HAL
open science

A Measure of the Smiling Synchrony in the Conversational Face-to-face Interaction Corpus PACO-CHEESE

Stéphane Rauzy, Mary Amoyal, Béatrice Priego-Valverde

To cite this version:

Stéphane Rauzy, Mary Amoyal, Béatrice Priego-Valverde. A Measure of the Smiling Synchrony in the Conversational Face-to-face Interaction Corpus PACO-CHEESE. Workshop SmiLa, Language Resources and Evaluation Conference, LREC 2022, Jun 2022, Marseille, France. hal-03675630

HAL Id: hal-03675630

https://hal.science/hal-03675630v1

Submitted on 23 May 2022

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.

A Measure of the Smiling Synchrony in the Conversational Face-to-face Interaction Corpus PACO-CHEESE

Stéphane Rauzy, Mary Amoyal, Béatrice Priego-Valverde
Aix Marseille Université, CNRS, Laboratoire Parole et Langage UMR 7309
5 avenue Pasteur, BP 80975, 13604 Aix-en-Provence, FRANCE
{stephane.rauzy, mary.amoyal, beatrice.priego-valverde}@univ-amu.fr

Abstract

The smiling synchrony of the French audio-video conversational corpora "PACO" and "Cheese!" is investigated. The two corpora merged altogether last 6 hours and are made of 25 face-to-face dyadic interactions annotated following the 5 levels Smiling Intensity Scale proposed by Gironzetti et al. (2016). After introducing new indicators for characterizing synchrony phenomena, we find that almost all the 25 interactions of PACO-CHEESE show a strong and significant smiling synchrony behavior. We investigate in a second step the evolution of the synchrony parameters throughout the interaction. No effect is found and it appears rather that the smiling synchrony is present at the very start of the interaction and remains unchanged throughout the conversation.

Keywords: Synchrony, convergence, smiling behavior, spontaneous interaction, face-to-face audio-video corpus

1. Introduction

It is now well established that participants involved in a conversational face-to-face activity exhibit similar patterns, the phenomenon having received various names including among them accommodation (Giles et al., 1991), entrainment (Brennan and Clark, 1996), alignment (Pickering and Garrod, 2004), convergence (Pardo, 2013), mimicry (Pentland, 2008) and synchrony (Edlund et al., 2009). This interactional behavior have been observed in different domains ranging from lexical adaptation (Brennan and Clark, 1996), pronunciation (Aubanel and Nguyen, 2010), prosodic patterns (De Looze et al., 2011), syntactic structures (Pickering and Ferreira, 2008) to facial expressions (Seibt et al., 2015). Herein we will focus on the synchrony analysis of smiles considered as interactive facial gesture (Bavelas and Gerwing, 2007). Convergence issues will be also examined by exploring the evolution of the synchrony parameters throughout the interaction. The synchrony of smiles and laughter have been previously addressed (Heerey and Crossley, 2013; Gironzetti et al., 2016b; Mui et al., 2018; El Haddad et al., 2019; Arnold and Piotr, 2020)). Our contribution is herein twofold: we first propose some new indicators for measuring synchrony and secondly we analyse the smiling synchrony of the PACO-CHEESE corpus.

2. Measuring synchrony

Synchrony can be essentially defined as the property for the participants to show temporally similar behaviours (Edlund et al., 2009). Various methods for measuring synchrony have been proposed in the literature depending on the timescale at which this similarity takes place. Given a variable observed along the time line (e.g. pitch, speech rate, smile intensity, ...), the Pearson's correlation between the two participant's time series is for example a popular indicator of synchrony (Edlund et al., 2009; De Looze and Rauzy, 2011; De Looze et al., 2014). If the match between the two series is not instantaneous but rather presents some time shift, Time-Lagged Cross Correlation techniques can be applied with benefits (Golland et al., 2019). For more complex time dependencies, alternative measurements relying on cross-spectral and relative phase approaches (Schmidt et al., 2012), mimicry detection (Feese et al., 2012; El Haddad et al., 2019) or cross-recurrence quantification analysis (Main et al., 2016; Paxton and Dale, 2017) have been build up.

