The effect of social roles on gaze cue utilisation in a real-world collaboration

Ross G. Macdonald (rgmacdonald@dundee.ac.uk)
School of Psychology, University of Dundee, Nethergate,
Dundee, DD1 4HN, UK

Benjamin W. Tatler (b.w.tatler@activevisionlab.org)
School of Psychology, University of Dundee, Nethergate,
Dundee, DD1 4HN, UK

Abstract

During collaboration, people communicate using verbal and non-verbal cues, including gaze cues. Social factors can affect gaze allocation, however most research on gaze cueing has not considered these factors. The presence of social roles was manipulated in a collaborative task whilst eye movements were measured. In pairs, participants worked together to make a cake. Half of the pairs were given roles ("Chef" or "Gatherer") and the other half were not. Across all participants we found, contrary to the results of static image experiments, that participants spent very little time looking at each other, challenging the generalisability of the conclusions from lab-based paradigms. When given spoken instructions, listeners in the roles condition looked at the speaker significantly more than listeners in the no roles condition. We conclude that our tendency to seek the gaze cues of collaborators is affected either by our social perceptions of the collaborator or their perceived reliability.

Keywords: eye movements; joint attention; real world; gaze cues; social interaction.

Introduction

When collaborating with another on a task, we need to communicate. As well as using spoken language, there are a number of non-verbal cues we can use, with the directional gaze cues given by the eyes being the most well-researched of these. Gaze cues are first used very early in life and continue to be given and followed throughout adulthood. People have a tendency to orient to and follow the gaze cues of others and can to do this with ease. However, there is evidence that the language accompanying a gaze cue and the social context of the cue can affect how people orient to and follow gaze cues. In the real world, gaze cues will always occur within a social context, yet this context is removed in most studies. The aim of the present study is to measure eye movements in a real-world setting to observe how the utilisation of gaze cues can be affected by social context in a natural collaboration.

When viewing images of faces, people have a tendency to look at the eyes (Yarbus, 1967) and when viewing images of social scenes people will seek out faces and eyes (Birmingham, Bischoff & Kingstone, 2007; 2009) even when the person being fixated is not visually prominent and has no role for understanding the scene (Zwickel & Võ, 2010). As well as orienting to these cues, people show a tendency to follow them. Friesen and Kingstone (1998) showed that incongruent gaze cues presented at fixation could slow down responses in a Posner (1980) task, suggesting that the artificial gaze cue stimuli automatically shifted attention away from the target. Variants of this study looking at eye movements have found that participants will also look in the direction of the distracting gaze cue, even though they know there is no reason to do so (Ricciardelli, Bricolo, Aglioti, & Chelazzi, 2002; Galfano et al, 2012). These findings have been used to suggest that humans are "hard-wired" to automatically follow the gaze cues of others (Emery, 2000).

The above research shows that people look at eyes and follow gaze cues when viewing isolated static images of others. However, in the real world, gaze cues usually occur alongside spoken language. There appears to be an intimate link between gaze allocation and spoken language, with people making anticipatory eye movements to objects that relate to what they hear (Altmann & Kamide, 1999). Gaze cue utilisation in particular has been shown to be affected by spoken language; changing the syntactic structure of a sentence, whilst maintaining meaning changes the timing of gaze following (Knoeferle & Kreysa, 2012). Reciprocally, Stuadte and Crocker (2011) showed the gaze cues can affect the understanding of spoken language; participants were shown videos of a robot describing the spatial and featural relations between a series of visible objects, whilst providing gaze cues. The robot made mistakes in his descriptions that could have been corrected in two different ways. The experimenters found that participants would correct in the way that was congruent with the gaze cue, suggesting that they were inferring meaning from the robot's gaze and assuming that the robot meant to refer to the object that it was gazing at. Given the effect gaze cues and language have on each other, it is important to use language in a paradigm investigating how gaze cues are used naturally in collaboration.

As well as mostly occurring alongside language, all gaze cues in the real world are provided in a social context. When interacting with another, where we look can be affected by our proximity to this other person (Argyle & Dean, 1965). Social effects specifically on gaze seeking were investigated by Laidlaw, Foulsham, Kuhn and Kingstone (2011), who found that participants sitting in a waiting room were significantly more likely to look at a person on a monitor than the same person present in the room. Gallup et al (2012) found similar results for gaze following rather than seeking. They observed people walking past an attractive

item in a hallway and found that people were more likely to look in the same direction as somebody walking in front of them than somebody walking towards them. The results of these studies were explained by their respective authors as being due to participants trying to avoid potential interactions with strangers, which might be triggered by any gaze seeking or following behaviour detected by the oncoming person. These findings indicate that social factors can affect the way we utilise the gaze cues of strangers, which suggests that social context may have an effect on gaze utilisation in one-to-one interactions and collaborations.

