A Deep Gaze into Social and Referential Interaction

Vidya Somashekarappa (vidya.somshekarappa@gu.se)
University of Gothenburg, Sweden

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

Asad Sayeed (asad.sayeed@gu.se)
University of Gothenburg, Sweden

Abstract

In this study, we explicitly code and study the social, referential, and pragmatic features of gaze in human-human spontaneous dyadic interaction, providing novel observations that can be executed in a machine in order to improve multimodal human-agent dialogue. Gaze is an important non-verbal social signal that contains attentional cues about where to look and provides information about others' intentions and future actions. In this work, various types of gaze behaviour are annotated in detail along with speech to explore the meaning of temporal patterns in gaze cues and their co-relations. Considering that 80% of the total stimuli perceived by the brain is visual, gaze behaviour is complex and challenging; hence, implementing human-human gaze cues to an avatar/robots could improve human-agent interaction and make it more natural.

Keywords: Gaze; Human-Robot Interactions; Multimodal Dialogue; Annotation; Attention

Introduction

Humans use gaze as a communicative signal in face-to-face interaction, and social gaze can be interpreted as communicative by observers. Humans have an unique ability, beyond that of non-human primates, to interpret others' attention through eye gaze (Emery, 2000). However, interactions with avatars and robots do not make full use of these cues.

Humans perceive robots differently from how they perceive other humans – even though robots cue higher-order responses to gaze, they do not trigger the face-specific pathways at very short timescales (Admoni & Scassellati, 2012; Ragni & Stolzenburg, 2015). However, it is also important to interpret and produce interactional cues, such as using gaze to indicate next speaker or a desire to take the next turn. Ho et al. (2015) established that gaze can be used to signal the beginning and end of a speaking turn in social interaction. To establish more natural dialogues with humans, a conversational agent's ability to direct attention to the most appropriate target in a multimodal interaction is important (Norris, 2004). Bousmalis et al. (2009) presented cues such as head nodding and smiles, and Hunyadi (2019) used gestures but did not include gaze while investigating the temporal patterns of non verbal cues to study agreement and disagreement.

Additionally, gaze aversions that may be intentionally expressive and designed to communicate thoughtfulness, are much more of a challenge to interpret (or produce) by avatars/robots (Andrist, Tan, Gleicher, & Mutlu, 2014). In the current research, we explore the social, referential and pragmatic aspects of gaze in human-human interaction with an eye towards their future implementation in human-robot interactions.

Importance of Gaze Interaction in Dialogue

Many of the difficulties in interacting with robots/avatars have been attributed to the "uncanny valley" effect which hypothesizes that there is a relationship between the degree of an object's resemblance to a human being and the emotional response to it (Mori, 2020). Several studies have been working on finding approaches that overcome the "uncanny valley" effect, but the focus on the appearance of a robot leaves a missing part that is the influence of non-verbal behaviour.

Thepsoonthorn et al. (2021) conducted an experiment to determine the relationship between the perceived human-likeness of a robot and participants' affinity towards it. Participants were asked to interact with a robot with different non-verbal behaviours ranging from no non-verbal behaviour (speaking only) to a model capable of gaze, head-nodding, and gestures. Results showed that for fixation duration there was a biphasic relationship between affinity and human likeness, with the longest fixation durations observed when the robot expressed non-verbal behaviours.

Terzioğlu et al. (2020) showed improvement in likeability and perceived sociability by using interactive gaze cues alone. Hence, in the design of social robots or understanding of human-robot interactions, the multifaceted robotic visual perception understanding of how to (and how not to) use social cues such as gaze becomes increasingly important.

Human-Robot Gaze Interaction

Chevalier et al. (2019) showed human and robot faces as stimuli to participants and demonstrated higher gaze attention for human faces compared to robot faces. These were paralleled by attention processes obtained from (Wykowska, Wiese, Prosser, & Müller, 2014) event-related potentials (ERP's) of Electroencephalography (EEG) as well as functional Magnetic Resonance Imaging (Ozdem et al., 2017). These prior findings indicated that the observed critical behavioural difference is mirrored by differential patterns of activation in the bilateral anterior temporo-parietal junction (TPJ) which is typically involved in attentional reorienting as well as mentalizing. It is important to note that the perceptional difference is because of lower gaze-cueing rather than the appearance of the robot itself.

