Routledge
Taylor & Francis Group

Research on Language and Social Interaction

ISSN: 0835-1813 (Print) 1532-7973 (Online) Journal homepage: http://www.tandfonline.com/loi/hrls20

Eye Blinking as Addressee Feedback in Face-To-Face Conversation

Paul Hömke, Judith Holler & Stephen C. Levinson

To cite this article: Paul Hömke, Judith Holler & Stephen C. Levinson (2017) Eye Blinking as Addressee Feedback in Face-To-Face Conversation, Research on Language and Social Interaction, 50:1, 54-70, DOI: 10.1080/08351813.2017.1262143

To link to this article: http://dx.doi.org/10.1080/08351813.2017.1262143

RESEARCH ON LANGUAGE AND SOCIAL INTERACTION
2017, VOL. 50, NO. 1, 54–70
http://dx.doi.org/10.1080/08351813.2017.1262143

Routledge
Taylor & Francis Group

Eye Blinking as Addressee Feedback in Face-To-Face Conversation

Paul Hömke[a], Judith Holler[a,b], and Stephen C. Levinson[a,c]

[a]Language and Cognition Department, Max Planck Institute for Psycholinguistics, The Netherlands; [b]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands; [c]Linguistics Department, Radboud University Nijmegen, The Netherlands

ABSTRACT
Does blinking function as a type of feedback in conversation? To address this question, we built a corpus of Dutch conversations, identified short and long addressee blinks during extended turns, and measured their occurrence relative to the end of turn constructional units (TCUs), the location where feedback typically occurs. Addressee blinks were indeed timed to the end of TCUs. Also, long blinks were more likely than short blinks to occur during mutual gaze, with nods or continuers, and their occurrence was restricted to sequential contexts in which signaling understanding was particularly relevant, suggesting a special signaling capacity of long blinks.

Language is primarily used in face-to-face conversation (Clark, 1996). The role of the addressee in conversation has been viewed from two main perspectives in the language sciences. In what Clark and Krych (2004) called "unilateral" views on conversation—widely adopted within linguistics and psycholinguistics—speaking and listening are individual processes. Speakers determine the course of their utterances on their own, and addressees try to understand those utterances on their own. In this view, the addressee is a passive receiver. In "bilateral" views on conversation—widely adopted by conversation analysts and some psychologists—speaking and listening is considered a joint activity in which speaker and addressee coordinate moment by moment to maintain mutual understanding (Brennan, Galati, & Kuhlen, 2010; Clark, 1996; Goodwin, 1981; Sacks, Schegloff, & Jefferson, 1974). According to this view, the addressee is an active collaborator. Observational as well as experimental evidence supports the bilateral account of conversation: While speakers are speaking, addressees provide vocal feedback like *mm-hm* and visual feedback like nods and smiles, which in turn affect the speakers' speaking (Bavelas, Coates, & Johnson, 2000; Brunner, 1979; Clark & Krych, 2004; Malisz et al., 2016; Schegloff, 1982; Stivers, 2008; Yngve, 1970).

In conversation analysis, a fine-grained classification of different types of addressee feedback has been established based on the specific functions they fulfill in specific sequential positions in conversation (Gardner, 2001). Addressees typically provide feedback at the end of speakers' turn constructional units (TCUs; Sacks et al., 1974). These are units with recognizable possible completions at which next speakers can appropriately start a turn or initiate repair. To pass up the opportunity to take a turn or initiate repair at the end of a TCU, small behavioral tokens (e.g., *mhm*) are produced by the addressee that treat the turn as still in progress, conveying "I'm with you, please continue," allowing the prior speaker to produce an extended turn consisting of multiple TCUs (Schegloff, 1982). Behaviors that serve this function in this position have been termed

CONTACT Paul Hömke paul.hoemke@mpi.nl Language and Cognition Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, Nijmegen, XD 6525 The Netherlands.

We thank Binyam Gebrekidan Gebre, Mart Lubbers, Emma Valtersson, and Sean Roberts for programming support, and Herb Clark, Lorenza Mondada, the Language & Cognition Department, and members of the Dialogue Project at the MPI Nijmegen for valuable discussions.