In De Looze and Rauzy (2011), the description of synchrony phenomena was tackled by drawing an analogy with the coupled oscillators model found in Physics. The model describes the dynamics of two oscillators (say two pendulums) coupled together by a spring. The spring plays here the role of a force coupling the respective oscillating trajectory x_1 and x_2 of the two pendulum masses. The general solution of the problem let emerge two oscillating *normal modes* associated with the sum and the difference of the trajectories:

$$x_{\text{sum}} = x_1 + x_2 \quad ; \quad x_{\text{diff}} = x_1 - x_2 \qquad (1)$$

The symmetric mode x_{sum} describes the motion of the system as a whole and is characterized by an oscillatory period T_{sum} determined by the two pendulum periods in absence of coupling. The asymmetric mode x_{diff} accounts for the internal oscillations of the two pendulums system and its characteristic period T_{diff} is necessarily shorter than T_{sum} if the system is coupled. This remark leads us to define a *coupling factor* k_c as:

$$k_c = \log(T_{\text{sum}}/T_{\text{diff}}) \quad ; \quad k_c > 0 \Longrightarrow \text{Coupling} \qquad (2)$$

This criterion allows in practice to detect the presence of a coupling between the two participants.

The dynamics of the coupled system is determined by a linear combination of the two oscillatory normal modes. It accounts for various coupling behaviours depending on the value of the amplitudes A_{sum} and A_{diff}

respectively associated with the symmetric and asymmetric modes. Pure synchrony corresponds for example to the case $A_{\text{diff}} = 0$ (i.e. $x_1 = x_2$) whereas $A_{\text{sum}} = 0$ depicts the situation of pure anti-synchrony (i.e. the pendulums are forced to move in the opposite direction). The degree of synchrony can be measured by evaluating the *coefficient of synchrony* ρ_S:

$$\rho_S = \frac{var(x_{\text{sum}}) - var(x_{\text{diff}})}{var(x_{\text{sum}}) + var(x_{\text{diff}})} \quad (3)$$

where the variance of the oscillating time series $var(x)$ is proportional to the square of its amplitude (e.g. $var(x_{\text{sum}}) \propto A^2_{\text{sum}}$). The coefficient of synchrony ρ_S varies from -1 to 1 and is indeed close to the Pearson's correlation coefficient $\rho(x_1, x_2)$ of the two observed participant's time series.

Estimation of the periods T_{diff} and T_{sum}

We denote by $W(t; x, \tau)$ the smoothed version of the time series $x \equiv x(t)$ smoothed at time scale τ. For example $W(t; x, \tau)$ can be the result of a Simple Moving Average operation with a window of size τ. Smoothing works herein as a low pass filter which removes from the signal fluctuations with frequency higher than the cut-off frequency. The variance $var(W(t; x, \tau))$, which measures the energy of the smoothed time series, varies from $var(x)$ when $\tau = 0$ (since $W(t; x, \tau = 0) \equiv x(t)$) to 0 when τ approaches infinity (in practice when τ is greater that the largest fluctuation present in the signal). We define the quantity $F(x, \tau)$ as the ratio of energy contained in the fluctuations with characteristic time scale lower than the smoothing time scale τ:

$$F(x, \tau) = 1 - \frac{var(W(t; x, \tau))}{var(x)} \quad (4)$$

The ratio $F(x, \tau)$ varies from 0 at time scale $\tau = 0$ and approaches 1 when τ is large enough. It represents the cumulative distribution function of the energy up to the time scale τ. The energy contained between two time scales τ_{inf} and τ_{sup} is given by $E(\tau_{\text{inf}}, \tau_{\text{sup}}) = F(x, \tau_{\text{sup}}) - F(x, \tau_{\text{inf}})$ and the energy density can be obtained by differentiating the cumulative energy distribution $F(x, \tau)$.