Macdonald and Tatler (2013) considered gaze seeking and following behaviour in a real world communicative task, involving one-to-one interaction between an instructor (the experimenter) and a participant. The instructor manipulated his use of gaze as well as the specificity of his instructions in a simple block-building task. Participants were found to only seek and follow gaze cues when the language was ambiguous (it did not specify which single block the participant was meant to pick up), suggesting that gaze cues are used flexibly, depending on other information that is available. It was also noted that even when gaze cues supplied the only unambiguous information about which block to pick up (because the spoken instructions were ambiguous) participants did not seek and follow these all of the time. It was speculated (Macdonald & Tatler, 2013) that social factors may have played a part in these results. More specifically, the social cost of looking at the instructor frequently in each trial may have deterred participants from seeking and following these gaze cues. Although this is speculation, these results make a case for manipulating social factors in a real-world gaze-cueing experiment.

One way to manipulate social factors in a gaze cueing task is to manipulate what the participant knows about the entity with which they are interacting. Participants carrying out a Posner (1980) task in Italy were shown distracter gaze cueing stimuli made from the faces of Italian political figures, including Silvio Berlusconi (Liuzza et al, 2011). The gaze of Berlusconi was found to cause significantly more interference in the task for right-wing voters (in-group) than left wing voters (out-group). These results suggest that people may be more prone to following the gaze cues of others with shared beliefs. Crosby, Monin and Richardson (2008) showed that participants were more likely to look at an individual on a monitor if they thought the individual could hear comments that were potentially offensive to that individual. These results show that social factors such as beliefs about another individual can affect how others look at them as well as how others look at external objects whilst communicating with them. Although, these results show effects of prior beliefs about others on gaze behaviour, it is still unclear how beliefs about the role or knowledge of another affect the use of gaze cues in natural collaboration.

The present study manipulates participants' perception of their collaborator by assigning them roles in a task. Participants, in pairs, were given a recipe to follow in order to make the batter for a cake. During this collaboration their eye movements were recorded using portable eye-trackers. When coding the data we were particularly interested in the time participants spent looking at each other (interpersonal gaze) or at the same object simultaneously (mutual gaze). Half of the pairs were given roles (chef or gatherer) to fulfil and the other half were not. By manipulating this we are able to investigate whether the perception of another's role in collaboration has any significant effect on the extent to which we seek and follow their gaze cues in a real-world interaction.

Methods

Participants

Twenty-four students from the University of Dundee participated in this experiment. They were split into twelve pairs to carry out the task. Six pairs were allocated to the roles condition and six were allocated to the no roles condition (see design).

Materials

The experiment took place in a kitchen area on the University of Dundee campus. The kitchen was fully equipped with standard kitchen appliances, but only the oven and microwave were used. All items and foodstuffs that could be removed were removed before testing and the experimental materials were arranged carefully around the kitchen. This included the items and foodstuffs that were to be used for the procedure as well as a selection of distractor items. All of these items were placed in the same location for each pair of participants. A Recipe Procedure sheet was provided for each pair. This sheet explained, step-by-step, how to make the batter for a Victoria Sponge. There was also a Chef Guidelines sheet and a Gatherer Guidelines sheet for those in the roles condition. These sheets explained the responsibilities and duties for participants in the chef and gatherer roles.

Design

This experiment had a between subjects design. The two independent variables for the analysis of mutual fixations and time participants spent looking at each other (interpersonal gaze) were the use of roles (roles or no roles) and the allocation of roles within the roles condition (chef or gatherer). For the analysis of the instruction statements the independent variables were the use of roles (roles or no roles) and the identity of participant (speaker or listener).

