Gaze direction signals response preference
in conversation

Kobin H. Kendrick
Max Planck Institute for Psycholinguistics

Judith Holler
Max Planck Institute for Psycholinguistics

Acknowledgments

Abstract

In this article, we examine gaze direction in responses to polar questions using both quantitative and conversation analytic (CA) methods. The data come from a novel corpus of conversations in which participants wore eye-tracking glasses to obtain direct measures of their eye movements. The results show that while most preferred responses are produced with gaze towards the questioner, most dispreferred responses are produced with gaze aversion. We further demonstrate that gaze aversion by respondents can occasion self-repair by questioners in the transition space between turns, indicating that the relationship between gaze direction and preference is more than a mere statistical association. We conclude that gaze direction in responses to polar questions functions as a signal of response preference. Data in American, British, and Canadian English.

The morphology of the human eye has evolved to allow those around us to recognize the direction of our gaze. The stark contrast between the white sclera and the darker pupil of the human eye exposes our gaze direction whereas the dark sclera of nonhuman primates serves to camouflage theirs (Kobayashi & Kohshima, 1997). Our unique sensitivity to gaze direction is evident in the behavior of infants who can discriminate between direct and averted gaze just two days after birth (Farroni, Csibra, Simion, & Johnson, 2002). In the behavioral repertoire of most species, direct gaze is predominantly a signal of aggression or threat (with the possible exception of mother-infant dyads amongst some of the great apes, Gomez, 1996) while gaze aversion tends to signal submission. In much of the animal kingdom, gaze thus functions as an antithetic signal (Darwin, 1872). The scope of functions that gaze fulfills in human interaction is vast in comparison (see Rossano, 2013).

The functional complexity of human gaze is perhaps most obvious in conversation. Goffman (1964) wrote of the "eye-to-eye ecological huddle" that is characteristic of human interaction, referring to the strong human tendency to mutually orient to one another, thus facilitating mutual visual perception. Despite this tendency, however, we do not constantly maintain eye contact. Rather, face-to-face conversation consists of a complex interactional dance, as it were, with frequently alternating periods of gazing at the other and gazing away. In his landmark study of gaze direction, Kendon (1967) identified, among other patterns, a general asymmetry in the behavior of speakers and recipients: whereas recipients tend to maintain gaze to speakers, speakers alternate their gaze towards and away from recipients as they speak (see also Argyle & Cook, 1976; Bavelas, Coates, & Johnson, 2002; Ho, Foulsham, & Kingstone, 2015).

There has been much speculation on what determines the occurrence and duration of mutual gaze in conversation. Kendon (1967) proposed three main functions of gaze direction. Firstly, he argued that gaze direction serves a *regulatory function*: speakers avert their gaze to signal that they are about to take a turn, may continue to look away in order to hold the turn, and may redirect gaze to the recipient in order to yield the turn. Kendon's observations had little impact on the development of the CA model of turn-taking by Sacks, Schegloff, and Jefferson (1974), in which gaze direction serves as resource for next speaker selection (Lerner, 2003) but does not otherwise organize transitions between speakers. Outside of conversation analysis, however, the regulatory function of gaze became a central part of the cue-based model of turn-taking developed by Duncan and colleagues (Duncan 1972; Duncan

& Fiske, 1977; Duncan, Brunner, & Fiske, 1979) and continues to garner attention (e.g., Bavelas et al., 2002; Ho et al., 2015).

Secondly, Kendon (1967) proposed that gaze direction serves a *monitoring function*: by looking at the recipient, speakers gather information about his or her attentional state, facial displays, intention to take a turn, and so on. As a corollary to this, he proposed that "paying attention to one's interlocutor and planning what to say are incompatible activities" (p. 34), suggesting that speakers avert their gaze in order to avoid interference while planning upcoming units of talk. In the subsequent literature, few have investigated the monitoring function of gaze per se (M. H. Goodwin, 1980), focusing instead on its cognitive costs (Beattie, 1978). In one line of experimental research, participants are asked not to avert their gaze as they normally would (e.g., to maintain eye-contact with the experimenter as they perform an experimental task) and consequently perform worse than those who are free to avert their gaze naturally (Beattie, 1981; Doherty-Sneddon, Bonner, & Bruce, 2001; Markson & Paterson, 2009). Such results have been understood to reveal evidence of cognitive interference caused by the automatic processing of the interlocutor's facial signals. The monitoring function of gaze is thus one element of what has become the "cognitive load hypothesis" of gaze aversion, namely that speakers avert their gaze to reduce the cognitive costs associated with the monitoring function.

Thirdly, Kendon (1967) argued that gaze direction serves an *expressive function*, whereby participants regulate the level of emotionality and arousal in the interaction. In one particular conversation, Kendon found that mutual gaze was inversely related to the rate of smiling, one measure of emotionality. If the level of emotionality becomes too high, Kendon maintained, participants may avert their gaze to express embarrassment and reduce arousal. According to Kendon, the expressive function of gaze also concerns the speaker's need for affiliation as mutual gaze appears to increase in affiliative and cooperative interactions (p. 48).

Here, we propose one specification of the expressive function of gaze direction, namely that gaze maintenance and aversion serve as resources for the construction of socially affiliative and disaffiliative actions, among their other uses. In conversation, speakers use the resources of language and the body to form recognizable actions (Schegloff, 1996; Levinson, 2013), which cohere to form sequences of action (Schegloff, 2007). One of the most basic and ubiquitous forms of action sequence is the adjacency pair, in which one action, a first

pair part, makes conditionally relevant another, a second pair part (Schegloff, 2007). For some adjacency pairs, the first and second pair parts are of the same type (e.g., greetings are followed by greetings), but for most, next speakers must select from a set of alternative seconds (e.g., invitations may be followed by acceptances or declinations). The construction of such alternative second pair parts typically reveals systematic asymmetries: while one may be produced quickly, take a simple form, and promote the accomplishment of the activity, a *preferred response*, the other may be delayed by silence or turn-initial particles, have a relatively complex design, and forestall the accomplishment of the activity, a *dispreferred response* (Schegloff, 2007; Stivers et al., 2009; Pomerantz & Heritage, 2013). The concept of preference in conversation analysis was developed to explain such systematic asymmetries, which serve to maximize opportunities for affiliative actions and minimize opportunities for disaffiliative ones (Heritage, 1984; Sacks, 1987).

We hypothesize that the systematic asymmetry in the construction of preferred and dispreferred responses should be visible in the gaze behavior of participants. Specifically, we predict (i) that gaze aversion should occur more frequently in dispreferred responses to polar questions than preferred ones and (ii) that the onset of gaze aversion should occur earlier for dispreferred responses than preferred ones. There are both ethological and interactional reasons to suspect such a hypothesis to be true. In some insect and fish species, ethologists have observed, eyespots on the skin serve as deimatic displays to startle predators (Edmunds, 1974). In nonhuman primates, the stare is "a component of all threat displays", such that its antithesis, gaze aversion, signals submission (Chevalier-Skolnikoff, 1973, p. 31). In humans, too, gaze maintenance can constitute a "threat stare" (Eibl-Eibesfeldt, 1989), whereas aversion or withdrawal can signal shame or embarrassment (Darwin, 1872; Kendon, 1967).

Interactional research also lends support to the hypothesis. Kidwell (2006) has shown that gaze aversion after directives issued by police officers signals resistance and non-compliance. Although Kidwell does not present her analysis in terms of preference, resistance is a dispreferred response to a directive. Haddington (2006) has argued that gaze patterns are resources for stance taking in conversation, observing that in assessment sequences mutual gaze occurs with convergent stances (i.e., agreements) and gaze aversion with divergent ones (i.e., disagreements), relating this to preference. Thus gaze aversion, like the mitigations and qualifications frequently found in the linguistic construction of dispreferred responses, may serve as a compensatory signal, working against the threat to the

other's face (Brown & Levinson, 1987) – and hence to social solidarity (Heritage, 1984; Clayman, 2002) – that disaffiliative actions pose. Whereas in the animal kingdom gaze aversion signals submission, in human interaction it may well serve to mitigate an otherwise threating action.

We do not argue, however, that gaze direction serves no other functions in conversation. To determine whether the status of a response as preferred or dispreferred influences gaze direction above and beyond other factors, we not only test our main hypothesis through a systematic analysis of gaze direction in responses to polar questions. We also consider whether the complexity of the response, as measured by its duration, has an effect on gaze direction. The complexity of the response could influence gaze direction for two reasons. First, gaze aversion could project that an incipient response will have a complex structure with multiple units, as Heritage (2015) found for "well"-prefaced responses. Second, gaze aversion could reflect the additional cognitive effort required to plan complex responses. Our interest here is not to assess such factors per se, but rather to control for them statistically and therefore to test whether preference has an independent effect on gaze direction.

