Conversational gaze in light and darkness

Elisabet Renklin\textsuperscript{1,2}, Fanny Cardell\textsuperscript{1,2}, Johanna Dahlbäck\textsuperscript{1,2}, Jens Edlund\textsuperscript{3} and Mattias Heldner\textsuperscript{1}
\textsuperscript{1}Linguistics, Stockholm University, Stockholm
\textsuperscript{2}CLINTEC, Karolinska Institutet, Stockholm
\textsuperscript{3}Speech, Music and Hearing, KTH Royal Institute of Technology, Stockholm

Abstract
The way we use our gaze in face-to-face interaction is an important part of our social behavior. This exploratory study investigates the relationship between mutual gaze and joint silences and overlaps, where speaker changes and backchannels often occur. Seven dyadic conversations between two persons were recorded in a studio. Gaze patterns were annotated in ELAN to find instances of mutual gaze. Part of the study was conducted in total darkness as a way to observe what happens to our gaze-patterns when we cannot see our interlocutor, although the physical face-to-face condition is upheld. The results show a difference in the frequency of mutual gaze in conversation in light and darkness.

Introduction
Face-to-face interaction (or conversation) has many components: it consists of words, prosody, head and hand gestures, facial displays, gaze, and so on. Some of these components are probably only meant to be heard, while others are only “made for the other person to see” (e.g. Bavelas, Chovil, Lawrie, & Wade, 1992, p. 483). Yet other components may be perceived by integrating information from the auditory and visual senses. This work is about gaze—a seemingly visual component—with a view to explore auditory perception of speaker’s facing direction as a correlate of gaze.

It is well attested that gaze, and in particular mutual gaze has an important function for social face-to-face interaction (e.g. Argyle & Dean, 1965; Kendon, 1967). The way we make and break eye contact is an important human behavior, and deviant gaze behavior makes us feel uncomfortable (Gibson & Pick, 1963). A typical gaze pattern, at least in Europe and in Northern America, is that the listener looks fairly constantly at the speaker, while the speaker looks at the listener (thereby establishing mutual gaze) in the vicinity of speaker changes or backchannels (e.g. Bavelas & Gerwing, 2011; Kendon, 1967). Thus, mutual gaze seems to be related to the coordination of speech and silence in conversation–to turn-taking.

Furthermore, gaze is closely tied to speaker’s facing direction. Edlund, Heldner & Gustafson (2012) demonstrated that blindfolded listeners are surprisingly apt at perceiving a speaker’s facing angle under normal conversational circumstances. This finding suggests that auditory perception of a speaker’s facing direction might provide a redundant correlate of gaze in visible conditions, and a correlate of gaze in non-visible face-to-face conditions, such as in the dark.

This work explores mutual gaze in dyadic conversation. It is a partial replication of the study by Kendon (1967), where gaze patterns (including mutual gaze) in light was studied, and extends that work in that it includes gaze behavior (or rather speaker’s facing direction) in darkness. We have focused on the occurrence of mutual gaze in the vicinity of situations where speaker changes or backchannels typically occur: joint silences where both participants are silent, and joint overlaps where they speak at the same time.

Conversation in darkness was included in search of evidence of the relevance of auditory perception of speaker’s facing direction for turn-taking. Following Bavelas, et al. (1992), we expect a drastic reduction in mutual gaze in darkness if gaze patterns are only ‘made to be seen’. If, on the other hand, auditory perception of speaker’s facing direction is used as a correlate of gaze, differences between light and darkness may not be as drastic.

Method
We recruited 14 participants (or 7 pairs of speakers) through social media (e.g. Facebook) and e-mail by offering a cinema ticket for the participation in a study about visual aspects of conversation involving recordings in light and in darkness. Five of the conversations took place between two women; one between a woman and a man and one between two men.
The participants all grew up in Sweden; they were all living in the Stockholm area; and they were between the ages 20-35.

The recordings took place in a studio with two chairs opposite of one another with 1.5 m between them and a table in the middle. Behind each chair there was a tripod with a video camera capturing the participant in the other chair (see Figure 1).

The studio was equipped with two Canon XA10 video cameras, recording in HD quality (1920 x 1080 pixels, 50 frames/s). To reduce the risk that the subjects would be distracted by light in the darkness condition, all diodes (e.g., the rec-lights) on the cameras were covered with black tape. Audio was recorded using individual close-talk microphones (Sennheiser ME 3-ew) connected to both cameras.

The participants were instructed to sit down in pairs and talk freely for 20 minutes. Ten minutes into the conversation we switched the cameras from normal to infrared mode with a built-in IR lamp and turned the lights off. Figure 2 illustrates recordings in the darkness condition. Seven dialogues were recorded and six of them analyzed. One dialogue was discarded as one of the participant’s eyes were hidden by her glasses.

The recordings were extracted from the memory cards into our computers and trans-coded into high-quality H.264 files.

We used ELAN (version 4.2.0, http://www.lat-mpi.eu/tools/elan/) to annotate gaze-direction and joint silences and overlaps. The files from each conversation were synchronized and then we annotated the gaze of the individual participants (in the ELAN segmentation mode). Throughout this paper, we will use a number of terms defined by Kendon (1967): ‘p’ stands for the person being observed and ‘q’ is the person interacting with ‘p’. A gaze is considered to be a q-gaze when ‘p’ is looking straight at the face of the other person. Mutual gaze is the situation where both participants look at each other’s faces. Mutual gaze was derived automatically from these annotations using the ELAN function Create Annotation From Overlaps.

