

Laughter dynamics in dyadic conversations

Bogdan Ludusan, Petra Wagner

Phonetics and Phonology Workgroup, Faculty of Linguistics and Literary Studies,
Bielefeld University, Germany

{bogdan.ludusan, petra.wagner}@uni-bielefeld.de

Abstract

Human verbal communication is a complex phenomenon involving dynamics that normally result in the alignment of participants on several modalities, and across various linguistic domains. We examined here whether such dynamics occur also for paralinguistic events, in particular, in the case of laughter. Using a conversational corpus containing dyadic interactions in three languages (French, German and Mandarin Chinese), we investigated three measures of alignment: convergence, synchrony and agreement. Support for convergence and synchrony was found in all three languages, although the level of support varied with the language, while the agreement in laughter type was found to be significant for the German data. The implications of these findings towards a better understanding of the role of laughter in human communication are discussed.

Index Terms: laughter, entrainment, convergence, conversational speech, paralinguistics

1. Introduction

A lot of recent work has emphasized the mutual influence between dialogue partners with respect to their verbal and non-verbal behaviours. More specifically, partners involved in conversation have been shown to align in different ways during the interaction, including on modalities not related to the spoken dimension, such as postures and facial expressions [1]. With respect to speech, alignment has been observed at various linguistic levels: lexical [2], syntactic [3], as well as temporal dynamics [4]. Furthermore, an alignment has been found also in the acoustic-prosodic dimension [5, 6, 7], with interlocutors being more likely to use similar acoustic characteristics (e.g., intensity, pitch, speech rate) or to converge towards the partner's speech targets.

Conversational speech includes also paralinguistic information and, among these phenomena, laughter seems to have an extensive presence in speech [8]. Laughter has been studied from different perspectives [9], such as the role it plays in human social interaction [10], as well as its acoustic realization [11, 12, 13, 14], not only at the level of entire events (bouts) but also at the level of its individual components (calls). The latter group of studies investigated durational and spectral characteristics of laughter, finding important inter- and intra-speaker variations along these dimensions. Despite this high variation, the characteristics of laughter across languages seem to be similar, as shown also in a study comparing laughter produced by German and Italian speakers [15].

Having seen how conversational partners align on various dimensions, we would like to investigate here if they would also align in their use of laughter. To our knowledge, only one previous study [16] has touched on this phenomenon. It considered three different views of alignment, in terms of the amount of laughter, the temporal alignment of the laughs and, what the au-

thors refer to as phonetic imitation. Based on data from two different corpora of conversational speech, as well as on the findings of previous investigations, they conclude that there is evidence for the proposed alignment aspects.

In this study we propose more fine-grained measures of alignment, describing three different facets of the phenomenon: synchrony, convergence and agreement. The former two views of alignment have been investigated also in the case of acoustic-prosodic alignment [5]. Similarity represents the degree of dynamic likeness between speakers throughout the interaction, while convergence characterizes the difference in the degree of similarity between speakers towards the end of the interaction, compared to its beginning. The last measure, agreement, relates to the phonetic imitation definition of [16] and it looks at the similarity in voicing of consecutive laughs produced by different speakers. Furthermore, the present research takes a cross-linguistic approach to the study of alignment, by employing a novel corpus of dyadic interactions (used previously only for studying the timing of laughter compared with the position of the laughable [17]), containing comparable materials in three typologically diverse languages: French, German and Mandarin Chinese. Based on the current knowledge in the field, no clear hypotheses can be formulated on whether speakers align their laughter in any of the aspects examined here. Nevertheless, taking into account that laughter seems to have similar roles and acoustic characteristics in the languages it was studied by now, we would expect our findings to be consistent across languages.

The paper is structured as follows: the corpus and the methodology employed in this study are presented in Sections 2 and 3, respectively. Next, the results obtained with the proposed alignment measures for the three examined languages will be illustrated. Finally, we will discuss the implications of these findings and we will consider the following steps to take in this investigation.

