Effects of Conversation Content on Viewing Dyadic Conversations

Jarkko Hautala[1,2], Otto Loberg[1], Piia Astikainen[1]
[1]Department of Psychology,
[2]Agora Center,
University of Jyväskylä, Finland
jarkko.v.hautala@jyu.fi

Lauri Nummenmaa
Turku PET Centre, and
Department of Psychology,
University of Turku, Finland

Jari K. Hietanen
School of Social Sciences and
Humanities/Psychology,
University of Tampere, Finland

People typically follow conversations closely with their gaze. We asked whether this viewing is influenced by what is actually said in the conversation and by the viewer's psychological condition. We recorded the eye movements of healthy (N = 16) and depressed (N = 25) participants while they were viewing video clips. Each video showed two people, each speaking one line of dialogue about socio-emotionally important (i.e., personal) or unimportant topics (matter-of-fact). In personal conversations, the viewers made more saccadic shifts between the discussants, but they looked more at the second than at the first speaker when watching matter-of-fact conversations. Higher depression scores were correlated with less looking at the currently speaking discussant. We conclude that subtle social attention dynamics can be detected from eye movements and that these dynamics are sensitive to the observer's psychological condition, such as depression.

Introduction

People routinely notice subtle dynamic cues in their social environment, yet cognitive mechanisms supporting real-world social attention remain poorly specified (for reviews, see Frischen, Bayliss, & Tipper, 2007; Graham & LaBar, 2012). Most previous studies have been conducted with static pictures, and research on the dynamic allocation of social attention has only begun fairly recently (for reviews, see Kingstone, 2009; Skarrat, Cole, & Kuhn, 2012). To our best knowledge, the present experiment is the first to study whether the semantic content of a conversation modulates the allocation of visual attention when an outsider follows conversations (i.e., a third-person perspective).

People tend to pay more attention to socially relevant vs. irrelevant visual stimuli (Frischen et al., 2007; Graham & LaBar, 2012). When viewing scenes with social content, human observers focus their gaze on people, especially on the latter's eye region (e.g., Birmingham, Bischof, & Kingstone, 2007), presumably to detect their emotions and intentions (Buchan, Parè, & Munhall, 2007). Various social features, such as signs related to social hierarchy (Maner, DeWall, & Gailliot, 2008), sexual content (e.g., Lykins, Meana, & Strauss, 2008; Nummenmaa, Hietanen, Santtila, & Hyönä, 2012), social approval (DeWall, Maner, & Rouby, 2009), and beliefs about other people's intentions (Garrod & Pickering, 2009; Nummenmaa, Hyönä, & Hietanen, 2009; Richardson, Dale, & Tomlinson, 2009), modulate viewers' attention. Both positive and negative emotional information attracts observers' attention efficiently (e.g., Isaacowitz, 2005; Kellough, Beevers,

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

Ellis, & Wells, 2008). These findings seem to hold well also in natural and dynamic situations as shown in recent studies (e.g., Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010; Foulsham, Walker, & Kingstone, 2011).

In natural situations, people are often involved in multiparty conversations and interactions. Conversants readily anticipate turn taking as evidenced by overlapping utterances between them (Riest, Jorschick, & Ruiter, 2015). This is predominantly based on the semantic structure of the conversation, as well as syntactic and prosodic cues for predicting turn taking (Riest et al., 2015). Turn taking also influences the gaze patterns of uninvolved viewers (Hirvenkari et al., 2013). On average, observers direct their gaze 74% of the time to the current speaker when viewing audio-visual recordings of two persons' conversations (Hirvenkari et al., 2013). The gaze shifts to the current speaker take place approximately 300 ms after the onset of speech (Hirvenkari et al., 2013). However, when following question-answer sequences with a more predictable structure, observers often shift their attention from one discussant to another even before the discussants start to talk or at possible turn completions (Holler & Kendrick, 2015).

However, on top of the turn-taking behavior, the semantic content of the conversation may modulate attention allocation in important ways. Studies on one-to-one interactions with virtual avatars have shown that the avatars are watched for a longer time when their speech expresses negative (as opposed to positive) contents (Choi et al., 2010; Schrammel, Pannasch, Graupner, Mojzisch, & Velichkovsky, 2009). For example, it can be expected that when following other people's conversations, an observer's attention may be guided toward the persons referred to in the current utterance or that socio-emotionally important statements may trigger complex viewing patterns across all involved discussants as the observer may be interested in seeing others' reactions to such statements. In this study, we examined for the first time how socio-emotionally important statements would modulate uninvolved observers' visual attention dynamics when following simple dialogues from video clips. We also investigated depression-related effects on this task. Depression markedly influences socio-emotional functioning (Libet & Lewinsohn, 1973; Segrin, 1992). Depression is also associated with an attention bias to negative emotional stimuli (e.g., Beck, 2008), a deficient activation of positive emotions, and problems in effortful high-level cognition of various types (Heller & Nitschke, 1997). Eye-tracking

studies have consistently reported a depression-related increase in maintaining the gaze on dysphoric contents (Caseras, Garner, Bradley, & Mogg, 2007; Eizenman et al., 2003; Kellough et al., 2008; Leyman, De Raedt, Vaeyens, & Philippaerts, 2011) and a reduced orientation toward positive contents (Ellis, Beevers, & Wells, 2011; Kellough et al., 2008; Sears, Bisson, & Nielsen 2011; for a review, see Armstrong & Olatunji, 2012). In real social situations, depressed people make substantially less eye contact with others (for a review, see Segrin, 1992). However, the effects of depression on attentive tracking of realistic social interactions remain elusive.

