



HAL
open science

Towards investigating gaze and laughter coordination in socially interactive agents

Vladislav Maraev, Chiara Mazzocconi, Christine Howes, Catherine Pelachaud

► **To cite this version:**

Vladislav Maraev, Chiara Mazzocconi, Christine Howes, Catherine Pelachaud. Towards investigating gaze and laughter coordination in socially interactive agents. The Eleventh International Conference on Human-Agent Interaction (HAI 2023), Dec 2023, Gothenburg (Sweden), Sweden. 10.1145/3623809.3623968 . hal-04415493

HAL Id: hal-04415493

<https://hal.science/hal-04415493>

Submitted on 24 Jan 2024

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.



Towards Investigating Gaze and Laughter Coordination in Socially Interactive Agents

Vladislav Maraev
vladislav.maraev@gu.se
University of Gothenburg
Sweden

Chiara Mazzocconi
chiara.mazzocconi@live.it
Laboratoire Parole et Langage, CNRS
Institute of Language Communication and the Brain
Aix-Marseille University
France

Christine Howes
christine.howes@gu.se
University of Gothenburg
Sweden

Catherine Pelachaud
catherine.pelachaud@sorbonne-universite.fr
Sorbonne Université, CNRS, ISIR
France

ABSTRACT

Gaze and laughter play a crucial role in managing miscommunication and coordinating social interactions. We hypothesise that models of laughter and gaze coordination in human dialogue extend to virtual entities. This paper describes methodology of the future experiment which involves a socially interactive agent (SIA) that incorporates previous theoretical findings.

CCS CONCEPTS

• **Human-centered computing** → *User models*.

KEYWORDS

gaze, laughter, virtual agents, socially interactive agents

ACM Reference Format:

Vladislav Maraev, Chiara Mazzocconi, Christine Howes, and Catherine Pelachaud. 2023. Towards Investigating Gaze and Laughter Coordination in Socially Interactive Agents. In *International Conference on Human-Agent Interaction (HAI '23)*, December 04–07, 2023, Gothenburg, Sweden. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3623809.3623968>

1 INTRODUCTION

Socially interactive agents (SIAs) are autonomous agents which have physical robotic bodies or virtual appearances and communicate both verbally and non-verbally with humans and other agents. With advances in computational linguistics and computer vision, SIAs that are able to hold conversations that are engaging and human-like have recently gained momentum.

There is some understanding of how humans use laughter and gaze in dialogue, and theoretical models are derived from humans. Nevertheless it is unclear how this knowledge extends to virtual entities. In a previous corpus study, [18] established that social incongruity laughs (i.e. laughs not related to humour, but rather aimed

at smoothing dialogue acts and conversations) are accompanied by distinct gaze patterns. In this paper we outline the forthcoming perception study of SIA's laughter and gaze behaviours associated with social incongruities.

2 BACKGROUND

2.1 Gaze and laughter in human dialogue

The role of gaze in maintaining conversational flow and coordinating dialogue acts is crucial in face-to-face interactions. While many works have argued for the importance of individual gaze for the fine regulation of turn-taking [6, 11], some scholars highlighted a lack of systematic relationship between gaze and turn-taking [3, 28, 33], proposing rather that gaze might function to elicit a response [2, 12]. Gaze patterns have also been found to differ depending on the speech act they accompany and on their pragmatic function [20, 27, 29]. Of interest in the study is also gaze aversion [14, 25, 26]. It has been proposed that gaze aversion could also be explained (or influenced) by social stress [31], with evidence from patients with social disorders [30]. Conversely, results from other studies suggest that cognitive load has the most impact on gaze aversion [5, 9].

The circumstances under which we can produce laughters are affected by the perceived context and expectations of the interlocutors regarding possible changes of context. [8] borrow the term *laughable* from Conversational Analysis [10] as denoting something that laughter *relates to* while being agnostic about whether it is humorous or not. [19] present a taxonomy of laughter functions from the perspective of the intended effect on the context and use it to annotate laughs in various corpora. It takes into account four types of incongruity that characterise a laughable and groups the functions accordingly. It is also necessary to mention that laughter has intrinsically important social effects, being crucial for bonding and managing relationships, while also being immensely influenced by social context [24].

2.2 Gaze and laughter in SIAs

SIAs benefit from a detailed analysis of multimodal input and output patterns observed during human-human interactions. Recently, there has been a growing research interest both on gaze and other

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

HAI '23, December 04–07, 2023, Gothenburg, Sweden

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0824-4/23/12.

