



How do you laugh in an fMRI scanner? Laughter distribution, mimicry and acoustic analysis

Chiara Mazzocchi, Benjamin O'Brien, Chaminade Thierry

► To cite this version:

Chiara Mazzocchi, Benjamin O'Brien, Chaminade Thierry. How do you laugh in an fMRI scanner? Laughter distribution, mimicry and acoustic analysis. Disfluency in Spontaneous Speech Workshop 2023, Universität Bielefeld, Aug 2023, Bielefeld, Germany. 10.21437/DiSS.2023-9 . hal-04155320

HAL Id: hal-04155320

<https://hal.science/hal-04155320>

Submitted on 7 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How do you laugh in an fMRI scanner?

Laughter distribution, mimicry and acoustic analysis

Chiara Mazzocchi^{1,2}, Benjamin O'Brien³, Thierry Chaminade^{1,4}

¹ Institute of Language, Communication, and the Brain, Aix-en-Provence, France

² Laboratoire Parole et Langage (UMR 7309), CNRS, Aix-Marseille University, France

³ Laboratoire Informatique d'Avignon (EA 4128), Avignon University, France

⁴ Institut de Neurosciences de la Timone, (UMR 7289), CNRS, Aix-Marseille University, France

chiara.mazzocchi@univ-amu.fr, benjamin.o-brien@univ-avignon.fr,
thierry.chaminade@univ-amu.fr

Abstract

The current study analyses laughter distribution, pragmatic use, and acoustic features in a multimodal conversation corpus, where one of the interlocutor had brain activity recorded with functional MRI. Particular focus is on the analysis of mimicking and non-mimicking laughter, as we hypothesized that the former might involve different neural pathways in comparison to latter, and a spectrotemporal modulation analysis was selected to evaluate the acoustic differences between them. The goals of our investigation are (1) to validate the corpus as ecological in terms of laughter production prior to evaluating neurophysiological correlates of laughter production and perception in conversational interactions, (2) to compare mimicking and non-mimicking laughter in terms of acoustic features, and (3) to evaluate which features need to be taken in account for our forthcoming analysis of fMRI data. Despite the unconventional recording conditions, laughter production and use results are comparable to those observed in face-to-face interactions. Mimicking laughter has significantly increased temporal modulations in comparison to non-mimicking laughter, however, no differences are observed in the spectral modulation domain. Our findings suggest that mimicking laughter might be related to different underpinning neuro-psychological processes on which we plan on gaining insight through future neural-correlate analysis.

Index Terms: laughter in conversation, fMRI conversational corpus, laughter mimicry, spectrotemporal modulation, generalized additive mixed models

1. Introduction

Laughter is a non-verbal vocalisation crucial in our conversations. It affects the unfolding of our dialogues and has important effects on the establishment and maintenance of relationships [1, 2, 3]. It conveys meaning, performs different pragmatic functions, and can be used to smooth face-threatening dialogue acts, such as criticism [4, 5, 6, 7]. Until now studies investigating the neural correlates of laughter have been mainly restrained to the study of laughs produced or perceived in isolation, produced in response to isolated humorous stimuli, or in response to electrical brain stimulation (e.g. [8, 9, 10, 11, 12, 13, 14]). Yet, very little is known as to the production and processing of laughter in interaction.

To fill this gap we annotated laughter in the human-human portion of the Convers corpus [15], where participants engaged in spontaneous loosely task-directed conversations with a confederate while their neural activity was recorded in an fMRI

scanner. The current paper presents a corpus study constituting an essential basis for future analysis of the neurophysiological data. We analysed laughter occurrences, laughter positioning in relation to speech and others' laughter, its pragmatic functions, and the acoustics of mimicking and non-mimicking laughter.