In this study, we investigated the conversation content's effects on the viewers' social attention by analyzing their eye movements. The participants viewed videos of conversations where the actors were instructed to avoid making nonverbal gestures or giving cues known to modulate visual attention (Rigolout & Pell, 2012; Schwartz & Pell, 2012). We varied the content so that 50% of the conversations were matter-of-fact, and the other 50% were personal in nature (see the Appendix). To clarify this category, all the personal conversations included statements related to the other discussant (e.g., "I think that your friend is boring"), while the matter-of-fact conversations referred to things in general (e.g., "I think that the friend is a barber"). Half of the personal conversations began with emotionally negative statements (e.g., "I think that your style is tasteless") and the other half with positive statements (e.g., "I think that your style is lovely"). We measured the temporal latency of the gaze shifts from one discussant to another, the rate that the viewers shifted their attention between the discussants, and the proportion of time they devoted to watching each of the discussants, with respect to turn taking and conversational content.

We expected that the observers' gaze behavior would be closely tied to the conversation's structure so that they would predominantly look at the current speaker (Andersen, Tiippana, Laarni, Kojo, & Sams, 2008; Buchan et al., 2007; Gullberg & Holmqkvist, 2006; Hirvenkari et al., 2013). The interpersonal dialogue was expected to activate social attention (Choi et al., 2010; Frischen et al., 2007; Graham & LaBar, 2012; Schrammel et al., 2009) as indexed by the increased rate of switching the attention between the discussants, faster shifts of attention from the first to the second speaker, or the higher proportion of fixations on the second speaker after the interpersonal comment of the first speaker. Finally, based on the findings that

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

depressed people showed reduced eye contact in social interaction (Segrin, 1992), we hypothesized that depressed individuals' gaze patterns would be less aligned with the discussants' turn-taking behavior as reflected by the observers' delayed gaze shifts in relation to turn taking, lower rate of attention shifting between the discussants and looking less at the discussants. Additionally, based on the negative bias in depression (Beck, 2008), the negative vs. the positive valence of the personal conversations might especially activate social attention among the depressed participants more than in the control group. In other words, it might induce a higher proportion of fixations on the receiver of the negative statement (the second speaker), and possibly also faster shift of attention to the second speaker or a higher rate of attention shifts between the discussants.

Methods

Participants

Forty-one adult volunteers (37 female, age range: 18–64, M age: 45.2) who were native Finnish speakers participated in the study. The depressed volunteers ($N = 25$, M age $= 38.9$, $SD = 14.6$) participated in a psychiatric interview administered by a physician independently from the study to confirm the diagnosis. The depression diagnosis was based on the criteria of the International Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) and the information available from each interviewee. Six participants met the criteria for mild depression (F32.0). Two were diagnosed with a mild dysthymic disorder (F34.1) and 11 with a recurrent depressive disorder with a mild current episode (F33.0). Two participants met the criteria for moderate depression (F32.1), and three were diagnosed with a recurrent depressive disorder with a moderate current episode (F33.1). All the depressed participants scored 12 or higher on Beck's Depression Inventory (BDI-II) (Beck, Steer, & Brown, 1996), with a mean score of 23.0 (range: 12–36, $SD = 7.2$).

Additionally, 16 participants (M age $= 48.6$, $SD = 11.2$) who neither had current depression nor a history of depression or other psychiatric disorders were recruited to form a control group. Their BDI scores were 2.5 on average (range: 0–8, $SD = 3.8$).

The exclusion criteria for all the participants included a history of neurological injury or disease, substance use or addiction, coexisting bipolar or psychotic disorder, and active suicidal ideation.

Written informed consent was obtained from the participants before their participation. The ethical committee of the University of Jyväskylä approved the research protocol. The experiment was undertaken in accordance with the Declaration of Helsinki.

Apparatus

Eye movements were recorded by using an Eyelink 1000 table-mount eye-tracking system (SR Research Ltd., Canada) at a 1000-Hz sample rate. A standard 5-point calibration-validation procedure allowing a maximum of 0.3° error was performed in the beginning of the experiment and after 32 trials at the halfway point of the experiment. The stimuli were presented on a Dell Precision T5500 workstation with an Asus VG-236 (1920 × 1080, 120 Hz, 52 × 29 cm) monitor. The participants viewed the stimuli at a distance of 60 cm.

Procedure and Stimuli

The participants were instructed to watch the video clips as if looking at any video or film, with no particular task assigned. A drift correction procedure was executed between the trials, requiring participant to fixate a target placed at the center of the screen until the experimenter started the next trial by a key press.

Figure 1 presents a screenshot from a stimulus video, overlaid with eye movements from a single trial. In total, 128 discussions were created, with each participant watching 64 videos in the experiment. The two actors were seated beside each other in front of a white wall. A high-definition (HD)-quality video recorder (Canon Legria HF200) was firmly fixed and zoomed so that only the discussants' heads were visible. Each video consisted of the following phases: 2 s of silence, the first line spoken by one discussant, 2 s of silence, the second line spoken by the other discussant, followed by 5 s of silence. There were four male actors in each series. All combinations of actors with both speaking sequences were recorded, and a mirrored version of each clip was created. The videos were edited with video editing software (Adobe Premiere Pro CS5.5). In the experiment, the videos were played in HD format (resolution 1920 × 1080, 25 frames/s). The stimulus size corresponded horizontally to 32 visual degrees and vertically to 18 degrees (Figure 1).

