To laugh or not to laugh? The use of laughter to mark discourse structure

Bogdan Ludusan
Phonetics Workgroup, CITEC &
Faculty of Linguistics and Literary Studies
Bielefeld University, Germany
bogdan.ludusan@uni-bielefeld.de

Barbara Schuppler
Signal Processing and
Speech Communication Laboratory
Graz University of Technology, Austria
b.schuppler@tugraz.at

Abstract

A number of cues, both linguistic and non-linguistic, have been found to mark discourse structure in conversation. This paper investigates the role of laughter, one of the most encountered non-verbal vocalizations in human communication, in the signalling of turn boundaries. We employ a corpus of informal dyadic conversations to determine the likelihood of laughter at the end of speaker turns and to establish the potential role of laughter in discourse organization. Our results show that, on average, about 10% of the turns are marked by laughter, but also that the marking is subject to individual variation, as well as effects of other factors, such as the type of relationship between speakers. More importantly, we find that turn ends are twice more likely than transition relevance places to be marked by laughter, suggesting that, indeed, laughter plays a role in marking discourse structure.

1 Introduction

Despite the spontaneous nature of human communication, turn-taking between conversational partners occurs rather smoothly (Sacks et al., 1978), with interlocutors negotiating control of the floor through the marking of so-called transition relevance places (points in the conversation where a speaker change may occur) by means of various cues. A significant amount of work has been dedicated on investigating the acoustic characteristics involved in speaker-turn marking (e.g., Wichmann and Caspers, 2001; Gravano and Hirschberg, 2009; Niebuhr et al., 2013; Zellers, 2017). Yet, discourse structure has been shown to be signalled by a combination of different features (Duncan, 1972), both linguistic (e.g., lexical, syntactic, semantic) and non-linguistic. The latter type includes body movements and gestures, such as posture shifts (Cassell et al., 2001) and gaze (Jokinen et al., 2013), but also non-verbal vocalizations, in the form of breathing sounds (Włodarczak and Heldner, 2016).

We examine here one of the most commonly encountered non-verbal vocalizations in spontaneous interaction, laughter. It plays various roles in human communication (Trouvain and Truong, 2017), including social and communicative (Glenn and Holt, 2013) as well as linguistic roles (Mazzocconi et al., 2020). Evidence from conversational analysis suggests a possible role of laughter in discourse structure, as a cue marking the edges of speaker-turns (Gavioli, 1995; Ikeda and Bysouth, 2013; Madden et al., 2002). Most of this evidence is of qualitative nature, but there are also quantitative findings that offer additional support for this hypothesis. Norris and Drummond (1998) found that about 30% of total produced laughter occurred with the beginning and end of discourse structures, in materials based on tasks eliciting laughter. In a distributional analysis of laughter in task-based dyadic interactions, Ludusan et al. (2020) reported that turns for which laughter occurred at turn-initial or turn-final represented up to 50% of all turns containing laughter, in the three studied languages (French, German and Mandarin Chinese). Turns marked by laughter at their edges made up between 13% and 20% of total turns in the same materials (Ludusan and Wagner, 2022). Also the fact that laughter entrainment effects have been found at the turn-level in conversation (Ludusan and Wagner, 2022), represents further indication of the potential role of laughter in marking turns.

The aforementioned studies, however, presented only descriptive statistics of laughter events co-occurring with turn edges, without showing a relationship between laughter and discourse structure. Thus, we aim to establish in this study the possible role of laughter in marking turn boundaries, by comparing laughter at speaker turn versus at transition relevance places and by determining whether turn-holds or turn-changes are more likely to be marked by laughter. Moreover, as some of these studies used materials from tasks that elicited

Proceedings of the SIGdial 2022 Conference, pages 76–82
Heriot-Watt University, Edinburgh, UK. 07-09, September, 2022 ©2022 Association for Computational Linguistics