2. Materials

We base our study on the DUEL corpus [18], which contains spontaneous conversations in French, German and Mandarin Chinese. For each language, ten speaker dyads were recorded, the dyads being composed of either friends/acquaintances (3/10, 5/10 and 5/10, respectively) or strangers. All were native speakers of their respective languages and all but three pairs of participants were students.

The dataset consists of three scenarios: "Dream Apartment", "Film Script" and "Border Control", with the total interaction for each pair of participants being around 45 minutes. In the first scenario, the participants are told they have a large amount of money to furnish and decorate a shared open-plan apartment. In the second scenario, they are asked to come up with a scene for a film in which something embarrassing happens to the main character, which may be inspired from their ex-

perience. The third scenario sees the participants enacting a dialogue between a border control official and a traveler which has a situation that disfavors them in the interview, with the two of them being in-laws. In the first two scenarios, the participants sit across each other in chairs, while in the last scenario they sit at a table. The corpus was manually transcribed and segmented at the turn and the utterance level. It was also annotated for disfluencies and laughter (including speech-laughs), the latter type of annotations being used in this study. Speech-laugh represent simultaneous productions of speech and laughter, in which neither of the two components is dominant, exhibiting speech-like fundamental frequency and laughter-like rhythm and amplitudes [19].

For a more straightforward cross-linguistic comparison, we employed the data from the Film Script scenario in all three languages, as it was the scenario containing the most laughter occurrences. From the ten dyads in each language, we eliminated the ones which laughed rarely. For this purpose, each recording was split into ten equal temporal bins, and the number of laughs of each participant in each of those ten time intervals was counted. The dyads in which one or both speakers had more than three bins without laughter were removed. Incidentally, the used measure also strongly correlated with the absolute number of laughs in the entire scenario. After this step, we retained seven French dyads, eight German dyads and six Chinese dyads. The resulting materials consisted of more than four hours of recordings (94, 107, and 78 minutes, respectively) and over 1900 laughs and speech-laugh (493, 849 and 580 occurrences, respectively).

3. Methods

Three measures of speech alignment were investigated here: *synchrony*, *convergence* and *agreement*. In all cases, we considered as laughter both laughs and speech-laugh produced by two conversation partners. For the first two measures, the entire conversation was divided into ten equal time bins and in each resulting bin the number of laughs of each of the two participants was counted. Although these measures have been previously employed in studies investigating alignment, there is no consensus on how to operationalize them. Thus, in the following paragraphs, we present the implementation used for these measures in our study.

Synchrony was defined as being the Spearman correlation between the counted laughs in the ten bins of each two interlocutors. We chose Spearman over Pearson correlation as it is more robust to outliers and skewed distributions. The average of the computed correlation coefficients is then reported for each language.

In order to calculate convergence, for each participant, we determined the number of laughs in the first half of the conversation (first five bins) and in the second half (last five bins), and computed the modulus of the difference in the number of laughs between two interlocutors for each half (see Equations 1, where L_{ij} represents the laughs count in bin i , for speaker j). Thus, we obtained two vectors for each language, one containing the absolute difference in laughs between each pair of speakers in the first half of the conversations and a second vector with the absolute difference in the second half. We then applied two-tailed paired t-tests to examine whether any of these differences are statistically significant.

$$\begin{aligned} diff_1 &= \left| \sum_{k=1}^5 L_{k1} - \sum_{k=1}^5 L_{k2} \right| \\ diff_2 &= \left| \sum_{k=6}^{10} L_{k1} - \sum_{k=6}^{10} L_{k2} \right| \end{aligned} \quad (1)$$

For a comparison with previous studies, we also calculated an additional measure of convergence, described in terms of the total amount of laughter, as employed in [16]. It was defined as the Spearman correlation between the total number of laughs produced by the two participants, across all dyads. The difference between this definition of convergence and the proposed synchrony measure is that, while the former is obtained over the whole dataset corresponding to a language, the latter is computed at the level of individual recordings.

