, which allows an L1 target to be identified. Nevertheless, this claim would again provide an incomplete characterisation in the present case.

Looking at by-dyad values for length, displayed in Figure 4, it appears that dyad-specific behaviour again provides a better explanation than proficiency. Learners' by-dyad values are very similar across their L1 and L2, with one exception, dyad FF, who did show an evident reduction in BC length when speaking in L2 German. However, one side note on this dyad is required. In their case, the long BC duration in L1 Italian is not due to complex or repeated BCs, but to an atypical use of prolonged 'okay' tokens by the instruction follower (only), who did not replicate this behaviour in L2 German.

Figure 3 - *Backchannel length operationalised as duration in ms (on the y-axis). Language groups are shown on the x-axis and are colour coded: violet for Italian learners' native speech, blue for beginner learners in L2 German, green for advanced learners in L2 German, and yellow for the native German control group*

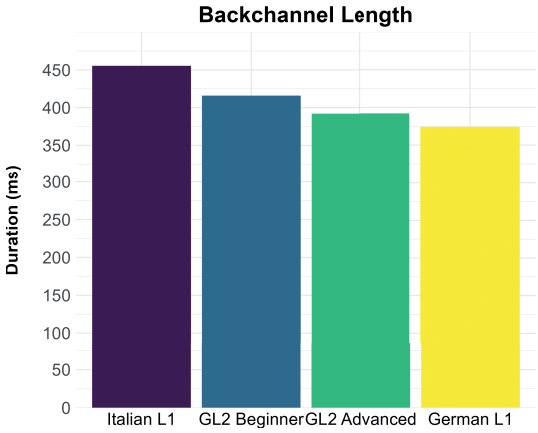
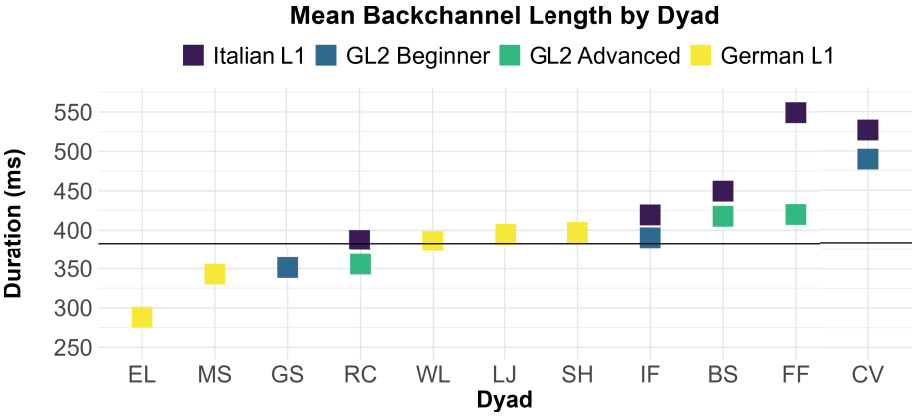


Figure 4 - *Backchannel length by-dyad operationalised as their duration in ms. Mean BC duration in milliseconds is displayed on the y-axis. Dyads are shown on the x-axis and the respective language groups are colour coded: violet for Italian learners' native speech, blue for beginner learners in L2 German, green for advanced learners in L2 German, and yellow for the native German control group. Italian learners of L2 German present two values corresponding to their L1 and L2 speech, distinguished by the colour of the square. The horizontal black line corresponds to the mean BC duration of the L1 German group pooled across all speakers, as a reference for learner productions*