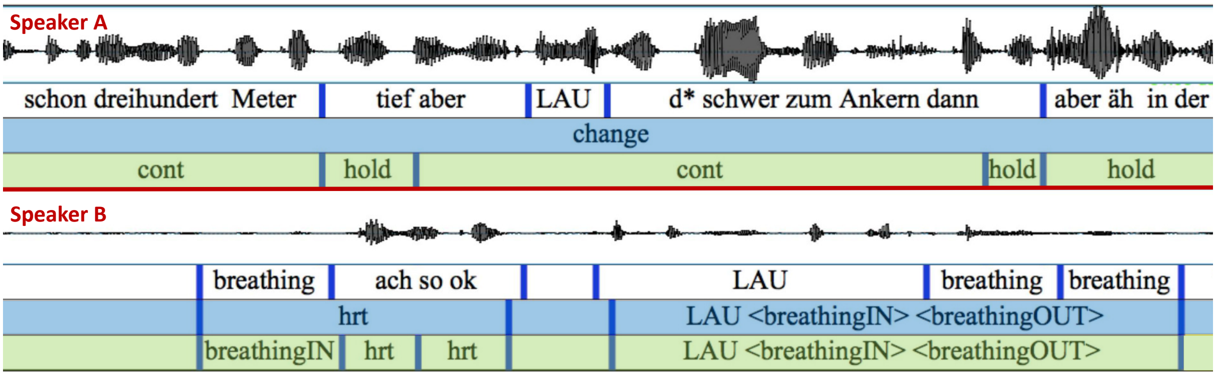

Figure 1: Conversation fragment from the GRASS corpus illustrating the discourse structure annotation. For each speaker (A and B), it shows the waveform of the recording, its orthographic transcription, the turn-level annotations (in blue), and the level of potential transition relevance places (in green). The laughter produced by the speakers is marked with *LAU*.

laughter and since laughter patterns in everyday conversations might differ from those produced in such tasks, we employ here informal conversations between friends/family members. We also evaluate the role of message-external factors, namely relation type between interlocutors and the gender composition of the dyad, as previous work has shown that they may play a role in the overall production of laughter (Smoski and Bachorowski, 2003).

2 Materials

The Graz Corpus of Read and Spontaneous Speech (GRASS) contains about 30 hours of Austrian German read and conversational speech, collected from 38 Austrian speakers (19 females, 19 males) (Schuppler et al., 2014). The conversational speech component contains speech from 19 pairs of speakers who had known each other for at least several years, and who were either friends, family-members, colleagues or couples, with a similar number of mixed-gender and same-gender dyads. These speaker pairs were recorded for one hour each, without interruption, in order to encourage a fluent, casual conversation. There were no restrictions in terms of topic or speaking behaviour, leading to the use of casual, partly dialectal pronunciation, frequent occurrence of overlapping speech, as well as laughter (laughs and speech-laughs) (Schuppler et al., 2017). This resulted in a wide variety of conversation topics, such as discussions about family or about public figures, travelling, relationship problems, or work-related issues.

The conversational speech component of GRASS is currently being manually annotated for discourse structure. As manual annotations are highly time consuming, in combination with limited resources, the manual annotation of the entire GRASS corpus is not possible. In order to capture as many different speakers and as many different communicative stages as possible, from each one-hour conversation, 5 minutes were annotated either from its beginning, its middle, or its end. So far, 14 dyads (5 f-m, 4 f-f, 5 m-m) were annotated, resulting in a total of 70 minutes of recordings available for this study.

Two independent discourse structure levels were annotated (cf., conversation example shown in Figure 1): one for turn management (based on inter-pausal units), further called *turn-level* (the blue tier in Figure 1), and one for potential transition relevance places (further called *TRP-level*), which were defined in terms of points of potential syntactic completion (the green tier in Figure 1). The turn-level labels were based on the four categories proposed in Zellers (2017): *hold* (the same speaker continues talking), *change* (a new speaker takes the floor), *question* (the speaker transfers the turn to another speaker), and *Hearer Response Tokens* (HRT, backchannel-like tokens, Sikveland, 2012). Three additional turn labels captured incomplete structures before pauses: *incomplete-hold* (the speaker makes a pause at a point of "maximum grammatical control", Schegloff, 1998: 241, and then continues speaking), *trail-off* (a syntactically incomplete speaker change, cf. Walker, 2012), and *self-interruption* (in the case of turn competition, one speaker interrupts themselves to cede the turn to the other speaker). The annotation at the TRP-level is more fine-grained, having the categories proposed by Zellers (2017) and six additional labels. For further details on the different labels used for annotating the TRP-level, we refer the reader to Schuppler and Kelterer (2021). All annotations were created while listening to the recordings and were not based on the orthographic transcription alone. Thus, for example, the token "ja" (yes) may be assigned the label *HRT* in one instance, where it was produced with the function of a backchannel (i.e., no interruption of the turn of the interlocutor), or the label *change* in another instance, where it was produced with a question-like intonation followed by a turn of the interlocutor.