The characteristic periods T_{diff} and T_{sum} associated with the two oscillating modes of the coupled system will be estimated from the energy distribution function of the two series. One can choose for example the time scale corresponding to the maximal peak of energy density as the characteristic period of the mode. The choice of the appropriate estimator will eventually depends on the form of the energy distribution function.

3. The PACO-CHEESE corpus

The PACO-CHEESE corpus results of the merge of the two French audio-video conversational corpora "PACO" (Amoyal et al., 2020) and "Cheese!" (Priego-Valverde et al., 2020; Priego-Valverde et al., 2018). The "Cheese!" corpus is composed of 11 dyadic interactions lasting between 15 to 20 minutes each. The two participants were recorded in an anechoic room with separate microphone and camera. The participants were asked to read each other a canned joke before freely conversing during the rest of the interaction. The corpus "PACO" contains 15 conversations and has been collected by following the same protocol as designed for "Cheese!". The main contrast between the two corpora is that the "Cheese!" participants were acquainted since they were students in the same class whereas "PACO" participants did not know each other. This condition is intended in practice to control the relationship factor between the two interlocutors (i.e. "acquainted" vs "initial interaction").

The smile intensity annotations

Smiles have been annotated thanks to the "Smiling Intensity Scale" (SIS) (Gironzetti et al., 2016a). The 5 levels of the scale start with level 0 (neutral face), contain three gradual intensities of smile (from 1 to 3) and end with level 4 encoding laughter. Each smile intensity category involves a specific combination of Action Units (AUs) detailed by the Facial Action Coding System (FACS) (Ekman and Friesen, 1978). The full description of the annotation procedure as well as a discussion concerning the benefits to adopt the 5 levels SIS system can be found in (Amoyal et al., 2020; Rauzy and Amoyal, 2020).

4. The smiling synchrony in PACO-CHEESE

4.1. Global synchrony

We investigate in this section the global smiling synchrony at the scale of the interaction for the 25 conversations of the PACO-CHEESE corpus. The starting canned jokes passage (see section 3) have been removed by cutting the first 3 minutes of each conversation.

For each interaction, the two times series x_1 and x_2 of the participants are extracted according to the smile intensity annotations presented section 3. The sum and the difference mentioned equation 1 are formed. An illustration of the trajectories of the 4 time series is presented figure 1.

The characteristic periods T_{sum} and T_{diff} are afterwards estimated. The top panel of figure 2 shows for the 4 time series the cumulative distribution function (CDF) of energy as defined equation 4. For the diad named ACMZ, the x_1, x_2 and x_{sum} present similar CDFs, with a median time scale around 9 seconds. The energy CDF of the asymmetric mode x_{diff} is by contrast shifted towards the low timescales (i.e. the median period is around 4 seconds).

A thorough analysis of the 25 PACO-CHEESE interactions reveals that the energy density distribution of the smile time series is well described by a lognormal distribution. The bottom panel of figure 2 presents the fitted lognormal models for the 4 time series of the ACMZ interaction. Our estimates of the characteristic periods T_{sum} and T_{diff} mentioned equation 2 will finally correspond to the peaks of the fitted energy densities for x_{sum} and x_{diff}.