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
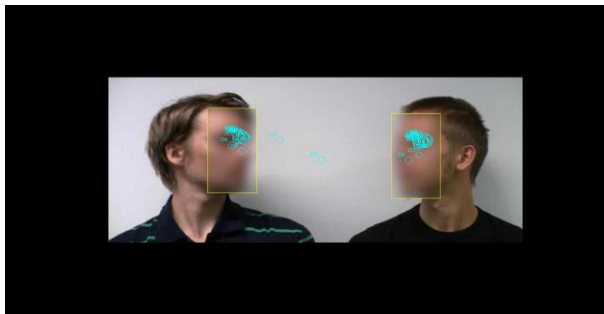Effects of conversation content on viewing dyadic conversations.

Figure 1. Screenshot from a stimulus video overlaid with area-of-interest and fixation position data from a single trial. Faces are blurred in this figure for making the actors unidentifiable. In the actual stimuli the faces were not blurred.

The Appendix presents the English translations of all 64 spoken lines used in the study. In each video, two discussants present their respective opinions ("I think that…") one at a time. Half of the conversations were related to personal issues (e.g., "I think that your style is lovely" or "I think that your style is tasteless"), followed by a reply (e.g., "I think that my style is lovely" or "I think that my style is tasteless"). The other half were related to matter-of-fact items (e.g., "I think that the vehicle is a truck" or "I think that the vehicle is a bus"), followed by a reply (e.g., "I think that the vehicle is a truck" or "I think that the vehicle is a bus"). Among the 32 personal conversations, 16 began with a positive emotional statement and the other 16 with a negative one. The effect of this valence factor was analyzed separately. Finally, to make the conversations appear more natural, unpredictable, and contain more meaningful content, in 50% of the conversations, the discussants agreed; in the other 50%, they disagreed.

The phonetic lengths of the sentences in different conditions were controlled. The presentation order of the videos was counterbalanced by the topics, actors, and mirrored vs. original video clips. The actors were instructed and practiced speaking the dialogues aloud without making any nonverbal gestures, and only such performances were selected for the set of stimuli videos. This was confirmed by a random performance (56% accuracy against the 50% guessing level) of three independent adult raters, who watched the videos with no sound and judged whether each conversation was personal, matter-of-fact, or they cannot be certain (the last option was selected for 46% of the cases).

Eye-movement Data Processing

Two areas of interest (AOIs) were assigned manually to each stimulus video, one corresponding to the left discussant's face and the other to the right discussant's face (Figure 1). These areas covered 82.4% of the fixations on the data. The timing of each spoken line was manually determined by setting the beginning and end times of each in the videos on a millisecond scale. In the analyses, the conversations were divided into three time periods, consisting of the first spoken line (period 1 from the start to the end), the duration between the lines (period 2 from the end of the first spoken line to the start of the second spoken line), and the second spoken line (period 3 from the start to the end). Four dependent measures were selected. Saccadic latency was analyzed from the start of each time period to the time point when the first saccade was launched from one discussant to another. If the time period changed during a fixation, the preceding saccade was included. The saccadic rate (per second) between the discussants in each time period indexed how often the viewer changed his or her gaze from one discussant to another. This measure should reflect the overall level of social attention activity at the interpersonal level when watching the videos. The proportion of fixation on the first vs. the second speaker in each time period indexed the comparative duration each discussant was looked at for the different time periods. These two measures revealed which of the discussants the viewer was more interested in at different phases of the conversation.

Analyses

Because of the novelty of the experimental paradigm, we first provide some descriptive analyses of the gaze behavior during the task. The descriptive data show what type of attention shifts occurred during the task and how they were aligned with the flow of the conversation.

Each dependent variable was analyzed with separate mixed-design analyses of variance (ANOVA) for repeated measures. For studying the effects of interpersonal content, the within-subject factors were the topic of the conversation (personal or matter-of-fact) and the time period (periods 1, 2, and 3, respectively corresponding to the first spoken line, the silence between the lines, and the second spoken line). For studying the effects of the conversation valence, the within-subject factors were the valence of the first spoken line (positive vs. negative) and the time period

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

(as defined in the first ANOVA model). Repeated contrasts between successive time periods were used. The interesting aspect of the interaction was whether there would be a significant difference between matter-of-fact and personal conversations in each of the periods (which were studied by paired t-tests), whereas the overall level of the dependent variables should conform to the main effect of the period in both types of conversations. Partial eta-squared (η_p^2) measures were used for effect size descriptions in ANOVA. Post-hoc analyses were performed with paired-sample t-tests, with a significance level of $p \leq 0.05$. Only significant effects are reported in this paper. A between-subject factor (group: depressed vs. control) was used in both analyses. Additionally, we calculated the correlations between BDI-II scores and eye movement measures in each time period. A false discovery rate of 10% was used to adjust the significance level of the multiple correlations, without being overly conservative (Benjamini & Hochberg, 1995).

Results

Descriptive Analysis of Gaze Behavior during the TaskTable 1 lists the frequencies of attention shifts from the first to the second speaker during the task. These data show that the majority of attention shifts were directed to the person who started (period 1) or would soon start to talk (period 2). Moreover, during the second spoken line, relatively few attention shifts were directed away from the speaker. Figure 2 shows the histograms of saccadic latencies of these attention shifts relative to the onset of the time periods. These data indicate that the attention shifts to the current speaker were closely aligned with the onsets of second spoken line, whereas the attention shifts from the current speaker were much more randomly scattered over time. Next, we report the results of the time-period locked analyses with mean values presented in Figure 3.