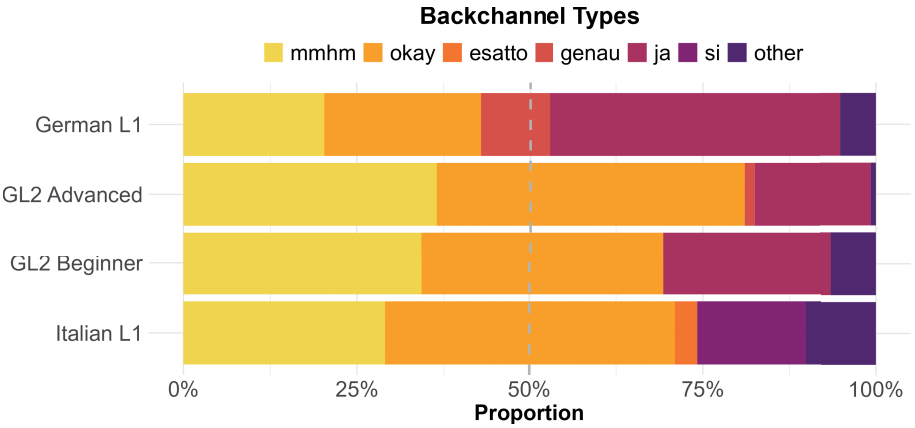


4.3 Backchannel Type

Figure 5 shows the proportions of BC types across groups. Comparing first the two L1s, it can be seen that they diverge regarding the preferred BC type. Both groups show a similar proportion of 'mmhm' (20% in L1 German and 29% in L1 Italian), but in L1 Italian there is a preference of 'okay' (41%) over 'sì' (15%), while the opposite is true in L1 German ('ja' 41%, 'okay' 22%). A type which seems to be typical for L1 German is 'genau' (10%), as the correspondent Italian 'esatto' is not used as much (3%).

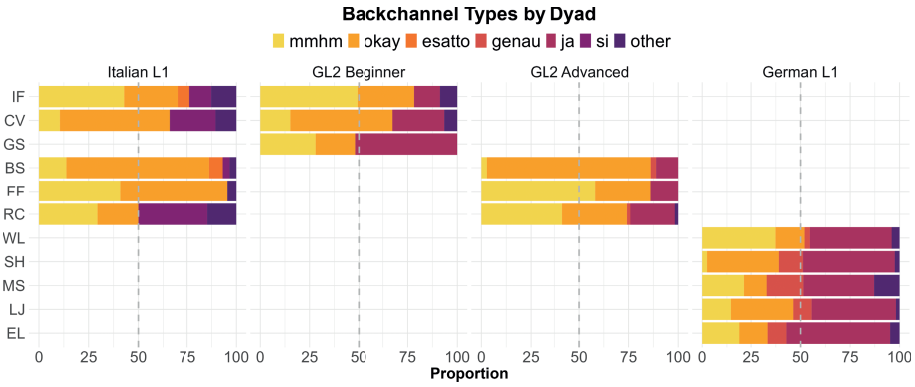
It is evident that L2 learners transfer their choice of BC type from the L1, showing similar proportions of 'mmhm', 'okay' and 'ja' across their L1 and L2 (beginners: 34%, 35% and 24% respectively; advanced: 36%, 44% and 16% respectively). Moreover, the word 'genau' is not used by beginners at all and only represents 1% of BC occurrences in the advanced learner group. This is probably a result of the fact that the Italian equivalent is used very rarely, meaning that learners need more experience and exposure to the L2 to start using this type of BC in the target language.

Figure 5 - *Backchannel Type. Proportions of BC types are shown in percentages on the x-axis. Language groups are shown on the y-axis and are each assigned a bar. The most frequently used BC types are listed in the legend and are colour-coded. The category "other" refers to types that were used only rarely*



The choice of BC type by dyad depicted in Figure 6 reveals highly similar patterns across the L1 and L2 within dyads, especially in the cases of IF, CV and BS, providing support for the transfer hypothesis and showing that dyad-specific patterns in the L1 tend to be reproduced in the L2. One difference between Italian L1 and German L1 concerns the proportions across types in a by-dyad comparison. In detail, it seems that the choice of BC type in L1 German is more consistent across dyads, whereas it appears more variable and dyad-dependent in Italian L1.