<https://doi.org/10.1145/3623809.3623968>

non-verbal behaviours, especially in the Affective Computing community, for the implementation of SIAs which are more competent from a pragmatic perspective and able to process and produce appropriate emotional responses [1, 15, 22, 32]. [1] established a basis for a context-aware eye-gaze generator for a SIA. To develop an improved gaze generator we should isolate the significant events detected in the multimodal scene that impact the gaze. [15] discuss the interpersonal role of gaze in interaction to signal feedback and direct conversation flow which current SIAs still lack. In a dynamic environment, even state-of-the-art SIAs struggle to direct gaze attention to peripheral movements. SIAs should therefore employ social gaze not only for interpersonal interaction but also to possess human attention attributes so that its eyes and facial expression portray and convey appropriate distraction and engagement behaviours.

Several models have been proposed to generate laughter [4, 13] and to decide when a SIA should laugh. For example, [4] created a laughter behaviour controller to generate face and body motions from laughter audio. [13] created a listening agent that predicts when to smile or to laugh depending on its interlocutor’s behaviours. [34] proposed a neural network learning model based on an attention mechanism to compute where to place a given type of laughter. Several recent studies have investigated laughter synthesis [16, 21]. Overall, they highlight that laughter synthesis in different contexts serving different functions still remains to be addressed. For instance, how can social laughter which accompanies an apology be synthesised together with speech in a naturalistic way? Laughter detection is a more developed topic compared to synthesis. State-of-the-art laughter detection is based on machine learning techniques: from support vector machines to deep learning approaches [7].

3 THEORY AND METHOD

3.1 Laughter and gaze coordination

In [18] we designed a study to determine gaze patterns in association with laughter. We showed that gaze plays a role in laughter coordination. For some cases, we observed that gaze is used to synchronise the onsets and offsets of laughter. Another important finding is that different types of laughter are accompanied by different gaze patterns both by the person laughing (the *laugher*) and their partner. In the case of pleasant (humorous) incongruity, laughter the *laugher* is more likely to look at their partner before the onset of laughter, whereas after the onset, social incongruity more likely to involve gazing at the partner (see Figure 1 and example 3.1). These findings suggest that any future SIA should not address all laughs in a uniform fashion – they should be sensitive to the precise coordination between laughter and gaze that are critically dependent on the types of laughables.

- (A) Good Housekeeping Institute (GHI) corpus, Pair 15 (00:01:25). Gaze at partner is marked with “x”, gaze anywhere else with “-”.

```
-----xxxxxxxxxxxxxxxxxxxx-----
A: Shall we say... No. Ta- tasty <laughs>
B:                                     <laughs>
-----
```

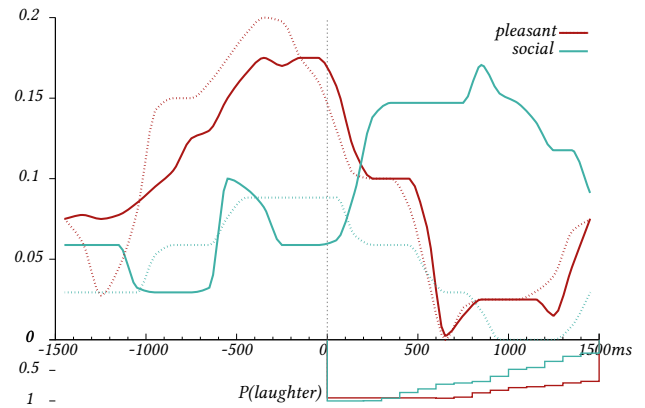


Figure 1: Probability of gaze at the interlocutor around the onset of laughter depending on laughable incongruity type. Solid line – laugher; dashed line – partner. The probability of laughter duration is shown at the bottom of the figure.

3.2 Method

The Greta platform¹ simulates SIAs able to communicate verbally and nonverbally. Given a set of intentions and emotions to be communicated, the platform instantiates them into sequences of synchronised nonverbal behaviors. It can be used to compute these multimodal behaviors when the SIA acts as a speaker or as a listener.

Regarding laughter in Greta, the main focus has been on synthesising the multimodal behaviours of “hilarious” laughter [4] and on measuring the impact of laughter on engagement and interaction quality during a human-agent interaction [17, 23]. Here, we will specifically focus on laughs associated with *social incongruities* [19]. The functions of such laughs include softening and trouble-telling (e.g., to accompany a criticism), benevolence induction (e.g., to accompany self-criticism), smoothing and showing sympathy.