In order to guarantee a high annotation quality, the same process was applied to both discourse structure levels: First, the conversations were annotated by one trained annotator, self-corrected at a later point in time and then corrected by another, second annotator. In order to estimate the inter-rater agreement for the two discourse structure annotation levels, we evaluated a set of 878 word tokens from 3 different conversations. The Cohen's kappa on whether a TRP was placed at a word boundary or not was $\kappa = 0.96$. The agreement between the two turn-level labels *change* and *hold* (the only two categories we discriminated between in this study) was $\kappa = 0.83$. Thus, both levels of discourse annotations used for this study showed a very high inter-annotator agreement.

3 Methods

Based on the annotations of GRASS, we determined the units (both at the turn- and at the TRP-level) which were marked by laughter at their end. For this, the speaker having the floor or their interlocutor should have produced laughter either at the end of the unit, overlapping with the end of the unit, or immediately following (within one second) the unit. If the interlocutor produced the laughter, they should not have produced any other speech between the end of the unit marked by laughter and the start of the laughter instance. For the labelling process, other non-verbal vocalizations, such as in- or out-breaths and coughs, were not considered as being speech. All units were labelled for the existence of laughter in the analysis, except for the HRT tokens, which do not represent an actual conversational turns. Although not included in the analysis, HRT were taken into account for the labelling of turn-units: If a speaker turn-end overlapped or was followed by an HRT of the conversational partner containing laughter, the turn was labelled as being marked with laughter. Statistics about the total number of units in our data, the ones analysed here (non-HRT) and the units marked with laughter can be found in Table 1.

Level	Total	Analysed	Laughter
Turn	1874	1313	125
TRP	3772	3071	64

Table 1: The number of units considered in this study. For each analysis level (turn/TRP), the total number of units, the number of analysed units (non-HRT), and the units marked by laughter are shown.

We then counted, for each speaker and each level, the number of units signalled by laughter and the number of units not signalled by laughter. These counts, representing together the odds of units having laughter (number of successes and failures), were used as dependent variable in a mixed effects logistic model, to determine whether a significant difference exists between the marking of two levels. The unit-level (turn/TRP) was employed as predictor in the model and the speaker was introduced as a random intercept. Three logistic models were then fitted on the data consisting of the turn-level counts, in order to determine the effect of several message-external factors on the signalling of turns with laughter. We considered the dyad identity (ranging between 1 and 14), its gender composition (f-f, f-m or m-m) or the relation between the conversational partners (colleagues, couples, family or friends), as the independent variables in those models. Finally, we checked whether turn-marking with laughter occurs more often for turn-change or for turn-hold. For this, we deemed all turns labelled as incomplete-hold and hold to represent a turn-hold and the remaining labels to represent a turn-change. We then tested the probability of having a turn-change marked by laughter, out of the total number of turns marked by laughter, by means of a binomial test. The R (R Core Team, 2019) software was used for all statistical analyses, with the mixed effects model being fitted by means of the lmerTest package (Kuznetsova et al., 2017), based on the lme4 package (Bates et al., 2015) functionalities.

4 Results

First, we examined the likelihood of laughter in marking turns. Figure 2 illustrates the proportion of speaker turns followed by laughter, out of the total number of turns produced by each speaker. Speakers were grouped based on the dyad they were part of and each speaker is represented by a point. On average, across dyads, 10.6% of all turns are marked by laughter (represented by a solid horizontal line), but there is significant variation across speakers (from a minimum of 0% for speaker B in dyad 3 to a maximum of 43.8% for speaker A in dyad 11). We checked whether the marking of turns by the various dyads differs significantly from mean value, by means of a logistic regression model with the dyad ID as predictor and employing a sum to zero contrast. Only three dyads (3, 11 and 13) showed significant differences from the overall mean.