Color versions of one or more of the figures in this article can be found online at www.tandfonline.com/HRLS.

"continuers" (Schegloff, 1982), and they have been contrasted with "epistemic tokens" (e.g., *oh*; Heritage, 1984), "activity-shift tokens" (e.g., *all right* or *okay*; Beach, 1993), and "assessments" (e.g., *oh wow*; Goodwin, 1986; see also "specific responses," Bavelas et al., 2000). While continuers have been primarily described in the vocal modality (e.g., *mhm, uh-huh*; Gardner, 2001; Schegloff, 1982), it has been shown that visual conduct—nods in particular (Heath, 1992; McClave, 2000; see also Stivers, 2008; Whitehead, 2011)—can also serve a continuer function if produced in the same sequential position (e.g., Mondada, 2011).

The goal of this study was to investigate blinking as one potential additional type of visual addressee feedback and more specifically, as visual conduct potentially serving a continuer function. Humans hardly blink at birth (Ponder & Kennedy, 1927), but blink rate increases until adulthood (Zametkin, Stevens, & Pittman, 1979). Adults blink more often than physiologically necessary for wetting the eyes (Doane, 1980), showing a blink rate of approximately 15 to 20 blinks per minute, with a mean blink duration of 300–400 ms (e.g., VanderWerf, Brassinga, Reits, Aramideh, & deVisser, 2003).

In terms of function, blink rate has been shown to index cognitive load. People blink less under high cognitive load and more under low cognitive load (e.g., Nakano, Yamamoto, Kitajo, Takahashi, & Kitazawa, 2009; Siegle, Ichikawa, & Steinhauer, 2008). Supporting these behavioral findings, a neuroimaging study has revealed that blinking activates the default-mode network while deactivating the attention network, suggesting an active involvement of blinking in attentional disengagement (Nakano, Kato, Morito, Itoi, & Kitazawa, 2013). At the same time, blinking has been linked to social-communicative functions. Looking at different activity types in humans—staring at a target, reading, having a conversation—the highest blink rate was found in conversation (Doughty, 2001). In nonhuman primates, blink rate is correlated with group size (Tada et al., 2013), a measure of social complexity that has been linked to neocortex size (evidence used to support the "social brain hypothesis," Dunbar, 1992). These findings suggest that in addition to peripheral physiological and central cognitive functions, blinking may serve a social-communicative function (see also Mandel, Helokunnas, Pihko, & Hari, 2014; Nakano & Kitazawa, 2010; Tada et al., 2013).

To our knowledge, there are only two studies specifically investigating *addressee* blinking in conversation. Sultan (2004) demonstrated that in American Sign Language addressees use blinks to signal understanding: "I'm with you. I've got it. You can continue." (Sultan, 2004, p. 50). She suggests that addressee blinking may have developed a feedback function in signed languages because of the need to control blinking to minimize visual information loss. However, addressee blinking as a signal of understanding has also been described in Yélî Dnye, a spoken language of Papua New Guinea (Levinson & Brown, 2004).

In the present study, we hypothesized that addressee blinking may serve a similar function in spoken Dutch, which also relies heavily on the visual channel—like most spoken languages do, at least in face-to-face contexts.[1] If addressee blinking does indeed fulfill a feedback function in spoken conversation, one should expect addressee blinks to be timed to speakers' talk at similar points in time as other addressee responses, namely, at the ends of syntactically, prosodically, and pragmatically complete units (Ford, Fox, & Thompson, 1996; Gardner, 2001; Schegloff, 1982; Selting, 2000; Yngve, 1970). If addressee blinking does not serve a feedback function in spoken conversation, one may expect addressee blinks to be randomly distributed across turns, as their occurrence should not be influenced by conversational context.

To address this question, we built a corpus of dyadic Dutch face-to-face conversations. We then quantified the timing of addressee blinks by measuring the temporal distance of each blink onset to the closest TCU end (focusing on syntactically, prosodically, and pragmatically complete units within extended turns). In addition, we examined the multimodal compositionality of addressee blinks as well as the placement of long addressee blinks in conversational context.