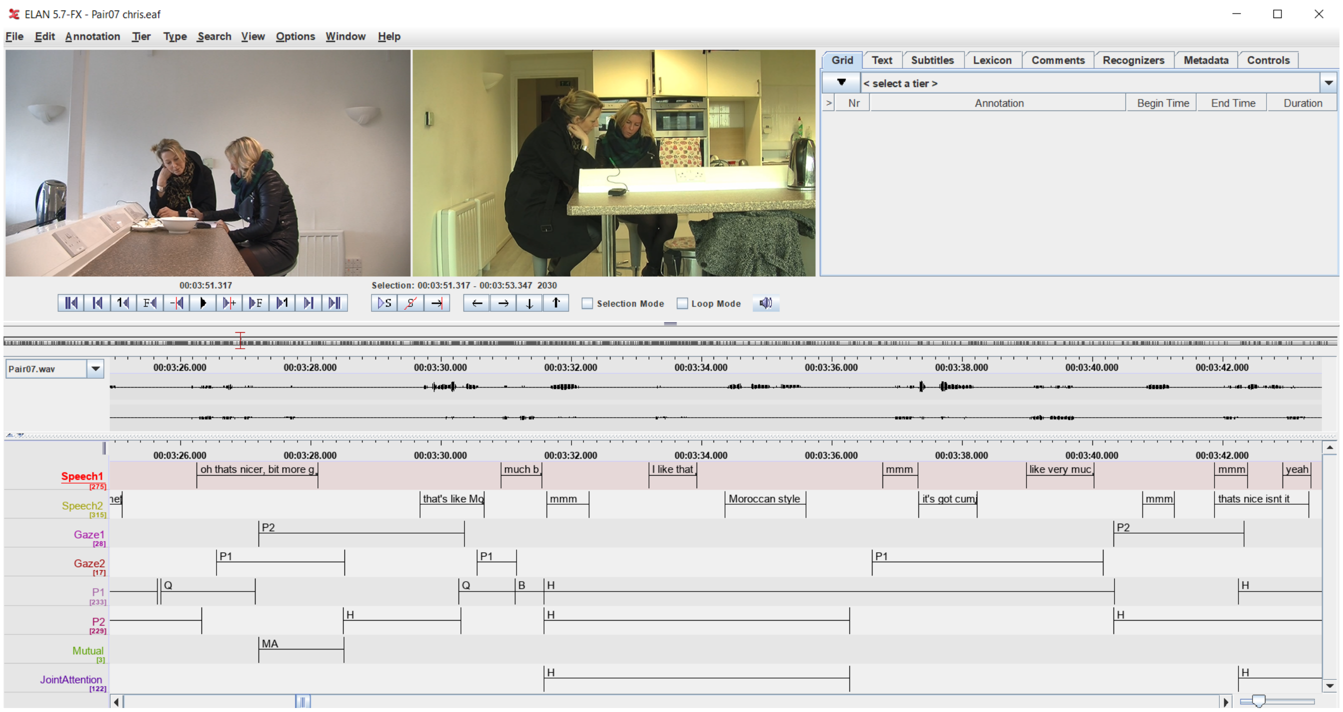

Figure 1: Speech and Gaze Annotation Labels

Despite, the development of the new generation perception devices such as Kinect and gaze control systems implemented on a FACE humanoid social robot, which also included multi-modal features like field of view, proxemics, verbal and non-verbal cues from the environment, the robot still does not direct its gaze appropriately and lacks the gaze-coordination required for smooth interaction (Zaraki, Mazzei, Giuliani, & de Rossi, 2014).

Yun (2016) proposed a hybrid approach, a computational model for selecting a suitable interlocutor for robots containing gaze control and perceptual measures for social cues in a multiparty situation. Physical space and conversational intimacy were the two factors that were added to the model calculation for controlling for the social gaze control effect. Although some research has been done to understand the attention processes and their implementation in smart devices, the effects of temporal difference between gaze cues such as joint attention, mutual attention, referential attention, gaze aversion and gaze transition in social setting of spontaneous interactions has not been studied before (Khan, Li, & Réhman, 2016).

Research Questions

1. Are there specific gaze patterns pertaining to different speech acts, and what influences them?
2. What are the temporal patterns of gaze, and what do they look like?
3. Can gaze inform or predict speech, and how does speech influence gaze?
4. What are the qualitative and quantitative findings that could help build a better gaze system in avatars/robots?

