Looking for Laughs: Gaze Interaction with Laughter Pragmatics and Coordination
Vladislav Maraev, Chiara Mazzocconi, Christine Howes, Catherine Pelachaud

To cite this version:
Vladislav Maraev, Chiara Mazzocconi, Christine Howes, Catherine Pelachaud. Looking for Laughs: Gaze Interaction with Laughter Pragmatics and Coordination. ICMI ’21: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION, Oct 2021, Montréal QC Canada, France. pp.636-644, 10.1145/3462244.3479947. hal-03878566

HAL Id: hal-03878566
https://hal.science/hal-03878566
Submitted on 29 Nov 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.
Looking for Laughs: Gaze Interaction with Laughter Pragmatics and Coordination

Chiara Mazzocconi  
chiara.mazzocconi@univ-amu.fr  
Institute of Language, Communication, and the Brain - Laboratoire Parole et Langage, Aix-Marseille University  
Aix-en-Provence, France

Vladislav Maraev  
Vidya Somashekarappa  
Christine Howes  
Centre for Linguistic Theory and Studies in Probability (CLASP), Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg  
Gothenburg, Sweden

Figure 1: Data collection setting [47]

ABSTRACT
Laughter and gaze have an important role in managing and coordinating social interactions. In the current work, using a multimodal corpus of dyadic taste-testing interactions, we explore whether laughs performing different pragmatic functions are accompanied by different gaze patterns towards the interlocutor, both from the point of view of the laughing participant and from her partner. We also investigate the role of gaze in laughter coordination between interactants. Our results show that laughs performing different pragmatic functions are related to different gaze patterns, both for the laugher and her partner, and that gaze is an important cue exploited by interactants when reciprocating laughter or laughing simultaneously. We discuss our data in relation to the literature about laughter and gaze functions in interaction, linking them to dialogic context. Our results stress the importance of laughter and gaze for modeling of multimodal meaning construction and coordination in interaction, and are therefore relevant for researchers designing human-like embodied conversational agents.

CCS CONCEPTS
• Human-centered computing → User models.

KEYWORDS
laughter, gaze, pragmatic function, coordination, multimodal communication

ACM Reference Format:

1 INTRODUCTION
Our conversations are highly coordinated, with synchronisation occurring even across modalities [11, 18]. Both laughter and gaze have been the object of in depth independent analyses and their crucial role in managing and coordinating interaction is not in doubt. Both laughter and gaze are perceivable actions [termed visible/audible acts of meaning in 4] which affect the unfolding of the upcoming dialogue [32]. While there is some work on the interaction of smiles, laughter and gaze in relation to humour [9, 20], less is known about the relation of laughter and gaze when this is not related to humour, but rather to what we call social incongruity. The only exception we are aware of is Romanik [42], who take a micro-analytic approach on the use of gaze to decline a laughter.
An example of the fine coordination between laughter and gaze is presented in (1), where we see the onset of gaze at the partner from A shortly before the onset of A’s laughter. The onset of A’s laughter is then shortly followed by B gazing at A, just before joining B’s laugh with her own.

(1) GHI Corpus [47], Pair03 (00:02:17)
A: It’s “like slightly”?...
B: yeah ((shrugs))
A: I like hummus |||laughter
B: yeah<laughter>

It is clear that both gaze and laughter are crucial elements to be taken into account when implementing algorithms for Embodied Conversational Agents (ECA) [7, 36], both for what concerns the interpretation of the users’ dialogue acts and for what concerns their own behaviour, in order to have ECAs more competent from a pragmatic perspective and also more human-like in terms of emotional displays, where this is desirable.

In the current work, we aim to fill this gap by investigating the following, to our knowledge, as yet unexplored questions. The answers are to provide insights into how meaning is constructed in interaction across modalities, as well as provide empirical data for the implementation of human-like ECA:

Q1 Does the laughing participant gaze at their partner, differ in terms of probability and timing, depending on the pragmatic function performed by the laughter?
Q2 Is the interlocutor’s gaze at the partner influenced, in terms of probability and timing, by the type of laughter produced by the partner?
Q3 Does gaze play a significant role in laughter coordination and alignment between participants?