[1]But see speakers of Mayan Tzeltal for an exception (e.g., Rossano, Brown, & Levinson, 2009).

Methods

Participants and corpus

The corpus consists of 10 dyads engaged in spontaneous Dutch face-to-face conversations for 1 hour each. All participants were native Dutch speakers (18–68 years; mean age = 30.7), they knew each other prior to the recording, and each participant participated only in one dyad. Four of the dyads were all female, four consisted of a female and a male participant, and two were all male. The recordings took place at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands.

Setup and equipment

The conversations were recorded in a sound-proof room, participants were positioned approximately 1 m from each other at a 45-degree angle, and each participant wore a head-mounted microphone (DPA-d:fine-88). Three HD video cameras (JVC GY-HM100) were used to record frontal views of each participant and a scene view (see Supplementary Material 1). An audio recorder (Roland R-44) recorded the two audio tracks in synchrony. Each recording session resulted in three videos and two audio files, which were then synchronized and exported in Adobe Premier Pro CS6 (MP4, 24 fps).

Procedure

Each recording session consisted of three 20-minute phases. To achieve the highest audio quality for this study, only the 20-minute phase in which participants wore head-mounted microphones was used. The whole session lasted about 90 minutes, and each participant was paid 16 euros. The study was approved by the Social Sciences Faculty Ethics Committee, Radboud University Nijmegen, and informed consent was obtained before and after filming.

Analysis

Turns analyzed. We selected tellings that occurred naturally in the conversations, and of these we selected only those that were elicited by the other speaker ("second-position tellings"; Mandelbaum, 2013). To address the question of whether *addressee* blinking may have a feedback function, we treated the whole telling, over several syntactically, prosodically, and pragmatically complete TCUs, as a single turn and excluded tellings with "recipient disruptions" (Mandelbaum, 2013), that is, nonminimal recipient responses (larger than three syllables), and other-initiations of repair or news receipts (e.g., *Oh did you?*) since they make a speaker response relevant before the telling continues.[2]

Turns and points of possible completion. Within all selected 46 turns, 456 points of possible completion were annotated, marking boundaries of *final* (230) and *nonfinal* (226) TCUs. At any point at which a turn was hearable to the coders as possibly complete in its context, focusing on convergence of syntactic, prosodic, and pragmatic completion, a *final* point of possible completion (Selting, 2000) was annotated, without this necessarily constituting the actual turn end. Final points of possible completion thus correspond to what Ford et al. (1996) called "Complex Transition Relevance Places," places at which speaker transition may occur.

Nonfinal points of possible completion were annotated if an utterance in progress was hearable to the coders as locally syntactically complete (e.g., a complete clause, whether syntactically

[2]Prior research suggests that blinking behaves differently in speakers as opposed to addressees (Cummins, 2012). Since we were interested in *addressee* behavior, we chose extended turns as a starting point because—compared to more turn-by-turn interaction—they provide a relatively clear distinction between speaker and addressee roles. The focus on extended turns in second position was based on plans for a follow-up experiment in which participant tellings would be initiated by an experimenter and for which this corpus study should serve as an observational basis.

independent or dependent) and locally prosodically complete (i.e., a complete intonational phrase) but further talk was made projectable syntactically (e.g., as in the case of an *if* clause), prosodically (e.g., rising intonation), or pragmatically (e.g., an incomplete answer to a question). Although at these nonfinal points of possible completion further talk is made projectable, these are typical points for addressee responses ("local pragmatic completion," Ford et al., 1996; Lerner, 1996). Whether addressee feedback is timed to nonfinal or final points of possible completion may have functional implications. For example, while feedback timed to nonfinal points may pass up the opportunity to initiate repair, feedback timed to final points may additionally pass up the opportunity to take a full turn (see also Goodwin, 1986).

The main coder (PH) and a second coder (EV) who was blind to the visual context in which the turns were produced identified the location of final and nonfinal points of possible completion by listening to the audio. Thirty-one extended turns were coded for training, and 11 extended turns (i.e., 24%) were coded to measure reliability. Initial discrepancies in coding regarded primarily prosodic completion points, and consensus was achieved through thorough discussion of cases where the two coders disagreed. Evaluation of their coding revealed high intercoder agreement on the identification of nonfinal and final TCU ends (84%).