Data

This multi-modal and multi-focal corpus contains twenty four videos of dyads tasting eight samples of hummus and rating them together in a single questionnaire. Participants were staff at the Good Housekeeping Institute (a consumer product testing organisation in the UK [1]) and are familiar with one another. The dyads were seated at right angles, each directly facing an HD camcorder, and each wore unidirectional microphones (see Figure 1). The data contains explicit information of social and referential gaze since the dialogue is open ended and task requires joint attention while performing the task. During the session, participants tasted and judged each product's appearance, aroma, flavour and texture, then provided a rating and described the product in three words by discussing with each other and coming to an agreement. Hummus was rated on a 9-point Hedonic scale ranging from dislike extremely to like extremely. Each session lasted for about 20-30 minutes.

Annotation

We used ELAN 5.7 (Berez, 2007) to annotate the speech and gaze for four sessions from the GHI corpus (see figure 1, (Somashekarappa, Howes, & Sayeed, 2020)).

[1]`https://www.goodhousekeeping.com/uk/the-institute/`

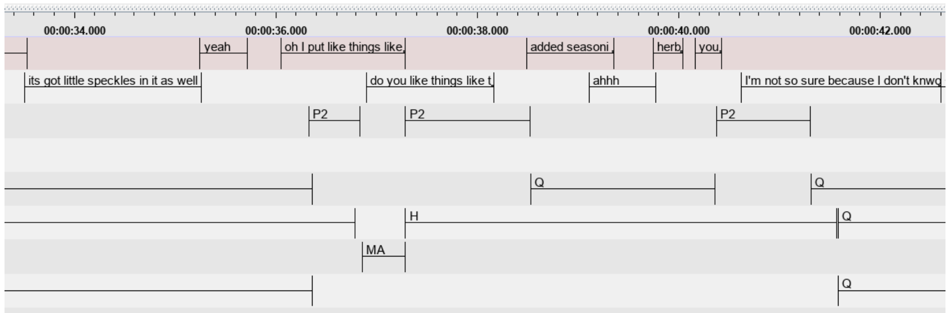

Figure 2: Excerpt 1

Gaze Annotation

Gaze was annotated according to the below schema:

1. **Referential gaze (P1, P2)** is gaze directed at an object or a location in space. Gaze information of each participant was annotated in separate tiers, as P1 and P2 (Participant1 and Participant2). The labels were the objects in their shared visual field such as bowl of hummus (H), Questionnaire (Q), breadstick (B), etc., or a location in space (Z).
2. **Joint attention (JA)** is sharing attention focus on a common object (Mundy, 2017). This tier was generated by temporal and object overlap of the P1 and P2 tiers.
3. **Gaze1 and Gaze2 (G1, G2):** For each participant, these encoded times they were looking at their partner.
4. **Mutual attention (MA)** is nothing but eye contact, it occurs when the gaze of both conversationalists is drawn to each other's face or eyes (Argyle, Cook, & Cramer, 1994). Mutual Attention is obtained by temporal overlap in G1 and G2.
5. **Gaze aversions** are the shifts in the the main direction of gaze that is away from the partner's face (Korkiakangas, 2018). This was obtained by *lack* of overlap in G1 and G2. This denoted P1 looking at P2, while P2 was looking at something else, and vice versa.

Analysis

We conducted a post-annotation qualitative and quantitative analysis. For qualitative measurement, we considered the onset and offset of gaze in relation to speech, and the numbers for quantitative analysis were exported directly from ELAN.

Qualitative analysis

The qualitative data contains excerpts (see figure 2) of speech segments from ELAN. As shown in figure 1 the speech information of each participant is followed by gaze tiers.

Pragmatic gaze assessment during disagreement

We assessed how the shifting of gaze on the partner helped to check if the statement made was being accepted or not.