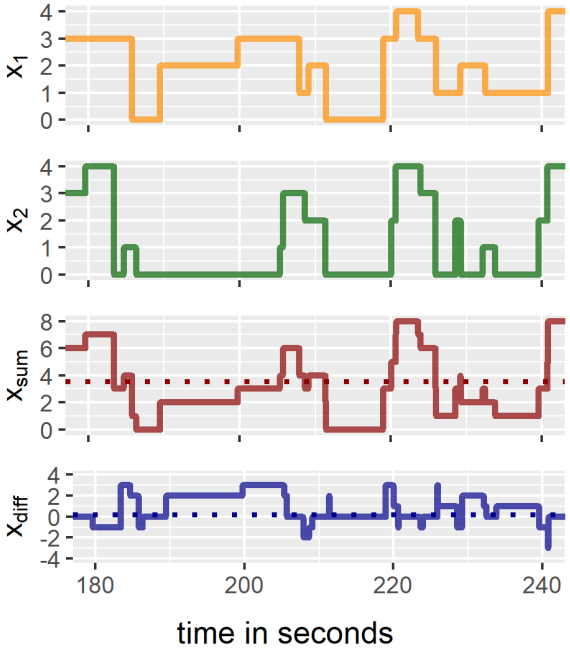

Figure 1: CHEESE-ACMZ interaction: A 60 seconds extract of the smile intensity time series (SIS encoded) for the two participants (panels x_1 and x_2) and their corresponding x_{sum} and x_{diff} variations.

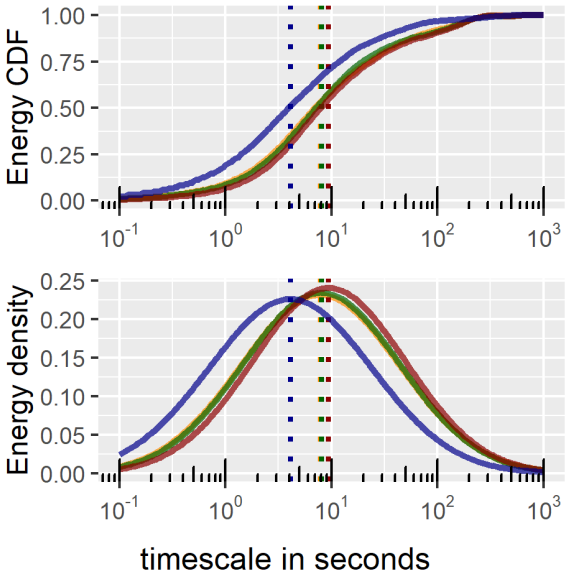

Figure 2: CHEESE-ACMZ interaction: (top panel) the cumulative distribution function of energy for the time series $x1$ (orange), x_2 (green), x_{sum} (red) and x_{diff} (blue) in function of the cut-off timescale (logarithmic scale). (bottom panel) The corresponding energy density models assuming a lognormal energy distribution.

At this stage, we observe that the asymmetric period T_{diff} is half as long as its symmetric counterpart T_{sum}. According to the criterion introduced equation 2, it suggests that the smile intensities of the ACMZ participants are in fact coupled. It remains however to show that this discrepency is statistically significant.

Standard errors associated to the estimates of the periods, the coupling factor k_c and the coefficient of synchrony are obtained by applying a random pairing strategy (Golland et al., 2019). A random pair of participants is created by pairing two participants not belonging to the same interaction. By construction there is no coupling for this fake interaction. Within the uncertainties due to statistical fluctuations, the values estimated from the fake interaction is thus the one expected for the no coupling condition.

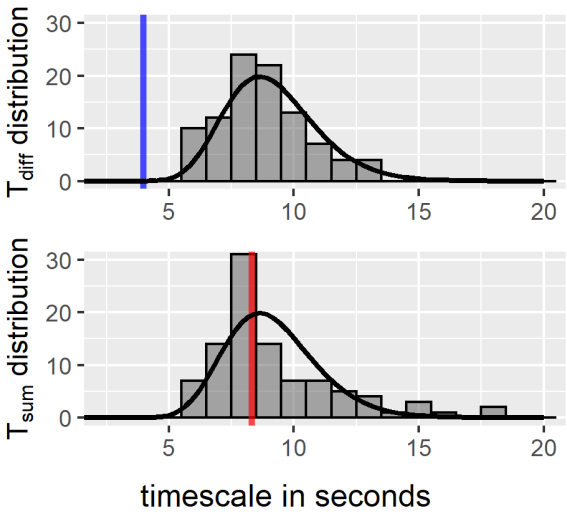

Figure 3: CHEESE-ACMZ interaction: (top panel) Histogram of the T_{diff} period estimate for the 96 random pairs. The black curve is a lognormal density fitted on the histogram distribution. Blue vertical line around 4 s indicates the real T_{diff} for the ACMZ pair. (bottom panel) Same plot for the T_{sum} period estimate.