The article is organized as follows. After a description of our data and methods, we present the results of quantitative and conversation analytic investigations into the relationship between gaze direction and response preference. We then examine the time course of gaze aversions for evidence of systematic differences between preferred and dispreferred responses. Before we discuss our results, we briefly consider whether response preference influences gaze direction above and beyond response complexity.

METHODS

Data

The present study examined seven triadic conversations, each approximately 20 minutes in duration, from the Eye-tracking in Multimodal Interaction Corpus (EMIC). The recordings were made in a soundproof room in which the participants sat in a triangular formation with the chairs arranged equidistantly (see Figure S1).[1] Each participant wore eye-tracking glasses

---

[1] The *S* designation refers to items in the Supplementary Materials.

to obtain direct measures of eye movements (SMI, sampling rate 30 Hz) and lightweight head-mounted microphones (Shure SM10A) to record the individual's speech. In addition each participant was recorded by high-definition video camera (Canon Legria HFG10, 25 fps), the placement of which also recorded the profiles of the other two participants (see Holler & Kendrick, 2015, for more information about the corpus). The participants had no objects to interact with physically during the recordings, but the environment did incidentally include objects to which they could attend (curtains, tables, boxes, and so on). The participants were friends or acquaintances, were recruited in groups of three to take part in the study, and were native speakers of English who lived in the Netherlands when the recordings were made. The participants were instructed to talk for the duration of the session but were not otherwise given specific instructions, tasks, or topics. All participants gave their permission for their data to be used.

Identification of question-response sequences

All polar questions were identified in the recordings. Both formal and functional criteria were used to identify the questions, in line with the coding scheme developed by Stivers and Enfield (2010). The collection thus includes polar interrogatives (e.g., "were you ever homesick?"), polar declaratives (e.g., "but here you're just doing random stuff"), declaratives with interrogative tags (e.g., "you support Man United, don't you?"), and non-clausal questions (e.g., "on this campus?"). The frequency and proportion of each question type is given in Table S1. Only polar questions that received a response in the next turn ($n = 205$) were included in the present analysis.

The response was operationalized as one or more polar response tokens, if present, and the first subsequent turn-constructional unit (TCU, Sacks et al., 1974) of the responding turn. This included lexical TCUs (e.g., "yeah", "no"); multiple productions of such units (e.g., "yeah yeah yeah"); lexical units immediately followed by an additional unit (e.g., "yeah Farsi"); and non-conforming responses without polar response tokens (e.g., "one more year"). If an "I don't know" response was immediately followed by an additional unit, both units together were analyzed as the response (e.g., "I don't know.=I don't remember."). Multiple instances of polar response tokens or relevant alternatives were included (e.g., "yeah on the radio yeah exactly"). For multimodal responses (e.g., nodding together with "yeah"), the linguistic response was analyzed. For visual responses without linguistic components (e.g.,

nodding alone), the duration of the embodied action was measured. If a noticeable pause occurred between a possibly complete unit and a subsequent unit, only the first unit was analyzed as the response.

For multi-unit responses, only the first unit beyond the response token, as defined above, was analyzed as the response. The boundaries of the response were operationalized in this way because the distinction between a relevantly multi-unit response and contingently multi-unit one could not be made with certainty for all sequences in the collection. In some cases, a polar question can make a multi-unit response relevant (e.g., a response to a topic initiation), whereas in others, the question does not necessarily make such a response relevant yet the response nonetheless includes multiple units (e.g., a response to a request for information that the speaker opportunistically expands). This distinction, while clear in theory, was ambiguous in many cases and was not applied for this reason. Further basic CA research on multi-unit responses is necessary to clarify the distinction and its application across a variety of cases.

Analysis of preference

CA research has shown that polar questions establish constraints on the response such that some responses are preferred over others (Sacks 1987; Raymond, 2003; Heritage & Raymond, 2012; Raymond 2010; Bolden, in press). The present study examined one particular aspect of the preference organization of polar question sequences, namely whether the response conformed to the expectations set by the grammatical format of the question. According to Boyd and Heritage (2006), the linguistic construction of a polar question – principally the selection between interrogative and declarative formats, affirmative and negative polarity, and the presence or absence of polarity items – establishes a preference for a "yes" or "no" answer. Heritage (2010) presents this as a set of associations between grammatical formats and preferences, given here with modification as Table 1 (see also Heritage & Clayman, 2010; Hayano, 2013). Non-clausal polar questions, which neither Boyd and Heritage (2006) nor Heritage (2010) discuss, can be understood to prefer "yes" in the affirmative and "no" in the negative. Responses that conformed to the preference set by the grammatical format of the question were analyzed as preferred and those that did not were analyzed as dispreferred.

Grammatical format	Example	Preference
affirmative interrogatives	"did you get it here?"	"yes"
declaratives with negative tags	"there was like a fifth one, wasn't there?"	"yes"
declarative questions	"so you have family here"	"yes"
negative interrogatives	"doesn't she live in Canada for half the year?"	"yes"
affirmative non-clausals	"in six months?"	"yes"
negative declaratives	"you haven't heard of it"	"no"
negative declaratives with positive tags	"you never had any regrets, have you?"	"no"
affirmative interrogatives with negative polarity items	"were you ever homesick?"	"no"
negative non-clausals	"no?"	"no"

Table 1: Grammatical formats for polar questions that prefer "yes" and "no" responses, adapted from Heritage (2010).

In addition to the analysis of preference by grammatical format, a small number of cases warranted special consideration. Clear cases of pro-forma agreement (Schegloff, 2007, p. 69-70; e.g., A: "you can get better accommodation though, can't you, if you're here next year" B: "uh::: yeah I don't know. I'll have to (.) look into it.") were analyzed as dispreferred. In two cases, next turns after preliminary questions did not respond to the question per se but rather apparently responded to a projected action (C: "have you seen the American version of the In Betweeners?" A: "oh it's aw:ful."; C: "have you seen Coach Trip as well," A: "Coach Trip is ama:zing."). Because these sequences were initiated by polar interrogatives, they were necessarily included in the collection. However, a binary analysis of preference presented two suboptimal choices: to analyze the responses as preferred would ignore the fact that they do not answer the questions per se, but to analyze them as dispreferred would ignore their cooperative and affiliative character. Rather than exclude such cases, we opted to analyze them as preferred because they are affiliative. Visual responses (e.g., head nods) were analyzed like their linguistic alternatives.

The analysis of preference did not address the preference for type-conformity in polar interrogative sequences (Raymond, 2003) because the collection of polar questions includes not only interrogatives but also declaratives for which the constraints on the form of the response may differ (see Raymond, 2010). The analysis of preference also did not address the

possible occurrence of cross-cutting preferences in which the action the question implements prefers a different response than its grammatical form (Schegloff, 2007). Most polar questions in the collection (of 184 cases, see below) implemented requests for information or confirmation (64.7%, n =119), in line with Stivers' (2010) study of naturally occurring conversation in which 62% of polar questions ($n = 133$) implemented these actions. Other-initiations of repair (12.5%, $n = 23$) and displays of surprise or news receipts (12.5%, $n = 23$) also occurred (see Table S2). Cross-cutting preferences were not clearly evident in these sequences. However, three questions implementing admonishments ("did you guys not read the thing they gave you:?") and self-depreciations ("my (0.5) .hhh English is is it was really getting: (.) quite ba:d, wasn't it?") arguably did involve cross-cutting preferences and were analyzed according to the preference set by the grammatical format of the question.

Measurements

The speech of each participant was individually segmented in Praat 5.3.82 (Boersma & Weenink, 2014) through inspection of the spectrogram and waveform to identify the precise beginnings and endings of all utterances. The segmentation was then imported into ELAN 4.8.1 (Wittenburg et al., 2006) for further analysis.

The timing of each response relative to the preceding question was measured based on three recognizable points in the sequence (cf. Kendrick & Torreira, 2015). Point 1: the last acoustic signal attributable to the question. Point 2: the first acoustic signal attributable to the response, including preparatory in-breaths and clicks. Point 3: the beginning of the first word or particle of the response after turn-initial in-breaths and clicks, if present. Based on these three points, two temporal offsets were calculated: Offset 1, the duration between Point 1 and Point 2, and Offset 2, the duration between Point 1 and Point 3.