We annotated joint silences and overlaps semi automatically in a process similar to that used in Heldner & Edlund (2010). First, we used the ELAN Silence Recognizer MPI-PL (in the ELAN annotation mode) to annotate intervals of silence and talk for the individual speakers on separate tiers. Next, we corrected these automatic annotations manually. Finally, joint silences and overlaps were derived from the individual tiers, again using the Create Annotation From Overlaps function in ELAN.

The last step in the annotation process was to manually annotate joint silences and overlaps with and without mutual gaze. We counted occurrences of mutual gaze within a two-second window ending at the offset of joint silences and overlaps. Joint silence and overlap annotations shorter than 120 ms were excluded.
Figure 2. ELAN annotation interface with snapshots from a recording in darkness.

Results

Table 1 shows the distribution of joint silences and overlaps with and without mutual gaze in light and darkness across all six recordings. A number of observations can be made.

Mutual gaze was relatively frequent in situations where speaker changes or backchannels typically occur when the participants could see each other. Mutual gaze occurred with about the same frequency as no mutual gaze, and mutual gaze was slightly more frequent in joint overlaps than in joint silences.

Table 1. Distribution of joint silences and overlaps (counts and percentages) with (+) and without (-) mutual gaze in light and darkness across all six recordings.

<table>
<thead>
<tr>
<th></th>
<th>LIGHT</th>
<th>DARK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+mutual</td>
<td>-mutual</td>
</tr>
<tr>
<td>Silences</td>
<td>188 (48%)</td>
<td>202 (52%)</td>
</tr>
<tr>
<td>Overlaps</td>
<td>189 (57%)</td>
<td>140 (43%)</td>
</tr>
<tr>
<td>Total</td>
<td>377 (52%)</td>
<td>342 (48%)</td>
</tr>
</tbody>
</table>

Furthermore, there was an apparent difference in the frequency of mutual gaze in conversations in light and darkness. Mutual gaze was considerably less frequent when the lights were off, in joint silences and overlaps alike. In four out of the six conversations, the mutual gaze frequency was close to zero or zero.

Analyses of the distribution of mutual gaze in the individual conversations revealed that one of the conversations deviated from the others: 27% of the overlaps in the light condition were preceded by mutual gaze, while the mean for the others was 60%. All the conversations except one were between two women or a woman and a man; the one between two men was the deviating one.

Discussion

The results support previous work in the area (e.g. Bavelas & Gerwing, 2011; Kendon, 1967): that mutual gaze often precedes situations where speaker changes and backchannels typically occur. Furthermore, the results indicate that there is a substantial difference in frequency of mutual gaze in light and darkness, with drastically fewer occurrences of mutual gaze in darkness. Thus, the results seem to suggest that (i) mutual gaze is “made for the other person to see” (Bavelas, et al., 1992); (ii) that gaze patterns is not just a learned behavior, but has an important function and meaning; and (iii) that auditory perception of speaker’s facing direction (Edlund, et al., 2012) is not used as a direct substitute for mutual gaze.

The structure of the conversations with respect to the frequencies of joint silences and overlaps was similar in the light and darkness.
conditions. (We have not considered the durations of the silences and overlaps, which we think would be an important addition in further studies). Thus, the fact that the participants could not see each other did not seem to affect their dialogue in terms of frequency of turn taking. This comes as no big surprise; we are after all perfectly able to speak with people we cannot see.

The fact that ‘p’ did not look at the face of ‘q’ as much in darkness as in light is not an unexpected result either, since it is not possible to make eye-contact or even know exactly where the other person’s face is in total darkness. Although the participants knew approximately where their interlocutor was positioned, our results might well depend on the fact that we only counted q-gazes in the strict sense of looking at the other person’s face. In future work, we will explore less strict criteria for q-gaze, requiring only that the participants are facing each other, instead of looking directly at each others’ faces. Such a re-analysis would also allow us to re-visit the idea of auditory perception of speaker’s facing direction as a correlate of gaze. Regardless of the results, we have now come to the conclusion that this is a better way to measure the effects of mutual gaze in darkness, as vertical changes of face direction ought to be less significant for the sound reaching the listeners ears (the sound waves reach both ears at the same time) than lateral changes.

We have collected a number of aspects to consider in future work. One such aspect is gender. Argyle and Dean (1965) stated that women tend to engage in mutual gaze more than men. This could be one explanation for the deviant result in the dialogue between two men. Technical aspects for future work include that the silence recognizer in ELAN was not always able to discriminate between silence and speech in an adequate way. In some cases laughter, breathing and background noises was recognized as speech. This led to errors that had to be adjusted manually, making the annotations less reliable than they could be. A better voice activity detector would likely result in less arbitrary annotations. Finally, to continue this study and to be able to find more specific results about turn-taking, it will be necessary to define which silences and overlaps that are speaker changes and to discriminate between speaker changes and backchannels.

Our results show that mutual gaze behaviors are easily abandoned in darkness, at least in dyadic dialogue. We note that many aspects that are crucial for negotiating the flow of the dialogue are reduced to formalities in dyadic situations: the addressee is always "the other person", and given that the interlocutors agree that a speaker change is coming up, the selection of the next speaker is trivial. In multi-party dialogue, decisions such as these, where gaze is likely to play a major role, require considerable attention. One interpretation of the data is that in dyadic conversation, the information we get from linguistic and phonetic aspects (e.g. prosody and sentence structure) provides sufficient clues to where the other person is going for successful conversation.

We can also speculate, in line with Argyle and Dean (1965), that mutual gaze may be of more importance when it comes to affiliation, emotional responses and social politeness, and that it is our need for feedback and some kind of connection that makes us look for mutual gaze rather than the structural aspect of turn-taking.

Acknowledgements

The research presented here was funded in part by the Swedish Research Council project Samtalets rytm (2009-1766).

References