The paper is structured as follows: in Section 2 we briefly present a literature review about laughter and gaze studies that constitute the background motivation of our questions, stressing how the study of laughter and gaze is increasingly important for ECA design. In Section 3 we explicitly state our hypotheses in relation to the background motivation of our questions, proposing rather that gaze might function to solicit a response [4, 25], which is not necessarily a speech turn [43]. More specifically, it has been argued that turn-taking is only a partial explanation for gaze behaviour in conversation, and that our study of gaze has to take into account both turn-taking and informational structure [8, 52].

Despite turn-taking not being the only function performed by gaze, and the fact that not all turn shifts are accompanied by gaze towards the listener, it has been consistently observed that there is a tendency for listeners to display more gaze at the speaker during the course of dyadic interaction, while the speaker tends to direct their gaze at the listener mainly towards the end of their speaking turn [16, 28]. In this way, when a speaker gazes at the listener mutual gaze is attained [23], a brief mutual gaze-window is established, and a change of floor may occur, having the previous listener looking away as they begin their speaking turn [28, 43, 47]. Gaze patterns to the interlocutor have also been found to differ depending on the speech act they accompany and on their pragmatic function [33, 44, 45].

Of interest in the study of gaze in interaction is not only gaze directed at one’s partner, but also its absence or avoidance [43]. For example, using a microanalytic analysis, Romaník [42] observed how gaze aversion can be used to decline laughter and terminate its relevance; while Kendrick and Holler [29] report that most preferred responses are produced with gaze toward the questioner, whilst most dispreferred responses are produced with gaze aversion. Moreover, it has been proposed that gaze aversion could also be explained (or influenced) by social stress [48], with evidence from patients with social disorders [46]. Conversely, results from other studies [14] suggest that cognitive load has the most impact on gaze aversion [21]. The latter hypothesis is based on the fact that visual cues are an important source of information and facilitate conversation, but cause higher cognitive load. This explanation

2.2 Gaze in interaction

The role of gaze in maintaining the conversational flow and coordinating dialogue acts is not in doubt. While many works have argued for the importance of individual gaze for fine regulation of turn-taking [15, 23], some scholars actually highlighted a lack of systematic relation between gaze and turn-taking [5, 12, 52], proposing rather that gaze might function to solicit a response [4, 25], which is not necessarily a speech turn [43]. More specifically, it has been argued that turn-taking is only a partial explanation for gaze behaviour in conversation, and that our study of gaze has to take into account both turn-taking and informational structure [8, 52].

In friendly conversation, interactants typically aim at an optimal level of cooperation and equilibrium avoiding direct disaffiliation as much as possible [39]. Nevertheless, social interactions often require the production of speech acts that can make this equilibrium unstable or at risk [40]. Following Mazzocconi et al. [32], we refer to any situation in which a clash is perceived between the current situation and a social norm and/or comfort as a social incongruity. Laughter, which can be used for bonding and showing friendliness, often comes in handy to cope with these situations. For example, in the case of embarrassment or awkward silence, laughter can smooth the situation; when putting forward a criticism, a laugh can soften the statement; or when asking a favour or advancing a proposal, a laugh can induce benevolence from the listener [22, 26, 27, 37].

Moreover laughter has been identified also as an attract-attention device, both in children [41, 50] and adults [38], especially for its emotional salience, making therefore extremely relevant to explore its effect on interactants’ gaze behaviour in natural conversation.

2 BACKGROUND

2.1 Laughter in interaction

Laughter production in conversation is not exclusively related to humour or to the appreciation of a pleasant incongruity. Many studies, particularly in conversation analysis, have shown its crucial role in managing conversations at several levels: dynamics (turn-taking and topic-change), lexical (signalling problems of lexical retrieval or imprecision in the lexical choice), pragmatic (marking irony, disambiguating meaning, managing self-correction) and social (smoothing and softening difficult situations or showing (dis)affiliation and marking group boundaries) [22, 27, 32].

Speech that overlaps with gaze at partner is shown in bold, with continuation of gaze marked by |||.
seems to be supported by results observing more gaze aversion in the initial phase of request formulations [28, 45], and by speakers showing less fluency when forced to constantly look at their listener [6], even though these results could also be explained by the social stress factor.