The respondent's gaze direction was analyzed manually frame-by-frame in ELAN using videos recorded by the eye-trackers that include a visual indication of the participant's gaze fixation at each frame of the video. For each question-response sequence, the respondent's gaze direction was analyzed into one of four categories. (1) Gaze maintenance: the respondent maintained his or her gaze on the questioner during the question and the response. If the respondent directed his or her gaze to the questioner partway through the question and maintained this gaze direction thereafter, this was categorized as gaze maintenance. (2) Gaze aversion: the respondent averted his or her gaze from the questioner

during the question, the transition space, or the response. For questions with multiple points of possible completion (see below), the gaze aversion was analyzed with respect to the first. (3) No gaze to the questioner: the respondent did not direct his or her gaze to the questioner during the question. (4) Insufficient data: the performance of the eye-tracker was poor such that the gaze fixation indicator was not visible or was only sporadically visible. Sequences with insufficient data ($n = 6$) and without gaze to the questioner ($n = 15$) were not included in the subsequent analyses. The collection thus included 184 question-response sequences.

The timing of each gaze aversion was measured from two recognizable points in the sequence, in addition to Point 3, defined above. Point 4: the gaze departure, defined as the first frame at which the respondent's gaze fixation, as indicated by the eye-tracker video, was no longer on, overlapping with, or directly adjacent to the questioner's head or technical headgear. Point 5: the first point of possible completion of the question. All questions were analyzed for the presence and location of points of possible completion, according to previous CA research (Sacks et al., 1974; Ford & Thompson, 1996; see Holler & Kendrick, 2015, p. 5-6 for details on this analysis). Based on these points, a measurement of the timing of each gaze aversion was made: the aversion-completion distance, defined as the duration in milliseconds from Point 4 to Point 5. Descriptive statistics for all continuous variables are given in Tables S3 and S4.
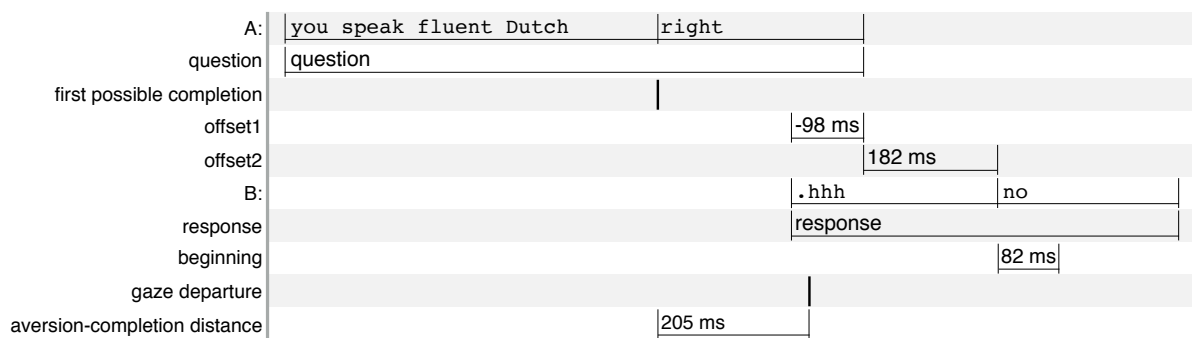

Figure 1: A visual representation of gaze and timing measurements for a single question-response sequence.

Statistics

The quantitative data were analyzed statistically using mixed effects logistic regression models (Raayen, Davidson, & Bates, 2008) generated by the *lme4* 1.1-12 package (Bates et al., 2015) in R 3.2.3 (R Core Team, 2015). To compare the effects of different predictors,

model comparisons were conducted, the details of which are reported in the supplementary materials (see Supplementary Methods).

RESULTS

Gaze direction and response preference

The analysis revealed a strong statistical relationship between gaze direction and preference. In total, 66.8% ($n = 123$) of responses were analyzed as preferred and 33.2% ($n = 61$) as dispreferred, and gaze aversion occurred in 53.8% ($n = 99$) of all responses. The proportion of responses with gaze aversion was significantly higher for dispreferreds (82.0%, $n = 50$) than for preferreds (39.8%, $n = 49$), as shown in Figure 2. Thus while most preferred responses were produced with gaze towards the questioner, most dispreferred responses were produced with gaze aversion. A mixed effects logistic regression with gaze as the dependent variable, preference as a fixed effect, and respondent and conversation as random effects showed that the relationship between gaze and preference was statistically significant ($\beta = 2.27$, $SE = 0.44$, $p < .001$).
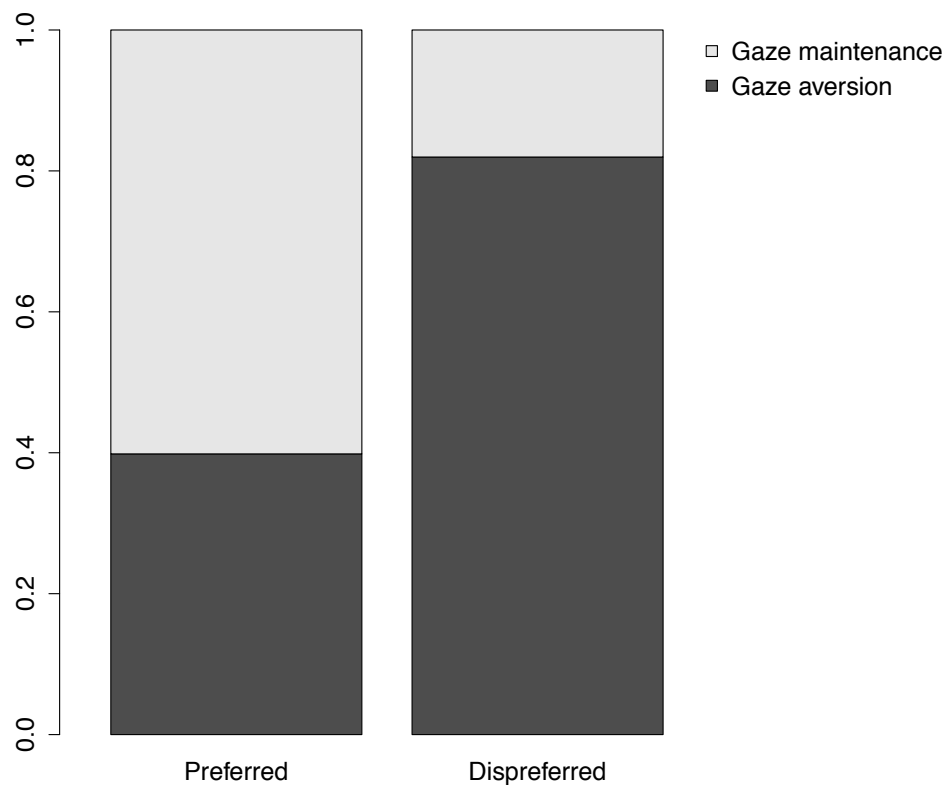

Figure 2: The proportion of gaze aversion and gaze maintenance for preferred and

dispreferred responses. Dispreferred responses are two times as likely as preferred responses to have gaze aversion.

An analysis of gaze direction at response onset specifically, defined as the first 82 ms (two video frames) of the response, yielded similar results as the analysis that included the entire response duration above. However, the proportion of preferred responses with gaze aversion decreased from 39.8% to 26.0% ($n = 32$) since those in which the aversion occurred later in the response were excluded. The proportion of dispreferred responses with gaze aversion also decreased slightly from 82.0% to 75.4% ($n = 46$; see Figure S4). Despite this, a mixed effects logistic regression with the same structure as the previous still showed a statistically significant effect ($\beta = 2.17$, $SE = 0.36$, $p < .001$). Overall, the vast majority of preferred responses begin with gaze to the questioner whereas dispreferred responses overwhelmingly begin with gaze aversion.

In the remainder of this section, we examine individual cases from the corpus in order to explore the relationship between gaze direction and preference in more detail. As with the quantitative analysis, the qualitative analyses in this section focus on one aspect of the preference organization of polar question sequences, namely the preference set by the grammatical format of the question. We begin with the most frequent types – preferred responses with direct gaze and dispreferred ones with gaze aversion – and then turn our attention to the less frequent ones.