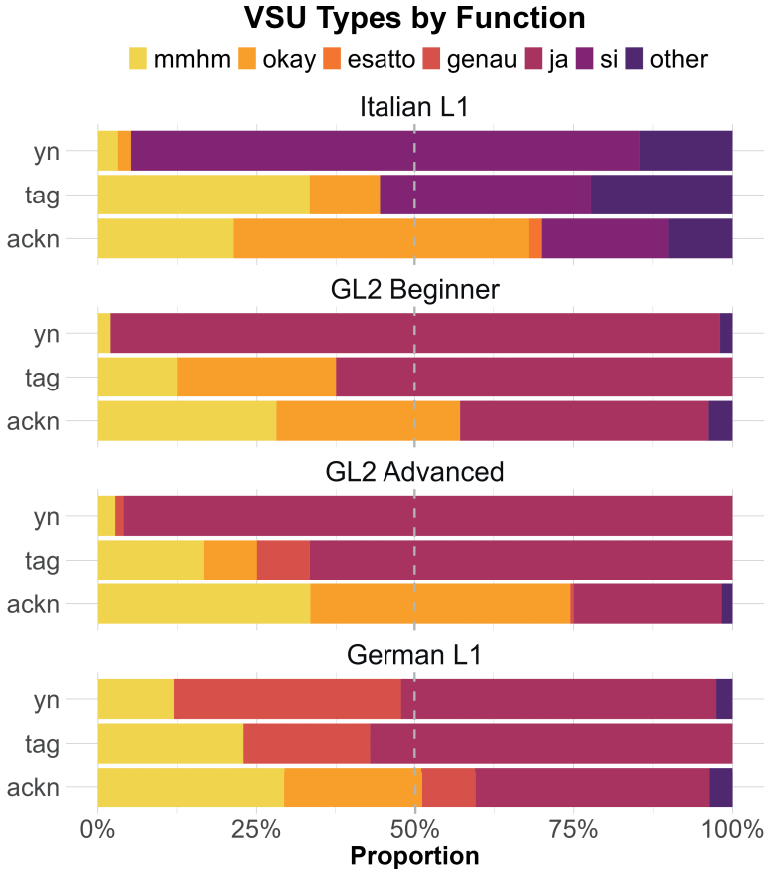
Figure 6 - *Backchannel type by dyad. Proportions of BC types are shown in percentages on the x-axis. Dyads arranged by language group are shown on the y-axis and are assigned a bar each. The most frequently used BC types are listed in the legend and are colour-coded. The category "other" refers to types that were used only rarely*



4.5 BC and VSU by type and function

Figure 5 shows the choice of BCs (acknowledgements) and other VSUs (replies to yes–no and tag questions) by functions and across language groups. The bars representing acknowledgments correspond to those in Fig. 7 and are repeated here to enable a direct comparison.

Figure 7 - *Very Short Utterance types by function for all language groups. Proportions of VSU types are shown in percentages on the x-axis. Functions are shown on the y-axis and are assigned a bar each: replies to yes–no questions, replies to tag questions and acknowledgements (BCs) to compare with. The most frequently used VSU types are listed in the legend and are colour-coded. The category "other" refers to types that were used only rarely.*



The two native languages differ in the proportion of types across the two VSU replies. For yes–no replies, Italians almost exclusively used 'sì' (80%), which is also transferred to the L2 by both beginner and advanced learners (96% respectively), while in L1 German there is more variety with only a few instances of "mmhm" (12%), more predominant 'ja' (50%) and many 'genau' (35%) utterances. For tag replies, Italians seem to prefer 'sì' and 'mmhm' (both 33%), while German

speakers clearly tended to use 'ja' (57%), a pattern which is, somewhat surprisingly, reproduced by L2 learners (beginners: 62%; advanced: 66%). Only advanced learners show some instances of the more typical German 'genau', especially as a response to tag questions (8%); "genau" is mostly used for yes–no and tag replies in the target language (36% yes–no and 20% tag).

Finally, comparing the two VSU categories to the acknowledgement category, it is evident that the choice of lexical type changes across the three functions, suggesting a relationship between type and function.

5. *Conclusion*

In this contribution, we conducted an exploratory analysis of BCs and other VSUs in dyadic interactions in German L2 spoken by Italian learners at two different proficiency levels, and across Italian and German as native languages to compare the learners' output with and assess possible transfer phenomena or the acquisition of target patterns. We took into account frequency, length and lexical type of BCs and of tokens presenting the same lexical types as BCs, but with functions other than acknowledgements, i.e. positive replies to yes-no and tag questions. We paid special attention to dyad-specific variability since individuals' behaviour in a conversation depends not only on idiosyncratic factors and speakers-specific speech style, but, more importantly, on the unique mechanisms that arise from the interaction between the two specific parties in the conversation.

With regards to frequency and length, we observed that dyad-specific behaviour is similar across learners' L1 and L2. Importantly, concentrating on group-level results could have led to the misleading conclusion that target-like patterns of BC frequency and length are achieved in the L2 along with increasing proficiency. A by-dyad analysis suggests, instead, that dyad-specific patterns are more important than proficiency levels when predicting the rate and the length of BCs produced. Moreover, the by-dyad variability found within the group of native German speakers calls into question the idea of target features to be acquired by learners regarding these aspects of BCs.