We will select “ground truth” samples of laughing behaviours from [18] dataset (3 female-female interactions, 23 minutes of video recordings annotated for laughter functions and gaze). We will reproduce short fragments of face-to-face dialogues from the corpus study but with a SIA, through fine-grained orchestrations of behavioural patterns of an agent. The stimuli will be created by replacing a human in a dyad by a SIA. Subsequently, the reproduced samples will be modified with alternative (to [18]) and random gaze patterns. We will then conduct an observation study to measure perception of a SIA and to compare the understanding of social laughters by observers in contrast with ground truth. Observers will be asked about the naturalness of laughter behaviours and their perception of a SIA in terms of warmth and competence. The samples will be annotated by experts for pragmatic functions of laughter, based on [19] protocol, as for the ground truth. We will then compare ground truth annotations and annotations of reproduced behaviours (including modifications). The experiments will be guided by the following research questions:

¹Greta is available on Github: <https://github.com/isir/greta>

- (1) How agents are perceived in two conditions – reproduced behaviour vs. modified behaviour.
- (2) How pragmatic functions of laughter are perceived in a virtual agent compared to ground truth.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council (VR 2023-00358).

REFERENCES

- [1] Gérard Bailly, Stephan Raidt, and Frédéric Elisei. 2010. Gaze, conversational agents and face-to-face communication. *Speech Communication* 52 (06 2010), 598–612. <https://doi.org/10.1016/j.specom.2010.02.015>
- [2] Janet Beavin Bavelas, Linda Coates, and Trudy Johnson. 2002. Listener responses as a collaborative process: The role of gaze. *Journal of Communication* 52, 3 (2002), 566–580.
- [3] Geoffrey W Beattie. 1978. Sequential temporal patterns of speech and gaze in dialogue. *Semiotica* 23, 1-2 (1978), 29–52.
- [4] Yu Ding, Jing Huang, and Catherine Pelachaud. 2017. Audio-driven laughter behavior controller. *IEEE Trans. on Affect. Comp.* 8, 4 (2017), 546–558.
- [5] Gwyneth Doherty-Sneddon and Fiona G Phelps. 2005. Gaze aversion: A response to cognitive or social difficulty? *Memory & cognition* 33, 4 (2005), 727–733.
- [6] Starkey Duncan. 1972. Some signals and rules for taking speaking turns in conversations. *Journal of personality and social psychology* 23, 2 (1972), 283.
- [7] Jon Gillick, Wesley Deng, Kimiko Ryokai, and David Bamman. 2021. Robust Laughter Detection in Noisy Environments. In *Proc. Interspeech 2021*. 2481–2485. <https://doi.org/10.21437/Interspeech.2021-353>
- [8] Jonathan Ginzburg, Ellen Breitholtz, Robin Cooper, Julian Hough, and Ye Tian. 2015. Understanding Laughter. In *Proceedings of the 20th Amsterdam Colloquium*. 137–146.
- [9] Arthur M Glenberg, Jennifer L Schroeder, and David A Robertson. 1998. Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition* 26, 4 (1998), 651–658.
- [10] Phillip Glenn. 2003. *Laughter in interaction*. Cambridge University Press, Cambridge.
- [11] Charles Goodwin. 1980. Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological inquiry* 50, 3-4 (1980), 272–302.
- [12] Marjorie Harness Goodwin and Charles Goodwin. 1986. Gesture and coparticipation in the activity of searching for a word. *Semiotica* 62, 1-2 (1986), 51–76.
- [13] Kevin El Haddad, Sandeep Nallan Chakravarthula, and James Kennedy. 2019. Smile and laugh dynamics in naturalistic dyadic interactions: Intensity levels, sequences and roles. In *2019 International Conference on Multimodal Interaction*. 259–263.
- [14] Kobin H Kendrick and Judith Holler. 2017. Gaze direction signals response preference in conversation. *Research on Language and Social Interaction* 50, 1 (2017), 12–32.
- [15] Jina Lee and Stacy Marsella. 2006. Nonverbal Behavior Generator for Embodied Conversational Agents. In *Intelligent Virtual Agents, 6th International Conference, IVA 2006, Marina Del Rey, CA, USA, August 21-23, 2006, Proceedings (Lecture Notes in Computer Science, Vol. 4133)*, Jonathan Gratch, Robert Michael Young, Ruth Aylett, Daniel Ballin, and Patrick Olivier (Eds.). Springer, 243–255. https://doi.org/10.1007/11821830_20