Procedure

This experiment required two participants. The experimenter began by fitting a portable eye tracker to the first participant. At this point in the roles condition the first participant was given the Chef Guidelines and the second participant was given the Gatherer Guidelines. They were both instructed to read over their sheet and make sure they understood their roles. The Chef Guidelines informed the chef that they were in charge of preparing the recipe and that the gatherer was there to assist them. The sheet explained that the chef was expected to mix and prepare ingredients, following a recipe which they could not show to the gatherer. The chef would not be expected to collect any items or foodstuffs, but to delegate those duties to the gatherer. The chef would also be able to ask the gatherer to assist them with any aspect of the preparation they wished. The Gatherer Guidelines explained that the gatherer would not be expected to make any decisions concerning the preparation, but should instead do as instructed by the chef. Once the participants declared they understood their roles the gatherer was asked to remain outside whilst the experimenter and the chef entered the kitchen. The experimenter then gave the chef the Recipe Procedure sheet and told the chef where all of the necessary items and foodstuffs were located. The chef was then told they would have approximately three minutes to familiarise themselves with the kitchen and the locations of the items. During these three minutes the experimenter fitted another portable eye-tracker to the gatherer. In the no roles condition the second eye-tracker was fitted straight after the first. At this point, in both conditions, both participants were brought into the kitchen and the eye-trackers were switched on.

The cameras were synchronised and the eye-trackers calibrated. Once calibration was complete, those in the no roles condition were directed to the Recipe Procedure sheet and informed that all of the items they would require were located around the kitchen. All participants were informed that the experimenter would be standing outside the kitchen, out of sight and that the participants must make no attempt to interact with him. The experimenter then told the participants that they may begin as soon as he was out of the room. The experimenter left and the procedure began. The procedure ended when the participants put the batter mixture in the oven.

Eye movement and sound recording

Participants' eye movements were tracked using two Positive Science LLC mobile eye trackers, which allowed free head movement. Each eye tracker has two cameras mounted on the frame of a pair of spectacles: one records the scene from the participant's point of view and the other records the right eye. Data from these cameras were captured on digital camcorders. For one of the eye-trackers these camcorders were stored, alongside a power supply for the eye-tracker, in a lumbar pack worn by the participant. The camcorders connected to the second tracker were again stored alongside a power supply, but were stored in a light backpack worn by the participant. This eye tracker also has a small microphone attached to the frame. This microphone recorded sound throughout the experiment and was able to pick-up the voices of both participants. Gaze direction was estimated off-line using Yarbus software provided by Positive Science, LLC, which tracks the pupil and corneal reflection. Calibration was carried out in two stages, one looking down at a counter and the other looking across the room. These two stages were used because by tracking one eye we are not able to directly measure the vergence of the eyes that occurs as participants focus on objects at different distances. Instead we fit the model to fixations on both proximal and distal points. If the tracker estimates in the scene video fell on the correct calibration positions the calibration was deemed adequate. Eye movement data were recorded at 30Hz with a spatial accuracy of about 1 degree. Once videos for both participants were rendered with the eye movement information, Quicktime Pro was used to synchronise both videos in to one movie file, ready for analysis.

Analysis

Eye tracking data were coded manually offline using Quicktime Media player and audio information was extracted using Audacity sound editing software. The first two dependent variables considered were (1) the proportion of time both participants fixated the same object (mutual fixations) and (2) the proportion of time a participant spent looking at their partner (interpersonal gaze). For these analyses, in each pair, one participant was labelled person A and the other was labelled person B. In the roles condition person A was the chef and B was the gatherer. Since there were not any defined roles in the no roles condition, participants in this condition were arbitrarily allocated as person A or B. The frame-by-frame coding of these data was split between the lead experimenter and three undergraduate volunteers from the School of Psychology. To begin, all four coders coded the same movie file and these were all compared by the lead experimenter to ensure a consistent and high quality of coding. Mutual fixations were compared across conditions by a t-test and the proportion of time spent on interpersonal gaze was analysed using a 2 (roles or no roles) by 2 (person A or person B) independent measures ANOVA.

The individual instructions were also coded and analysed. These were coded by the lead experimenter alone, using audacity sound editing software and the Quicktime movie files. For each pair, each instruction statement was numbered and transcribed, noting the speaker. The time that the speaker first looked (if at all) at the listener and vice versa was coded for each instruction statement. In the roles condition, the speaker was always the chef and the listener always the gatherer. In the no roles condition the participant who gave the instruction was considered to be the speaker. Therefore the identity of the speaker and listener would

switch throughout each movie in the no roles condition. From coding these data we considered the percentage of instructions in which the participant looked at the other participant. This was analysed using a 2 (role or no role) by 2 (speaker or listener) ANOVA.

