9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Conversational quality is affected by and reflected in prosodic entrainment

Jan Michalsky[1], *Heike Schoormann*[2], *Oliver Niebuhr*[3]

[1,2]Institute for German Studies, University of Oldenburg, Germany
[3]SDU Electrical Engineering, Mads Clausen Institute, University of Southern Denmark
j.michalsky@uol.de, heike.schoormann@uol.de, olni@sdu.dk

Abstract

Prosodic entrainment is connected to various forms of communicative success. One possibility to assess successful communication in non-task-oriented everyday conversations is through the participants' perception of conversational quality. In this study we investigate whether a speaker's degree of prosodic entrainment reflects the perceived conversational quality in dating conversations. Furthermore, we ask whether prosodic entrainment can influence the perceived conversational quality in return. Based on 98 spontaneous mixed-sex dating conversations we find that conversational quality has a significant effect on a speaker's degree of pitch level entrainment. Furthermore, pitch entrainment also has a significant effect on how one's interlocutor perceives the conversational quality. However, we find differences between the two effects which suggests that what a speaker does in reaction to increased perceived conversational quality is not necessarily what an interlocutor perceives as increased conversational quality. Accordingly, while we find a bidirectional influence of prosodic entrainment and perceived conversational quality, this connection is not reciprocal in nature.

Index Terms: prosodic entrainment, conversational quality, social prosody, speaker sex differences, dating conversations

1. Introduction

When it comes to the connection of prosodic entrainment, i.e. speakers matching their prosodic features during a conversation [1, 2], and the social relationship within a conversation, we can distinguish two different theories. According to the *communication model* by [3] as well as the *perception behavior link* [4], entrainment of any form constitutes an unconscious and largely automatic adaptation within human communication. The potential function of such an automatism may be to facilitate intelligibility between interlocutors by bridging phonetic differences in speech behavior. In contrast, approaches such as the *Communication Accommodation Theory* [5] postulate that entrainment reflects the social relationship of interlocutors with an increase in linguistic distance resembling an increase in social distance. Several studies found entrainment in communication situations which were characterized by closer relationships between the interlocutors and/or a higher degree of communicative success. For example, entrainment showed a correlation with task success in learning tasks with student engineering groups [6], in map task experiments [7], and in student tutoring programs [8]. According to the *theory of alignment* [9], entrainment is crucial to communicative success. Furthermore, [10] found prosodic entrainment to be connected to collaboration and rapport in learning tasks. This connection between task success and entrainment could be explained by a higher demand in a common situational model and coordination of behaviour which may be facilitated through more similar speaking styles [11, 12].

However, not every type of conversation is inherently task-oriented. Although [13] found that speakers entrain syntactical structures more in task oriented dialogues than in spontaneous speech, entrainment occurs in spontaneous speech as well. One general goal of everyday non-task-oriented conversations for example is to establish or further a social relationship. In this context, conversations that are overall perceived as pleasant by both interlocutors may be regarded as successful. This can be transferred to the communicative setting of dating conversations investigated in this study. Although the long term goal of dating is often to find a suitable partner, the more immediate goal of flirting/dating conversations is to establish a social relationship. Accordingly, we are interested in whether conversational quality i.e. perceiving a conversation as more or less positive, good, or pleasant, affects prosodic entrainment. Furthermore, we investigate a so far largely disregarded component of prosodic entrainment by asking whether prosodic entrainment also affects conversational quality in return.

Most of the studies on task success can be linked to conversational quality. [10] for instance measured task success in relation to the collaboration between interlocutors which was then related to prosodic entrainment. Other studies specifically investigated conversational quality in connection to entrainment. [14] found prosodic entrainment to be connected to overall dialogue quality. [15] found that entrainment is a predictor for the probability of initiating romantic relationships as well as the stability of existing relationships. Moreover, prosodic entrainment was greater in marriage counselling dialogues when the interlocutors talked about positive topics [16]. Furthermore, [17] found prosodic entrainment to be connected to variables of social relationships such as giving encouragement or trying to be liked as well as resulting in smoother conversations with respect to turn latencies and decreased number of interruptions also found by [18]. Lastly, [19-22] found perceived attractiveness and general likability to be connected to prosodic entrainment in dating conversations.