Agreement represents the proportion of consecutive laughs exhibiting speaker changes that agree in voicing (voiced or unvoiced laughs), out of the total amount of consecutive laughs. A laugh was considered to be consecutive if it either overlapped with or began within one second of the end of the laugh produced by the other speaker. For determining the voicing of each laugh, the pitch of each recording was extracted with Praat [20], using the default parameters, and checked whether the produced laughs contained any voiced frames. If so, they were marked as voiced laughs, otherwise as unvoiced laughs (as in [21]). For each pair of consecutive laughs, a non-consecutive pair was considered, formed by the first laugh of the consecutive pair and a different, randomly sampled laugh among the laughs produced by the other participant during the conversation. The process of randomly sampling non-consecutive laughs was run for 10,000 times and the average agreement measure obtained after these runs was compared to the agreement calculated for consecutive laughs.

Synchrony and convergence were computed for all three languages in our dataset, but we examined agreement only in the German data, as both French and Mandarin recordings contained a high degree of cross-talk, which would have confounded our measure of agreement.

4. Results

We present here the results obtained with the proposed measures of alignment. In terms of synchrony, we observed weak to medium correlations in our data, varying with the language: $\rho = 0.255$ for French, $\rho = 0.508$ for German and $\rho = 0.456$ for Mandarin. These values, along with their corresponding confidence intervals are reported in Figure 1.

The convergence results are illustrated in Table 1. One can see that there is a trend present in all three languages: the dif-

Table 1: Average absolute differences in the amount of laughter between the first half and the second half of the conversation, respectively, in the three studied languages: French, German and Mandarin.

Language	Difference	
	1st half	2nd half
French	11.1	7.3
German	9.4	8.5
Mandarin	13.5	11.8

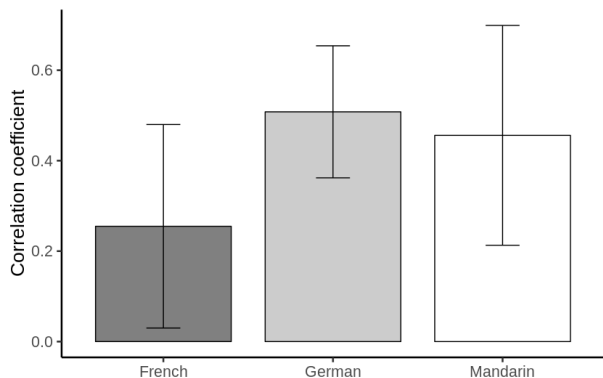


Figure 1: Spearman correlation coefficient for the synchrony measure, in the three languages investigated: French, German and Mandarin. The error bars represent 95% confidence intervals.

ference in the amount of laughter between the two speakers appears to decrease in the second part of the recording, suggesting a certain level of convergence. Nevertheless, none of the differences were found to be statistically significant, probably due to our low sample size (between 6 and 8 dyads per language). When looking, instead, at convergence from the point of view of the total amount of laughs, as in [16], we obtained the following Spearman correlation coefficients: 0.847, 0.524 and 0.058 for French, German and Mandarin, respectively.

Finally, the results for agreement, in the German data, revealed a higher voicing agreement in consecutive laughs (83.1%) than in randomly-sampled non-consecutive laughs (77.4%). Testing this difference with a paired two-tailed t-test, it was found statistically significant ($t = 2.64$; $df = 7$; $p < 0.05$). The agreement difference for each of the eight dyads is illustrated in Figure 2 and shows that the effect is consistent across speakers, with only one dyad showing a small difference in the opposite direction. This result is even more significant, considering that around 85% of the laughs produced by the dyads included in our analysis were voiced, which increases the likeliness of having randomly sampled non-consecutive laughs agreeing in voicing.