Preferred responses with direct gaze. In face-to-face conversation questions constitute a systematic locus for mutual gaze between current and next speakers. In general, questioners use gaze direction to address a question to a particular participant and thereby allocate the next turn to the gazed-at individual or party (Sacks et al., 1974; Lerner, 2003). In the course of the question, the gazed-at participant normally returns the questioner's gaze, displaying his or her recipiency to it (Kendon, 1967; Rossano, Brown & Levinson, 2009). For preferred responses, next speakers tend to maintain this state of mutual gaze in the transition space between turns and throughout the response.

```
Extract 1    [EMIC_10t_02:36]
B:    +*was it really poorly done?
      *gaze to C------------------>
c     +gaze to B------------------>
C:    yes.
      --->>
```

```
b      --->>
```

```
Extract 2    [EMIC_05t_06:41]
A:    +*do you not have a car here yet?
      *gaze to B---------------------->
b     +gaze to A---------------------->
B:    no.
      -->>
a     -->>
```

```
Extract 3    [EMIC_06t_04:00]
A:    +d*id- did you stu-+ uh learn Dutch or
         *gaze to C----------------------->
c     +away-->          +gaze to A-------->
C:    yeah I did.=but (.) I was a lot younger.
      ------------------------------------>>
a     ------------------------------------>>
```

In each case, the current speaker uses gaze direction to addresses a question to a particular participant and thereby selects him or her as next speaker, and the next speaker returns the current speaker's gaze such that the two establish a state of mutual gaze. [2] As affirmative polar interrogatives, the questions prefer affirmative responses that match the polarity of the question, and in each case the next speaker produces a preferred response, without hesitation or delay. In Extracts 1 and 2, the responses are minimal type-conforming "yes" and "no" tokens (Raymond, 2003), and in Extract 3, the response is more complex, including both particle and repetitional responses as well as an additional turn component. In each case, the respondent maintains his or her gaze to the questioner throughout the question or midway through the question, the transition space, and the response. Extract 3 shows how speakers can use self-repair to solicit the recipient's gaze if it is not already present (Goodwin, 1980; Kidwell, 2006). In the collection, questioners gaze at respondents in 87.8% of questions ($n =$ 159). For simplicity, from here on the transcripts include only the gaze direction of the respondents.

Dispreferred responses with gaze aversion. In the majority of cases, dispreferred responses were produced with gaze aversion. As an initial specimen, consider the following sequence between Catherine and Ann, two English women who have lived in the Netherlands (where the recording was made) for many years. When Catherine first moved, she lived in a small town in a conservative region of the country. Here Ann asks about her experience there.

---

[2] Extended transcripts for all data extracts are available in the Supplementary Materials.

```
Extract 4   [EMIC_02t_06:53]

A:    +#but did you feel at home among+st tho+#se pe+:opl:#::e e-
c                                      +.............+gesture-->
c     +gaze to A-------------------------+away-->
fig   #fig.1a                                 #fig.1b     #fig.1c
A:    could you+ find a way+ of# feeling at +home amongs[t# them.
c         -->+,,,,,,,,,,,+
c                                           -->+gaze to A------------>>
fig                           #fig.1d                 #fig.1e
C:                                                    [#°not really°
```


Figure 3: Catherine shown in the top row averts her gaze from Ann before she responds. The images in the bottom row, from the eye-tracker, correspond temporally to those in the top row. The small circle indicates the direction of her gaze at each moment.

The grammatical form of Ann's question at line 1, a polar interrogative, prefers a "yes" response, which would affirm that Catherine indeed "felt at home" in that part of the country. As the question nears its first possible completion, the place at which a response from her would be due (Sacks et al., 1974), Catherine lifts her hand from her lap and begins to rub her lip as she turns and averts her gaze from Ann (see Figure 3). In reaction to this, Ann first slows her speech, stretching the word "pe:opl:::e" to over a second. This postpones the relevance of the transition between speakers and thereby orients to Catherine's actions as a sign of delay, an indication that the response will not be immediately forthcoming. A mechanism for this is the "home-away-home" organization of her movements (Sacks & Schegloff, 2002). The physical and attentional alignment between the participants at this place in the sequence constitutes a home position. To avert one's gaze at this place projects a course of action in which the gaze will return to the home position. As Catherine averts her gaze, Ann initiates self-repair in the transition space (Schegloff et al., 1977), refashioning her

question. Whereas the first version asked Catherine to affirm a blanket assessment of that period of her life, the second asks whether she eventually came to feel at home, perhaps after some difficulty. In this way, the repair orients to Catherine's actions not only as an indication of delay but also as a sign that a preferred response is not forthcoming. And indeed Catherine goes on to produce a dispreferred response, using turn-constructional practices – qualification and sotto voce prosody – commonly associated with dispreference (Kendrick & Torreira, 2015). In this case, the gaze aversion precedes the dispreferred response, but as we will see, the two actions can also occur together.

That gaze aversion can occasion self-repair in the transition space indicates not only that current speakers monitor the gaze behavior of next speakers but also that they understand gaze aversion as a sign of trouble. This provides endogenous evidence that the relationship between gaze aversion and dispreference is not merely a statistical one, but one that participants themselves orient to and use in the interaction. As we have seen, however, gaze aversion can co-occur with other signs of trouble, such as an expansion of the transition space and the production of visible actions by the next speaker within it. In Extract 2, an audible in-breath precedes gaze aversion by approximately 170 ms and is the leading indicator that the response is likely to be dispreferred. A transition space self-repair (">or PlayStation.<") thus appears to have been occasioned by a complex signal with both auditory and visual components.

```
Extract 5    [EMIC_05t_19:25]
A:    +#do you have an Xbo[x he+r#e.]
b     +gaze to A---------------+away-->
fig   #fig.4a                  #fig.4b
B:                        [. h h h  ] h h #[ h h  [no, back +home.
                                              -->+gaze to A->>
fig                                      #fig.4c
A:                                       [>or Pl[ayStation.<
```

The self-repair orients to a basis on which Bijan would be unable to provide a preferred response, namely that he may not own the video game console named in question. The self-repair thus indicates that Aston has understood Bijan's actions – taking a breath, averting his gaze – as signs of a dispreferred response.

Figure 4: Bijan shown in the top row averts his gaze in the transition space before Aston initiates self-repair.

The co-occurrence of gaze aversion with other indicators of dispreference raises a critical question: does gaze aversion alone project that the response will be dispreferred? In the collection, explicit orientations to incipient responses as dispreferred through transition space self-repairs are rare ($n = 6$), and in all such cases gaze aversion co-occurs with other indicators, as the preceding cases demonstrate (see Extract S1 for an additional example). Thus, using CA methods, one cannot answer this question with the available data. However, using quantitative methods, one can tease apart these indicators by comparing statistical models that include different combinations of predictors to determine whether alternative or additional predictors improve the predictive power of the model. In order to determine which indicator best predicts whether the response will be preferred or dispreferred, we compared a series of mixed-effects logistic regression models using the likelihood ratio test. We began with a full model with gaze direction, in-breath, and gap duration (Offset 2) and then removed predictors one by one, comparing at each step (see Supplementary Materials for details). The results showed that statistical models that include in-breath and/or gap duration are not significantly better than a simpler model with gaze direction alone. While we do not know whether such models correspond to the methods that participants use to understand

each other's conduct, the statistical evidence suggests that a participant who wishes to predict the preference status of an incipient response could do so based on gaze direction alone.

Although transition space self-repairs provide crucial evidence that participants orient to gaze aversion as indicative of a dispreferred response, they are rare. More typically gaze aversion precedes or co-occurs with a dispreferred response without any orientation by the questioner, as in the following cases.

```
Extract 6    [EMIC_02t_12:14]
B:    +all your family's in England+ I expec[t.
c     +gaze to B------------------+away-->
C:                                           [no Aust+ralia(h)
                                                -->+gaze to B-->>


Extract 7    [EMIC_05t_06:41]
A:    +you gonna get one?
b     +gaze to A-------->
B:    +.hh I don't know it's like s+o: expensive.
      ->+away-->              -->+gaze to A-->>
```

In Extract 6, the next speaker averts her gaze just as the question comes to a possible completion and then returns it as she responds. That she breaks mutual gaze when she produces the dispreferred type-conforming token "no" is typical. Most dispreferred type-conforming tokens occur with gaze aversion (68.2%, $n = 15$), in contrast to only 17.5% ($n = 17$) of preferred ones. And in Extract 7, gaze aversion co-occurs with other turn-constructional features – a turn-initial in-breath, a non-answer response, and an account – commonly found in dispreferred responses.