Differently, for lexical type we observed preferred type-function relations which are language-specific, thus representing a target for learners. In most cases, L2 learners tend to prefer types that are shared with their L1 Italian over specifically German ones, such as 'genau'. The use of specifically German BCs is only present in advanced learners, indicating a positive effect of proficiency for this aspect of BC production.

This first exploration was based on a limited sample of speakers. Therefore, an extension of the analysis to the whole corpus will clarify whether the trends observed are robust, including statistical testing using Bayesian linear regression modelling. However, we set a groundwork from which a few suggestions for further studies can be derived. First, we found that there seem to be preferential co-occurrences among single aspects of BCs, so future studies should address the relation among them. With respect to this point, one further aspect that we did not take into

account and should be addressed in the future is the prosodic realisation of BCs and how it might interact with lexical type and function. Secondly, our preliminary results suggest that dyad-specific patterns are more important than proficiency level when predicting some BC aspects in L2. For this reason, it is important to consider learners' L1 as baseline and investigate dyad-specific behaviour to set apart individual variability from the transfer or acquisition of patterns. Finally, in line with the literature, we showed that there are aspects of BC use which are similar across-languages, but also that some language-specific aspects are not correctly reproduced in L2. Therefore, more comparative studies of varied language pairs can be useful for L2 pedagogy applications.

Acknowledgement

Financial support for carrying out the research is gratefully acknowledged from a.r.t.e.s Graduate School 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.

References

Amador-Moreno, C.P., Mccarthy, m., & O'Keeffe, A. (2013). Can English provide a framework for Spanish response tokens?. In *Yearbook of Corpus Linguistics and Pragmatics 2013*. Dordrecht: Springer. 175-201.

Anderson, A.H., Bader, M., Bard, E.G., Boyle, E.H., Doherty, G.M., Garrod, S.C., & Weinert, R. (1991). The HCRC map task corpus. In *Language and speech*, 34(4), 351-366.

Běnuš, Š, Gravano, A., & Hirschberg, J. (2007). The prosody of backchannels in American English. In J. Trouvain & W.J. Barry (Eds.), *16th International Congress of Phonetic Sciences* (pp. 1065–1068). Saarbrücken: Universität des Saarlandes.

Berry, A. (1994). Spanish and American turn-taking styles: A comparative study. In L.F. Bouton, (ed.), *Pragmatics and Language Learning Monograph Series*, Urbana: University of Illinois, 180-190.

Bögels, S., Torreira, F. (2015). Listeners use intonational phrase boundaries to project turn ends in spoken interaction. In *Journal of Phonetics*, *52*, 46-57.

Boersma, P., Weenink, D. (2022). PRAAT*:* doing phonetics by computer. [software] Version 6.2. https://www.praat.org/

Castello, E., Gesuato, S. (2019). Holding up one's end of the conversation in spoken English: Lexical backchannels in L2 examination discourse. In *International Journal of Learner Corpus Research,* 5, 231–252.

Clancy, P.M., Thompson, S., Suzuki, R. & Tao, H. (1996) The conversational use of reactive tokens in English, Japanese and Mandarin. In *Journal of Pragmatics,* 26, 355-387. https://doi.org/10.1016/0378-2166(95)00036-4

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge, U.K: Press Syndicate of the University of Cambridge.

Cutrone, P. (2005). A case study examining backchannels in conversations between Japanese–British dyads. In *Multilingua. Journal of Cross-Cultural and Interlanguage Communication*, 24(3), 237-274. https://doi.org/10.1515/mult.2005.24.3.237

Cutrone, P. (2014). A cross-cultural examination of the backchannel behavior of Japanese and Americans: Considerations for Japanese EFL learners. *Intercultural Pragmatics, 11*(1), 83-120. doi: 10.1515/ip-2014-0004

De Jong, N.H. (2016). Fluency in second language assessment. In Tsagari, D., Banerjee, J. (Eds.), *Handbook of second language assessment.* Mouton de Gruyter, 203-218.

Drummond, K., Hopper, R. (1993). Back channels revisited: Acknowledgement tokens and speakership incipiency. In *Research on Language and Social Interaction*, 26(2), 157-177.

Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. In *Language in Society,* 2, 161–180.

Duncan, S., Fiske, D.W. (1977). *Face to face interaction: research, methods and theory*. New Jersey: Lawrence Erlbaum.

Edlund, J., Heldner, M. & Pelcé, A. (2009). Prosodic features of very short utterances in dialogue. In *Nordic Prosody: Proceedings of the Xth Conference*, Helsinki, Finland, August 2008, 57–68.

Figueras, N., North, B., Takala, S., Van Avermaet, P. & Verhelst, N. (2009). *Relating language examinations to the common European framework of reference for languages: learning, teaching, assessment (CEFR): a manual.* Strasbourg, France: Council of Europe, Language policy division.

Fellegy, A.M. (1995). Patterns and functions of minimal response. In *American Speech International Journal of Educational Best Practices*, 2(1), 186-199.

Fries, C.C. (1952). *The structure of English*. London: Longmans, Green and Company.

Galaczi, E.D. (2014). Interactional Competence across Proficiency Levels: How do Learners Manage Interaction in Paired Speaking Tests?, In *Applied Linguistics*, 35(5), 553–574.

Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*. Amsterdam: J. Benjamins Publishing Company.

Goodwin, C. (1986). Audience diversity, participation and interpretation. In *Text & Talk, 6*, 283–316.

Grice, M., Savino, M. (2003). Map tasks in Italian: Asking questions about given, accessible and new information. In *Catalan journal of linguistics*, *2*, 153-180.

Ha, K.P. (2012). *Prosody in Vietnamese – Intonational Form and Function of Short Utterances in Conversation*. Asia-Pacific Linguistics 002 (SEAMLES 001). PhD thesis. Canberra: The Australian National University.

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. *In Proceedings of Speech Prosody 8*, Boston, USA, 31 May-3 Jun 2016, 801-805.

Hall, J.K. (1993). The Role of Oral Practices in the Accomplishment of Our Everyday Lives: The Sociocultural Dimension of Interaction with Implications for the Learning of Another Language. In *Applied Linguistics*, 14, 145-166.

Hall, J.K. (1995). (Re)creating Our Worlds with Words: A Sociohistorical Perspective of Face-to-face Interaction. In *Applied Linguistics*, 16, 206-232.

Hasegawa, Y. (2014). *Japanese: A linguistic introduction.* Cambridge: Cambridge University Press.

He, A.W., Young, R. (1998). Language proficiency interviews: A discourse approach. In R. Young & A.W. He (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency.* Amsterdam: John Benjamins, 1-24.

Heinz, B. (2003). Backchannel responses as strategic responses in bilingual speakers' conversations. In *Journal of Pragmatics*, 35, 1113–1142.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In Atkinson, J.M. and Heritage, J. (Eds.) *Structures of Social Interaction: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, 299-345.

Jacoby, S., Ochs, E. (1995). Co-construction: An introduction. *Research on language and social interaction*, *28*(3), 171-183.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "yeah"; and "mm hm". In *Paper in Linguistics*. 17, 197–216. https://doi.org/10.1080/08351818409389201

Kendon, A. (1967) Some functions of gaze-direction in social interaction. In *Acta Psycholoigia* 26, 22-63. https://doi.org/10.1016/0001-69186790005-4

Kjellmer, G. (2009). Where do we backchannel? On the use of mm, mhm, uh huh and such like. In *International Journal of Corpus Linguistics*, 14, 81–112.

Kraaz, M. & Bernaisch, T. (2022). Backchannels and the pragmatics of South Asian Englishes. In *World Englishes*, 41(2), 224-243.

Kraut, R.E., Lewis, S.H. & Swezey, L.W. (1982). Listener responsiveness and the coordination of conversation. *Journal of Personality and Social Psychology*, 43(4), 718–731. https://doi.org/10.1037/0022-3514.43.4.718

Lambertz, K. (2011). Back-channelling: The use of yeah and mm to portray engaged listenership. In *Griffith Working Papers in Pragmatics and Intercultural Communication*, 4, 11-18.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. In *Language Learning*, 40, 387-412.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.). *Perspectives on fluency*. Michigan: The University of Michigan Press, 25-42.

Levinson, S. (2015). Turn-taking in Human Communication – Origins and Implications for Language Processing. Trends in Cognitive Sciences. 20. 10.1016/j.tics.2015.10.010.