2.3 Gaze and Laughter in ECA

Recently, there has been a growing research interest both on gaze and other non-verbal expression, especially in Affective Computing community, for the implementation of ECAs which are more competent from a pragmatic perspective and able to process and produce appropriate emotional responses [2, 30, 34, 49]. Virtual agents benefit from a detailed analysis of multimodal input and output patterns observed during human-human interactions and from the interplay with their cognitive interpretation. Bailly et al. [2] established a basis for a context-aware eye-gaze generator for an ECA. In order to develop an improved gaze generator we should isolate the significant events detected in the multi-modal scene that impact the closed-loop control of gaze. Lee and Marsella [30] discuss the interpersonal role of gaze in interaction to signal feedback and direct conversation flow which current ECAs still lack. Simultaneously, in a dynamic environment, even the state-of-the-art ECAs struggle to direct gaze attention to peripheral movements. An embodied conversational agent 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. ECA simulations for face-face conversation are mainly dyadic and convey appropriate distraction and engagement behaviours. ECA simulations for face-face conversation are mainly dyadic and convey appropriate distraction and engagement behaviours.

2.3 Gaze and Laughter in ECA

Recently, there has been a growing research interest both on gaze and other non-verbal expression, especially in Affective Computing community, for the implementation of ECAs which are more competent from a pragmatic perspective and able to process and produce appropriate emotional responses [2, 30, 34, 49]. Virtual agents benefit from a detailed analysis of multimodal input and output patterns observed during human-human interactions and from the interplay with their cognitive interpretation. Bailly et al. [2] established a basis for a context-aware eye-gaze generator for an ECA. In order to develop an improved gaze generator we should isolate the significant events detected in the multi-modal scene that impact the closed-loop control of gaze. Lee and Marsella [30] discuss the interpersonal role of gaze in interaction to signal feedback and direct conversation flow which current ECAs still lack. Simultaneously, in a dynamic environment, even the state-of-the-art ECAs struggle to direct gaze attention to peripheral movements. An embodied conversational agent 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. ECA simulations for face-face conversation are mainly dyadic and convey appropriate distraction and engagement behaviours. ECA simulations for face-face conversation are mainly dyadic and convey appropriate distraction and engagement behaviours.

4 MATERIALS AND METHODS

4.1 Corpus data

Our data consist of 23 minutes taken from three female-female dyadic interactions from the Good Housekeeping Institute (GHI) Corpus [47]. The GHI corpus contains video and audio of pairs of participants discussing and rating different kinds of hummus on a paper questionnaire (see Figure 1). We annotated the interactions for laughter and gaze as described in the following sections.

4.2 Laughter Annotation

Our annotations have been conducted using the software ELAN [10]. Coding was carried out by the first author watching and listening to a video until a laugh occurred. The coder then marked the onset and offset of the laugh, and, following the annotation scheme proposed in Mazziocconi et al. [32], annotated the laughter’s form, temporal sequence in relation to speech and others’ laughs, context of occurrence, laughable it was related to (i.e. the argument of the laughter), and pragmatic function. In the current paper we focus on two of these features: (1) the type of incongruity present in the laughable, (2) the positioning of laughter in relation to the partner’s laughter (laughation coordination). We assessed the agreement on laughter identification and segmentation (start-time and end-time boundaries) using the Staccato algorithm implemented in ELAN [31], having two annotators marking 70% of the data. We run the analysis with 1000 Monte Carlo Simulations, a granularity for annotation length of 10, and α = 0.05. The degree of organisation is 0.8386.

4.2.1 Laughable classification. Following Mazziocconi et al. [32] we consider laughter as an event predicate, the meaning of which is constituted by two dimensions: the laughable and arousal, which we do not consider in the current paper. By laughable we mean the argument the laughter predicates about, an event or state referred to by an utterance or exophorically [22]. Different kinds of laughable can be distinguished based on whether they contain an incongruity or not, and if so, which kind of incongruity (see [19] 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. Common examples are 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. This kind of incongruity can be identified, for example, in the case of irony, scare-quoting, hyperbole etc. Typically in such cases laughter is used by the speaker in order to signal changes of meaning within their own utterance.