Numerous acoustic features have been used to describe the perceptual characteristics associated with laughter (see [16, 17, 18, 19]). These descriptors span both temporal and spectral dimensions. Research in speech production and perception has focused on the changes across these dimensions by transferring the time-frequency representations of acoustic signals into the *spectrotemporal modulation* domain (hereafter STM) [20]. Numerous studies have focused on unpacking modulation representations of speech signals in the context of assessing intelligibility [21, 22] and voice pathology [23, 24], and more recently laughter. [25, 26] showed that modulations of the amplitude envelope and F0 could be used to distinguish speech, laughter, and speech-laughter. [27] relied on such measures to examine whether mimicking and non-mimicking laughter are characterised by specific modulation rates. Since previous work suggests mimicking laughter might involve different pathways in comparison to non-mimicking laughter [28], a spectrotemporal modulation analysis was selected to analyse the acoustic differences between the two.

The current work has three main goals: (1) to evaluate whether annotated interactions are ecologically valid in terms of laughter occurrences, laughter mimicry and laughter pragmatics; (2) to explore whether mimicking and non-mimicking laughter are significantly different in terms of acoustic features, which might be informative about deeper neuropsychological processes; and (3) to evaluate which distributional and acoustic trends should be taken into account when applying neurophysiological analysis to the corpus in a future analysis of laughter neural correlates during spontaneous conversation.

2. Methods

2.1. Corpus

The Convers corpus [15] is constituted by spontaneous loosely task-directed French dyadic conversations between participants and a confederate of the experiment (in and outside of a fMRI scanner, respectively). 24 minutes of multimodal recording are available for each participant: 12 min while interacting with a human confederate and 12 min while interacting with a robot (Furhat) controlled by a confederate. The current analysis is exclusively focused on the human-human interactions (total: 276 minutes). Our corpus contains 23 dyadic conversation: 23

Table 1: *Inter-annotator agreement*

Feature	% Agreement	Krippendorff's α
L-SL	97.3%	0.94
Arousal	68.9%	0.66
Laughable origin	86.5%	0.75
Incongruity Branch	81.1%	0.54
Overall	84.9%	0.73

participants (16 females) unfamiliar with a human confederate of the same sex. The conversation was loosely task-directed, where dyads were presented graphic drawings intended to be a part of advertising campaign and, through conversation with the confederate, the goal was to discover the key message behind the campaign. The participants could see the interlocutor recorded live with a webcam during the the conversation.

2.2. Laughter Annotation

The corpus was annotated by author CM using the software ELAN [29]. The laughter audio-visual identification criteria follows the procedures outlined in [30, 31]. Following the annotation scheme proposed in [7], the level of arousal displayed by the laughter's form (low/medium/high), temporal sequence in relation to speech (laughter/speech-laughter), and others' laughs (non-/mimicking laughter) were annotated. We define *mimicking laughter* as laughter that shortly follows a laughter from the partner, either overlapping or beginning within one second from the end of the initiating laughter (see [30, 27] for calculation methods). The term *non-mimicking laughter* refers to any laughter not meeting this criteria.

Based on [7], we consider laughter as an event predicate relating to a *laughable* in the context (i.e., the laughter argument). Different types of laughables can be distinguished based on whether they contain an incongruity or not, and, if so, which kind of incongruity (see [4] for a formal definition of incongruity). The annotation categories are as follows:

1. *Pleasant incongruity* is a clash between the laughable and certain background information perceived as witty, rewarding and/or somehow pleasant (e.g., jokes, puns, goofy behaviour and conversational humour).
2. *Social incongruity* is a clash between social norms and/or comfort and the laughable.. Examples include social discomfort (e.g., embarrassment or awkwardness), violation of social norms (e.g., invasion of another's space, asking a favour), or an utterance that clashes with the interlocutor's expectations concerning one's behaviour (e.g., criticism).
3. *Pragmatic incongruity* arises when there is a clash between what is said and what is intended (e.g., in the case of irony, scare-quoting, hyperbole etc.). In such cases laughter is used by the speaker in order to signal the need to opt for a less probable meaning interpretation within their own utterance.
4. *Pleasantness* refers to cases where no incongruity can be identified. In these cases what is associated with the laughable is a sense of closeness that is either felt or displayed towards the interlocutor, e.g., while thanking or receiving a pat on the shoulder.