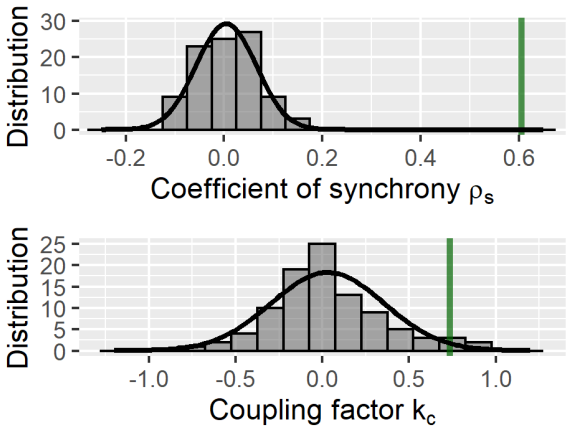

Figure 4: CHEESE-ACMZ interaction: Same plot as in figure 3 for the estimate of the coefficient of synchrony ρ_S (top panel) and the coupling factor k_C (bottom panel).

For each of the 25 interactions of the PACO-CHEESE corpus, we formed the 2x48 random pairs and computed for each pair the parameter estimates. The results are illustrated figures 3 and 4 for the ACMZ interaction. The distribution of the estimates for the random pairs allows to compute the standard deviation associated with the estimator and the expected value in the no coupling condition. Figure 3 shows that T_{sum} and

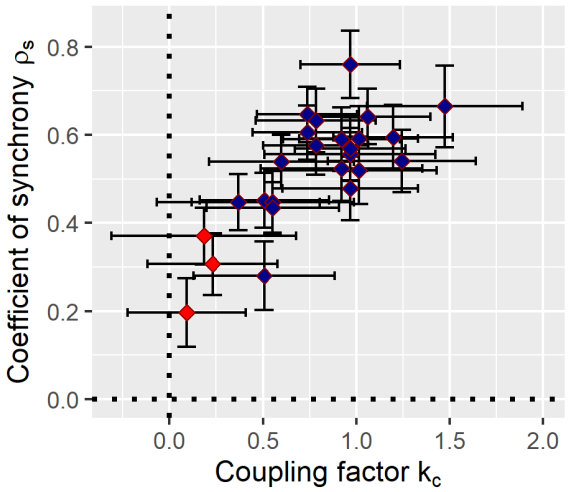

Figure 5: The coefficient of synchrony and the coupling factor for the 25 interactions of the PACO-CHEESE corpus. Red points denote T_{diff} periods greater than 9 seconds.

T_{diff} are expected to be identical in the absence of coupling and that the value of T_{diff} for the real pair AC-MZ (the blue vertical line) is clearly shorter than the one expected by chance. One can also see figure 4 that in absence of coupling the expected values for the coupling factor and the coefficient of synchrony are centered on 0 and that the observed values for the true pair AC-MZ are far above this threshold within the standard deviation.

The final result is presented figure 5 for the 25 interactions. The 1σ error bars are computed for each interaction using the random pairs strategy mentioned above. After removing the 3 outlying interactions with T_{diff} greater than 9 seconds (the red points on the graph), the mean T_{diff} is 5.41 s with a standard dispersion of 1.23 s to compare with 13.65 s and 5.66 s for the T_{sum} period. The mean T_{diff} and T_{max} define respectively timescales below which the participants are locally not aligned and above which the synchrony is observed. For all the interactions of PACO-CHEESE the participants show a strong synchrony in their smiling behaviour, this property is revealed both by the significant measurements of the coupling factor and the coefficient of synchrony.

4.2. Synchrony and evolution

Since smiling synchrony appears as a general behaviour adopted by participants, the question arises whether the synchrony strength evolves throughout the conversation or instead remains constant. We sliced each interaction in 5 time windows of equal duration, from a starting time at 180 seconds to the end of the interaction (the mean bin duration is 182 seconds containing around 30 smile changes in average). Each window bins contains thus several periods of the T_{sum} symmetric mode which warrants in practice the safe evaluation of the synchrony parameters.