Then, with regards to the effect of message-external factors on the laughter-marking of turns, we examined the purposely built logistic models, having either the relation between speakers or the gender composition of the dyads as independent variable. Logistic models estimate the effect of the predictors on the log odds ratio of success vs. failure (here, the probability of a turn to be marked vs. not be marked by laughter). Higher odds indicate a higher probability of turns being marked by laughter. For the relation status, the highest odds were seen for the dyads made up of couples (the intercept of the model, $\beta = -1.998$), followed by family ($\beta = -0.301$, $p = .270$), friends ($\beta = -0.460$, $p = .051$), and the lowest odds for colleagues ($\beta = -0.725$, $p = 0.008$). Regarding the gender composition, the highest odds were observed for the female-female dyads (intercept, $\beta = -2.187$), with similar odds for mixed gender dyads ($\beta = -0.025$, $p = .913$) and lower odds for all-male dyads ($\beta = -0.551$, $p = .028$). The difference between mixed-gender and all-male dyads was also found significant ($p = .018$).

Next, we estimated whether there is an effect of the discourse level where laughter is used for marking the structure (turn/TRP). Employing the mixed effects model described in the Methods section, we obtained a significant effect of the level ($p = 1.3e^{-6}$), with the odds of a laughter-marked structure increasing by 107% (95% confidence interval: [0.54, 1.78]) at the turn-level compared to at the TRP-level. While the intercept of the model showed that the probability of a TRP to be signalled by laughter is about 4%, it increases more than twice in the case of turn boundaries.

Finaly, we looked in more detail at which types

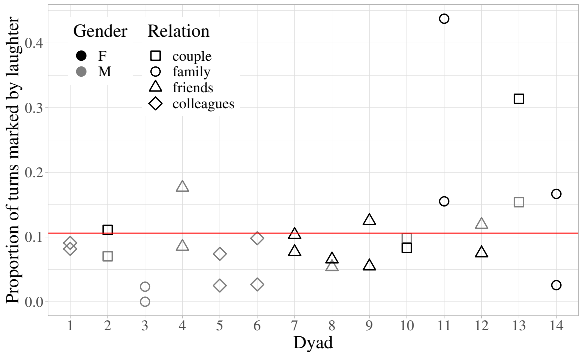

Figure 2: The proportion of turns marked by laughter, out of the total number of turns produced by each speaker. The results are illustrated on a per-dyad basis, with each dyad being represented by two data points, one for each dyad member. Each speaker is coded by a colour, representing their gender, and a shape, encoding their relation with the interlocutor. The horizontal line represents the average proportion across all dyads.

of turns were more likely marked by laughter. The conducted binomial test showed a significant preference for turn-changes ($p = .007$), with a probability of 0.62 (95% confidence interval [0.53, 0.71]).[1]

5 Discussion

Based on our data from casual conversations between family members or friends, we have found that turn boundaries tend to be signalled by laughter, on average in 10% of the cases. This represents a lower value than those reported by Ludusan and Wagner (2022), in which between 13% and 20% of turns were marked by laughter, across the three studied languages. Moreover, in the latter case, backchannels were counted as turns, thus a higher proportion of turns might be marked by laughter if one were not to consider backchannels, as in our case. These differences may well reflect the different data elicitation methods. The data employed by Ludusan and Wagner (2022) consisted of recordings in which a significant amount of laughter was expected, due to the nature of the considered task (coming up with an idea for a film script based on an embarrassing moment). This emphasizes the role of the type of data employed in the investigation: In a context consisting of casual conversations between individuals that are close to each other, a lower proportions of turns are signalled by laughter. The observed laughter-marking behaviour seems to

[1] The fact that turns signalled by partner laughter were included in the analysis did not bias these results, as there was a higher proportion of turn-holds (0.40) than turn-changes (0.18) marked by partner laughter, in our data.

be consistent in our data, with 11 out of the total 14 dyads showing no significant difference from the mean.