Table 1

Frequencies of attention shifts during each time period.

| | Type of attention shift | |

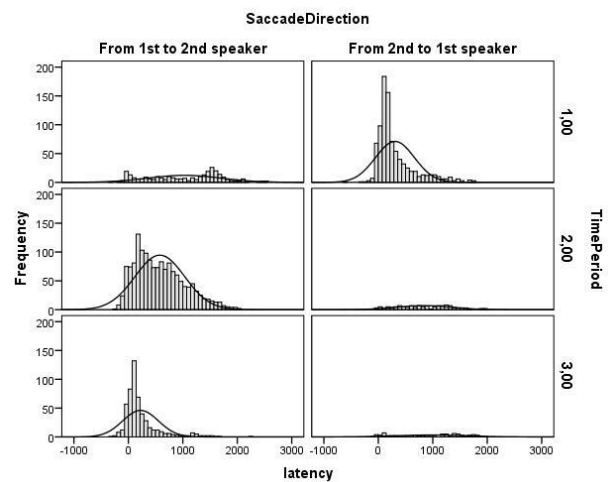| Time period	From 1st to 2nd speaker	From 2nd to 1st speaker	Total
1: First spoken line	288	890	1,178
2: Between	1,517	123	1,640
3: Second spoken line	529	74	603
Total	2,334	1,087	3,421

Figure 2. Histograms of saccadic latencies from 1st to 2nd (left panels) and from 2nd to 1st (right panels) speaker in each time period.

Saccadic latency.

Conversation content. An ANOVA for topic x period x group revealed a main effect of the period; $F(2, 35) = 14.32$, $p < 0.001$, $\eta_p^2 = 0.450$. The saccadic latencies were 381 ms, 567 ms, and 292 ms for periods 1, 2, and 3, respectively ($ps < 0.001$). The two-way interaction of topic x period, $F(2, 35) = 5.06$, $p = 0.012$, $\eta_p^2 = 0.224$, was significant in both planned contrasts between period 1 vs. 2, $F(1, 36) = 7.00$, $p = 0.012$, $\eta_p^2 = 0.163$, and period 2 vs. 3, $F(1, 36) = 6.90$, $p = 0.013$, $\eta_p^2 = 0.161$. However, the latencies for personal vs. matter-of-fact statements were only trend-like longer in periods 1 and 3, whereas in period 2, this was reversed, with latencies being trend-like longer in matter-of-fact conversations, all $p > 0.050$. Additionally, there was a trend for the three-way interaction of topic x period x group, F(2, 35) = 2.77, $p = 0.077$, $\eta_p^2 = 0.136$, indicating that the latencies in period 2 for matter-of-fact vs. personal conversations were delayed only in the control group.

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

Valence. An ANOVA for valence x period x group only showed a trend toward a main effect of valence, $F(1, 31) = 3.18$, $p = 0.084$, $\eta_p^2 = 0.093$, and valence x period, $F(2, 30) = 3.21$, $p = 0.055$, $\eta_p^2 = 0.176$, with a significant contrast between periods 1 and 2, $F(1, 31) = 5.78$, $p = 0.021$, $\eta_p^2 = 0.159$. The latencies were delayed for negative vs. positive conversations during period 1 (463 vs. 327 ms; $F(1, 40) = 6.20$, $p = 0.017$, $\eta_p^2 = 0.134$). There were no significant effects for the group or any of its interactions.

Post-hoc analysis. Figure 2 shows that in period 1, the saccades directed from the second to the first speaker were more common and time locked to the onset of the first spoken line, whereas the saccades from the first to the second speaker were fewer, and their latency was distributed more evenly in time. In a post-hoc analysis, we studied whether the latencies in both of these types of saccades would be subject to semantic modulation. Only the saccade latencies from second to first speaker seemed to be affected by the semantic content of the conversation, that is, a 150-ms delay on personal over matter-of-fact statements and a 200-ms delay on negative over positive personal statements. When these more infrequent types of attention shifts were removed from the data, the effects of conversations on saccadic latencies disappeared in period 1, for topic, $t(40) = -606$, $p = 0.548$, and for valence, $t(40) = 0.693$, $p = 0.493$.

Saccadic rate between discussants.

Conversation content. In the ANOVA for topic x period, a main effect of topic, $F(1, 39) = 5.27$, $p = 0.027$, $\eta_p^2 = 0.119$, resulted from the higher saccadic rate for personal (mean = 0.297) vs. matter-of-fact (mean = 0.277) conversations. The main effect of the period, $F(2, 38) = 75.7$, $p < 0.001$, $\eta_p^2 = 0.799$, resulted from the higher saccadic rate during period 2 (0.465) vs. period 1 (0.263) or period 3 (0.134) ($ps < 0.001$). The result of the two-way interaction of topic x period, $F(2, 38) = 6.05$, $p = 0.005$, $\eta_p^2 = 0.241$, was significant in both planned contrasts between period 1 vs. 2 and period 2 vs. 3, $F(1, 39) = 4.95$, $p = 0.032$, $\eta_p^2 = 0.113$. The saccadic rate was .040 higher in personal vs. matter-of-fact conversations during both periods 1, $F(1, 40) = 7.27$, $p = 0.010$, $\eta_p^2 = 0.543$, and 2, $F(1, 40) = 4.24$, $p = 0.046$, $\eta_p^2 = 0.096$, whereas during period 3, this difference was not present ($p > 0.050$).