(4) **Friendliness** refers to cases where no incongruity can be identified. In many of 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.

In the current work we focus on the observation of gaze patterns accompanying laughs related to pleasant incongruity compared to social incongruity, as these are the most frequent kinds of laughable across contexts of interaction and languages [32] (see also Table 1). These categories are also the furthest apart in terms of pragmatic function, since pleasant incongruities are related to something pleasant and rewarding, whilst social incongruities are related to potential discomfort and unpleasantness.

Our dataset is therefore constituted of 74 laughs: 40 related to pleasant and 34 to social incongruity. 60% of the data (50 laughs) were annotated by two of the authors. The inter-annotator percentage agreement was 82%, with Krippendorff’s $\alpha = 0.69$.

### 4.2.2 Laughter coordination
In our annotation we distinguish 3 classes pertinent to the sequential distribution of the laughter in relation to laughs produced by the partner:

1. **Isolated laughter**: a laugh not preceded by or co-occurring with another laughter;
2. **Antiphonal laughter**: a laugh shortly following a laugh from the partner, starting during the partner’s laugh, or within one second after its offset;
3. **Coactive laughter**: a laugh with the same onset time as a laughter from the interlocutor. We did not give an exact time definition for shared laughter onset, rather we relied on annotators’ intuitions. We tested whether this intuitive notion was appropriate by calculating inter annotator agreement, which was high. Laughs which were considered to be coactive had a relative onset time of less than 100ms.

Inter-annotator agreement for this variable conducted over 60% of the data (50 laughs) reached 85.7%; Krippendorff’s $\alpha = 0.76$.

### 4.3 Gaze Annotation
Following Somashekarappa et al. [47], the gaze annotation was coded for four aspects:

1. **Participant1 and Participant2 (P1, P2)**: The gaze of each participant to an object, for example, Hummus (H), Questionnaire (Q), Breadstick (B) etc
2. **Joint attention (JA)**: Looking at the same object, obtained by temporal and object overlap in P1 and P2.
3. **Gaze1 and Gaze2 (G1, G2)**: For each participant, these encoded whether they were looking at their partner.
4. **Mutual Attention (MA)**: Looking at each other, obtained by temporal overlap in G1 and G2.

In the current work we explore only gaze at each other (G1, G2), leaving a more fine grained analysis of whether the gaze reciprocity (MA), and questions about gaze to objects including joint attention (JA), for future work.

### 4.4 Data extraction
In order to perform our analysis we made use of the ELAN Analysis Companion (EAC) software [1] to conduct event-related analysis. Our dependent variables is Gaze at partner (G1 and G2), both from the laughter and to the laughter from the other participant. In order to address questions (1) and (2) we used Laughter Type (whether a laughter was related to a pleasant or social incongruity) as predictor (sec. 4.2.1); while in order to address question (3) the predictive variable is Laughter coordination (isolated, antiphonal, or coactive) (Sec. 4.2.2). Each analysis is centered either on the onset or the offset of the laughter. Following Andersson and Sandgren [1] we considered a time window of 3000 milliseconds (i.e. 1500 seconds before and after the laughter onset/offset). We selected 10ms resolution, using a “first come first served” overlap handling and binned the data at intervals of 100ms, rounding up any fractions to 1.

Given the type of gaze $G$ and the type of laughable $L$ (or the laughter coordination classes for question (3)), the probability of gaze before or after the event for a given time window “bin” $b$ is calculated as follows:

$$P_b(G|L) = \frac{\sum_{i=1}^{N} P(g_b|l_i)}{N}$$  \hspace{1cm} (1)
where $N$ is the total number of laugh events, and $P(g_b|l_i) \in \{0, 1\}$ is the probability of gaze for a single bin $b$ for a given event $l_i$.

For each of our models, reported below, we ran a mixed-effect logistic regression in R, using the `glmer` function from the `lme4` package [3], with subjects as a random factor.

2 Including dyads as a random effect did not improve the models.

3 For any bin when gaze shift was occurring (the raw value therefore being between 0 and 1) we rounded up the value to one.