Addressee blinks. Participants' blinks were detected automatically using motion tracking software ("IntraFace"; Xiong & De la Torre, 2013; see Supplementary Material 2). Within second-position tellings, all detected addressee blinks were manually corrected in terms of false positives, false negatives, and blink duration. Blink duration annotations included the first frame in which a downward movement of the eyelids was observable and the last frame in which an upward movement of the eyelids was observable.[3] Voluntary blinks have been shown to have longer durations (Kaneko & Sakamoto, 1999) and according to Levinson and Brown (2004), and a neuroimaging study (Mandel et al., 2014), longer blinks have a special communicative salience. We therefore decided to categorize blinks into short and long blinks. Previous research categorized blinks based on duration using a threshold of 250 ms (Levinson & Brown, 2004), 240 ms (and 400 ms; Cummins, 2012), and 420 ms (Hermann, 2010). We used a similar threshold as Hermann (2010), who, as we did in the present study, used the first observable downward movement of the eyelids to determine the blink onset and the last observable upward movement to determine the blink offset. We used a threshold of 410 ms, which separated the longest 25% from the rest (splitting them at the upper quartile[4]). This resulted in 350 short blinks (<410 ms) and 61 long blinks (≥410 ms).

Multimodal compositionality. We assessed for each addressee blink co-occurring addressee behaviors. We focused our analysis on the most salient addressee responses, namely, on nods (vertical head movement including at least one upward and one downward movement), vocal continuers (e.g., *mm-hm*), and combinations of these. Blinking was considered to be co-occurring if the blink overlapped with a nod or a vocal continuer or if it preceded or followed the nod or vocal continuer without perceived interruption, such that the behaviors together formed a multimodal *Gestalt* (Mondada, 2014). Blinking was *not* considered to be co-occurring if there was a temporal distance ≥ 250 ms between the blink and the nod or the vocal continuer. Secondly, blinking was considered with respect to the interactants' gaze direction. Blinking was considered as having occurred during a period of mutual gaze if mutual gaze existed at the onset of the blink, leading

[3]Our criterion mirrors Hermann (2010) but contrasts with Cummins (2012), who did not annotate a blink onset before the visible part of the cornea was not "substantially occluded" (p. 8), resulting in shorter blink durations overall.

[4]We split them at the upper quartile rather than the median because using the median as a threshold would have included too many blinks that were impressionistically short blinks in the category of "long blinks." What seems to differentiate impressionistically short from impressionistically long blinks is not so much the duration of the opening phase, nor the duration of the closing phase, but rather the duration of the cornea being completely occluded.

to a disruption of mutual gaze as the blinking was executed. All annotations were created in ELAN 4.8.1 (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

Results

We focus our analysis first on all addressee blinks and then examine potential differences in timing in different subsets of our data (short versus long blinks, blinks timed to nonfinal versus final TCUs). We then compare the multimodal compositionality of short and long blinks, followed by quantitative analyses with qualitative analyses of long addressee blinks in conversational context.

Addressee blinks timed relative to the end of TCUs

First, we looked at the frequency of addressee blinks. Addressees blinked approximately every other second on average, although there was substantial interindividual variability (mean blinks per minute = 30.53; *SD* = 20.43). Then we measured the temporal distance (in ms) between each blink onset and the closest end of a TCU (collapsing across short and long blinks and nonfinal and final TCU ends). If blink timing was random and TCU duration was constant, one would expect a uniform distribution of blink timings, that is, a flat horizontal line, indicating equal likelihood of blink occurrence at any point during a TCU. However, our analysis revealed that the most typical timing of blinks was very close to the TCU end (estimated mode = 52 ms, median = 20 ms, mean = −20 ms; see Figure 1), especially when considering the average TCU length of 1754 ms. This suggests that addressees tend to coordinate the onset of their blinks with the end of a speaker's TCU (see Supplementary Materials 3a and 3b for video examples).