Excerpt 1.1: referring to the appearance of the hummus

P2: it's got little speckles in it as well
P1: ————- H ————–
P2: ————- H ————–
P1: yeah, oh I put like things like that
P1: — H — P2 —- MA ————
P2: —– H ———MA————–
P2 : do you like things like that
P1: —- MA ——- P2———
P2: —–MA —— H———-
P1 : added seasoning, herbs, you?
P1: ——— P2————–
P2: ———– H————–
P2 : I'm not so sure cause I don't know
P1: ———P2 ——— Q—————
P2: ——— H——— Q—————-

When P2 says "it's got little speckles in it as well", the gaze attention of both P1 and P2 is on the hummus, as they are evaluating its appearance. In the next utterance, P1 responds in favour of the appearance, during which their gaze transitions to their partner. Their partner shows understanding by reciprocating the gaze, leading to the establishment of Mutual Attention by the end of P2's utterance. This is followed by P2 asking a question to re-assess P1's opinion, and here we can see that P1 continues to look at P1, but P2's gaze shifts back to the hummus, in an example of active gaze avoidance during disagreement. P1 pursues a verbal answer by asking "you?", while still maintaining their gaze on their partner, which changes to looking at the questionnaire after obtaining a negative response from P2, whose gaze attention has either been on the hummus or the questionnaire, avoiding looking at P1.

This is one of many examples that suggests that gaze can reveal information pertaining to negation or disagreement well before declaring it verbally.

Pragmatic gaze assessment during agreement

Excerpt 1.2: referring to the texture of the hummus

P1: plain
P1: —- Q —-
P2: —- H —-
P2: yeah, its pretty normal isn't it

P1: — Q ——— MA ——— Q —–
P2: — H ——— MA ——— Q —-
P2 : do you like things like that
P1: —- MA ——- P2———
P2: —–MA ——- H—————
P1 : yeah
P1: — Q —-
P2: — Q —-

While discussing about the texture of the hummus, P1 says "plain", where the gaze attention of P1 is on the questionnaire and P2 on the hummus. Right before P2 utters "yeah, its pretty normal isn't it", gaze attention is on the partner and P2 looks at P1. It is followed by an agreeing utterance from P1 "yeah". This excerpt shows that mutual gaze was established which reveals information about agreement well before declaring it verbally.

Gaze Check

Excerpt 2: tasting the hummus
P1 : feel a bit of pepper in there
P1 : ——-H————P2———-
P2 : ——-H————————-
P2 : Um-m
P1 : –P2–Q–
P2 : —-H—–

In excerpt 2, the gaze from the joint attention of P1 shifts to P2 before completing the sentence, representing a "Gaze Check" phenomenon. Here, P2 still continues to look at the hummus while responding. But P1's attention shifts from P2 to a different entity after receiving agreement in response.

Gaze Prediction

Excerpt 3: rating the hummus
P1 : like it moderately
P1 : —-Q——-P2——-
P2 : ———H————
P2 : yeah I would say like it moderately
P1 : —-P2—————H——————-
P2 : ——————Q———————-

This is another example of 'Gaze Check' from P1 to understand if P2 agrees to rate the hummus, "like moderately". Another important observation is that the gaze shift of P1 after looking at P2, is toward the object that P2 has fixated on for the entire duration when P1 was looking at P2. This type of gaze attention could further help in predicting the next movement of gaze towards the intended object of interest without any assistance from speech.

Overlapping Gaze Transition

The P1 and P2 tiers give us most of the gaze information of each participant except Mutual Attention and attention on the partner. It is interesting to think about what factors might influence the subsequent gaze shift. Does it depend on speech? If there is no speech then does that mean that there is no shift in gaze at all? Here is an example of one such phenomenon:

Excerpt 4: gaze shift in the absence of speech

P1: —-Q———-H———B—————H——
P2: —Q———H———Q——B——H———

P1's gaze attention follows P2's just 2 ms after P2 looks at the Questionnaire. Once P2's attention is on Hummus, approximately 3 ms later, P1 joins again before looking at the breadstick. P2 briefly focuses on the questionnaire before joining P1, and when P2 shifts gaze, P1 continues to join. This is a very interesting pattern seen consistently throughout the experiment, where the overlap occurs just a few milliseconds before joint attention.

Referential gaze

Excerpt 5: reaching for a breadstick

P1 : okay breadstick
P1 :——-B————-
P2 :———-Q———-
P2 : sure here it is
P1 :—-B————-
P2 :—-B————-
P2 : do you mind my hands on it
P1 :————B——————
P2 :————B——————

The example above explains how speech can influence gaze transition. Particular speech utterances especially ones mentioning the objects in the shared space drives gaze attention to the particular object away from the initial point of interest. For example, the phrase "okay breadstick", shifted the gaze of the partner to look away towards the location of the bowl containing breadsticks. These would enable gaze prediction based on referential speech and gaze attention. It is also important to note that P1 looked at the breadstick before verbally uttering the word.