5 RESULTS

5.1 Laugher’s gaze × Laughable Type

5.1.1 Onset laughter. Figure 2a shows the probability of the laugher gazing at her partner around the onset of their own laugh depending on whether the laughter is related to a pleasant or a social incongruity. We observe a contrasting pattern of Gaze, especially after the onset of the laughter. The laugher is more likely to be looking at the partner before the onset of the laughter when it is related to a pleasant incongruity (as in (1)), the opposite is true after the onset of the laughter, when they are more likely to gaze at her partner when the laughter is related to a social incongruity (as in (2)).

(1) Pleasant incongruity [Pair15 (00:01:25)]

A: I quite like it ’cause I like it when there’s the little chickpeas on top|||
B: ((gazes at hummus)) yeah

A: ’cause it’s quite posh <laughs>
B: ((gazes at hummus)) yeah <laughs>

(2) Social incongruity [Pair03 (00:02:59)]

A: Shall we say... No. Ta- tasty <laughs>|||<laughs>
B: <laughs>
A: ((returns gaze at hummus))

5.1.2 Offset laughter. Fig. 2b shows the probability of the laugher gazing at the partner around the offset of her own laugh depending on whether the laughter is related to a pleasant or a social incongruity. We observe a significant main effect of Time ($CE = -0.81, SE = 0.22, z = -3.56, p < .001$) and Time in relation to the onset of the laughter ($CE = -0.70, SE = 0.20, z = -3.51, p < .001$), and a significant interaction between the two factors ($CE = 1.60, SE = 0.30, z = 5.31, p < .001$). While the laugher is more likely to gaze at her partner before the onset of the laughter when it is related to a pleasant incongruity (as in (1)), the opposite is true after the onset of the laughter, when they are more likely to gaze at her partner when the laughter is related to a social incongruity (as in (2)).
whether the laughter was produced in relation to a social or a pleasant incongruity. We observe a significant main effect of Time (CE = −2.83, SE = 0.2, z = −4.13, p < .001) and Laughter type (CE = −0.99, SE = 0.2, z = −4.43, p < .001), and no significant interaction (CE = 0.31, SE = 0.35, z = 0.87, p = 0.38), meaning that the partner is more likely to look at the laugher before the onset of the laughter, and in general more likely to look at the partner if the laughter was related to a pleasant incongruity.

5.2 Partner’s gaze \times Laughter Type

5.2.1 Onset Laughter. Figure 3a shows the probability of the partner gazing at the laugher at the onset of the laugh depending on whether the laughter was produced in relation to a social or a pleasant incongruity. We observe a significant main effect of Time (CE = −0.31, SE = 0.17, z = 1.18, p = 0.23), nor a significant interaction (CE = −0.12, SE = 0.17, z = −0.72, p = 0.47). This means that regardless of the Laughter Type, the laugher is more likely to look at her partner before the offset of her own laughter rather than after the offset.

5.2.2 Offset Laughter. Fig. 3b shows the probability of the partner gazing at the laugh at the offset of the laugh depending on whether the laughter was related to a social or a pleasant incongruity. We observe the opposite pattern to Fig. 2a, with the partner more likely to gaze at the laugher if the laugh was related to a pleasant incongruity rather than a social one.

We observe a significant effect of Time (CE = 0.51, SE = 0.22, z = 2.27, p = 0.02), while the main effect of Laughter Type is not significant (CE = 0.27, SE = 0.25, z = 1.07, p = 0.28). Of particular interest is the significant interaction (CE = −1.43, SE = 0.38, z = −4.31, p < .001). This shows that after the offset of the laughter the partner is much less likely to be looking at the laugher if the laughter was related to a social incongruity rather than a pleasant one, while the opposite pattern is observed before the offset.

5.3 Laugher’s gaze \times Laughter coordination

5.3.1 Onset Laughter. Fig. 4a shows the probability of the partner looking at the laugher at the onset of the laugh depending on whether the laugh produced was an isolated one (chosen as reference level), an antiphonal or a coactive one. We do not observe any significant difference in the probability of the laugher gazing at her partner before or after the onset of the laugh (CE = −0.37, SE = 0.21, z = −1.78, p = 0.07), while we do observe a main effect of Laughter Coordination, having isolated laughter as a reference level (Antiphonal-Isolated: CE = −0.76, SE = 0.31, z = −2.46, p = 0.01; Coactive-Isolated: CE = 1.24, SE = 0.25, z = 4.79, p < .001). No interaction was significant (Time \times Antiphonal-Isolated: CE = −0.24, SE = 0.45, z = −0.54, p = 0.58; Time \times Coactive-Isolated: CE = 0.57, SE = 0.35, z = 1.63, p = 0.10).