In the previous studies relating communicative success and prosodic entrainment, communicative success has been assessed through post-hoc ratings of external listeners or related to an evaluation of the interlocutor, e.g. regarding his/her likability. For both approaches, conclusions may be drawn regarding conversational quality. However, prosodic entrainment has rarely been linked to conversational quality as judged by the immediate participants of a conversation.

We arrive at the following research questions. 1) Does perceived conversational quality affect a speaker's degree of prosodic entrainment? 2) Does the degree of prosodic entrainment affect an interlocutor's perception of conversational quality?

10.21437/SpeechProsody.2018-79

2. Method

2.1. Subjects

10 female and 10 male paid volunteers from the University of Oldenburg participated in the study. All subjects were aged between 19 and 28 years, monolingual speakers of High German and grew up in Lower Saxony. Only heterosexual singles were included in this study. The subjects were unacquainted and did not interact prior to the experiment.

2.2. Procedure

All subjects were informed about the dating setting prior to the experiment. Each participant was paired with each participant of the opposite sex for a total of 100 opposite-sex pairs. The subjects were seated in a quiet room and participated in short spontaneous conversations of 15 to 20 minutes each without topic restrictions. All participants judged the visual attractiveness as well as the general likability of their interlocutor immediately before and after each conversation. Following each conversation, subjects were also asked to give their general impression of the conversation on the whole, i.e. to rate the conversation on a 10-point Likert scale. These ratings were not revealed to the respective interlocutors. Recordings were made in stereo using a portable digital recorder (Tascam HD P2) at a sampling rate of 48 kHz and 24-bit resolution with head-mounted microphones (DPA 4065 FR).

2.3. Acoustic analysis

The acoustic analysis was carried out with *Praat* [23]. The speech parts of all speakers were segmented into interpausal units (IPU) (cf. [2]). Pauses were annotated manually but defined mechanically by an interruption of speech of at least 500 ms making no difference between actual pauses and hesitation pauses in favor of annotator consistency. The corpus consisted of a total of 14.687 IPUs extracted from 98 conversations since two conversations had to be excluded for technical reasons. We calculated the F0 mean for the interpausal units in immediate adjacency to a turn break inducing speaker change for both interlocutors. The measurements were converted to semitones with a reference of 50 Hz. The F0 mean values of all speakers were normalized around the group mean per speaker.

2.4. Statistical analysis

We conducted linear mixed effects models using *R* [24], the *lme4*-package [25], and the *lmerTest*-package [26]. Model fit was determined by maximum likelihood ratio tests. *P*-values were calculated using the Satterthwaite approximation. We calculated four different models separated first by whether the effects of entrainment on conversational quality or the effects of conversational quality on entrainment were tested and second by speaker sex. For the effects of conversational quality on entrainment, we used F0 mean of the speaker who took the turn (*taker F0*) as the dependent variable. The F0 mean of the speaker who passed the turn (GIVER F0) as well as the CONVERSATIONAL QUALITY perceived by the speaker taking the turn, were included as fixed factors. We added *speaker* to the model as a random factor. Furthermore, we included perceived VISUAL ATTRACTIVENESS and LIKABILITY as additional fixed factors. The effects of those factors are not reported in this paper (see [19-22]) but their effects are included in the model to exclude their influence on conversational quality from the analysis. For the effects of entrainment on conversational quality, we used *conversational quality* as the dependent variable, the GIVER F0 and the TAKER F0 as the fixed factors, and *speaker* as a random factor. Two-way interactions between fixed factors were included for all four models.

3. Results

Table 1 and 2 show the statistical results for the effects of conversational quality on the degree of entrainment from the perspective of the speaker taking the turn. Both sexes show significant effects for the main factors as well as the interaction but with differing coefficients. We find that female speakers tend to converge, i.e. approach the F0 mean of their interlocutor in absolute terms. However, this effect is dependent on the conversational quality. Female speakers show stronger convergence in conversations that were evaluated positively than in conversations that were judged as average. Furthermore, the effect nearly diminishes when conversational quality is very low. We find a general effect of synchrony for the male speakers, meaning that male speakers raised their F0 mean when taking the turn if the female interlocutor's F0 mean in the previous turn was already raised and vice versa. The interaction effects, however, suggest that this effect becomes slightly smaller if conversational quality is increased.