5. Discussion

The results presented in the previous section shed further light on the phenomenon of conversational alignment, bringing complementary knowledge to the findings of Trouvain and Truong [16]. In particular, we were able to confirm that paralinguistic communication is also subject to conversational alignment, but in language, or maybe culture-specific ways. Furthermore, we could elaborate on the previous findings by means of a more fine-grained analysis, using an enriched set of metrics and languages to study the various dimensions of cross-speaker alignment.

In the following, we will compare our analyses and findings with those of Trouvain and Truong more in depth. First, their measure of alignment on the amount of laughter characterizes the degree of similarity in the overall use of laughter in the conversation. The synchrony measure we proposed here captures a more fine-grained phenomenon, namely the dynamics of laughter used during the conversation. These two measures are not necessarily correlated, as one could find high levels of agree-

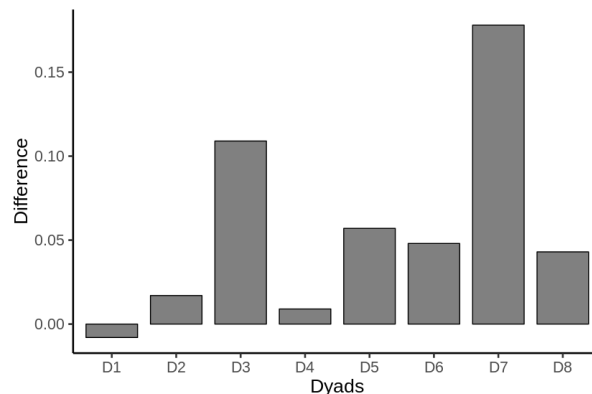


Figure 2: Difference in voicing agreement between consecutive laughs and randomly sampled non-consecutive laughs for the eight dyads considered from the German data.

ment in the total number of laughs, but little or no synchrony in the distributions of laughs (as in our French data). This is usually the case when speakers laugh more in different parts of the conversation. Also, the opposite case is possible: a relatively synchronous distribution of laughs may coincide with a low correlation for the number of overall laughs, if one of the speakers tends to laugh more than her or his partner (as in the Mandarin recordings).

Our measure of synchrony relates, in a way, also to their measure of temporal alignment. Having a high synchronicity between the two partners might involve also a high degree of temporal alignment. We tested this hypothesis by examining the proportion of consecutive laughs produced by the two partners, out of the total number of laughs in the conversation (the same definition for consecutive laughs was used as for the agreement measure). We found the highest proportion of consecutive laughs in German (41%), followed by Mandarin (23%) and then French (17%), which seems to confirm our assumption.

Lastly, Trouvain and Truong found evidence for phonetic imitation, in the form of stronger prosodically marked overlapping laughs, while we observed a higher degree of voicing agreement in consecutive than non-consecutive laughs between same conversation partners.

An important aspect that has to be mentioned here is the amount and the distribution of laughs and speech-laughs in the data included in the analysis. While, in general, speech-laughs tend to be discarded from the analysis of laughter in conversation, they are pervasive in human communication. For example, in a study looking at mother-infant interactions [19], between 5% and 50% of the laughter tokens produced by the mothers were speech-laughs (18.6%, on average). In our dataset, 22%, 33% and 44% of the total amount of laughter were speech-laughs, in the German, French and Mandarin parts, respectively. While the higher incidence of speech-laughs might be due to the nature of the task the participants were performing, there seems to be an even stronger influence of the language on its production frequency. We believe that the ubiquity of this phenomenon warrants its inclusion in future studies examining the role of laughter in conversation.