Interannotator agreement was conducted on 20% of the corpus by a native French graduate student. The agreement on laughter identification and segmentation (start/end boundaries) was assessed using the Staccato algorithm implemented

in ELAN [32]. We ran the analysis with 1000 Monte Carlo Simulations, a granularity for annotation length of 10, and $\alpha = 0.05$. The average degree of organisation between annotators is 0.83. The results in terms of percentage agreement and Krippendorff's α [33] are given in Table 1 for each variable annotated. Overall we have a percentage agreement of 84.86% and a Krippendorff's α of 0.73.

2.3. Acoustic Analysis

2.3.1. Acoustic pre-processing

To remove unwanted noise generated by the fMRI machines, the *noisereduce* [34] Python function was used with the stationary noise reduction setting (1.5 standard deviations above mean signal:noise ratio). All recordings were then normalised by adjusting the maximal amplitude to a target of 100% of the signal dynamic (MATLAB 2021b, MathWorks Inc, USA).

2.3.2. Spectrotemporal modulation processing

The energy distribution across the spectrotemporal modulation domain can be expressed as the Modulation Power Spectrum (MPS). Studies by [21, 35, 36] define MPS as a two-dimensional Fourier transform of the time-frequency representation of an audio signal. In general Equation 1 defines the MPS, where s and r are spectral and temporal modulations, respectively, and $Y(t, f)$ is the amplitude extracted from the Fourier transform:

$$\text{MPS}(s, r) = \int \int |Y(t, f)| e^{-2\pi i s f} e^{-2\pi i r t} df dt \quad (1)$$

Scripts developed in MATLAB by [22] were adapted to obtain the MPS of laughs. Laughter recordings were first downsampled to 16 kHz. Time-frequency representations were obtained by using a gammatone filter bank summation method (128 full-width half-maximum gaussians with center frequencies logarithmically spanning the frequency domain). Hilbert transforms were used to extract the analytical amplitudes from filter outputs. The *fft2* MATLAB function transformed the time-frequency representations into the modulation domain. It has been shown that speech is embedded with temporal and spectral modulations ranging from 0 to 32 Hz and 0 to 4 cyc/oct, respectively. Thus, these ranges were used to compare STM profiles associated with mimicking and non-mimicking laughter.

2.3.3. Statistics for acoustic analysis

Generalized additive mixed models (GAMMs) [37] were used to evaluate the efficiency of temporal modulation (TM) and spectral modulation (SM) metrics to model mimicking (M) and non-mimicking (NM) laughter. GAMMs were selected for statistical analysis as they can handle time-varying data with non-linear relationships and have been shown to be effective at evaluating amplitude and F0 modulations [25]. Following the procedures suggested by [38, 25] the R-package *mgcv* was used. Thin plate regression splines were used as smoothing functions to model the non-linear variation present in the data. The R-package *itsadug* [39] was used to estimate an AR-1 correlation parameter ρ and pairwise differences between the non-linear smooths of the factor levels.

Formula 2 describes the GAMM used in the current study. Amplitudes A (in dB) corresponding to modulations v (the unit for v is Hz for TM and cyc/oct for SM) were used as dependent variables. The term L represents the laughter type (mimicking

Table 2: *Incongruity type contained in the laughables*

Incongruity type	N	Overall %	Participant	Confederate
Pleasant	219	62.93%	56.97%	68.31%
Social	124	35.63%	41.21%	30.60%
Pragmatic	3	0.86%	1.21%	0.55%
Pleasantness	2	0.57%	0.61%	0.55%

vs. non-mimicking). Participant p was set to a non-linear random factor. The ρ -value described in Formula 2 was estimated from the data and included to control for auto-correlation in the time series ($\rho_{TM} = 0.85$; $\rho_{SM} = 0.88$). As proposed in [25], each model was first tested against a base model not containing the fixed factor via the *compareML* function.