Valence. There were no significant effects in the ANOVA for the valence or the group.
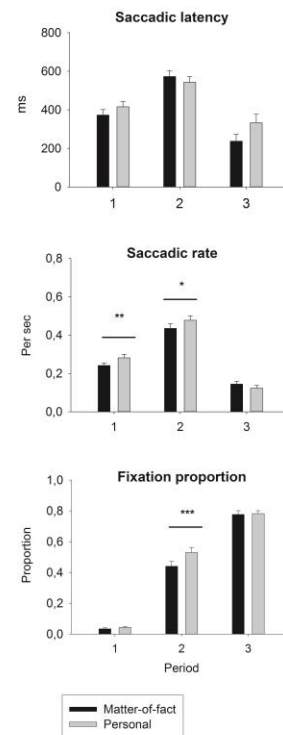

*Figure 3. Condition-wise means and standard errors of eye-movement measures of social attention during each time period. Saccadic latency indicates how fast, after the beginning of a time period, the participants shift their attention to the other discussant. Saccadic rate describes how often the participants shift their attention between the discussants. The fixation proportion on the second speaker indicates how large is the portion of fixations directed to the second speaker in each time period. ***p < 0.001, **p < 0.01, *p < 0.05*

Fixation proportions on first speaker.

Interpersonal content. In the ANOVA for topic x period, a main effect of the topic, $F(1, 39) = 7.41$, $p = 0.010$, $\eta_p^2 = 0.160$, was due to a slightly higher probability of looking at the first speaker during matter-of-fact vs. personal conversations (the probabilities were 0.332 and 0.311, respectively). The main effect of the period, $F(2, 38) = 693$, $p < 0.001$, $\eta_p^2 = 0.973$, reflected less fixations on the current speaker in the later time periods (from 0.726 during period 1 to 0.225 during period 2 and to 0.013 during period 3; $ps < 0.001$). The two-way interaction of topic x period, $F(2, 38) = 7.39$, $p = 0.002$, $\eta_p^2 = 0.280$, was significant between period 2 vs. 3, $F(1, 39) = 10.97$, $p = 0.002$, $\eta_p^2 = 0.220$. The first speaker was looked at more during matter-of-fact vs.

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

personal conversations in period 2 than in period 3 (0.244 vs. 0.206; $F(1, 40) = 11.17$, $p = 0.002$, $\eta_p^2 = 0.218$).

Valence. There were no significant effects in the ANOVA for the valence x period or the group.

Fixation proportions on second speaker.

Conversation content. The ANOVA for topic x period indicated a main effect of the topic, $F(1, 39) = 23.45$, $p = 0.0001$, $\eta_p^2 = 0.376$, which resulted from the higher probability of looking at the second speaker during personal (0.455) vs. matter-of-fact (0.419) conversations. The main effect of the period, $F(2, 38) = 597.7$, $p < 0.001$, $\eta_p^2 = 0.969$, resulted from the increased probability of looking at the second speaker from period 1 to period 3 (the means for periods 1, 2, and 3 were 0.036, 0.487, and 0.788, respectively, $ps < 0.001$). The two-way interaction of topic x period, $F(2, 38) = 8.80$, $p = 0.001$, $\eta_p^2 = 0.317$, resulted from significant planned contrasts between period 1 vs. 2, $F(1, 39) = 18.06$, $p < 0.001$, $\eta_p^2 = 0.316$, and period 2 vs. 3, $F(1, 39) = 9.09$, $p = 0.005$, $\eta_p^2 = 0.189$. This was because the difference between personal (0.533) and matter-of-fact (0.441) conversations was present only during period 2, $F(1, 40) = 24.4$, $p < 0.001$, $\eta_p^2 = 0.379$. There were no significant effects in the ANOVA for the group.

Valence. The ANOVA for valence x period showed a significant main effect of valence, $F(1, 39) = 4.5$, $p = 0.040$, $\eta_p^2 = 0.103$, resulting from the higher probability of looking at the second speaker during conversations that had a negative (0.456) vs. a positive (0.435) first spoken line. There was a nearly significant two-way interaction of valence x period, $F(2, 38) = 3.08$, $p = 0.057$, $\eta_p^2 = 0.140$, resulting from significant contrasts between period 1 vs. 2, $F(1, 39) = 5.71$, $p = 0.022$, $\eta_p^2 = 0.128$, and period 2 vs. 3, $F(1, 39) = 5.41$, $p = 0.025$, $\eta_p^2 = 0.122$, collectively indicating that it was the period between the spoken lines, $F(1, 40) = 6.95$, $p = 0.012$, $\eta_p^2 = 0.148$, when the second speaker was looked at relatively more if the first line had a negative (0.555) vs. a positive (0.485) valence. The two-way interaction of period x group was approaching significance, $F(2, 38) = 3.08$, $p = 0.057$, $\eta_p^2 = 0.140$. The nature of this trend was that the members of the control group tended to look more at the second speaker during period 3 in comparison to depressed individuals.

Correlations between Eye-movement Measures and BDI-II Scores

Table 2 presents the correlations between BDI-II scores and eye movement measures in each time period and Figure 4 presents the scatterplots of selected significant correlations. The fixation proportions on the first speaker in period 1, $r(41) = -0.431$, $p = 0.005$, and period 3, $r(41) = 0.374$, $p = 0.016$, and the fixation proportions on the second speaker in period 3, $r(41) = -0.396$, $p = 0.010$, correlated significantly with the BDI scores. These correlations indicated that individuals with higher BDI scores looked less at the current speaker. The positive correlation with the BDI score and the saccadic latency in period 1, $r(41) = 0.324$, $p = 0.039$, indicated a slower shift of attention to the second speaker by individuals with higher BDI scores. The BDI correlation with the saccadic rate in period 2, $r(41) = -0.358$, $p = 0.021$, indicated a smaller number of attention shifts between the discussants in period 2.

Table 2

Pearson correlations between eye-movement measures and BDI-II score.