Preferred responses with gaze aversion. In a minority of cases, however, preferred responses are produced with gaze aversion. An examination of such cases revealed two recurrent types: (i) delayed preferred responses and (ii) complex preferred responses. In the first type, gaze aversion co-occurs with an expansion of the transition space before the response.

```
Extract 8    [EMIC_03t_12:42]
A:    +how long did it go on for.=years?
c     +gaze to A---------------------->
      (0.3)+(0.2)
c     -----+away-->>
C:    yea:h.
```

```
Extract 9   [EMIC_06t_12:25]
B:    +did you- +did you take it as w[ell?
a     +away-->  +gaze to B-------------->
A:                                      [mmm:+::::: yeah.+
a                                           +away-->    +B-->>
```

In Extract 8, the respondent gazes to the questioner for approximately 0.3 seconds in the transition space before she averts her gaze and produces a minimal preferred response. But, as Extract 9 demonstrates, gaze aversion can also occur together with other forms of delay, such as the non-lexical particle *mmm* which delays the response by over 0.8 seconds. Cases such as these suggest a general relationship between gaze aversion and delay, such that averting one's gaze in the absence of a verbal response may signal that the response will not be immediately forthcoming.

In the second type, gaze aversion co-occurs not with an expansion of the transition space before the response, but with an expansion of the response itself. Responses to polar questions can be minimal, containing only a type-conforming "yes" or "no" token (Raymond, 2003), or they can be more complex, repeating components from the question (Heritage & Raymond, 2012) or including additional components after a type-conforming token (Raymond, 2013). In the following case, gaze aversion occurs at the beginning of such a complex preferred response. Prior to this extract, Aston and Christopher have asked Bijan about his background, and Bijan, who is Persian but grew up in England, has reported that he knows Farsi. Here Christopher checks whether Bijan is fluent in the language.

```
Extract 10  [EMIC_05t_12:40]
C:    +and you're fluent in that.
b     +gaze to C-------------->
B:    +yea:h.=just speaki+ng though?=it's +like .hhh reading and
      +away-->          -->+gaze to C-------+away-->>
      writing I've never been taught? ((turn continues))
```

Just as Bijan begins to respond, he averts his gaze from Christopher. The response he produces first confirms that he is fluent and then qualifies this through an additional component ("just speaking though"). He then expands the response, producing a multi-unit turn that details his knowledge of the language. As Raymond (2013) observes, such responses must overcome a systematic obstacle to their production: "because the type-conforming token with which the turn begins may deliver a possibly complete response, speakers must project more talk in order to create the space to produce it" (p. 196). Raymond describes how

speakers can manipulate the prosody of type-conforming tokens (e.g., using rising intonation) to project more talk. Extract 10 suggests that gaze direction is another resource that speakers use to achieve this outcome: gaze aversion at the beginning of a response projects that the response will be more than minimal and thus allows the recipient of the response to see that the turn will not be complete after the type-conforming token.

   Dispreferred responses with direct gaze. In the minority of cases, however, dispreferred responses occur without gaze aversion. While generalizations are difficult to make with only a small number of cases ($n = 11$), an intriguing possibility concerns the proportion of dispreferred responses with an explicit "no" rather than more mitigated forms (e.g., "not really"). For dispreferred responses with gaze aversion, only 28.8% ($n = 15$) contain an explicit "no" component, but for dispreferred responses with direct gaze, 63.6% ($n = 7$) do. Thus the following cases are representative of the statistical trend.

```
Extract 11  [EMIC_02t_18:16]
A:    +but do you speak English +with your daughters?
c     +away-->                  +gaze to A---------->
C:    no.
      -->
```

```
Extract 12  [EMIC_05t_03:01]
A:    +I thought (.) you:r (0.2) studies (.) is in (0.2) Dutch.
b     +gaze to A------------------------------------------->
B:    no it's in English.
      ----------------->>
```

In each case, the response includes none of the turn-constructional features that frequently occur in dispreferred responses: there are no delays, no audible in-breaths, no qualifications, and no accounts (though Extract 12 does include a correction, replacing "Dutch" with "English"). The responses thus exhibit features of preferred turn formats even though they perform dispreferred actions (see Kendrick & Torreira, 2015). While it would be premature to draw a firm conclusion, the data suggest that gaze direction correlates with other turn-constructional features such that direct gaze belongs to a preferred turn format and averted gaze to a dispreferred one, a point to which we return in the discussion.

The timing of gaze aversion

In addition to a categorical relationship between gaze direction and response preference, we also hypothesized that the timing of gaze aversions should differ systematically by preference, with an earlier gaze aversion onset for dispreferred responses. Such an early indication that a dispreferred response is forthcoming would be interactionally advantageous. It would allow the current speaker to self-repair his or her turn, reversing the preference of the response and thereby minimizing the opportunity for disaffiliation, even before the next turn becomes relevant, that is, before the transition space, not within it. To investigate this possibility, we examined the *aversion-completion distance*, the duration from the respondent's first observable eye movement away from the questioner to the first possible completion of the question (see Methods).
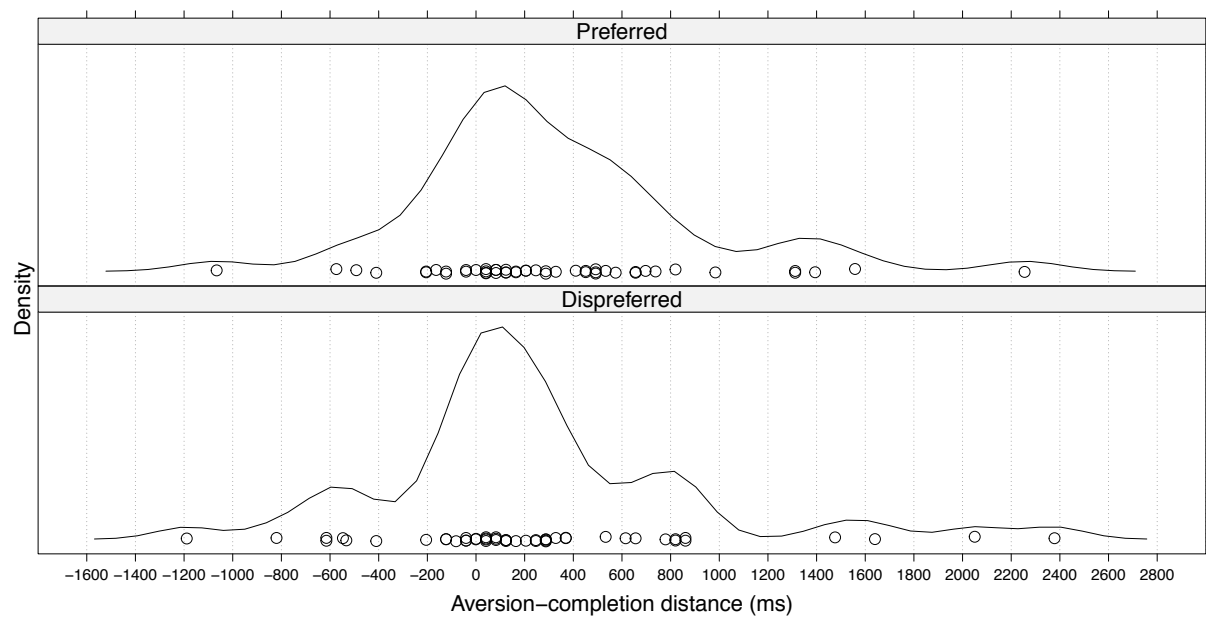

Figure 5: The duration of the aversion-completion distance in milliseconds for preferred and dispreferred responses. Zero represents the first possible completion of the question; negative values indicate gaze aversions that began before this point; and positive values indicate those that occurred after it. One dispreferred response with value of 5371 ms is not shown.

The density plots in Figure 5 show the distributions of aversion-completion distances for preferred and dispreferred responses. An initial observation is that the two distributions have similar shapes and overlap substantially. The modal aversion-completion distances, represented by the peaks of the curves, are the same for preferred and dispreferred responses, approximately 100 ms. No statistically significant difference was found in a mixed effects

model. This indicates that respondents most frequently begin to avert their gaze 100 ms after the first possible completion of the question, regardless of the preference status of the response. Experimental research has shown that observable eye movements require on average 200 ms to plan and execute (see, e.g., Fischer & Ramsperger, 1984). This suggests that the initiation of an observable gaze aversion would have occurred approximately 100 ms before the first possible completion of the question. To put this into perspective, the average duration of a syllable in English conversation is 200 ms (Greenberg, 1999), and turn-final syllables are known to be lengthened (see, e.g., Local, Kelly, & Wells, 1986). The initiation of gaze aversion would therefore have occurred within the final syllable preceding the first possible completion of the question.