5.3.2 Offset Laughter. Fig. 4b shows the probability of the partner looking at the laugher at the onset of the laugh depending on whether the laugh produced was an isolated one (chosen as reference level), an antiphonal or a coactive one. The laugher is more likely to be gazing at her partner before rather than after the offset of her own laughter regardless of the variable Laughter Coordination (Time: CE = 1.62, SE = 0.32, z = 5.04, p < 0.001). We observed gaze at the partner to be significantly more likely when the laugher is producing an antiphonal laughter in comparison to an isolated one (Antiphonal-Isolated: CE = 1.6, SE = 0.39, z = 4.08, p < .001), and even more likely when producing a coactive laughter (Coactive-Isolated: CE = 2.49, SE = 0.4, z = 6.19, p < .001). We observe a significant interactions of Time and Laughter Coordination (Time \times Antiphonal-Isolated CE = −1.62, SE = 0.46, z = −3.51, p < .001; Time \times Coactive-Isolated CE = −0.94, SE = 0.47, z = −2.0, p = 0.04).

5.4 Partner’s gaze \times Laughter Coordination

5.4.1 Onset Laughter. Fig. 5a shows the probability of the partner looking at the laugher at the offset of the laugh depending on whether the laugh produced was an isolated one (chosen as reference level), an antiphonal or a coactive one. We observed significant main effects of all the predictors included in the model, but no significant interactions: gaze at the laugher is more likely before the onset of the laugh (Time: CE = −0.69, SE = 0.31, z = −2.28, p = 0.02), significantly more likely at the onset of an antiphonal laughter than an isolated one (CE = 1.06, SE = 0.26, z = 3.81, p < .001), and even more likely before the onset of a coactive laughter (CE = 2.22, SE = 0.29, z = 7.53, p < .001).

5.4.2 Offset Laughter. Figure 5b shows the probability of the partner gazing at the laugher at the offset of the laugh depending on whether the laughter was produced in relation to a social or a pleasant incongruity. We observed a significant main effect of Time (CE = −0.72, SE = 0.28, z = −2.55, p = 0.01), no significant difference between Antiphonal and Isolated laughter (CE = 0.14, SE = 0.27, z = 0.51, p = 0.6), but a significant difference between Coactive and Isolated laughter (CE = 1.19, SE = 0.31, z = 3.81, p < .001). We also observe a significant interaction between Time and Laughter Coordination (Time \times Antiphonal-Isolated: CE = 1.59, SE = 0.38, z = 4.14, p < .001; Time \times Coactive-Isolated: CE = 0.89, SE = 0.43, z = 2.04, p = 0.04).

6 DISCUSSION

Our data show that laughter related to different types of laughables, performing different pragmatic functions in interaction, is characterised by different accompanying gaze patterns both from the laugher and her partner. These observations confirm that both laughter and gaze play a crucial pragmatic role in the unfolding of dialogue; providing further evidence for the stance that to model gaze behaviour one needs to consider not only turns, but also propositional content and dialogue acts performed [52]. Furthermore our results validate the laughter taxonomy proposed in Mazzocconi et al. [32] showing that laughs belonging to different classes are produced and perceived as performing different pragmatic functions, eliciting different multimodal behaviours from the interactants. Below we discuss our results concerning laugher’s and partner’s gaze in relation to the type of laughable, and the results for gaze in relation to laughter coordination between interlocutors.

6.1 Laugher’s gaze

We observe that the laugher is more likely to gaze at the participant before the onset of a laugher related to pleasant incongruity rather than social incongruity, while the opposite is true after the onset. This result clashes with our hypothesis 1, according to which we expected an absence of gaze to the partner both before and after the onset of a laugher related to social incongruity. Our hypothesis was
Looking for Laughs: Gaze Interaction with Laughter Pragmatics and Coordination

Figure 4: Probability of Laugher’s gaze at Partner according to Laughter Coordination. The probability of laughter duration is shown at the bottom of the figures.