Table 1: *Effects of conversational quality on prosodic entrainment for the female speakers.*

	b	*SE*	*df*	*t*	*p*
GIVER F0	0.55	0.08	7266	6.52	<.001
CONVERSATIONAL QUALITY	-2.94	0.88	7270	-3.32	<.001
INTERACTION	0.04	0.01	7270	3.26	<.01

Table 2: *Effects of conversational quality on prosodic entrainment for the male speakers.*

	b	*SE*	*df*	*t*	*p*
GIVER F0	0.12	0.05	7369	2.54	<.05
CONVERSATIONAL QUALITY	6.47	1.08	7370	5.96	<.001
INTERACTION	-0.07	0.01	7370	-5.85	<.001

Table 3 and 4 show the statistical results for the effects of the F0 mean of the turn passing and the turn taking speaker as well as their interaction. Both tables thus give the effects of prosodic entrainment on conversational quality as perceived by the turn speaker passing the turn. For the female speakers, we find that conversations in which the male interlocutor takes over the turn with a higher F0 mean are judged as more positive in general which results in a positive main effect for TAKER F0. Furthermore, the interaction can be interpreted as a preference for convergence. Conversations in which the male speaker approached the female speaker's F0 mean in absolute terms after a turn transition are judged as more positive. These effects are intertwined since raising the F0 mean of a male speaker often leads to a decreased difference between the two, unless the female speaker also raised her F0 mean to a similar amount at turn transition. For the male speakers, we find the same effects but with an inherent contradiction. Male speakers also rated conversations higher in which their interlocutor raised her F0 mean. However, this did not result in a significant main effect. The interaction shows that conversations in which the female speaker took the turn with a lowered F0 mean were likewise evaluated more positively, if this resulted in a

noticeable decrease between the differences in F0 mean, i.e. pitch convergence. While convergence was an additional effect observed for the female speakers, it is an incompatible effect found for the male speakers.

Table 3: *Effects of prosodic entrainment on conversational quality for the female speakers.*

	b	*SE*	*df*	*t*	*p*
GIVER F0	9.14e-03	6.98e-03	7280	1.31	n.s.
TAKER F0	3.43e-02	6.51e-03	7278	5.27	<.001
INTERACTION	-8.65e-03	1.38e-03	7279	-6.27	<.001

Table 4: *Effects of prosodic entrainment on conversational quality for the male speakers.*

	b	*SE*	*df*	*t*	*p*
GIVER F0	3.00e-02	6.10e-03	7303	4.92	<.001
TAKER F0	1.07e-02	5.71e-03	7299	1.87	n.s.
INTERACTION	-5.46e-03	1.51e-03	7298	-3.62	<.001

4. Discussion

We investigated whether the degree of prosodic entrainment of participants in dating conversations depends on the subjectively judged quality of the conversation and whether prosodic entrainment in return affects the conversational quality as perceived by an interlocutor. We find that subjects showed significant effects of conversational quality on the degree of prosodic entrainment. In return, the degree of prosodic entrainment significantly affected the perception of conversational quality. However, the reciprocity of this relationship has to be interpreted with caution as the way entrainment reflects and affects conversational quality differs both in itself as well as across speaker sexes.

Regarding the influence of conversational quality on prosodic entrainment, we found a general effect of convergence (cf. [1, 2]) for the female subjects. Female speakers lowered their F0 mean, thereby approaching the generally lower F0 mean of the male interlocutor in absolute terms. This effect is connected to the relative movement of the male interlocutor's mean. Female speakers decreased their F0 mean to a larger extent when taking over the turn from a male interlocutor who already showed a lower F0 mean himself. We suggest that this additional lowering of the F0 mean in females serves to mitigate the difference and thus cause convergence instead of synchrony (cf. [1, 2]). However, the observed convergence is highly dependent on the speaker's perception of the quality of the conversation. The degree to which the female speakers lowered their F0 mean increased substantially with a higher degree of perceived conversational quality. Consequently, the female speakers showed less convergence with a lower degree of perceived conversational quality up to the point of showing close to no variation in their F0 mean in conversations to which they assigned poor quality.