Quantitative Analysis

The number of times the variables were annotated for different conditions is noted here. The total annotated data (Table 1) contains approximately 45 minutes (2700 seconds) of 8 participants in pairs, and the annotation comprises of 1,700 spontaneous speech utterances and 2,300 annotations for various types of gaze. A total of five dependant variables were measured across various gaze behaviours.

Table 1: Annotations summary. The tiers are Speech (S), Gaze at partner (G), Participants referential gaze (P), Mutual Attention (MA), Joint Attention (JA). The minimum, maximum, mean and total duration are in seconds

	Duration of annotations					Number of
Tier	Min	Max	Mean	Total	%	annotations
S	0.15	3.98	0.97	271.61	24.97	1706
G	0.27	4.90	1.44	48.55	4.29	269
P	0.17	20.83	2.98	527.62	41.84	1529
MA	0.41	2.73	1.09	17.66	1.99	51
JA	0.15	20.43	2.90	379.56	32.08	519

As shown in Table 1, the fewest gaze annotations (51) were coded in mutual attention (MA), with most in referential gaze (P: 1529). Interestingly, when we look at the total annotation duration for each individual compared to their partner, although there is an extremely high correlation between the amount of speech of each participant in a dyad ($r = 0.97, p < 0.001$) and the amount of gaze at reference objects between participants ($r = 0.96, p < 0.001$), showing the symmetry of the task, there is no such correlation between participants gaze patterns to each other ($r = -0.02$). There were also no correlations between amount of speaking and gaze at partner (self; $r = 0.009$ or other; $r = 0.002$).

Fixation Duration The duration of fixation is the total amount of time participants looked at different entities. Data from the separate tiers of Speech and Gaze were combined. Hence, all eight participants devoid of their interaction partner were taken into account to measure the co-relation between different entities.

Joint Attention For Joint attention, minimum and maximum fixation durations were under 2 miliseconds to 20 seconds with a mean of 2.9 seconds. Another aspect here is that the joint attention on one particular object did not last long (avg 2.9 sec); instead, there was a constant gaze transition and interaction with the surroundings. Overall the total annotation duration was approximately 380 seconds (out of 2700). This showed that the participants spent nearly as much time (JA) looking at the same object together, as they did looking at objects their partner was not looking at (P, 527 seconds) with equivalent average durations, showing how coordination of gaze is critical in a task requiring coordination with a partner in other respects, such as coordinating on which aspect of the task was being undertaken on a moment to moment basis.

Mutual Attention Duration of mutual attention accounted for as little as 0.4 seconds to 2.7 seconds. On average participants looked at each other for approximately one second at a time which is extremely short compared to Joint Attention (JA) or Gaze at the partner (G). In total, the annotation duration on Mutual attention was only 17.6 seconds (of 2700 seconds). Looking at the partner eye-to-eye for an extended amount of time can lead to an uncomfortable situation. This could result in the participants spending the least amount of time making eye contact and avoiding uncomfortable eye contact, which could be defined as “eerie mutual attention”. This is a novel observation that needs to be taken into consideration while improving gaze interaction in robots.

Individual Attention The attention on the partner while the partner looked somewhere else accounted for 4.29% (see row G in table 1), for which the annotation duration was 48.5 seconds (of 2700 seconds). Mutual attention where the participants looked at each other was notably rare in the data.

Discussion

Our annotated data, a detailed coding of gaze in conversation, adds another dimension in understanding multimodal dyadic interaction, showing that gaze is a complex non-verbal mechanism that still follows a very coordinated pattern in interaction upon analysis. These patterns are especially evoked in social and referential scenarios (Somashekarappa et al., 2020). Gaze behaviours shift in split seconds; researchers have been able to study these behaviours through psychological techniques, but without large amounts of observational data.

The partner establishes mutual gaze before making remarks. But while denying or not sharing the same opinion as the partner, the gaze is in the shared environment as noted in *Excerpt 1*, Similar pattern was observed across scenarios pertaining to agreement/disagreement scenarios. This hypothesis of existence of a unique correlation between gaze patterns, is seconded by Grynszpan and Nadel (2015).