Results

Overview of eye movements in collaboration

The first set of results is focused on the general eye movement behaviour of participants in the roles and no roles conditions. To investigate this behaviour we measured the proportion of time participants spent mutually fixating objects and the proportion of time spent fixating on the co-participant (interpersonal gaze).

The mean proportion of time in which both participants fixated on the same item (mutual fixation) is shown in Figure 1.

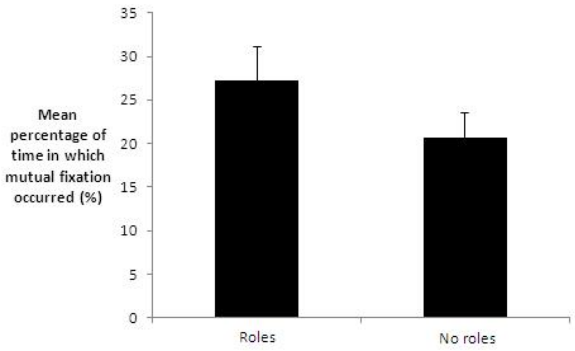

Figure 1: The mean percentage of time in which mutual fixation occurred for participant pairs in the roles and no roles condition (with standard error bars).

A larger mean percentage of time was spent on mutual fixation in the roles condition (27.23%) than the no roles condition (20.69%). However this difference was not found to be significant ($t(10) = 1.37$, $p = 0.200$)

The mean percentage of time spent engaged in interpersonal gaze is shown in Figure 2. This plot shows the percentage of time that A spends looking at B and vice versa for the roles and no roles conditions. The amount of time when participants A and B simultaneously looked at each other is also shown in Figure 2, for the roles and no roles conditions.

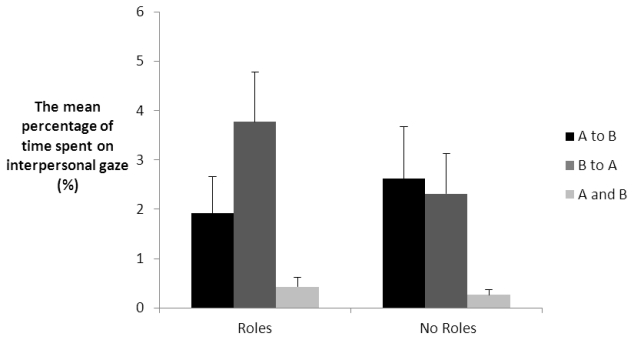

Figure 2: The mean percentage of time participants spent on interpersonal gaze for Person A, Person B and A and B simultaneously for both the roles and no roles conditions (with standard error bars)

It can be seen from Figure 2 that on average person B spent more time looking at person A (3.77%) than vice-versa (1.92%) in the roles condition, whilst participants spent only 0.43% of the total time on simultaneous interpersonal gaze. In the no roles condition, Person A was found to spend slightly more time looking at person B (2.62%) than vice versa (2.31%) and only 0.27% of the time was spent simultaneously looking at one another. A two (roles, no roles) by two (person A, person B) ANOVA was carried out on these results. No main effects of role condition ($F(1,20) = 0.171$, $p = 0.683$) or participant ($F(1,20) = 0.701$, $p=0.412$) were found, nor was there any significant interaction ($F(1,20) = 1.381$, $p = 0.254$).

Analysis of eye movements during instructions

These results consider the eye movement behaviour during the periods when one of the participants was giving spoken instructions to the other. For the roles conditions the spoken instructions were always provided by the chef. For the no roles conditions, any instructions could have been provided by either participant. We investigated the mean percentage of (spoken) instructions in which interpersonal gaze occurred. For each of the roles and no roles conditions, we considered cases when the speaker looked at the listener, the listener looked at the speaker or both speaker and listener looked at each other at the same time (Figure 3).

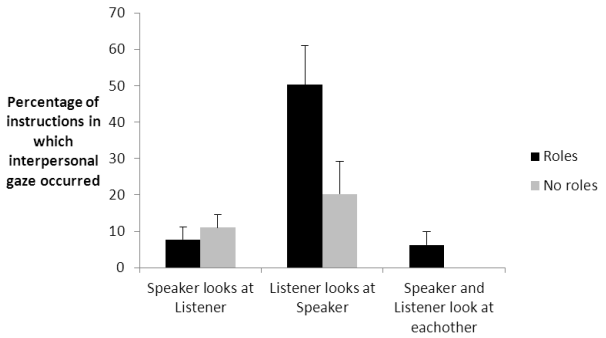

Figure 3. The mean percentage of instructions in which interpersonal gaze occurred for speakers, listeners and both speakers and listeners in the roles and no roles conditions (with standard error bars).