- [16] Hieu-Thi Luong and Junichi Yamagishi. 2021. LaughNet: synthesizing laughter utterances from waveform silhouettes and a single laughter example. *CoRR abs/2110.04946* (2021). arXiv:2110.04946 <https://arxiv.org/abs/2110.04946>
- [17] Maurizio Mancini, Beatrice Biancardi, Florian Pecune, Giovanna Varni, Yu Ding, Catherine Pelachaud, Gualtiero Volpe, and Antonio Camurri. 2017. Implementing and evaluating a laughing virtual character. *ACM Trans. on Internet Technology (TOIT)* 17, 1 (2017), 1–22.
- [18] Chiara Mazzocconi, Vladislav Maraev, Vidya Somashekarappa, and Christine Howes. 2021. Looking for Laughs: Gaze Interaction with Laughter Pragmatics and Coordination. In *Proceedings of the 2021 International Conference on Multimodal Interaction (Montréal, QC, Canada) (ICMI '21)*. Association for Computing Machinery, New York, NY, USA, 636–644. <https://doi.org/10.1145/3462244.3479947>
- [19] Chiara Mazzocconi, Ye Tian, and Jonathan Ginzburg. 2020. What’s your laughter doing there? A taxonomy of the pragmatic functions of laughter. *IEEE Trans. on Affective Computing* (2020), 1–1. <https://doi.org/10.1109/TAFFC.2020.2994533>
- [20] Patricia Mirenda, Anne Donnellan, and David Yoder. 1984. Gaze behavior: A new look at an old problem. *Journal of Autism and Developmental Disorders* 13 (01 1984), 397–409. <https://doi.org/10.1007/BF01531588>
- [21] Tomohiro Nagata and Hiroki Mori. 2020. Defining Laughter Context for Laughter Synthesis with Spontaneous Speech Corpus. *IEEE Trans. Affect. Comput.* 11, 3 (2020), 553–559. <https://doi.org/10.1109/TAFFC.2018.2813381>
- [22] Radosław Niewiadomski, Elisabetta Bevacqua, Maurizio Mancini, and Catherine Pelachaud. 2009. Greta: an interactive expressive ECA system. In *Proc. AAMS 2009*. International Foundation for Autonomous Agents and Multiagent Systems, 1399–1400.
- [23] Magalie Ochs, Catherine Pelachaud, and Gary Mckeown. 2017. A User Perception–Based Approach to Create Smiling Embodied Conversational Agents. *Trans on Interactive Intelligent Systems* 7, 1 (2017).
- [24] Robert R Provine. 2004. Laughing, tickling, and the evolution of speech and self. *Current Directions in Psychological Science* 13, 6 (2004), 215–218.
- [25] Tanya Romaniuk. 2009. The ‘Clinton Cackle’: Hillary Rodham Clinton’s Laughter in News Interviews. *Crossroads of Language, Interaction, and Culture* 7 (2009), 17–49.
- [26] Federico Rossano. 2013. Gaze in Conversation. In *The handbook of conversation analysis*, Jack Sidnell and Tanya Stivers (Eds.). John Wiley & Sons, Chapter 15, 308.
- [27] Federico Rossano, Penelope Brown, and Stephen C Levinson. 2009. Gaze, questioning and culture. In *Conversation Analysis: Comparative Perspectives*. 187–249.
- [28] JP De Ruiter. 2005. If eye-gaze frequency drops, its relationship with turn-taking disappears. In *Poster presented at AMLAP*.
- [29] Olof Sandgren, Richard Andersson, Joost van de Weijer, Kristina Hansson, and Birgitta Sahlén. 2012. Timing of gazes in child dialogues: A time-course analysis of requests and back channelling in referential communication. *International Journal of Language & Communication Disorders* 47, 4 (2012), 373–383.
- [30] Franklin R Schneier, Thomas L Rodebaugh, Carlos Blanco, Hillary Lewin, and Michael R Liebowitz. 2011. Fear and avoidance of eye contact in social anxiety disorder. *Comprehensive psychiatry* 52, 1 (2011), 81–87.
- [31] Gordon Stanley and Donald S Martin. 1968. Eye-contact and the recall of material involving competitive and noncompetitive associations. *Psychonomic Science* 13, 6 (1968), 337–338.
- [32] Catherine J. Stevens, Bronwyn Pinchbeck, Trent Lewis, Martin Luerksen, Darius Pfitzner, David M. W. Powers, Arman Abrahamyan, Yvonne Leung, and Guillaume Gibert. 2016. Mimicry and expressiveness of an ECA in human-agent interaction: familiarity breeds content! *Computational Cognitive Science* 2, 1 (06 2016). <https://doi.org/10.1186/s40469-016-0008-2>
- [33] Obed Torres, Justine Cassell, and Scott Prevost. 1997. Modeling gaze behavior as a function of discourse structure. In *First International Workshop on Human-Computer Conversation*. Citeseer.
- [34] Kaibin Xu, Junpei Zhong, and Kristiina Jokinen. 2021. It is Time to Laugh: Discovering Specific Contexts for Laughter with Attention Mechanism. In *2021 IEEE 4th International Conference on Information Systems and Computer Aided Education (ICISCAE)*. 211–215. <https://doi.org/10.1109/ICISCAE52414.2021.9590766>