For the male subjects we found a general effect of synchrony (cf. [1, 2]), i.e. male speakers raised their F0 mean when taking the turn from a female interlocutor who showed a raised F0 mean and vice versa. However, the interaction showed that this effect is smaller if the degree of perceived conversational quality increases, meaning that male speakers entrained slightly less, if they perceived a conversation as positive.

For the effects of prosodic entrainment on conversational quality, we found two seemingly connected effects for the female speakers. Female speakers showed a tendency to judge conversations more positively, if the male interlocutor took the turn with a relatively higher F0 mean. This effect interacted with the female speakers' own F0 mean in a way that conversations were judged as significantly better if the raised male interlocutor's F0 mean also resulted in a decreased absolute distance to the F0 mean of the female speaker, i.e. convergence. In cases where the male interlocutors took over the turn with a mean raised by no more than the previous female speaker, this had no significant effect on the female speaker's' impression of conversational quality. However, a lowering of the F0 mean resulted in a categorical decrease of perceived conversational quality. Overall, female speakers preferred F0 means that were high in relative terms and close to their own in absolute terms.

We find the same effects for the male speakers, where they are contradicting each other. Firstly, we find the same tendency to judge conversations in which the female interlocutor takes over the turn with a higher F0 mean more positively for the male speakers. Accordingly, conversations in which the female interlocutors systematically lowered their F0 mean were generally perceived negatively. However, the significant interaction effect shows that this is not always the case. If the male speaker passed the turn with a raised F0 mean and the female interlocutor lowered her F0 mean when taking the turn, the resulting convergence increased the perceived conversational quality. Overall, male speakers also preferred F0 means that are high in relative terms and close in absolute terms. As these two effects contradict each other, we expect a trade-off between the two effects in natural conversations.

The four different effects combined draw an odd picture. Both male speakers and female speakers perceived conversations in which the interlocutor's F0 mean was higher in relative terms or closer to their own as more positive. At the same time, female speakers who perceived a conversation as more positive were found to lower their F0 mean. This F0 alteration only had a positive effect on conversational quality, if their male interlocutors' F0 mean had already been raised itself. In general, male speakers showed a high degree of synchrony. Accordingly, if a female interlocutor responded to a good conversation by lowering her mean, this affected the male speaker's perception of conversational quality in a predominantly negative way. Additionally, it led to synchronous entrainment by the male speaker which in turn negatively affected the female interlocutor's perceived conversational quality.

5. Conclusion

While the effects of perceived conversational quality on prosodic entrainment are hard to interpret, the results of prosodic entrainment on the conversational quality are in line with the *Conversation Accommodation Theory* [5] and previous studies on task success [6-8] and conversational quality [14-22] and consistent, although varying in their implications for speaker sex. We suggest that the relationship between prosodic entrainment and conversational quality is not reciprocal. Entrainment as a consequence of perceiving a conversation as positive, does not necessarily lead to an increase in conversational quality through entrainment. Further research is needed to show whether listeners cannot perceive pitch convergence as a cue to a good conversation, whether they do not use it, or whether a complex interaction of different strategies is the cause of the effects observed.

6. References

[1] J. Edlund, M. Heldner and J. Hirschberg, "Pause and gap length in face-to-face interaction," *Proceedings of INTERSPEECH 2009*, 2009.

[2] R. Levitan. *Acoustic-prosodic entrainment in human-human and human-computer dialogue*. Columbia University. PhD thesis, 2014.

[3] M. Natale, "Convergence of mean vocal intensity in dyadic communication as a function of social desirability," *Journal of Personality and Social Psychology*, vol. 32, no. 5, pp. 790–804, 1975.