$$\text{bam}(A \sim L + s(v, \text{by} = L) + s(v, p, \text{by} = L, \text{bs} = \text{"fs"}), \text{rho} = \rho) \quad (2)$$

3. Results

3.1. Laughter distribution, mimicry and laughable features

Overall 348 laughs were identified: 165 from Participants in the fMRI scanner (P) and 183 from Confederates (C). The overall frequency was 1.16 laughs per minute. The mean duration of the laughs was 1.35 ± 0.79 s, ranging from 0.24 to 4.79 s. 23.28% of laughter events in the dataset (N=81) was constituted by speech-laughter. We observed a high inter-individual variability in terms of laughter production regardless of the conversational role (P: 7.86 ± 6.76 laughs; C: 8.32 ± 6.54 laughs). Similarly we observed high inter-individual variability in the percentages of speech-laughter over the total laughs produced by each individual (P: $21 \pm 23\%$; C: $18\% \pm 22\%$). Overall we observed more speech-laughter in P (30.91%) than in C (16.39%).

28% of laughter events in the dataset (N=97) was constituted by mimicking laughter (P: 25%; C: 31%). Regarding arousal, we observed 73.85% of the laughs to be low, 22% medium, and very few (4%) high arousal. 58.05% of laughs was related to laughables produced by the laugher her/himself, 37.36% was related to laughables produced by the partner, while much lower percentages was related to laughables jointly constructed (2.3%) or external (2.3%). Table 2 reports occurrences and percentages of the different types of laughables.

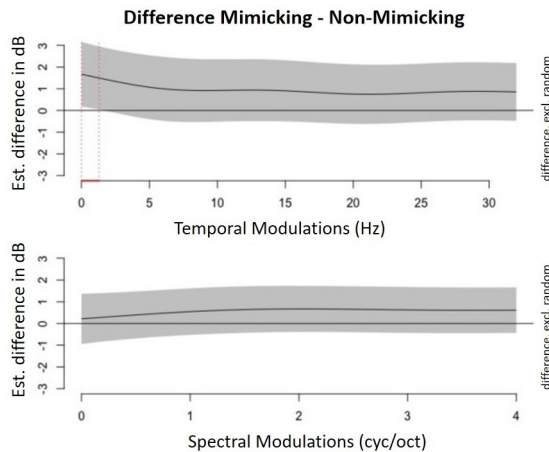


Figure 1: *Differences between fitted class models for temporal (top) and spectral (bottom) modulations between mimicking and non-mimicking laughter. The red interval represents the range of modulations for which two models differed significantly.*

3.2. Acoustic analysis

Table 3 describes the results of mimicking and non-mimicking laughter GAMMs based on temporal modulations (adjusted $r^2 = 0.75$). Figure 1-Top illustrates the difference between the two TM-based models. Mimicking laughter had significantly increased temporal modulation amplitudes (≈ 1.5 dB) in the 0-1.3 Hz range in comparison to non-mimicking laughter. Table 4 describes the results of mimicking and non-mimicking laughter GAMMs based on spectral modulations (adjusted $r^2 = 0.93$). No significant differences were observed between mimicking and non-mimicking laughter (Fig. 1-Bottom). Aware that mimicking laughter was more frequent when the laughable was constituted by a pleasant incongruity [7] (79% of mimicking laughter in the dataset), we ran the same analysis excluding laughs related to social incongruity and found no significant differences from results reported here.