However, an alternative explanation needs to be ruled out. TCU duration was variable, with many relatively short and much fewer long TCUs. Thus, even if blinks were timed randomly with respect to the end of TCUs, the temporal distance between blink onset and TCU end would necessarily be smaller for shorter than for longer TCUs. This, too, could lead to a distribution that peaks around zero.

To rule out this alternative explanation, we standardized our data by dividing the temporal distance between each blink onset and the closest TCU end by the duration of the respective

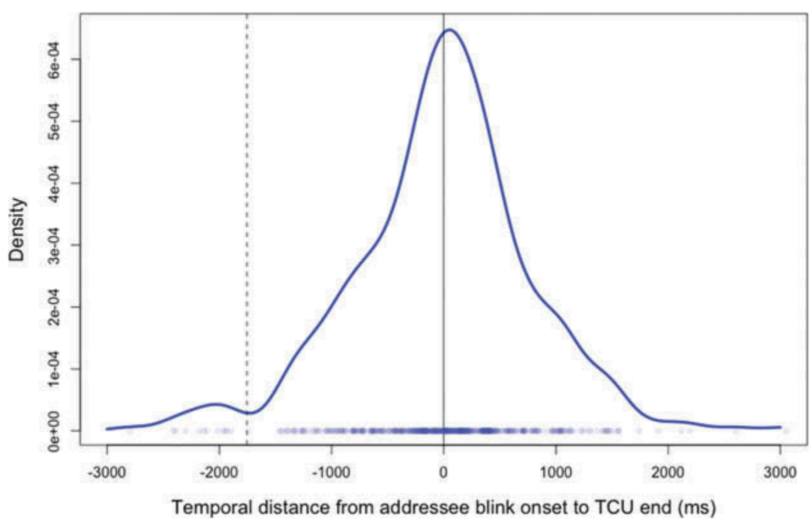

Figure 1. Addressees' blink onset (*N* = 411) relative to the closest TCU end. The vertical line at the zero point of the *x*-axis marks the TCU end. The distance between this line and the dashed vertical line marks the average TCU duration (1,754 ms). The peak of the distribution represents the estimate of the mode (52 ms).

TCU. This provided a measure indicating how closely each blink was timed to the TCU end *relative to the duration of that TCU*. A value of 0 means that the blink onset coincided with the TCU end, and a value of 1 means that the temporal distance of blink onset to TCU end was as large as the TCU itself, that is, the blink onset occurred at the beginning of the TCU. Visual inspection of the data reveals that taking into account the variability in TCU duration increased the spread of the distribution (see Figure 2) compared to the unstandardized data, which shows a relatively tighter distribution around zero (see Figure 1). However, measures of central tendency of the standardized data revealed that blinks were again most typically timed close to the end of TCUs (estimated mode = .00, median = .01, mean = .20). Taken together, this points to a clear tendency of addressees to coordinate their blinking with the end of TCUs.

Short versus long addressee blinks timed relative to nonfinal and final TCU ends. We used R (R Core Team, 2012) and *lme4* (Bates, Maechler, & Bolker, 2012) to test in a mixed effects model whether standardized blink timing differed depending on blink duration (short versus long blinks) and type of unit end (*nonfinal* versus *final*). Outliers deviating more than two standard deviations from the mean (n = 11, 2.6 % of the data) and three participants contributing less than five data points (requirement for *lme4*) were excluded. As fixed effect, we entered blink duration (short, long) and finality of TCUs (nonfinal, final), plus the interaction term. Participants were modeled as nested inside conversations, and as random effects we entered intercepts for participants into the model. Main effects and interaction effects were calculated using the ANOVA function of the *car*-package (Fox & Weisbert, 2011). Although based on visual inspection of Figure 3, long blinks seem to have occurred later than short blinks relative to TCU ends, this difference is not statistically significant (main effect of blink duration, $\chi^2(1) = 2.45$, $p = .11$). Moreover, blinks were not timed significantly differently relative to *nonfinal* versus *final* TCU ends (main effect of type of unit end, $\chi^2(1) = .32$, $p = .56$, and neither were short and long blinks timed significantly differently at *nonfinal* versus *final* TCU ends (interaction of blink duration and type of unit end, $\chi^2(1) = .10$, $p = .74$).