Gaze cueing influences joint attention, but during verbal interaction, joint attention can occur simultaneously without the influence of a verbal utterance, as reported in this paper; this could be due to the development of social cognition (Beaudoin & Beauchamp, 2020). Considering the task required coordination, the maximum gaze duration during the session was spent in Joint attention, but interestingly, on average, each of these fixations lasted for about 2.9 seconds. The gaze-shifts from one object to another is quick and is not directly influenced by other factors such as speech. From a conversational robotics perspective, short duration fixations in robotic gaze could in turn make interaction more natural.

Eye gaze signals reduce cognitive effort and balance attention with intimacy when the speaker wants to maintain or relinquish the floor and the gaze aversions we observe (lack of overlap between G1 and G2) signal cognitive effort that is looking away or toward the speaker while beginning to answer a question depending on whether they were in agreement or not, which in turn, suggests that these gaze aversions are influenced by the purpose of the direction shift (Andrist et al., 2014).

Gaze-following without speech relies on following the motion based on observation of static stimuli, but it acts differently during dialogue acts (Shepherd, 2010). As noted in *Excerpt 4*, gaze transition occurs a few seconds before the actual implication of movement because of the added assistance from the speech stimuli. As reported by Admoni and Scassellati (2012) people can process gaze information, such as direction, really quickly, as shown in overlapping gaze transition (see qualitative analysis). Interactive eye gaze improves fluency and smooths task performance and subjective experience.

Establishing mutual attention can assist in reducing pragmatic overload in perceiving cues (Zhang, Beskow, & Kjellström, 2017). However, in our study mutual attention gaze occurrences are the least common, even though they still act as powerful conversational reinforcements.

Peoples' personalities can also affect gaze duration in a conversation. People are more likely to speak when their conversation partner looks at them more often (Vertegaal & Ding, 2002). The interpersonal dynamics between partners i.e the familiarity is correlated to the amount of mutual gaze not only on each partner's individual gaze behaviour.

From an adaptive evolutionary perspective, attending where others attend can provide information about behaviourally relevant events in the surroundings, particularly action plans, intentions and successive action plans, through the means of joint attention (Shuai, 2012).

Conclusion and Future Work

Detailed gaze annotation helps to unearth hidden layers in human interactions which can further help build automated dialogue systems. For future work, inter-rater reliability will be measured for all the videos that are annotated by the first author followed by automatic annotation of the manually-coded data will be conducted which would allow us to expand the corpus for multimodal interactions. The rapid advancements in the field of robotic technologies increases the importance of social robots that are built for interacting with people and are designed for various contexts such as therapy, education, and industrial applications. Depending on the degree to which they would need the autonomous capacity to display socially acceptable behaviour for human comfort, the results of this paper can lead to the implementation of gaze cues in avatars/robots such as Furhat (Domingo, Gómez-García-Bermejo, & Zalama, 2021).

Acknowledgements

This research was supported by a grant from the Swedish Research Council for the establishment of the Centre for Linguistic Theory and Studies in Probability(CLASP) at the University of Gothenburg.

References

Admoni, H., & Scassellati, B. (2012, 03). Robot gaze is different from human gaze: Evidence that robot gaze does not cue reflexive attention..

Andrist, S., Tan, X. Z., Gleicher, M., & Mutlu, B. (2014, 03). Conversational gaze aversion for humanlike robots.. doi: 10.1145/2559636.2559666

Argyle, M., Cook, M., & Cramer, D. (1994, 12). Gaze and mutual gaze. *British Journal of Psychiatry*, *165*, 848-850. doi: 10.1017/S0007125000073980

Beaudoin, C., & Beauchamp, M. (2020, 01). Social cognition. *Handbook of clinical neurology*, *173*, 255-264. doi: 10.1016/B978-0-444-64150-2.00022-8

Berez, A. (2007, 01). Eudico linguistic annotator (elan). *Lang. Document. Conserv.*, *1*.

Bousmalis, K., Mehu, M., & Pantic, M. (2009). Spotting agreement and disagreement: A survey of nonverbal audiovisual cues and tools. In *2009 3rd international conference on affective computing and intelligent interaction and workshops* (pp. 1–9).