Table 3: *Temporal modulation model overview*

Intercept	estimates	SE	t	p
M	0.41	0.27	1.55	n.s
NM	-37.83	0.27	-141.04	***
Fixed smooth terms	edf	ref.df	F	p
s(v):M	8.80	8.93	262.77	***
s(v):NM	8.92	8.98	821.48	***
Random smooth terms	edf	ref.df	F	p
s(v,p):M	48.79	125.0	2.36	***
s(v,p):NM	65.40	207.0	2.41	***

Table 4: *Spectral modulation model overview*

Intercept	estimates	SE	t	p
M	0.38	0.22	1.74	n.s
NM	-42.74	0.22	-195.98	***
Fixed smooth terms	edf	ref.df	F	p
s(v):M	8.84	8.86	952.73	***
s(v):NM	8.90	8.92	2104.88	***
Random smooth terms	edf	ref.df	F	p
s(v,p):M	117.65	125.0	31.45	***
s(v,p):NM	199.51	202.0	42.53	***

4. Discussion

4.1. Laughter occurrences, mimicry and laughable features

Laughter frequency in the Convers corpus (1.16 per minute) was comparable to more ecological corpora, such as in-home friendly conversations [40] (5.8 ± 2.5 per 10 min) and fully ecological and diverse contexts (BNC: 5 per 10 min) [7]. On the other hand, laughter was less frequent than in other contexts, such as speed-dating (21 ± 9.28 per 5 min) [41] and friendly, loosely-controlled conversations recorded in a laboratory setting (DUEL French: 45 per 10 min; DUEL Chinese: 26 per 10 min) [7]. Laughter duration was similar to findings those reported in other works (e.g., [30, 42]). Laughter arousal levels were similar to those observed in face-to-face corpora recorded in laboratory settings (e.g., DUEL corpus [7]). The percentage of mimicking laughter was similar to findings in ecological settings (BNC: 26%[7]; [43]: 32%), and lower than in corpora recorded in laboratory settings (DUEL Fr: 39%; Ch: 37% [7]).

Consistent with the literature, more than half of the laughs produced (58%) related to utterances produced by the laugher her/himself (DUEL French and Chinese 57%; BNC 46% [7]; 43.5 - 55.3% [40]). Regarding which type of laughable the laughter related to, we observed similar patterns to those observed in other multilingual corpora [7]. Most of laughs related to *pleasant incongruity*, about one third related to *social incongruity*, while rarely related to *pragmatic incongruity* and *pleasantness* laughables. Participants exhibited significantly higher *social incongruity* laughter in comparison to *pleasant incongruity* compared to Confederates ($\chi^2_1 = 4.05$, $p < 0.05$). Participants' percentages were more similar to those observed in face-to-face natural dyadic conversations, where the relationship between interlocutors is more balanced. We can speculate that confederates were in a more natural position, sitting in front of a screen outside of the fMRI scanner. Moreover they were aware of the experimental design and potentially less concerned about *social incongruity* laughables.

4.2. Acoustic analysis

The temporal modulation models yielded significant differences between mimicking and non-mimicking laughter. Given the approximate 1:4 ratio of mimicking to non-mimicking laughter, temporal modulations were quite effective at discriminating laughter types, as evidenced by the temporal modulation GAMM 75.2% of deviance explained by the models (Sec. 3.2). The models demonstrated that, in comparison to non-mimicking laughter, mimicking laughter exhibited an increase in temporal modulations below 2 Hz. As studies have reported temporal modulations in the 1-2 Hz range are associated with prosody [22, 44], one interpretation of this finding suggests that participants might prosodically mark their responsive (mimicking) laughter, crucial to showing affiliation and bonding [2], to compensate for the experimental setting, i.e., one partner in an fMRI machine, communication via computer-mediation, etc. A similar trend was observed in [27], which reported children communicating with their parent via similar audiovisual means, i.e., interaction through a screen, produced mimicking laughter with increased temporal modulations in the 0-4 Hz range. Although the authors reported no differences in temporal modulations between laughter types for adults, mimicking laughter produced by parents interacting with their child were increased in comparison with other adults. Taking together these findings suggests low frequency temporal modulations are sensitive to differences between mimicking and non-mimicking laughter. This acoustic distinction observed between mimicking and non-mimicking laughter might prove important when considering future analysis of the neural correlates linked to laughter production. Although slightly more deviance was explained by the spectral modulation GAMM (93%), no significant differences between laughter types were observed. These findings are consistent with those reported in [27], where no significant differences between mimicking and non-mimicking laughter produced by adults were observed. It is possible that mimicking laughter is more strongly associated with modulations of amplitude envelopes rather than changes in spectral features, such as formants and fundamental frequency, which have been correlated with arousal in humans [45, 46, 43] and in non-human animals [47, 48]. Although the frequency of arousal levels were similar to those observed in face-to-face interactions, it is possible that, as the dyads were composed of strangers (see: [45, 46]) conversing in unnatural conditions, the expression of high arousal states might have been constrained.