A two (roles, no roles) by two (speaker, listener) ANOVA showed a main effect of identity of participant (speaker or listener) ($F(1,20) = 12.00$, $p = 0.002$). The main effect of role condition was not significant ($F(1,20) = 3.21$, $p = 0.089$), however, there was a significant interaction ($F(1,20) = 4.92$, $p = 0.038$). Post-hoc t-tests showed that listeners looked at speakers during significantly more instructions in the roles condition (50.20%) than the no roles condition (20.28%, $p = 0.010$), but there was no significant difference found between speakers' looks to listeners across the roles (7.78%) and no roles (10.97%) conditions ($p = 0.766$).

Discussion

The aim of this study was to investigate the effect of manipulating social context on the utilisation of gaze cues in a real world collaborative social interaction. Using portable eye trackers we were able to measure the eye movements of both collaborators for the duration of the task. The time participants spent looking at each other in this real world paradigm was much less than expected, given the results of experiments using static social scenes (Birmingham et al, 2007; 2009). Social context was actively manipulated in this paradigm by the presence or absence of roles as there is evidence from lab based studies (Crosby et al, 2008; Liuzza et al, 2011) that beliefs about a collaborator can affect gaze behaviour. The amount that listeners looked at speakers during instructions was affected by our manipulation of the roles of the two participants, providing evidence that the tendency to look at another individual during a real world interaction may be influenced by the social context provided by the roles of the individuals. This result is consistent with previous suggestions that gaze seeking and following may depend on the social context of the gaze cues (Gallup et al, 2012; Laidlaw et al, 2011; Macdonald & Tatler, 2013).

There was no significant difference found between the percentage of time in which mutual fixations occurred in the roles and no roles conditions, with collaborators spending approximately one-quarter of task time mutually fixating on the same objects. There was also no significant difference between the percentage of time that interpersonal gaze occurred across roles conditions. However, participants spent far less time (between 2-4%) looking at each other than they spent mutually fixating other objects. This is notable as it appears to be at odds with the results of some previous lab-based studies. People have been shown to have a preference for looking at eyes when viewing pictures of people (Yarbus, 1967) or social scenes (Birmingham et al, 2009; Zwickel & Võ, 2010), however in this task participants spent very little time looking at their partners. Given the potential informativeness of the eyes (Tomasello et al, 2007) and the ease with which people can interpret gaze direction (Anderson, Risko & Kingstone 2011) this finding may seem surprising. However, studies using real people as stimuli may offer an explanation. Laidlaw et al (2011) showed that people were less likely to look at a present confederate than the same confederate on a video monitor and Gallup et al (2012) found that people were less likely to follow the gaze of strangers that could see them than strangers who could not. They concluded that this was due to there being potential consequences (social interaction) to looking at the present confederate or the on-coming stranger. A collaborator in the present study could potentially react to the looks of a participant, whereas the static and video images in lab based paradigms could not. Therefore, these lab-based studies may have over-estimated the tendency of people to look at eyes and faces in social settings.

These results present an obvious question; if people rarely look at each other in an interaction, can they still utilise gaze cues? Although our results cannot lead us to a definite answer, there are three main arguments for the ability to utilise gaze cues in these circumstances. Firstly, it has been shown that gaze cues can be followed and affect language comprehension, even when they are not directly fixated (Knoeferle & Kreysa, 2012). Secondly, when gaze cues are fixated, the fixations do not necessarily involve long periods of time viewing the eyes. Looks to gaze cues may be very brief, but very informative. Thirdly, it may be the case that eyes are generally not sought out during a task, but are used effectively when required, for example, during instructions.

From our findings it is possible to speculate about the third possibility. Listeners were found to look at the speaker during significantly more instructions in the roles condition than the no roles condition. This finding shows that our preference for looking at others can be affected by social context. In the roles condition the listener was always the gatherer, following instructions given by an informed chef, who was in charge. In the no roles condition the identity of the listener would switch between the two equal partners, depending on who was giving the instruction. Macdonald and Tatler (2013) found that the degree of informativeness of gaze cues affected the extent to which the cues were sought out, with highly informative cues being sought most often. One possible interpretation of the present findings could be that our manipulation of the roles of the participants effectively manipulated the perceived

informativeness of the cues provided by the chef: listeners in the roles condition may consider the gaze cues of the chef to be highly informative, whereas the gaze cues of the speaker in the no roles condition may be considered less informative.