Based on Schneier et al. [46], Stanley and Martin [48] proposing that social stress makes gaze aversion more likely. Kendrick and Holler [29] also suggest that gaze aversion is more likely while producing a dispreferred answer, which is a dialogue act that belongs to the social incongruity laughable class in Mazzocconi et al. [32]’s taxonomy.

However, we can explain this data considering that we are looking at gaze during the laugher rather than during the laughable production (e.g. a dispreferred answer). Most of laughs follows the production of the laughable (e.g. [51]). Our data might therefore still be consistent with [29]. Indeed, we observe a lower probability of gaze at the partner before the onset of laughter – a time that often coincides with laughable production. We might therefore imagine a scenario where the expression of a dispreferred answer is not accompanied by gaze at the partner, and only while laughing the laugher looks at the partner, to monitor that the laughter has smoothed the disagreement or the unmet expectation, having the desired positive effect. Our data cannot therefore neither confirm nor disconfirm the social stress hypothesis of gaze aversion. It is possible that the “social stress” component is what influences the laugher to not look at her partner before the onset of a laughter related to social incongruity, while at the same time being the motivation to check her partner’s appraisal of the laughter (produced to ease the situation) during the laughter production.

The lower probability of gaze after the onset of laughter related to pleasant incongruity mirrors results reported by Gironzetti [20], who observed a lower probability of attention directed towards smiling facial expressions (including laughter) when it occurred in the context of humorous exchanges in comparison to non-humorous ones. These observations can be interpreted in the light of Becker-Asano and Ishiguro [7]: they observed that when their robot was directing its gaze at the partner while laughing, the laughter was interpreted negatively as being directed at the participant, rather than being produced cooperatively. We can therefore speculate that the tendency to avoid looking at the partner while producing a laughter related to pleasant incongruity is a way to disambiguate the laugher’s intention and social attitude towards the partner.

The opposite pattern observed before the laughter onset (i.e. higher probability of laugher gazing at the partner when about to produce a laughter related to pleasant incongruity), might be a result of the fact that the laugher is “careful” to assess whether laughter is an appropriate contribution, before producing it. There are indeed judgemental, moral, and cognitive aspects related to laughter production (e.g., not everything can be a subject for laughter, it is silly to laugh at some things, some laughter can be offensive for someone).

The consistency of our results with Gironzetti [20] is interesting considering the differences between corpora in terms of physical arrangement, data considered, and task. In Gironzetti [20], participants were opposite each other, conversing freely (without the goal-directed task of our study), and they also considered smiling, suggesting that similar dynamics are at play in the multimodal integration of smiling and laughter pragmatic processing. In our

Figure 5: Probability of Partner’s gaze at Laugher according to Laughter Coordination. The probability of laughter duration is shown at the bottom of the figures.
corpus, in contrast, participants are engaged in a task which requires them to pay (and share) attention on objects on the table in front of them, and they are seated at a 45° angle (a setting similar to the referential communication task in [45]). This means that gaze at the partner and mutual gaze are rarer than in other corpora (engagement in competing activities allows interactants to look away from their partner more frequently [43]), which leads us to speculate that when gaze at the partner instead does occur it is specifically motivated by pragmatic functions.

6.2 Partner’s gaze

Partner’s gaze at the laugher is significantly more likely to occur before the laughter onset rather than after. This is compatible with the idea of gaze being a cue for soliciting a response [4, 25], and that such a response does not have to be a verbal speech turn [43]. The main effect of laughable type (i.e. that gaze at the laugher is more likely before the onset of laughter related to pleasant incongruity) has to be considered together with the data represented in Fig. 3a about gaze and laughter coordination. In Sec. 5.4 we report that antiphonal and coactive laughs are significantly more likely to be preceded by gaze from the partner, meaning that gaze can be interpreted as an invitation to join in. The distribution of antiphonal laughs is though skewed towards pleasant incongruity (pattern replicated in several corpora e.g. [32]), which therefore constitutes a confounding variable. Due to the small data set we cannot consider such factor in our statistical modelling. We leave this exploration to further work when a larger dataset will be available.