5. Conclusions

The current study examined laughter production in a corpus of semi-structured conversation while one interlocutor underwent fMRI scanning. We conducted laughter annotations to study laughter distribution, its pragmatic use, and to perform acoustic comparisons between mimicking and non-mimicking laughter. These analyses and results are foundational in order to conduct successive neuro-imaging analysis exploring the neural correlates of different types of laughter in spontaneous conversations. The following were major take-away messages from the study: (1) despite the peculiar condition in which the conversations took place, laughter behaviours are comparable to that observed in face-to-face corpora, (2) in comparison to non-mimicking laughter, mimicking laughter had increased temporal modulations below 2 Hz, however, no significant difference were observed between the two in the spectral modulation domain, and (3) significant individual variability emerges in laughter production. These observations will inform future analysis of neuro-physiological correlates relating of laughter perception and production.

6. Acknowledgements

This work was supported by the ILCB (ANR-16-CONV-0002) and the Excellence Initiative of Aix-Marseille University (A*MIDEX AAP-ID-17-46-170301-11.1). We are grateful to Charlie Hallart for her work on the inter-annotator agreement.

7. References

- [1] S. K. Scott, N. Lavan, S. Chen, and C. McGettigan, "The social life of laughter," *Trends in cognitive sciences*, vol. 18, no. 12, pp. 618–620, 2014.
- [2] M. Davila-Ross and E. Palagi, "Laughter, play faces and mimicry in animals: evolution and social functions," *Philosophical Transactions of the Royal Society B*, vol. 377, no. 1863, p. 20210177, 2022.
- [3] R. Dunbar, "Laughter and its role in the evolution of human social bonding," *Philosophical Transactions of the Royal Society B*, vol. 377, no. 1863, p. 20210176, 2022.
- [4] J. Ginzburg, C. Mazzocconi, and Y. Tian, "Laughter as language," *Glossa*, vol. In resubmission, 2020.
- [5] P. Glenn, *Laughter in interaction*. Cambridge University Press, 2003, vol. 18.
- [6] J. Raclaw and C. E. Ford, "Laughter and the management of divergent positions in peer review interactions," *Journal of pragmatics*, vol. 113, pp. 1–15, 2017.
- [7] C. Mazzocconi, Y. Tian, and J. Ginzburg, "What's your laughter doing there? A taxonomy of the pragmatic functions of laughter," *IEEE Transactions on Affective Computing*, 2020.
- [8] G. Lombardi, M. Gerbella, M. Marchi, A. Sciutti, G. Rizzolatti, and G. Di Cesare, "Investigating form and content of emotional and non-emotional laughing," *Cerebral Cortex*, vol. 33, no. 7, pp. 4164–4172, 2023.
- [9] C. McGettigan, E. Walsh, R. Jessop, Z. Agnew, D. Sauter, J. Warren, and S. Scott, "Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity," *Cerebral Cortex*, pp. 220–227, 2013.
- [10] D. P. Szameitat, B. Kreifelts, K. Alter, A. J. Szameitat, A. Sterr, W. Grodd, and D. Wildgruber, "It is not always tickling: distinct cerebral responses during perception of different laughter types," *Neuroimage*, vol. 53, no. 4, pp. 1264–1271, 2010.