For both investigated message-external factors, the relation type between the conversational partners and the gender mix of the dyad, we observed significant effects on the laughter-marking of turns. Couples exhibited higher odds of turns marked with laughter than family members, friends and colleagues, although only the difference between couples and colleagues was found to be significant. Previous work looking at the effect of interlocutors' relation on laughter production (e.g. Smoski and Bachorowski, 2003; Jansen et al., 2021) considered two cases: familiar/unfamiliar, and the results were mixed, either showing a significant effect (Smoski and Bachorowski, 2003), or a lack of it (Jansen et al., 2021). Looking at the marking of turn edges by laughter, Ludusan and Wagner (2022) found no effect of familiarity (defined as the number of years the speaker knew each other). However, we employed here a definition based on the relationship between speakers, which may be more appropriate. With respect to the gender mix, we saw no difference between all-female and mixed-gender dyads, but significantly lower odds for all-male dyads compared to the other two groups. Our results partially align with work reporting more laughter in mixed-gender dyads composed of friends (Smoski and Bachorowski, 2003) (although a different behaviour may be seen for mixed-gender dyads composed of strangers Grammer and Eibl-Eibesfeldt, 1990; Smoski and Bachorowski, 2003). The observed differences may stem from the types of laughter considered in each study (laughter at turn boundaries here, all laughter instances in previous studies).

How does the marking of turns by means of laughter compare to the signalling of turns by other cues? Niebuhr et al. (2013) observed differences in speech reduction phenomena between turn-final and turn-internal positions of up to more than four times, while Cassell et al. (2001) found that posture shifts at turn boundaries were five times more likely than turn-internal. We have seen here that laughter turns are twice more likely to be signalled by laughter, than transition relevance places. While laughter may seem, therefore, a weaker cue to the marking of turns, one must take into account that we compared here turn-final laughter with laughter produced only at TRPs (not any turn-internal position). When comparing turn-final with phrase-final positions, also Niebuhr et al. (2013) showed that the difference in likelihood between these two levels is lower than between turn-final and any turn-internal location.

Among the types of considered turn-units, we observed a higher probability of turn-changes than turn-holds being marked with laughter. This finding indicates that laughter is one of the cues that speakers employ to signal the end of their turn or the taking of the floor from their interlocutor. While the current study did not examine the characteristics of the various turn-final laughter instances, it might be that giving/taking the turn may use different types of laughter (laughs vs. speech-laughs, snorts vs. grunts, etc) or laughs with different acoustic properties (voiced vs. unvoiced, etc.). Further investigations in this direction would be necessary to better understand the role of laughter in turn-taking. Moreover, studies on larger datasets as well as on other languages are welcome, in order to test the generalizability of these findings.

6 Conclusions

We investigated the role of laughter in the marking of speaker turns in a corpus of informal conversations between family members and friends. Besides establishing the frequency of occurrence of laughter at turn-ends, in a dataset not composed of task-based interactions, we also showed that laughter is twice more likely to occur at the end of turn-units than at TRPs. Next, we found that the probability of laughter-marked turn-changes was higher than for turn-holds, suggesting a possible role of laughter as a cue signalling turn-change. Finally, our study revealed that this laughter function is modulated by message-external factors, such as the nature of the relationship between speakers and the dyad gender composition. These results represent one step further in understanding the various functions that non-verbal phenomena and laughter, in particular, play in human communication.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 461442180 to BL. BS was supported by grant P-32700-N from the Austrian Science Fund (FWF). We would like to thank Anneliese Kelterer for her support with the discourse structure annotations of GRASS, as well as for the computation of the inter-rater agreement.

References

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Justine Cassell, Yukiko Nakano, Timothy Bickmore, Candace Sidner, and Charles Rich. 2001. Non-verbal cues for discourse structure. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, pages 114–123.

Starkey Duncan. 1972. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292.

Laura Gavioli. 1995. Turn-initial versus turn-final laughter: Two techniques for initiating remedy in English/Italian bookshop service encounters. *Discourse Processes*, 19(3):369–384.

Phillip Glenn and Elizabeth Holt. 2013. *Studies of laughter in interaction*. London: Bloomsbury Academic.

Karl Grammer and Irenaus Eibl-Eibesfeldt. 1990. The ritualisation of laughter. In Walter A. Koch, editor, *Natürlichkeit der Sprache und der Kultur*, pages 192–214. Brockmeyer Bochum, Germany.

Agustín Gravano and Julia Hirschberg. 2009. Turn-yielding cues in task-oriented dialogue. In *Proceedings of the SIGDIAL 2009 Conference*, pages 253–261.