	Period 1	Period 2	Period 3
Saccadic latency	0.324*	-0.235	-0.063
Saccadic rate	-0.117	-0.358	-0.173
Fixation prop. 1st speaker	-0.431*	-0.074	0.374*
Fixation prop. 2nd speaker	0.288	-0.036	-0.396*

$p < 0.1$, N = 41

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
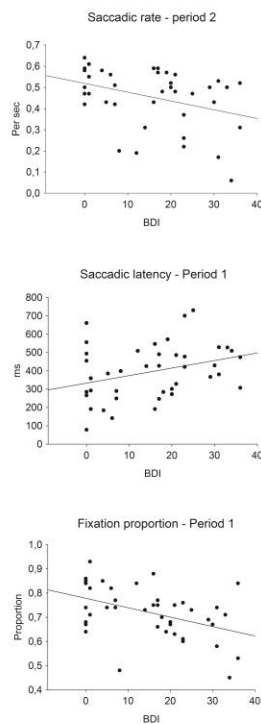Effects of conversation content on viewing dyadic conversations.

Figure 4. Scatter plots of selected significant correlations between BDI-II scores and means of eye-movement measures in different time periods. The bottom panel shows fixation proportions on the first speaker when he talks.

Discussion

This study's results show that the semantic content of the conversations dynamically modulates the viewers' gaze behavior in important ways. This is in line with recent evidence that semantic information is also used for predicting discussants' turn-taking behavior (Riest et al., 2015). We suggest that the underlying principle explaining this modulation is to predict (Kilner, Friston, & Frith, 2007) the social consequences of what is being said.

During the first spoken line, the majority of the attention shifts were directed from the second to the first speaker as soon as he started to talk. Our post-hoc analysis revealed that the latencies of these saccades were not affected by the conversation's content. Thus, it can be concluded that these saccades are determined by the discussants' turn-taking behavior, as found in previous studies

(e.g., Hirvenkari et al., 2013; Riest et al., 2015). In contrast, in some cases, the attention was already shifted to the second speaker during the first speaker's statement, and it was these saccades that seemed to be delayed if the statement was personal and especially if it was negative. The underlying cognitive and/or emotional mechanisms explaining this finding are unknown. Future work should pursue whether such transient attention capture is caused merely by emotional reactions induced by such statements or by the intent to maximize the visual information uptake (lip movements and nonverbal gestures) about the person saying such things. On the other hand, faster latencies on the second speaker in response to personal and provocative statements could also be expected upon seeing the earliest emotional responses of the receiver, but based on our data, allocating social attention to the current speaker seems to be prioritized for the very first, before making attentional shift to another discussant.

However, such attention capture was very short-lived. During the same first spoken line and between the lines, the saccadic rate between the discussants already increased during personal vs. matter-of-fact conversations. Between the spoken lines, the second speaker was fixated proportionally more in personal conversations, especially when the first spoken line had a negative valence, while the first speaker was fixated proportionally more in the matter-of-fact conversations. Together, these findings showed that in personal vs. matter-of-fact conversations, the viewers were relatively more interested in the second speaker. This is understandable because in these two-line conversations, the first speaker takes the initiative, while the second speaker is expected to react. This expectation then seems to be especially high when the initiative is highly provocative, as is the case during the negative personal conversations. One possible reason for paying attention to the second speaker could be that such a provocation might be responded to aggressively; a threat is known to elicit fast attention responses for protective purposes (e.g., Kellough et al., 2008).

The overall implication of these findings is that socio-emotionally important discussions induce generally more social attention. After the highly transient focus of attention on a person saying important things, people then rapidly seem to shift their attention to the receiver of such statements. This sequence of visual attention allocation may maximize both the linguistic and the social information uptake about both the speaker and the listener. In a

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

real social situation, it may be crucial to look carefully at the speaker to determine if his or her linguistic message is associated with some nonverbal gestures, potentially even altering the meaning of the statement. However, at the very next moment, observers are already interested about other people's reactions to such an important message. Fundamentally, cognition involves a continuous prediction of the future based on the memory of the past, so typically, all cognitive processing is biased toward prediction instead of a detailed analysis of current and past events, that is, the so-called predictive coding of the environment (Kilner et al., 2007). In the social context, such predictions are most urgent concerning severely negative statements directed to other people as these may be responded to aggressively, requiring possible fast responses from others (Kellough et al., 2008). Deeper processing of what happened can then be reflected on, relying on people's own memories and those of others.

A descriptive analysis of gaze behavior showed that attention shifts between the discussants followed the conversational turn taking in a closely timed fashion; most of the viewers continued looking at the current speaker instead of shifting their attention frequently between the discussants. These findings are consistent with those of previous studies in which people mostly watched the current speaker (Andersen et al., 2008; Buchan et al., 2007; Gullberg & Holmqkvist, 2006; Hirvenkari et al., 2013). Then, between the lines, our study's uninvolved observers predictably shifted their attention to the second speaker, as was recently found in another study (Holler & Kendrick, 2015). Such anticipation was not observed in a study exploring gaze behavior when watching videos of real conversations (Hirvenkari et al., 2013). This result suggests that such anticipation occurs only in conditions when the turn taking is predictable or when the viewer is highly engaged in the conversation, for example, participating in a real dialogue or required to react to it somehow (see Holler & Kendrick, 2015).

Individuals with psychiatric disorders show biases when watching socio-emotional pictures. One particularly well-documented example is depression-related negativity bias (e.g., Beck, 2008), in which greater attention is allocated to emotionally negative content. Our correlational analyses generally indicated that more depressed individuals' gaze behavior was less aligned with the conversational flow, demonstrated by looking less at the current speaker, a slower shift of attention to the other speaker than

to the currently fixated discussant during the first spoken line, and less saccadic shifting between the spoken lines. In the statistical analysis, we only found the depressed individuals' tendency to look less at the second speaker while he was talking, in comparison to the control participants. A possible explanation for the effects' failure to reach a significant level in the statistical analysis may be that only a subgroup or the most seriously affected individuals show clear disturbances in their dynamic control of social attention. Overall, these findings are in line with the bulk of the studies showing that depressed people are socially less active and make less eye contact with other people (for a review, see Scherer et al., 2013; Segrin, 1992). More specifically, our results indicate that severe depression may weaken people's attention capabilities to follow even simple conversations, which in turn may contribute to their tendency to withdraw from social interactions.