- [11] E. Wattendorf, B. Westermann, K. Fiedler, E. Kaza, M. Lotze, and M. R. Celio, "Exploration of the neural correlates of ticklish laughter by functional magnetic resonance imaging," *Cerebral Cortex*, vol. 23, no. 6, pp. 1280–1289, 2012.
- [12] B. Wild, F. A. Rodden, W. Grodd, and W. Ruch, "Neural correlates of laughter and humour," *Brain*, vol. 126, no. 10, pp. 2121–2138, 2003.
- [13] M. N. Neely, E. Walter, J. M. Black, and A. L. Reiss, "Neural correlates of humor detection and appreciation in children," *Journal of Neuroscience*, vol. 32, no. 5, pp. 1784–1790, 2012.
- [14] F. Caruana, P. Avanzini, F. Gozzo, S. Francione, F. Cardinale, and G. Rizzolatti, "Mirth and laughter elicited by electrical stimulation of the human anterior cingulate cortex," *Cortex*, vol. 71, pp. 323–331, 2015.
- [15] B. Rauchbauer, Y. Hmamouche, B. Bigi, L. Prevot, M. Ochs, and T. Chaminade, "Multimodal corpus of bidirectional conversation of human-human and human-robot interaction during fmri scanning," in *Proceedings of LREC*, 2020, pp. 668–675.
- [16] J.-A. Bachorowski and M. J. Owren, "Vocal acoustics in emotional intelligence," *Emotions and social behavior. The wisdom in feeling: Psychological processes in emotional intelligence*, pp. 11–36, 2002.
- [17] D. P. Szameitat, K. Alter, A. J. Szameitat, D. Wildgruber, A. Sterr, and C. J. Darwin, "Acoustic profiles of distinct emotional expressions in laughter," *The Journal of the Acoustical Society of America*, vol. 126, no. 1, pp. 354–366, 2009.
- [18] N. Lavan, S. K. Scott, and C. McGettigan, "Laugh like you mean it: Authenticity modulates acoustic, physiological and perceptual properties of laughter," *Journal of Nonverbal Behavior*, vol. 40, no. 2, pp. 133–149, 2016.
- [19] A. Wood, J. Martin, and P. Niedenthal, "Towards a social functional account of laughter: Acoustic features convey reward, affiliation, and dominance," *PloS one*, vol. 12, no. 8, 2017.
- [20] T. Chi, Y. Gao, M. C. Guyton, P. Ru, and S. Shamma, "Spectrotemporal modulation transfer functions and speech intelligibility," *The Journal of the Acoustical Society of America*, vol. 106, no. 5, pp. 2719–2732, 1999.
- [21] T. M. Elliott and F. E. Theunissen, "The modulation transfer function for speech intelligibility," *PLOS Computational Biology*, vol. 5, no. 3, pp. 1–14, 03 2009. [Online]. Available: <https://doi.org/10.1371/journal.pcbi.1000302>
- [22] A. Flinker, W. Doyle, A. Mehta, O. Devinsky, and D. Poeppel, "Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries," *Nature Human Behaviour*, vol. 3, 04 2019.
- [23] L. Moro-Velázquez, J. Gómez-García, J. Godino Llorente, and G. Andrade-Miranda, "Modulation spectra morphological parameters: A new method to assess voice pathologies according to the grbas scale," *BioMed Research International*, vol. 2015, 10 2015.
- [24] A. Marczyk, B. O'Brien, P. Tremblay, V. Woisard, and A. Ghio, "Correlates of vowel clarity in the spectrotemporal modulation domain: Application to speech impairment evaluation," *The Journal of the Acoustical Society of America*, vol. 152, pp. 2675–2691, 2022.
- [25] B. Ludusan and P. Wagner, "Speech, laughter and everything in between: A modulation spectrum-based analysis," 05 2020, pp. 995–999.
- [26] —, "An Evaluation of Manual and Semi-Automatic Laughter Annotation," in *Proc. Interspeech 2020*, 2020, pp. 621–625.
- [27] C. Mazzocconi, B. O'Brien, K. El Haddad, K. Bodur, and A. Fourtassi, "Differences between mimicking and non-mimicking laughter in child-caregiver conversation: A distributional and acoustic analysis," in *Proceedings of CogSci*, 2023.