The above-mentioned findings might stem from language or cultural differences, as well as from interpersonal dynamics. Our results on language- or culture-specific differences

in alignment are fully in line with previous phonetic investigations on the cross-cultural perception of attitudinal speech: [22] confirmed both a universality of the processing of attitudinal expression, and cross-cultural differences in the processing of highly specific attitudinal expressions. A potential interpretation of our findings would be that the universality of laughter alignment stems from a general tendency of mimicking positive emotional expressions such as smiles [23], while the differences may be the result of language-specific discourse strategies [24], i.e., unique ways of employing paralinguistic alignment in the organization of spoken interaction for serving different functions (e.g. social laughter, laughter for face saving or suppressing embarrassment). Naturally, such functions may occur more or less often depending on the task, and could also be responsible for different distributions of laughter and its alignment. Also, previous studies on phonetic alignment have found it to vary strongly as a function of interpersonal dynamics, such as the status of the relationship, friendship or mutually perceived closeness [25]. As these factors were not controlled in our study, they may have influenced our results, and given the small sample for each language, may be the result of a data sparsity problem.

Thus, one limitation of our study might be the size of the employed dataset, in particular the low number of speaker dyads that we could consider from each language. Although larger datasets exist, it is a known fact that laughter annotations in conversational corpora have reliability issues, not only due to inconsistencies in the definition of laughter and its labelling, but also because of technical errors [21]. Since conversational laughter cannot be studied in more controlled environments (such as a laboratory setup), it is essential that annotation of large conversational corpora be performed more reliably.

6. Conclusions and future work

To summarize, we examined in this study several measures of laughter alignment in conversation, obtaining promising results in this direction. Our measure of synchrony, which captures the laughter dynamics between conversational partners, received weak to moderate support, while the metric used for laughter agreement was significant for the German data. Lastly, although our convergence measure did not reach significance, probably due to the small sample size, it showed a consistent trend towards convergence. With the cross-linguistic dimension being an important aspect of this study, it was interesting to see that the two measures tested cross-linguistically, synchrony and convergence, returned similar patterns.

Our findings have implications not only for the field of paralinguistics, by uncovering the patterns of laughter use in conversation, but also for a better understanding of how laughter dynamics may be used in human communication. This, in turn, could have consequences in several other fields, such as human-machine interaction, by improving the naturalness of the communication, as well as in social rehabilitation programmes for clinical populations.

In the future, we would like to extend our investigation to other aspects of laughter alignment. With laughter having a high intra-speaker variation (including for the vowel quality of the calls; e.g. [26]), it would be interesting to explore whether laughter alignment between speakers occurs also at the segmental level, by making conversational partners more similar in the quality of the produced vowel.

7. Acknowledgements

Bogdan Ludusan's work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 799022. We are grateful to David Schlangen for making the DUEL corpus available to us.