Alternatively, our pattern of results could arise from a social effect of authority. Liuzza et al (2011) found that right-wing voters were more heavily influenced by the gaze cues of their political leader than the gaze cues of the opposition leader. In the roles condition, the chef is in charge of the procedure and is therefore the leader of the gatherer. It is possible that, as well as being more inclined to follow the gaze cues of a leader, people are also more inclined to orient to the leader's gaze cues. Although the results do not allow us to favour one explanation over the other, these findings provide good evidence that the social context of collaboration can affect the extent to which collaborators look at each other during communication. A more controlled future experiment may be able to distinguish between the effects of the perceived reliability of a person and the perceived social role of a person.

This experiment investigated the effect of social roles on eye movement behaviour in a natural collaboration by using dual portable eye-trackers. We manipulated the roles of the participants to investigate the effect on gaze behaviour. Listeners were found to look more at a speaker providing verbal instructions if the speaker was playing the role of a chef. This suggests that our tendency to look at others is either affected by our social perceptions of a person or by our perception of their reliability. Additionally, we found that in this real social collaborative setting, people spent very little time looking at each other, challenging the generalisability of the conclusions from lab-based paradigms (Birmingham et al, 2007; 2009; Zwickel and Võ, 2010). Our results provide a strong case for investigating gaze cueing behaviour in highly naturalistic environments as well as providing evidence for the effect of social context on the utilisation of gaze cues.

Acknowledgments

This paper was supported by an EPSRC Studentship awarded to Ross Macdonald. The authors would like to thank Anne-Joanna MacGregor, Julia McVean and Nicole Spittle for assistance with video coding.

References

Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247-264.

Anderson, N.C., Risko, E.F., & Kingstone, A. (2011). Exploiting human sensitivity to gaze for tracking the eyes. *Behavioural Research,43*(3), 843-52.

Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry, 28*(3), 289-304.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2007). Why do we look at people's eyes? *Journal of Eye Movement Research, 1*(1), 1-6.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Get Real! Resolving the debate about equivalent social stimuli. *Visual Cognition, 17* (6), 904-924.

Crosby J. R., Ronin B., & Richardson D. R. (2008). Where do we look during potentially offensive behaviour? *Psychological Science. 19*(3): 226-228.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioural Reviews, 24*, 581-604.

Friesen, C.K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review, 5*, 490-495.

Galfano, G., Dalmaso, M., Marzoli, D., Pavan, G., Coricelli, C., & Castelli, L. (2012). Eye gaze cannot be ignored (but neither can arrows). *Quarterly Journal of Experimental Psychology, 65*(10), 1895-1910.

Gallup, A. C., Chong, A., & Couzin, I. D. (2012). The directional flow of visual information transfer between pedestrians. *Biology Letters, 8*(4), 520-522.

Knoeferle, P., & Kreysa, H. (2012). Can speaker gaze modulate syntactic structuring and thematic role assignment during spoken sentence comprehension? *Frontiers in Psychology, 3*, 538.

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences, 108*, 5548-5553.

Liuzza, M.T., Cazzato, V., Vecchione, M., Crostella, F., Caprara, G. V., & Aglioti, S.M. (2011). Follow my eyes: The gaze of politicians reflexively captures the gaze of ingroup voters. *PloS ONE 6*(9), e25117.

Macdonald, R.G., & Tatler, B.W. (2013). Do as eye say: Gaze-cueing and language in a real-world social interaction. *Journal of Vision, 13*(4):6, 1-12.

Posner, M.I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*, 3-25.

Ricciardelli, P., Bricolo, E., Aglioti, S. M., & Chelazzi, L. (2002). My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. *Neuroreport, 13* (17), 2259-2264.

Staudte, M., & Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human-robot interaction, *Cognition, 120*, 268-291.

Tomasello, M., Hare, B., Lehmann, H. & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution, 52*, 314-320.

Yarbus, A.L. (1967). *Eye Movements and Vision.* New York: Plenum Press.

Zwickel, J., & Võ, M. L. H. (2010). How the presence of persons biases eye movements. *Psychonomic Bulletin & Review, 17*, 257-262.