The present study also pinpoints some methodological challenges in naturalistic studies of conversations. To ensure that the effects would stem from people's internal social representations (the semantic relationship between the lines in a dialogue), we did our best to control all other possible sources of variation between conditions, resulting in stereotypic and inexpressive conversations. Moreover, the subjects viewed the videos passively without being involved in the conversations in any way. In future studies, researchers may consider adding comprehension control questions after some of the videos to ensure that viewers are paying attention to what is being said during the conversations. These factors may have attenuated the present effects of conversational content on gaze. The effects of conversational content also seem very transient, so future studies may aim to analyze the responses at the level of single-word timing. We fully acknowledge that real-life social interaction is much more complex than studied here. However, if these findings are also replicated when watching more natural and less controlled conversations, researchers can begin to grasp what aspects of social interaction are important dynamic determinants of social attention.

Conclusions

We conclude that eye movements are meaningfully associated with the semantic content of speech while following dynamic conversations and that a depressed patient's

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

gaze is less aligned with the conversational flow. Eye movements while watching naturalistic social interactions thus constitute a reliable index of the social attention and psychological condition of a viewer.

Acknowledgments

References

Andersen, T.S., Tiippana, K., Laarni, J., Kojo, I., Sams, M. (2008). The role of visual spatial attention in audiovisual speech perception. *Speech Communication*, *51*(2), 184-193.

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review*, *32*(8), 704-723.

Beck, A.T. (2008) The Evolution of the Cognitive Model of Depression and Its Neurobiological Correlates. *American Journal of Psychiatry, 165*, 969–977.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

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

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*, *2*, 1-13.

Caseras, X., Garner, M., Bradley, B. P., & Mogg, K. (2007). Biases in visual orienting to negative and positive scenes in dysphoria: An eye movement study. *Journal of abnormal psychology*, *116*(3), 491.

Choi, S., Ku, J., Han, K., Kim, E., Kim, S. I., Park, J., Kim, J. (2010). Deficits in eye gaze during negative social interactions in patients with schizophrenia. *The Journal of Nervous and Mental Disease, 198*, 829-835.

DeWall, C. N., Maner, J. K. & Rouby, A. D. (2009). Social Exclusion and Early-Stage Interpersonal Perception: Selective Attention to Signs of Acceptance. *Journal of Personality and Social Psychology*, *96*, 729–741.

Eizenman M., Yu, L. H., Grupp, L., Eizenman, E., Ellenbogen, M., et al. (2003). A naturalistic visual scanning approach to assess selective attention in major depressive disorder. *Psychiatry Research., 118*, 117–28.

Ellis, A. J., Beevers, C. G., & Wells, T. T. (2011). Attention allocation and incidental recognition of emotional information in dysphoria. *Cognitive Therapy and Research*, *35*(5), 425-433.

Foulsham, T., Cheng, J.T., Tracy, J.L., Henrich, J. & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition, 117*, 319-331.

Foulsham, T., Walker, E. & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*, 1920-1931.

Frischen, A., Bayliss, A. P., Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychogical. Bulletin. 133*, 694-724.

Garrod, S., & Pickering, M. (2009). Joint action, interactive alignment, and dialog. *Topics in Cogitive Science. 1,* 292-304.

Graham R, & LaBar K. S. (2012). Neurocognitive mechanisms of gaze-expression interactions in face processing and social attention. *Neuropsychologia*, *50,* 553-566.

Gullberg, M., & Holmquist, K. (2006). What speakers do and what addresses look at—Visual attention to gestures in human interaction live and on video. *Pragmatics & Cognition*, *14*, 53–82.

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

Hayes, S. C. (2004). Acceptance and commitment therapy and the new behaviour therapies. In S. C. Hayes, V. M. Follette, & M. M. Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive-behavioural tradition* (pp. 1-29). New York: Guilford.

Hirvenkari, L., Ruusuvuori, J., Saarinen, V-M., Kivioja, M., Peräkylä, A., et al. (2013). Influence of turn-taking in a two-person conversation on the gaze of a viewer. *PLoS ONE*, *8*, e71569.

Ho, S., Foulsham, T., & Kingstone, A. (2015). Speaking and listening with the eyes: gaze signaling during dyadic interactions. PloS one, 10(8), e0136905.

Holler, J., & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: optimizing recipiency. *Frontiers in psychology*, *6, 98.*

Isaacowitz, D. M. (2005). The gaze of the optimist. *Personality and Social Psychology Bulletin*, *31*, 407-415.

Kellough, J. L., Beevers, C. G., Ellis, A. J. & Wells, T. T. (2008). Time course of selective attention in clinically depressed young adults: an eye tracking study. *Behaviour Reearch and Therapy. 46,* 1238–1243.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive processing*, *8*(3), 159-166.

Kingstone, A. (2009). Taking a real look at social attention. *Current Opinion in Neurobiology, 19*, 52–56.

Leyman, L., De Raedt, R., Vaeyens, R., & Philippaerts, R. M. (2011). Attention for emotional facial expressions in dysphoria: An eye-movement registration study. *Cognition and Emotion*, *25*(1), 111-120.