8. References

- [1] 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, p. 893, 1999.
- [2] S. E. Brennan and H. H. Clark, "Conceptual pacts and lexical choice in conversation," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 22, no. 6, p. 1482, 1996.
- [3] H. P. Branigan, M. J. Pickering, and A. A. Cleland, "Syntactic coordination in dialogue," *Cognition*, vol. 75, no. 2, pp. B13-B25, 2000.
- [4] J. Edlund, J. B. Hirschberg, and M. Heldner, "Pause and gap length in face-to-face interaction," in *Proceedings of 10th Annual Conference of the International Speech Communication Association*, 2009, pp. 2779-2782.
- [5] R. Levitan and J. Hirschberg, "Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions," in *Proceedings of the 12th Annual Conference of the International Speech Communication Association*, 2011, pp. 3081-3084.
- [6] C. De Looze, S. Scherer, B. Vaughan, and N. Campbell, "Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction," *Speech Communication*, vol. 58, pp. 11-34, 2014.
- [7] C.-C. Lee, A. Katsamanis, M. P. Black, B. R. Baucom, A. Christensen, P. G. Georgiou, and S. S. Narayanan, "Computing vocal entrainment: A signal-derived PCA-based quantification scheme with application to affect analysis in married couple interactions," *Computer Speech & Language*, vol. 28, no. 2, pp. 518-539, 2014.
- [8] N. Campbell, "Approaches to conversational speech rhythm: Speech activity in two-person telephone dialogues," in *Proceedings of the XVth International Congress of Phonetic Sciences*, 2007, pp. 343-348.
- [9] J. Trouvain and K. P. Truong, "Laughter," in *The Routledge Handbook of Language and Humor*. Routledge, 2017, pp. 340-355.
- [10] P. Glenn, *Laughter in interaction*. Cambridge University Press, 2003, vol. 18.
- [11] D. E. Mowrer, L. L. LaPointe, and J. Case, "Analysis of five acoustic correlates of laughter," *Journal of Nonverbal behavior*, vol. 11, no. 3, pp. 191-199, 1987.
- [12] J.-A. Bachorowski, M. J. Smoski, and M. J. Owren, "The acoustic features of human laughter," *The Journal of the Acoustical Society of America*, vol. 110, no. 3, pp. 1581-1597, 2001.
- [13] J. Vettin and D. Todt, "Laughter in conversation: Features of occurrence and acoustic structure," *Journal of Nonverbal Behavior*, vol. 28, no. 2, pp. 93-115, 2004.
- [14] B. Bigi and R. Bertrand, "Laughter in French spontaneous conversational dialogs," in *Proceedings of the 10th Language Resources and Evaluation Conference*, 2016, pp. 2168-2174.
- [15] H. Rothgänger, G. Hauser, A. C. Cappellini, and A. Guidotti, "Analysis of laughter and speech sounds in Italian and German students," *Naturwissenschaften*, vol. 85, no. 8, pp. 394-402, 1998.
- [16] J. Trouvain and K. P. Truong, "Convergence of laughter in conversational speech: effects of quantity, temporal alignment and imitation," in *Proceedings of the International Symposium on Imitation and Convergence in Speech*, 2012, pp. 37-38.
- [17] Y. Tian, C. Mazzocconi, and J. Ginzburg, "When do we laugh?" in *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2016, pp. 360-369.

- [18] J. Hough, Y. Tian, L. de Ruyter, S. Betz, S. Kousidis, D. Schlangen, and J. Ginzburg, "DUEL: A multi-lingual multi-modal dialogue corpus for disfluency, exclamations and laughter," in *Proceedings of the 10th Language Resources and Evaluation Conference*, 2016, pp. 1784–1788.
- [19] E. E. Nwokah, H.-C. Hsu, P. Davies, and A. Fogel, "The integration of laughter and speech in vocal communication: A dynamic systems perspective," *Journal of Speech, Language, and Hearing Research*, vol. 42, no. 4, pp. 880–894, 1999.
- [20] P. Boersma, "Praat, a system for doing phonetics by computer," *Glott International*, vol. 5, pp. 341–345, 2002.
- [21] K. P. Truong and J. Trouvain, "Laughter annotations in conversational speech corpora—possibilities and limitations for phonetic analysis," *Proceedings of the 4th International Workshop on Corpora for Research on Emotion Sentiment and Social Signals*, pp. 20–24, 2012.
- [22] H. Mixdorff, A. Hönemann, A. Rilliard, T. Lee, and M. Ma, "Audio-visual expressions of attitude: How many different attitudes can perceivers decode?" *Speech Communication*, vol. 95, pp. 114–126, 2017.
- [23] E. G. Krumhuber, K. U. Likowski, and P. Weyers, "Facial mimicry of spontaneous and deliberate Duchenne and non-Duchenne smiles," *Journal of Nonverbal Behavior*, vol. 38, no. 1, pp. 1–11, 2014.
- [24] D. Tannen, "The pragmatics of cross-cultural communication," *Applied Linguistics*, vol. 5, no. 3, pp. 189–195, 2017.
- [25] J. S. Pardo, R. Gibbons, A. Suppes, and R. M. Krauss, "Phonetic convergence in college roommates," *Journal of Phonetics*, vol. 40, pp. 190–197, 2012.
- [26] D. P. Szameitat, C. J. Darwin, A. J. Szameitat, D. Wildgruber, and K. Alter, "Formant characteristics of human laughter," *Journal of Voice*, vol. 25, no. 1, pp. 32–37, 2011.