Chevalier, P., Kompatsiari, K., Ciardo, F., & Wykowska, A. (2019, 12). Examining joint attention with the use of humanoid robots-a new approach to study fundamental mechanisms of social cognition. *Psychonomic Bulletin & Review*, *27*. doi: 10.3758/s13423-019-01689-4

Domingo, J., Gómez-García-Bermejo, J., & Zalama, E. (2021, 01). Optimization of a robotics gaze control system. In (p. 213-226). doi: 10.1007/978-3-030-62579-5_15

Emery, N. (2000, 09). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and biobehavioral reviews*, *24*, 581-604. doi: 10.1016/S0149-7634(00)00025-7

Grynszpan, O., & Nadel, J. (2015, 01). An eye-tracking method to reveal the link between gazing patterns and pragmatic abilities in high functioning autism spectrum disorders. *Frontiers in Human Neuroscience*, *8*. doi: 10.3389/fnhum.2014.01067

Ho, S., Foulsham, T., & Kingstone, A. (2015, 08). Speaking and listening with the eyes: Gaze signaling during dyadic interactions. *PloS one*, *10*, e0136905. doi: 10.1371/journal.pone.0136905

Hunyadi, L. (2019, 06). Agreeing/disagreeing in a dialogue: Multimodal patterns of its expression. *Frontiers in Psychology*, *10*. doi: 10.3389/fpsyg.2019.01373

Khan, M., Li, H., & Réhman, S. (2016, 05). Gaze perception and awareness in smart devices. *International Journal of Human-Computer Studies*, *92*. doi: 10.1016/j.ijhcs.2016.05.002

Korkiakangas, T. (2018, 05). Gaze aversion and the progress of interaction. In (p. 208-239). doi: 10.4324/9781315621852-7

Mori, M. (2020, 01). The uncanny valley. In (p. 89-94). doi: 10.5749/j.ctvtv937f.7

Mundy, P. (2017, 09). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, *47*. doi: 10.1111/ejn.13720

Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. doi: 10.4324/9780203379493

Ozdem, C., Wiese, E., Wykowska, A., Müller, H., Brass, M., & Van Overwalle, F. (2017, 01). Believing androids - fmri activation in the right temporo-parietal junction is modulated by ascribing intentions to non-human agents. *Social Neuroscience*, *12*.

Ragni, M., & Stolzenburg, F. (2015, 05). Higher-level cognition and computation: A survey. *KI - Künstliche Intelligenz*, *29*. doi: 10.1007/s13218-015-0375-y

Shepherd, S. (2010, 03). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in integrative neuroscience*, *4*, 5. doi: 10.3389/fnint.2010.00005

Shuai, L. (2012, 09). Modelling the coevolution of joint attention and language. *Proceedings. Biological sciences / The Royal Society*, *279*, 4643-51. doi: 10.1098/rspb.2012.1431

Somashekarappa, V., Howes, C., & Sayeed, A. (2020, 05). An annotation approach for social and referential gaze in dialogue..

Terzioğlu, Y., Mutlu, B., & Sahin, E. (2020, 03). Designing social cues for collaborative robots: The role of gaze and breathing in human-robot collaboration. In (p. 343-357). doi: 10.1145/3319502.3374829

Thepsoonthorn, C., Ogawa, K.-i., & Miyake, Y. (2021, 01). The exploration of the uncanny valley from the viewpoint of the robot's nonverbal behaviour. *International Journal of Social Robotics*, 1-13. doi: 10.1007/s12369-020-00726-w

Vertegaal, R., & Ding, Y. (2002, 01). Explaining effects of eye gaze on mediated group conversations: Amount or synchronization? In (p. 41-48). doi: 10.1145/587078.587085

Wykowska, A., Wiese, E., Prosser, A., & Müller, H. (2014, 04). Beliefs about the minds of others influence how we process sensory information. *PloS one*, *9*, e94339. doi: 10.1371/journal.pone.0094339

Yun, S.-S. (2016, 10). A gaze control of socially interactive robots in multiple-person interaction. *Robotica*, *35*, 1-17. doi: 10.1017/S0263574716000722

Zaraki, A., Mazzei, D., Giuliani, M., & de Rossi, D. (2014, 04). Designing and evaluating a social gaze-control system for a humanoid robot. *IEEE Transactions on Systems Man and Cybernetics - Part A Systems and Humans*, *44*, 157-168. doi: 10.1109/THMS.2014.2303083

Zhang, Y., Beskow, J., & Kjellström, H. (2017, 10). Look but don't stare: Mutual gaze interaction in social robots. In (p. 556-566). doi: \url{https://doi.org/10.1007/978-3-319-70022-9_55}