Figure 6 illustrates the variation of the synchrony measurements throughout the conversation. For each bin position, we computed the averages and 1σ error bars over the interactions of each corpus. The estimates of the coupling factor and the coefficient of synchrony do not reveal any evolution trend. For both corpora, it appears that the smiling synchrony is rather present at the very start of the interaction and remains unchanged throughout the conversation.

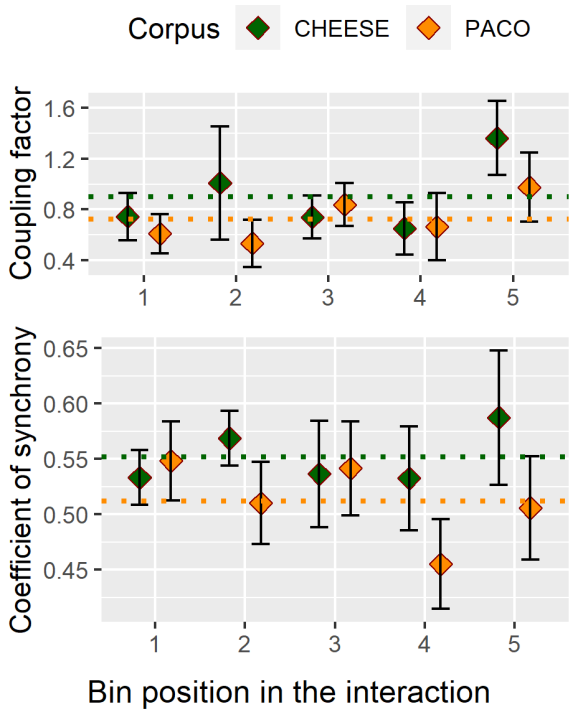

Figure 6: Variation of the synchrony parameters in function of the window bin position in the conversation. The averages and 1σ error bars are computed individually for each corpus (11 interactions for CHEESE and 12 for PACO).

5. Conclusions

We performed a synchrony analysis of the smile annotations of the PACO-CHEESE corpus encoded following the Smiling Intensity Scale. We introduced new indicators allowing to define two timescales associated with the synchrony phenomenon, one period around 5 seconds below which participant's smiling are locally not aligned and a second period around 14 seconds above which the similarity between the two smiling behaviors takes place. That period also settles in practice the minimal timescale required to study smiling synchrony. As expected from previous study on face-to-face conversations (Heerey and Crossley, 2013), the results reveal that almost all the 25 interactions of PACO-CHEESE show a strong and significant smiling synchrony behavior. In a second step, the question of the convergence was investigated by measuring the evolution of the synchrony parameters throughout the interaction. We did not found such an effect, the smiling synchrony is indeed detected at the outset of the conversation and its strength does not increase along time.

6. Bibliographical References

Amoyal, M., Priego-Valverde, B., and Rauzy, S. (2020). PACO: A corpus to analyze the impact of common ground in spontaneous face-to-face interaction. In *Language Resources and Evaluation Conference*, Marseille, France, May.

Arnold, A. J. and Piotr, W. (2020). Smile (but only deliberately) though your heart is aching: Loneliness is associated with impaired spontaneous. *Social Neuroscience*.

Aubanel, V. and Nguyen, N. (2010). Automatic recognition of regional phonological variation in conversational interaction. *Speech Communication*, 52:577–586, 06.

Bavelas, J. B. and Gerwing, J. (2007). Conversational hand gestures and facial displays in face-to-face dialogue. *Social communication*, pages 283–308.

Brennan, S. and Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22:1482–1493.

De Looze, C. and Rauzy, S. (2011). Measuring speakers' similarity in speech by means of prosodic cues: methods and potential. In *Interspeech 2011*, Florence, Italy.