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

Li, H.Z., Cui, Y. & Wang, Z. (2010). Backchannel responses and enjoyment of the conversation: The more does not necessarily mean the better. In *International journal of psychological studies*, *2*(1), 25.

Maynard, S. (1997). Analyzing interactional management in native/non-native English conversation: A case of listener response. In *IRAL- International Review of Applied Linguistics in Language Teaching*, 35(1), 37-60. https://doi.org/10.1515/iral.1997.35.1.37

Mccarthy, M. (2009). Rethinking spoken fluency. In *ELIA, 9, 11-29*.

Nurjaleka, L. (2019). Backchannel Behavior in Interview Discourse: A contrastive study between Japanese and Indonesian. In *Eleventh Conference on Applied Linguistics (CONAPLIN 2018)*, Atlantis Press, 451-457.

Rühlemann, C. (2007). *Conversation in context: A corpus-driven approach*. London: Longman. International Journal of Educational Best Practices, 2, 39–53.

Sacks, H., Schegloff, E.A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735. https://doi.org/10.2307/412243

Saito, K., Ilkan, M., Magne, V., Tran, M.N. & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low-, mid- and high-level second language fluency. In *Applied Psycholinguistics*, 39(3), 593-617.

Sbranna, S., Cangemi F. & Grice, M. (2020). Quantifying L2 interactional competence. In Romito, L. (Eds.), *Language change under contact conditions: acquisitional contexts, languages, dialects and minorities in Italy and around the world*, Collana Studi AISV 7, Milano: Officinaventuno, 383-405. DOI: 10.17469/O2107AISV000018

Schegloff, E.A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.), *Analysing discourse: Text and talk* (pp. 71-93). Washington, D.C.: Georgetown University Press.

Shelley, L., Gonzalez, F. (2013). Back Channeling: Function of Back Channeling and L1 Effects on Back Channeling in L2. In *Linguistic Portfolios*, 2(1), 9.

Simon, C. (2018). The functions of active listening responses. In *Behavioural Processes,* 157, 47–53. https://doi.org/10.1016/j.beproc.2018.08.013

Stocksmeier, T., Kopp S. & Gibbon, D. (2007). Synthesis of prosodic attitudinal variants in German backchannel ja. In *Interspeech*, 1290-1293.

Tao, H., Thompson, S.A. (1991). English backchannels in Mandarin conversations: A case study of superstratum pragmatic 'interference'. In *Journal of Pragmatics*, 16(3), 209-223. https://doi.org/10.1016/0378-2166(91)90093-D

Tolins, J., Fox Tree, J.E. (2014). Addressee backchannels steer narrative development. In *Journal of Pragmatics*, 70, 152–164.

Tottie, G. (1991). Conversation style in British and American English: The case of backchannels. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics: Studies in honour of Jan Svartvik*. London: Longman, 254–271.

Ward, N., Tsukuhara, W. (2000). Prosodic features which cue back-channel responses in English and Japanese. In *Journal of Pragmatics*, 32, 1177–1207. https://doi.org/10.1016/s0378-2166(99)00109-5

Wehrle, S., Grice, M. (2019). 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.

White, S. (1989). Backchannels across cultures: A study of Americans and Japanese. In *Language in Society*, 18(1), 59-76. https://doi.org/10.2307/4168001

Wood, D.A. (2001). In Search of Fluency: What is it and How Can we Teach It? Canadian Modern Language Review-revue Canadienne Des Langues Vivantes, 57, 573-589.

Wolf, J.P. (2008) The effects of backchannels on fluency in L2 oral task production, *System*, Volume 36, Issue 2, Pages 279-294, ISSN 0346-251X

Wong, D., Peters, P. (2007). A study of backchannels in regional varieties of English, using corpus mark-up as the means of identification. In *International Journal of Corpus Linguistics*, 12, 479–509.

Yngve, V.H. (1970). On getting a word in edgewise. In *Chicago Linguistics Society*, 6, 567–578.

CRediT authorship contribution statement

Simona Sbranna
Conceptualisation
Data curation
Visualisation
Writing – original draft

Simon Wehrle
Conceptualisation
Methodology
Visualisation
Writing – review and editing

Martine Grice
Conceptualisation
Methodology
Supervision