Lykins, A. D., Meana, M., & Strauss, G. P. (2008). Sex differences in visual attention to erotic and non-erotic visual stimuli. *Archives of Sexual Behaviour, 37*, 219–228.

Maner, J. K., DeWall, C. N., & Gailliot, M. T. (2008). Selective attention to signs of success: Social dominance and early stage interpersonal perception. *Personality and Social Psychology Bulletin*, *34,* 488–501.

Nummenmaa, L. Hyönä, J., & Hietanen, J.K. (2009). I'll walk this way: Eyes reveal the direction of locomotion and make passers-by to look and go the other way. *Psychological* Science, *20,* 1454-1458.

Nummenmaa, L., Hietanen, J. K., Santtila, P., & Hyönä, J. (2012). The body of evidence: gender and visibility of sexual cues modulate visual scanning of faces and bodies. *Archives of Sexual Behaviour, 41,* 1439-1451.

Rigoulot, S., & Pell, M. D. (2012). Seeing emotion with your ears: emotional prosody implicitly guides visual attention to faces. *PLoS One*, *7*(1), e30740.

Richardson, D.C., Dale, R., & Tomlinson, J.M. (2009) Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science, 33*, 1468-1482.

Riest, C., Jorschick, A. B., & De Ruiter, J. P. (2015). Anticipation in turn-taking: mechanisms and information sources. *Frontiers in Psychology*, *6, 89.*

Scherer, S., Stratou, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A., & Morency, L. P. (2013). Automatic behavior descriptors for psychological disorder analysis. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on* (pp. 1-8). IEEE.

Schrammel, F., Pannasch, S., Graupner, S. T., Mojzisch, A., Velichkovsky, B. M. (2009). Virtual friend or threat? The effects of facial expression and gaze interaction on psychophysiological responses and emotional experience. *Psychophysiology, 46*, 922–931.

Schwartz, R., & Pell, M. D. (2012). Emotional Speech Processing at the Intersection of Prosody and Semantics. *PloS One*, *7*, e47279.

Sears, C. R., Bisson, M. S., & Nielsen, K. E. (2011). Dysphoria and the immediate interpretation of ambiguity: Evidence for a negative interpretive bias in error rates but not response latencies. *Cognitive Therapy and Research*, *35*(5), 469-476.

Skarratt, P. A., Cole, G. G., & Kuhn, G. (2012). Visual cognition during real social interaction. *Frontiers in Human Neuroscience, 6*, 196.

Journal of Eye Movement Research
9(7):5, 1-12

Hautala, J., Loberg, O., Hietanen, J. K, Nummenmaa, L. & Astikainen, P. (2016)
Effects of conversation content on viewing dyadic conversations.

Appendix: Spoken Lines in the Dyadic Conversations

The conversations included all four order-specific combinations of the two spoken lines presented on each row (i.e., AA, BB, AB, and BA) for the matter-of-fact conversations. The personal conversation always started with one variant of spoken line 1 by the first speaker (either positive or negative adjective valence), followed by one variant of spoken line 2 (either positive or negative adjective valence, which was either congruent or incongruent with spoken line 1) as the reply of the second speaker.

Matter-of-fact	Spoken line A	Spoken line B
relationship	I think that the relationship is a marriage.	I think that the relationship is common law.
dog	I think that the dog is a hound.	I think that the dog is a retriever.
friend	I think that the friend is a barber.	I think that the friend is a baker.
wife	I think that the wife is an aunt.	I think that the wife is a godmother.
achievement	I think that the achievement is collective.	I think that the achievement is his.
work	I think that the work is washing dishes.	I think that the work is cleaning up.
child	I think that the child wants to sleep.	I think that the child wants to eat.
idea	I think that the idea is Markku´s.	I think that the idea is Paavo´s.
car	I think that the vehicle is a truck.	I think that the vehicle is a bus.
sister	I think that the sister is dark.	I think that the sister is blond.
style	I think that the style is comedy.	I think that the style is drama.
father	I think that the father is retired.	I think that the father is working.
customary	I think that it is customary to rise.	I think that it is customary to sit.
family	I think that the family comes from Savonia.	I think that the family comes from Lapland.
brother	I think that the brother is a pilot.	I think that the brother is a doctor.
mother	I think that the mother is at home.	I think that the mother is outside.
Personal	**Spoken line 1**	**Spoken line 2**
relationship	I think that your marriage is happy/unhappy.	I think that my marriage is happy/unhappy.
dog	I think that your dog is kind/angry.	I think that my dog is kind/angry.
friend	I think that your friend is nice/boring.	I think that my friend is nice/boring.
wife	I think that your wife is beautiful/ugly.	I think that my wife is beautiful/ugly.
achievement	I think that your achievement is significant/insignificant.	I think that my achievement is significant/insignificant.
work	I think that your work is important/futile.	I think that my work is important/futile.
child	I think that your child is well/not well.	I think that my child is well/not well.
idea	I think that your idea is brilliant/rubbish.	I think that my idea is brilliant/rubbish.
car	I think that your car is great/a wreck.	I think that my car is great/a wreck.
sister	I think that your sister is nice/difficult.	I think that my sister is nice/difficult.
style	I think that your style is lovely/tasteless.	I think that my style is lovely/tasteless.
father	I think that your father is safe/unstable.	I think that my father is safe/unstable.
customary	I think that your behavior is polite/annoying.	I think that my behavior is polite/annoying.
family	I think that your family is rich/poor.	I think that my family is rich/poor.
brother	I think that your brother is successful/in trouble.	I think that my brother is successful/in trouble.
mother	I think that your mother is gentle/strict.	I think that my mother is gentle/strict.