- [28] F. Caruana, "Two simulation systems in the human frontal cortex? disentangling between motor simulation and emotional mirroring using laughter," *Cortex*, vol. 148, pp. 215–217, 2021.
- [29] H. Brugman and A. Russel, "Annotating multi-media/ multi-modal resources with elan," in *LREC*, 2004.
- [30] K. El Haddad, S. N. Chakravarthula, and J. Kennedy, "Smile and laugh dynamics in naturalistic dyadic interactions: Intensity levels, sequences and roles," in *2019 International Conference on Multimodal Interaction*, 2019, pp. 259–263.
- [31] C. Mazzocconi and J. Ginzburg, "A longitudinal characterisation of typical laughter development in mother-child interaction from 12 to 36 months: formal features and reciprocal responsiveness," *Journal of Nonverbal Behavior*, 2022.
- [32] A. Lücking, S. Ptock, and K. Bergmann, "Assessing agreement on segmentations by means of staccato, the segmentation agreement calculator according to thomann," in *International Gesture Workshop*. Springer, 2011, pp. 129–138.
- [33] K. Krippendorff, *Content analysis: An introduction to its methodology*. Sage, 2012.
- [34] T. Sainburg, M. Thielk, and T. Q. Gentner, "Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires," *PLoS computational biology*, vol. 16, no. 10, p. e1008228, 2020.
- [35] N. C. Singh and F. E. Theunissen, "Modulation spectra of natural sounds and ethological theories of auditory processing," *The Journal of the Acoustical Society of America*, vol. 114, no. 6, pp. 3394–3411, 2003.
- [36] E. Thoret, B. Caramiaux, P. Depalle, and S. Mcadams, "Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre," *Nature Human Behaviour*, vol. 5, 03 2021.
- [37] S. Wood, *Generalized Additive Models: An Introduction With R*, 01 2006, vol. 66.
- [38] M. Wieling, "Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between 11 and 12 speakers of english," *Journal of Phonetics*, vol. 70, 06 2018.
- [39] J. van Rij, M. Wieling, R. H. Baayen, and D. H. van Rijn, "itsadug: Interpreting time series and autocorrelated data using gamms," 2015.
- [40] 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.
- [41] S. Fuchs and T. Rathcke, "Laugh is in the air?" in *Proceedings of Laughter Workshop 2018, Paris, France*, 2018, pp. 21–24.
- [42] S. Petridis, B. Martinez, and M. Pantic, "The mahnob laughter database," *Image and Vision Computing*, vol. 31, no. 2, pp. 186–202, 2013.
- [43] M. Smoski and J.-A. Bachorowski, "Antiphonal laughter between friends and strangers," *Cognition & Emotion*, vol. 17, no. 2, pp. 327–340, 2003.
- [44] A.-L. Giraud and D. Poeppel, "Cortical oscillations and speech processing: Emerging computational principles and operations," *Nature neuroscience*, vol. 15, pp. 511–7, 03 2012.
- [45] G. A. Bryant and C. A. Aktipis, "The animal nature of spontaneous human laughter," *Evolution and Human Behavior*, vol. 35, no. 4, pp. 327–335, 2014.
- [46] C. Oveis, A. Spectre, P. K. Smith, M. Y. Liu, and D. Keltner, "Laughter conveys status," *Journal of Experimental Social Psychology*, vol. 65, pp. 109–115, 2016.
- [47] J. W. Schwartz, M. M. Sanchez, and H. Gouzoules, "Vocal expression of emotional arousal across two call types in young rhesus macaques," *Animal Behaviour*, vol. 190, pp. 125–138, 2022.
- [48] K. Pisanski, V. Cartei, C. McGettigan, J. Raine, and D. Reby, "Voice modulation: A window into the origins of human vocal control?" *Trends in Cognitive Sciences*, vol. 20, no. 4, pp. 304–318, 2016.