[4] T. L. Chartrand and J. A. Bargh, "The chameleon effect: The perception-behavior link and social interaction," *Journal of Personality and Social Psychology*, vol. 76, no. 6, pp. 893–910, 1999.

[5] H. Giles, N. Coupland and J. Coupland, "Accommodation theory: Communication, context, and consequence. Contexts of accomodation," *Developments in applied sociolinguistics*, vol. 1, 1991.

[6] H. Friedberg, D. Litman and S. Paletz, "Lexical entrainment and success in student engineering groups," *Spoken Language Technology Workshop (SLT), 2012 IEEE*, pp. 404–409, 2012.

[7] D. Reitter and J. D. Moore, "Predicting success in dialogue," *Annual Meeting - Association for Computational Linguistics*, vol. 45, pp. 808, 2007.

[8] J. Thomason, H. V. Nguyen and D. Litman, "Prosodic entrainment and tutoring dialogue success," *Artificial Intelligence in Education*, pp. 750–753, 2013.

[9] M. J. Pickering and S. Garrod, "Alignment as the basis for successful communication," *Research on Language and Computation*, vol. 4, pp. 203–228, 2006.

[10] N. Lubold and H. Pon-Barry, "Acoustic-Prosodic Entrainment and Rapport in Collaborative Learning Dialogues," *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge, November 12-12, 2014, Istanbul, Turkey*, 2014.

[11] J. G. Taylor, "Cognitive computation," *Cognitive Computation*, vol. 1, pp. 4–16, 2009.

[12] S. Beňuš, "Social aspects of entrainment in spoken interaction," *Cognitive Computation*, vol. 6, no. 4, pp. 802–813, 2014.

[13] D. Reitter, F. Keller and J. D. Moore, "Computational modelling of structural priming in dialogue," *Proceedings of HLT/NAACL*, pp. 121–123, 2006.

[14] A. L. Gonzales, J. T. Hancock and J. W. Pennebaker, "Language style matching as a predictor of social dynamics in small groups," *Communication Research*, 2009.

[15] M. E. Ireland, R. B. Slatcher, P. W. Eastwick, L. E. Scissors, E. J. Finkel and J. W. Pennebaker, "Language style matching predicts relationship initiation and stability," *Psychological Science*, vol. 22, pp. 39–44, 2011.

[16] C. C. Lee, M. P. Black, A. Katsamanis, A. C. Lammert, B. R. Baucom, A. Christensen, P. G. Georgiou and S. S. Narayanan, "Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples," *Proceedings of Interspeech*, pp. 793–796, 2010.

[17] R. Levitan, A. Gravano, L. Willson, S. Beňuš, J. Hirschberg and A. Nenkova, "Acoustic-prosodic entrainment and social behavior," *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 11–19, 2012.

[18] A. Nenkova, A. Gravano and J. Hirschberg, "High frequency word entrainment in spoken dialogue," *Proceedings of the 46th Annual Meeting of the Association of Computational Linguistics on Human Language Technologies: Short Papers*, pp. 169–172, 2008.

[19] J. Michalsky, "Pitch synchrony as an effect of perceived attractiveness and likability," *Proceedings of DAGA 2017*, 2017.

[20] J. Michalsky and H. Schoormann, "Effects of perceived attractiveness and likability on global aspects of fundamental frequency," *Proceedings of P&P12*, pp. 120–124, 2016.

[21] J. Michalsky and H. Schoormann, "Pitch convergence as an effect of perceived attractiveness and likability," *Proceedings of INTERSPEECH 2017*, pp. 2253–2256, 2017.

[22] J. Michalsky, H. Schoormann and O. Niebuhr, "Turn transitions as salient places for social signals – Local prosodic entrainment as a cue to perceived attractiveness and likability," *Proceedings of P&P13*, in print.

[23] P. Boersma and D. Weenink. *Praat: Doing phonetics by computer*, 2016.

[24] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.

[25] D. Bates, M. Maechler, B. Bolker and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[26] A. Kuznetsova, P. B. Brockhoff and R. H. B. Christensen. *lmerTest: Tests in linear mixed effects models. R package version 2.0-30*, 2016.