Keiko Ikeda and Don Bysouth. 2013. Laughter and turn-taking: Warranting next speakership in multiparty interactions. In Phillip Glenn and Elizabeth Holt, editors, *Studies of laughter in interaction*, pages 39–64. London: Bloomsbury Academic.

Michel-Pierre Jansen, Khiet Truong, and Dirk Heylen. 2021. How familiarity influences the frequency, temporal dynamics and acoustics of laughter. In *Proceedings of the 9th International Conference on Affective Computing and Intelligent Interaction*, pages 1–8.

Kristiina Jokinen, Hirohisa Furukawa, Masafumi Nishida, and Seiichi Yamamoto. 2013. Gaze and turn-taking behavior in casual conversational interactions. *ACM Transactions on Interactive Intelligent Systems*, 3(2):1–30.

Alexandra Kuznetsova, Per Brockhoff, and Rune Christensen. 2017. lmertest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26.

Bogdan Ludusan and Petra Wagner. 2022. Laughter entrainment in dyadic interactions: Temporal distribution and form. *Speech Communication*, 136:42–52.

Bogdan Ludusan, Maik Wesemann, and Petra Wagner. 2020. A distributional analysis of laughter across turns and utterances. In *Proceedings of the Workshop on Laughter and Other Non-verbal Vocalisations*, pages 28–31.

Mary Madden, Mary Oelschlaeger, and Jack Damico. 2002. The conversational value of laughter for a person with aphasia. *Aphasiology*, 16(12):1199–1212.

Chiara Mazzocconi, Ye Tian, and Jonathan Ginzburg. 2020. What's your laughter doing there? A taxonomy of the pragmatic functions of laughter. *IEEE Transactions on Affective Computing*, pages 1–19.

Oliver Niebuhr, Karin Görs, and Evelin Graupe. 2013. Speech reduction, intensity, and F0 shape are cues to turn-taking. In *Proceedings of the SIGDIAL 2013 Conference*, pages 261–269.

Melissa Norris and Sakina Drummond. 1998. Communicative functions of laughter in aphasia. *Journal of Neurolinguistics*, 11(4):391–402.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Harvey Sacks, Emanuel Schegloff, and Gail Jefferson. 1978. A simplest systematics for the organization of turn taking for conversation. In *Studies in the Organization of Conversational Interaction*, pages 7–55. Elsevier.

Emanuel Schegloff. 1998. Reflections on studying prosody in talk-in-interaction. *Language and Speech*, 41(3-4):235–263.

Barbara Schuppler, Martin Hagmüller, Juan Andres Morales-Cordovilla, and Hannes Pessentheiner. 2014. GRASS: the Graz corpus of Read And Spontaneous Speech. In *Proceedings of LREC*, pages 1465–1470.

Barbara Schuppler, Martin Hagmüller, and Alexander Zahrer. 2017. A corpus of read and conversational Austrian German. *Speech Communication*, 94:62–74.

Barbara Schuppler and Anneliese Kelterer. 2021. Developing an annotation system for communicative functions for a cross-layer ASR system. In *Proceedings of the Integrating Perspectives on Discourse Annotation Workshop*, pages 1–5.

Rein Ove Sikveland. 2012. Negotiating towards a next turn: Phonetic resources for 'doing the same'. *Language and Speech*, 55(1):77–98.

Moria Smoski and Jo-Anne Bachorowski. 2003. Antiphonal laughter between friends and strangers. *Cognition and Emotion*, 17(2):327–340.

Jürgen Trouvain and Khiet Truong. 2017. Laughter. In *The Routledge Handbook of Language and Humor*, pages 340–355. Routledge.

Gareth Walker. 2012. Coordination and interpretation of vocal and visible resources: 'trail-off' conjunctions. *Language and Speech*, 55(1):141–163.

Anne Wichmann and Johanneke Caspers. 2001. Melodic cues to turn-taking in English: evidence from perception. In *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*, pages 1–6.

Marcin Włodarczak and Mattias Heldner. 2016. Respiratory turn-taking cues. In *Proceedings of INTERSPEECH 2016*, pages 1275–1279.

Margaret Zellers. 2017. Prosodic variation and segmental reduction and their roles in cuing turn transition in Swedish. *Language and Speech*, 60(3):454–478.