We also observe that after the offset of the laughter, partner is less likely to look at the laugher if the laugh was related to social incongruity. We interpret this as a “choice” from the partner to avoid direct gaze in order to not put extra pressure on the laugher (who has appraised some situation or dialogue act as potentially discomforting) and maybe choose to give feedback (reassuring the laugher) in another modality, signalling that the issue should be declined as not important [42].

The higher probability of gaze at the laugher after the offset of laughter related to pleasant incongruity, may be explained as a partner’s strategy to check whether it would be appropriate to join the laughter. This is consistent with the results reported in Fig. 2a: i.e. laughers being more likely to gaze at the partner before the production of laughter related to pleasant incongruity. Laughing can indeed be a “no laughing matter”: not all laughs should be reciprocated [27].

Our hypothesis 2 is therefore partially invalidated in the setting of our corpus. We do not observe a significant higher probability to look at the laughing partner after the onset of the laughter (Fig. 3a), but rather before the onset. This data can be explained considering that the participants are sitting very close, engaged in a task, and mutual attention is already granted without the need to be signalled through gaze. Our data, on the other hand, highlight the role of laughter to elicit a (laughter) response from the partner [4].

6.3 Laughter Coordination

H3 was partly confirmed: we observed higher probability of gaze from the partner at the onset of an antiphonal or coactive laughter, but we did not observe a higher probability of looking at the partner from the laugher preceding the production of an antiphonal laughter. We therefore do not observe the need for a “gaze window” in order to respond to a laughter with a laughter. It would be nonetheless interesting in future work to control whether the laughter was produced by the participants in turn-initial position or not, in order to see if the observations mirror the patterns observed for speech (e.g. [4]). Nevertheless, our data show the important role played by gaze in the coordination of laughter production between participants and its role for eliciting responses from the partner, not just in terms of speech turns [43]. The role of gaze for laughter coordination is particularly striking in the case of coactive laughter (i.e. both interlocutors start laughing at the same onset time), where participants seem to look at each other not only to synchronise on the simultaneous onset but also to terminate the laughter. This kind of gaze may not only serve the purpose of syncing the response, but also monitoring other non-verbal cues about the partner’s disposition towards the laughable. This is an open question for future work.

7 CONCLUSION, LIMITATIONS AND FURTHER WORK

Our work provides evidence for the important link that gaze behaviour has for coordination in interaction, and also stresses the interaction between gaze, non-verbal behaviour and dialogue acts. Our data therefore offer new material for modelling multimodal meaning construction in interaction – important not only from a theoretical linguistic perspective, but also for the implementation of ECAs able to be more pragmatically adequate and to read non-verbal cues from the user in order to refine their own behaviour (e.g. if the user laughed and looked at the ECA, a likely adequate response might be to laugh back). More complex models for semantic processing are needed in order to tune ECAs behaviour to the pragmatic functions performed by the laugher, but our results suggest that gaze might be one of the cues to be considered in order to classify the type of laughter.

Our study is limited by the small sample size analysed. We are in the process of extending our dataset, which will allow us to employ more complex statistical models able to account for several variables at the same time (e.g. laughter position in relation to the speech-turn, to the laughable placement, arousal). Cross-cultural studies (e.g. [44]) showed differences in gaze behaviours between different communities (consideration relevant also for the implementation of ECAs appropriate to the user’s culture). Our results therefore should not be taken as absolute, but open up the possibility for interesting comparative studies.

ACKNOWLEDGMENTS

This study has been supported by the “Investissements d’Avenir” French government program managed by the French National Research Agency (reference: ANR-16-CONV-0002) and from the Excellence Initiative of Aix-Marseille University—“A*MIDEX” through the Institute of Language, Communication and the Brain. Open access funding was supported by grant 2014-39 from the Swedish Research Council, which funds the Centre for Linguistic Theory and Studies in Probability (CLASP) in the Department of Philosophy,
Linguistics, and Theory of Science at the University of Gothenburg. We would like to thank Andy Lücking, Emanuele Franceschi, and our anonymous reviewers for their insightful comments.

REFERENCES