De Looze, C., Oertel, C., Rauzy, S., and Campbell, N. (2011). Measuring dynamics of mimicry by means of prosodic cues in conversational speech. In *ICPhS 2011*, Hong-Kong, China.

De Looze, C., Scherer, S., Vaughan, B., and Campbell, N. (2014). Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Commun.*, 58:11–34, mar.

Edlund, J., Herldner, M., and Hirschberg, J. (2009). Pause and gap length in face-to-face interaction. In *Interspeech*, pages 2779–2782.

Ekman, P. and Friesen, W. V. (1978). *Facial Action Coding System: Manual*, volume vol. 1-2. Consulting Psychologists Press.

El Haddad, K., Chakravarthula, S. N., and Kennedy, J. (2019). Smile and laugh dynamics in naturalistic dyadic interactions: Intensity levels, sequences and roles. In *2019 International Conference on Multimodal Interaction*, pages 259–263. ACM.

Feese, S., Arnrich, B., Tröster, G., Meyer, B., and Jonas, K. (2012). Quantifying behavioral mimicry by automatic detection of nonverbal cues from body motion. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, pages 520–525.

Giles, H., Coupland, N., and Coupland, J., (1991). *Accommodation theory: Communication, context, and consequence*, page 1–68. Studies in Emotion and Social Interaction. Cambridge University Press.

Gironzetti, E., Attardo, S., and Pickering, L. (2016a). Smiling, gaze, and humor in conversation: A pilot study. In Leonor Ruiz-Gurillo, editor, *Metapragmatics of Humor: Current research trends*, pages 235 – 254.

Gironzetti, E., Pickering, L., Huang, M., Zhang, Y., Menjo, S., and Attardo, S. (2016b). Smiling synchronicity and gaze patterns in dyadic humorous conversations. *HUMOR*, 29(2):301–324.

Golland, Y., Mevorach, D., and Levit-Binnun, N. (2019). Affiliative zygomatic synchrony in co-present strangers. *Scientific Reports*, 9(1):3120, February.

Heerey, E. A. and Crossley, H. M. (2013). Predictive and reactive mechanisms in smile reciprocity. *Psychological Science*, 24(8):1446–1455.

Main, A., Paxton, A., and Dale, A. (2016). An exploratory analysis of emotion dynamics between mothers and adolescents during conflict discussions. *Emotion*, 16(6):913–928.

Mui, P. H. C., Goudbeek, M. B., Roex, C., Spierts, W., and Swerts, M. G. J. (2018). Smile mimicry and emotional contagion in audio-visual computer-mediated communication. *Frontiers in Psychology*, 9.

Pardo, J. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology*, 4.

Paxton, A. and Dale, R. (2017). Interpersonal movement synchrony responds to high- and low-level conversational constraints. *Frontiers in Psychology*, 8.

Pentland, A. S. (2008). *Honest Signals: How They Shape Our World*. The MIT Press.

Pickering, M. and Ferreira, V. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3):427–459, May.

Pickering, M. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *The Behavioral and brain sciences*, 27:169–90; discussion 190, 05.

Priego-Valverde, B., Bigi, B., Attardo, S., Pickering, L., and Gironzetti, E. (2018). Is smiling during humor so obvious? A cross-cultural comparison of smiling behavior in humorous sequences in American English and French interactions. *Intercultural Pragmatics*.

Priego-Valverde, B., Bigi, B., and Amoyal, M. (2020). "Cheese!": A corpus of face-to-face French interactions. A case study for analyzing smiling and conversational humor. In *Language Resources and Evaluation Conference*, pages 460–468, Marseille, France.

Rauzy, S. and Amoyal, M. (2020). SMAD: A tool for automatically annotating the smile intensity along a video record. In *HRC2020, 10th Humour Research Conference*.

Schmidt, R. C. P., Morr, S., Fitzpatrick, P., and Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior*, 36(4):263–279.

Seibt, B., Mühlberger, A., Likowski, K., and Weyers, P. (2015). Facial mimicry in its social setting. *Frontiers in Psychology*, 6.