Gaze direction and response complexity

The statistical relationship between gaze direction and preference suggests that gaze functions as a signal of preference: maintaining one's gaze as one begins to respond indicates that the response will likely be preferred whereas averting one's gaze indicates a greater likelihood that it will be dispreferred. Moreover, self-repairs by questioners in the transition space between turns ground this relationship in the reality of the participants (Schegloff, 1996), revealing it to be more than a mere statistical association between variables.

What, then, should one conclude about the other functions of gaze direction that Kendon (1967) and others have identified? If gaze direction serves a regulatory function, gaze aversion at the beginning of a response could be a powerful resource for projection, signaling that a multi-unit turn is underway, as we have suggested (see Extract 10). If it reduces cognitive load, as respondents momentarily withdraw from the "eye-to-eye ecological huddle" to plan a response, then this, too, could indicate that more than a simple "yes" is in the works. And if gaze aversion signals response preference, as we maintain, might not the operative mechanism be the projection of a relatively simple or a relatively complex response? In the collection, the mean response duration for preferred responses is 1252 ms whereas that for dispreferreds is 1702 ms, evidence that the two are not equally complex. The functions of gaze direction in conversation resist isolation, but using statistical methods the effects of different functions can be compared.

In order to determine whether the preference status of a response influences gaze direction above and beyond the complexity of the response, we compared a series of mixed

effects logistic regression models with gaze direction as the dependent variable, preference and response duration as a continuous variable, each alone and both together, as fixed effects, and respondent and conversation as random effects (see Supplementary Materials for details). The results show that a model with both preference and response duration as predictors performs better than models with just one or the other. The final model revealed statistically significant effects for both preference and response duration (preference: $\beta = 2.16$, $SE = 0.47$, $p < .001$; response duration: $\beta = 2.06$, $SE = 0.40$, $p < .001$). This indicates that the complexity of the response, as measured by its duration, does indeed influence gaze direction, but importantly, also that the preference of the response has a significant effect independent of response duration.

The complex relationship between gaze direction, preference, and response complexity can also be seen in Figure 6, which presents an different measure of response complexity than the continuous one used in the mixed effects models, namely whether the response included a single TCU (e.g., "yes" in Extract 1) or multiple TCUs (e.g., "yeah I did.=but (.) I was a lot younger" in Extract 3).[3] Figure 6 shows the proportion of gaze aversion and gaze maintenance for preferred and dispreferred responses split between single and multiple TCU responses. The figure clearly shows the effects of both preference and complexity. First, for both single and multiple TCU responses, the proportion of responses with gaze aversion is higher for dispreferred responses than for preferreds. Second, for both preferred and dispreferred responses, the proportion of gaze aversion is higher for multiple TCU responses than single ones. We therefore conclude that both response preference and response complexity influence gaze direction, and thus that the relationship between gaze direction and preference that we observe, both in our statistical models and in the participants' orientations, does not depend on or merely derive from the relationship between gaze direction and complexity.

---

[3] For the purpose of this figure visual responses without verbal components (e.g., affirmative nods) were analyzed as single TCUs though the status of such actions as TCUs is equivocal. For a similar graph with responses split at the median response duration into long and short responses see Figure S5.
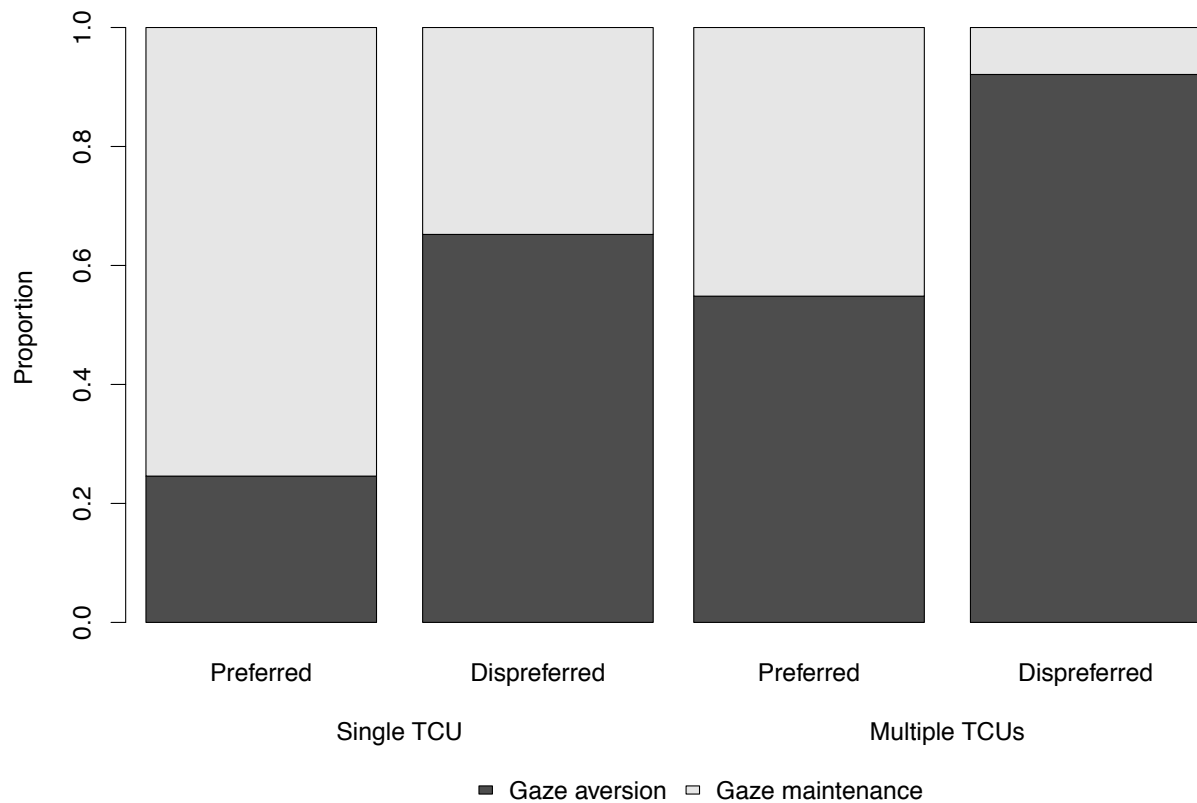
Figure 6: The proportion of gaze aversion and gaze maintenance for preferred and dispreferred responses, split into single and multiple TCU responses.

DISCUSSION

We have proposed one specification of the expressive function of gaze direction identified by Kendon (1967), namely that gaze direction serves as a resource for the construction of affiliative and disaffiliative actions in conversation. More specifically, we hypothesized that gaze aversion should be more frequent in dispreferred responses than preferred ones. A quantitative analysis of gaze direction in responses to polar questions confirmed this hypothesis. The vast majority of dispreferred responses are produced with gaze aversion whereas only a minority of preferred ones are. An analysis that examined only the very beginnings of responses yielded a similar result, indicating that the first visible moments of a response furnish crucial resources for early action recognition, and may thereby facilitate rapid turn-taking (see Schegloff, 1987; Levinson, 2013). In addition to the quantitative analyses, an analysis of self-repairs by questioners in the transition space between turns showed that participants also orient to gaze aversion – at the moment of its production – as a signal of a dispreferred response. Thus the statistical association between gaze direction and

response preference appears to reflect the methods that participants use not only to produce preferred and dispreferred responses but also to recognize them.

We also hypothesized that the timing of gaze aversion would differ systematically by response preference. In particular, we predicted that gaze aversion should begin earlier for dispreferred responses than preferred ones. This hypothesis was not confirmed. Rather than reflect the preference of the incipient response, the timing of gaze aversions appears to reflect a basic principle of the turn-taking system, namely that the first possible completion of a turn at talk constitutes a place where the transition between speakers becomes relevant (Sacks et al., 1974). In general, respondents maintained gaze with questioners until a transition in speakership became relevant, at which point they averted their gaze, about half of the time. This resonates with Holler and Kendrick's (2015) observation that unaddressed participants maintain gaze to questioners until the first possible completion of the question where they then frequently shift their gaze to addressed participants, optimizing recipiency between the two. In the present study, the fact that respondents look away indicates a greater likelihood that the incipient response will be dispreferred, but *when* they do so does not.

Our results provide strong evidence that gaze direction serves as a signal of response preference in conversation. However, we do not argue that gaze direction has no other functions, even in responses to polar questions. Indeed, a statistical model that included both preference and complexity, as measured by response duration, showed that both factors significantly influence gaze direction. But crucially the comparison also showed that preference has an independent effect above and beyond complexity. Even for relatively short and simple responses, gaze aversion is far more frequent for dispreferred responses than preferred ones. Evidence for the regulatory function of gaze direction was also observed in preferred responses with gaze aversion. In complex preferred responses, for example, gaze aversion at the possible completion of the initial TCU (e.g., "yea:h" in Extract 10) visibly displays that the response is still "in progress" and thereby both interdicts the relevance of transition between speakers and projects an additional unit. This points to a provocative hypothesis, namely that body-behavioral indications of turn continuation, such as gaze aversion, may supersede linguistic indications of turn completion.

How does gaze direction work as a signal of preference?

Thus far we have argued that in responses to polar questions gaze direction functions as a signal of preference, but we have yet to describe the mechanism whereby gaze aversion comes to be recognizable, firstly, as a meaningful action and, secondly, as a sign of dispreference. Two accounts naturally present themselves, one more speculative than the other.

The first account concerns the normative organization of gaze direction in question-response sequences. As we have seen, question-response sequences are a systematic locus for mutual gaze between current and next speakers (see also Rossano et al., 2009). In the production of a question, gaze direction to a particular participant serves to address the question to the gazed-at participant and thereby to select him or her as the next speaker (Sacks et al., 1974; Lerner, 2003). The addressed participant normally reciprocates, establishing a state of mutual gaze, and in so doing tacitly ratifies his or her selection as next speaker. As Goodwin (1980) and Kidwell (2006) have demonstrated, current speakers have practices, such as forms of self-repair, to elicit the gaze of addressed participants if it is directed elsewhere. The use of such practices orients to, and thus provides evidence of, a normative expectation of mutual gaze (see, e.g., Extracts 3 and 9).

Gaze aversion in the transition space or at the beginning of a response, we propose, constitutes a meaningful departure from this norm. The quantitative evidence indicates that in most cases respondents maintain gaze to questioners up to the beginning of the response, if not longer, and that when they avert their gaze, the response is more likely to be dispreferred. Furthermore, the observation that gaze aversion also occurs in a subset of preferred responses, specifically those which are delayed or contain multiple TCUs (see Extracts 8 to 10), suggests that respondents only avert their gaze "for cause" (see Raymond, 2003, p. 950). That is, gaze aversion occurs in circumstances in which the response departs from normative expectations, such as the expectation that a response should be produced on time, should contain only one TCU (Sacks et al., 1974), or should implement a preferred action. The mechanism whereby gaze aversion comes to signal dispreference thus includes (i) a normative expectation of mutual gaze in the transition space and at the beginning of the response; (ii) a visible departure from this; (iii) a search for a possible "cause" for this departure; and (iv) the selection of dispreference (i.e., the anticipation of a non-preferred response alternative) as the possible cause. A more complete specification would also

describe the methods that participants use to select among alternative interpretations. To give but one example, if the gaze aversion coincides temporally with the prompt production of a preferred type-conforming token (e.g., Extract 10), then it can serve neither as a signal of delay nor as a signal of dispreference, but it can still signal that the response will contain multiple TCUs. Although the data show that such alternative interpretations are indeed possible, the robust statistical correlation between gaze direction and response preference, together with orientations by participants to the specific relationship between gaze aversion and dispreference, suggest that the methods of selection, whatever their composition, favor dispreference over other interpretations.

The second account, which is the more speculative of the two, concerns the possibility that the combination of direct gaze with a disaffiliative action, such as a dispreferred response, could amplify the disaffiliative import of the action and is therefore avoided. An analogy to the linguistic construction of preferred and dispreferred responses is instructive here. The construction of dispreferred responses normally includes features – such as prefaces, qualifications, and accounts – that work to mitigate such uncooperative and disaffiliative actions (Clayman, 2002). In contrast, the construction of preferred responses, which embody a cooperative and affiliative stance towards the sequence, tend to lack such features, taking relatively simple forms. However, the normal associations between preferred and dispreferred actions and turn formats can be inverted, such as when a speaker rejects an invitation without hesitation or qualification (e.g., A: ".hh Do you want to go:" B: "Oh no:"; Kendrick & Torreira, 2015, p. 23). Dispreferred actions can be produced in preferred turn formats (or vice versa), that is, without an orientation to the action as uncooperative or disaffiliative. Rather than *mitigate* the threat to the other's face (Brown & Levinson, 1987) or to social solidarity (Clayman, 2002) that such actions pose, dispreferred responses in preferred turn formats *enhance* this effect. In vernacular terms, they are blunt, even rude. To return to gaze direction, given that gaze maintenance most frequently occurs in preferred actions, maintaining eye-contact as one produces a dispreferred response may have a similar effect, amplifying the disaffiliative import of the action. The tendency for dispreferred responses with direct gaze to include explicit "no" tokens (e.g., Extracts 12 and 13), rather than more mitigated forms (e.g., "not really"), is consistent with this possibility. That said, such cases were rare overall and the analysis focused on the preference set by the grammatical format of the question, not on other aspects of response design, such as the

congruity or incongruity of epistemic and evaluative stances (see Bolden, in press), which may also prove relevant.

Do the results generalize to naturally occurring conversation?

In contrast to laboratory and experimental studies of human behavior, conversation analytic studies generally enjoy an exemption from questions about ecological validity. Because CA uses naturalistic observation, few researchers doubt that the results generalize to everyday life. The present study is not exempt from such questions, however. To determine whether our results generalize to naturally occurring conversation, we examined polar question sequences across six video recordings made under more natural conditions. In most video recordings, the participants' eyes were not clearly visible or were only sporadically so, making a systematic analysis of gaze direction impossible. That said, one factor did appear to influence the occurrence of gaze aversion: the F-formation that the participants adopt (Kendon, 1990). For gaze aversion to occur, the respondent must first have his or her gaze directed to the questioner. While this occurred in almost all sequences in the present study (92.5% of the 199 cases with sufficient data, see Methods), in the more naturalistic video recordings, the gaze behavior of the participants varied across different F-formations. Mutual gaze in question-response sequences appeared to be more common in face-to-face configurations than in side-by-side or L-shaped ones. In a side-by-side configuration, for example, the respondent may gaze forward (e.g., at an object) while he or she answers the question. The F-formation used in the corpus may therefore have promoted mutual gaze and hence the possibility of gaze aversion.

This does not mean, however, that the results of the present study do not generalize beyond the lab. An analysis of one dyadic conversation in which the participants sat across from each other in a face-to-face configuration (i.e., RCE15a) suggests that the results do generalize. In the approximately 15 minute conversation, 21 polar questions received responses, with 16 (76.2%) preferred and five (23.8%) dispreferred. All dispreferred responses (100%) had gaze aversion whereas only two preferred responses (12.5%) did. The examination of gaze direction and response preference in naturally occurring conversation thus suggests that the results generalize most clearly to those occasions in which participants arrange themselves into F-formations that promote mutual gaze such as face-to-face and circular configurations, or in Goffman's (1964) terms, an "eye-to-eye ecological huddle".

CONCLUSION

A recent thread of research on gaze behavior by Rossano (2012, 2013; Rossano et al., 2009) has led to the conclusion that gaze direction is primarily organized in relation to a single order of interaction, namely the organization of sequences (Schegloff, 2007). An important discovery to emerge from this work is that at the possible completion of a sequence (e.g., after a response to a question) gaze aversion treats the sequence as complete whereas gaze maintenance orients to the relevance of expansion (Rossano, 2012). However, to conclude on the basis of this and other findings that gaze direction "is mainly organized in relation to sequences of talk" (Rossano et al., 2009, p. 191) is perhaps premature. As we have shown, gaze direction in responses to polar questions – that is, *within* sequences – is organized in relation to preference. Rather than argue that gaze direction has a primary organizational locus, we propose that gaze direction is a positionally-sensitive signal the function of which depends on the specific context in which it occurs and the organizationally-relevant order(s) of interaction at that position. At the possible completion of a sequence, where the relevant organizational problem for participants is whether to close or expand the sequence, gaze direction operates at the level of sequence organization. However, in the transition space between questions and responses, where a relevant organizational problem for participants is whether the response will conform to or depart from normative expectations, the gaze direction of respondents operates at the level of preference organization. As formal signals, gaze maintenance and gaze aversion have no primary organizational locus, but rather become recognizable as meaningful actions in relation to different interactional organizations at different positions in the interaction.

CONVENTIONS FOR MULTIMODAL TRANSCRIPTION

Embodied actions are transcribed according to the following conventions developed by Mondada (2014).

* * + +	Gestures and descriptions of embodied actions are delimited between ++ two identical symbols (one symbol per participant) and are synchronized with correspondent stretches of talk.
---> --->	The action described continues across subsequent lines until the same symbol is reached.

>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
.....	Action's preparation.
,,,,,	Action's retraction.
ali	Participant doing the embodied action is identified when (s)he is not the speaker.
fig #	The exact moment at which a screen shot has been taken is indicated with a specific sign showing its position within turn at talk.

REFERENCES

Argyle, M. and Cook, M. (1976). *Gaze and mutual gaze*. London: Cambridge University Press.

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication, 52,* 566-580.

Beattie, G. W. (1978). Sequential Temporal Patterns of Speech and Gaze in Dialogue. *Semiotica*, *23*(1-2), 29–52.

Beattie, G. W. (1981). A further investigation of the cognitive interference hypothesis of gaze patterns during conversation. *British Journal of Social Psychology*, *20*, 243-248.

Boersma, Paul & Weenink, David (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.82, retrieved 26 July 2014 from http://www.praat.org/

Bolden, G. B. (in press). A simple da?: Affirming responses to polar questions in Russian conversation. *Journal of Pragmatics*. http://doi.org/10.1016/j.pragma.2015.07.010

Boyd, E., & Heritage, J. (2006). Taking the history: Questioning during comprehensive history taking. In J. Heritage & D. Maynard (Eds.), *Communication in medical care: Interaction between physicians and patients* (pp. 151–184). Cambridge, England: Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Clayman, S. E. (2002). Sequence and solidarity. In S. R. Thye & E. J. Lawler (Eds.), *Group Cohesion, Trust and Solidarity* (Vol. 19, pp. 229–253). Oxford: Elsevier Science Ltd.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Doherty-Sneddon, G., Bonner, L., & Bruce, V. (2001). Cognitive demands of face monitoring: Evidence for visuospatial overload. *Memory & Cognition*, *29*, 909-919.

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*, 283-292.

Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory*. Lawrence Erlbaum Associates.

Duncan, S., Brunner, L. J., & Fiske, D. W. (1979). Strategy signals in face-to-face interaction. *Journal of Personality and Social Psychology*, *37*, 301-313.

Edmunds, M. (1974). *Defence in animals: a survey of anti-predator defences*. Longman Publishing Group.

Eibl-Eibesfeldt, I. (1989). *Human Ethology*. New York: Aldine Transaction.

Ekman, P., & Chevalier-Skolnikoff, S. (1973). Facial expression of emotion in nonhuman primates. In *Darwin and Facial Expression: A Century of Research in Review* (pp. 11–90). Cambridge, MA: Malor Books.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*, 9602-9605.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, *57*, 191-195.

Ford, C. E., & Thompson, S. A. (1996). Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 134–184). Cambridge: Cambridge University Press.

Goffman, E. (1964). *Behavior in public places*. New York: The Free Press.

Gomez, J. C. (1996). Ostensive behavior in great apes: The role of eye contact. In A. E. Russon, K. A. Bard, and S.T. Parker (Eds.), *Reaching into thought: The minds of the great apes*, pp. 131-151. Cambridge University Press.

Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, *50*(3- 4), 272–302.

Goodwin, M. H. (1980). Processes of Mutual Monitoring Implicated in the Production of Description Sequences. *Sociological Inquiry*, *50*(3-4), 303–317.

Greenberg, S. (1999). Speaking in shorthand – A syllable-centric perspective for understanding pronunciation variation. *Speech Communication*, *29*(2–4), 159–176.

Haddington, P. (2006). The organization of gaze and assessments as resources for stance taking. *Text & Talk*, *26*(3), 281–328.

Hayano, K. (2013). Question design in conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 395–414). Malden: Blackwell Publishing Ltd.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Heritage, J. (2010). Questioning in medicine. In A. F. Freed & S. Ehrlich (Eds.), *'Why do you ask?':
The function of questions in institutional discourse* (pp. 42–68). Oxford: Oxford University
Press.

Heritage, J. (2015). Well-prefaced turns in English conversation: A conversation analytic perspective.
Journal of Pragmatics, *88*, 88–104.

Heritage, J., & Clayman, S. E. (2010). *Talk in Action: Interactions, Identities, and Institutions* (1
edition). Wiley-Blackwell.

Heritage, J., & Raymond, G. (2012). Navigating Epistemic Landscapes: Acquiescence, Agency and
Resistance in "Repetitive" Responses to Polar Questions. In J. P. de Ruiter (Ed.), *Questions*
(pp. 179–192). Cambridge: Cambridge University Press.

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.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta psychologica*, *26*, 22-
63.

Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*.
Cambridge: Cambridge University Press.

Kendrick, K. H., & Torreira, F. (2015). The Timing and Construction of Preference: A Quantitative
Study. *Discourse Processes*, *52*(4), 255–289.

Kidwell, M. (2006). "Calm down!": the role of gaze in the interactional management of hysteria by
the police. *Discourse Studies*, *8*(6), 745–770.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *387*, 767-768.

Lerner, G. H. (2003). Selecting next speaker: The context-sensitive operation of a context-free
organization. *Language in Society*, *32*(02), 177–201.

Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The
Handbook of Conversation Analysis* (pp. 101–130). Malden: Blackwell Publishing Ltd.

Local, J., Kelly, J., & Wells, W. H. (1986). Towards a phonology of conversation: turn-taking in
Tyneside English. *Journal of Linguistics*, *22*(2), 411–437.

Markson, L., & Paterson, K. B. (2009). Effects of gaze aversion on visual spatial imagination. British
Journal of Psychology, 100, 553-563.

Mondada, L. (2014). Conventions for multimodal transcription. Available from
https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.
pdf.

Pomerantz, A., & Heritage, J. (2013). Preference. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 210–228). Malden: Blackwell Publishing Ltd.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raymond, G. (2003). Grammar and Social Organization: Yes/no Interrogatives and the Structure of Responding. *American Sociological Review*, *68*(6), 939–967.

Raymond, G. (2010). Grammar and Social Relations: Alternative Forms of Yes/No Type Initiating Actions in Health Visitor Interactions. In A. F. Freed & S. Ehrlich (Eds.), *"Why Do You Ask?": The Function of Questions in Institutional Discourse* (pp. 87–107). New York: Oxford University Press.

Raymond, G. (2013). At the intersection of turn and sequence organization: On the relevance of "slots" in type-conforming responses to polar interrogatives. In B. Szczepek Reed & G. Raymond (Eds.), *Units of talk - units of action* (pp. 169–206). Amsterdam: John Benjamins Publishing Co.

Rossano, F. (2012). *Gaze Behavior in Face-to-Face Interaction*. Radboud University, Nijmegen.

Rossano, F. (2013). Gaze in Conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 308–329). Malden: Blackwell Publishing Ltd.

Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning, and culture. In J. Sidnell (Ed.), *Conversation Analysis: Comparative Perspectives* (pp. 197–249). Cambridge: Cambridge University Press.

Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button, & J. R. E. Lee (Eds.), *Talk and Social Organization* (pp. 54–69). Clevedon, UK: Multilingual Matters.

Sacks, H., & Schegloff, E. A. (2002). Home position. *Gesture*, *2*(2), 133–146.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735.

Schegloff, E. A. (1987). Recycled Turn Beginnings: A Precise Repair Mechanism in Conversation's Turn-taking Oragnisation. In *Talk and Social Organisation* (pp. 70–93). Phiadelphia: Multilingual Matters.

Schegloff, E. A. (1996). Confirming Allusions: Toward an Empirical Account of Action. *American Journal of Sociology*, *102*(1), 161–216.

Schegloff, E. A. (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, *53*(2), 361–382.

Stivers, T. (2010). An overview of the question-response system in American English conversation. *Journal of Pragmatics*, *42*(10), 2772–2781.

Stivers, T., & Enfield, N. J. (2010). A coding scheme for question–response sequences in conversation. *Journal of Pragmatics*, *42*(10), 2620–2626.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H. (2006). ELAN: A Professional Framework for Multimodality Research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.