Multimodal compositionality of short vs. long addressee blinks

While short and long blinks were not timed differently relative to TCU ends, they might differ in the extent to which they co-occur with other addressee responses. We focused on the most

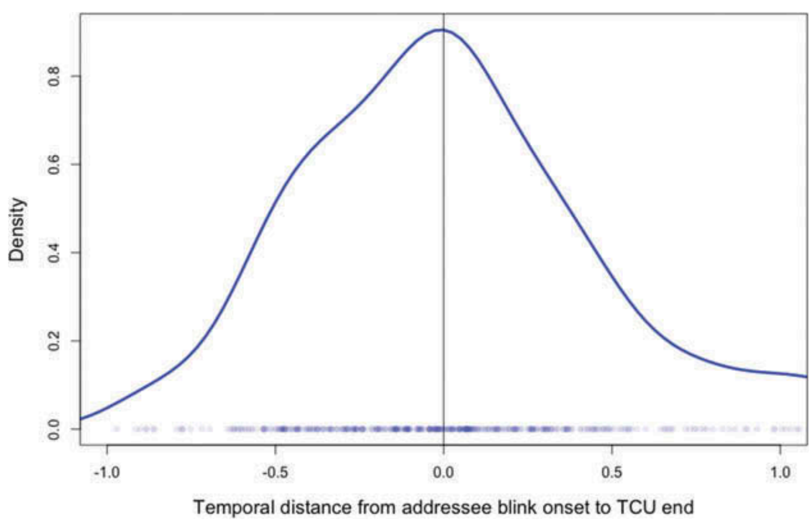

Figure 2. Timing of addressee blink onset relative to TCU ends based on standardized TCU duration.

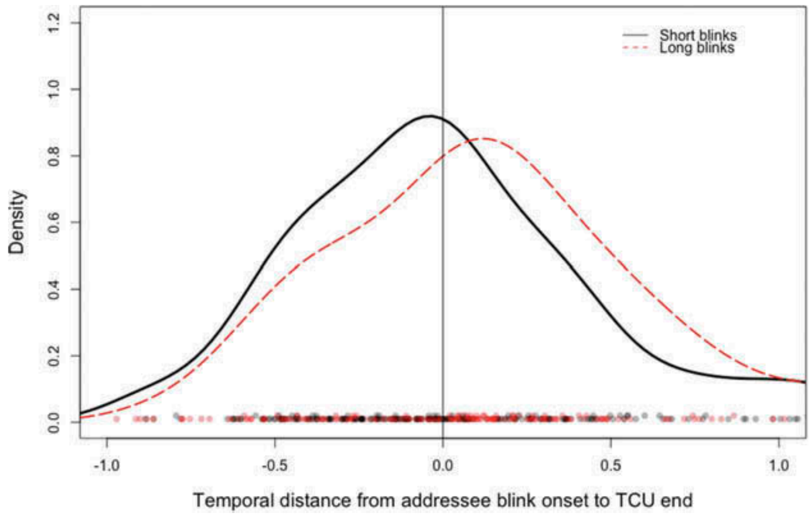

Figure 3. Addressee blink onset of *short* blinks versus *long* blinks relative to the ends of TCUs standardized in duration.

salient co-occurring addressee responses in our corpus, namely, on nods, vocal continuers (e.g., *mm-hm*), and combinations of these (see Method section). As one can see in Figure 4, while most short blinks (73% [n = 256]) and approximately half of the long blinks (54% [n = 33]) occurred without other responses (i.e., neither with nods nor with vocal continuers), long blinks were approximately twice as likely to co-occur with nods (18% [n = 11]) *or* vocal continuers (18% [n = 11]) than short blinks (9% [n = 32] and 8% [n = 29] respectively), supporting impressionistic observations by Cummins (2012). Predicting blink duration (short, long) by co-occurrence of other addressee responses (nod, vocal continuer, nod plus vocal continuer, absence of nod and vocal continuer) in a binomial logistic regression analysis confirmed this observation

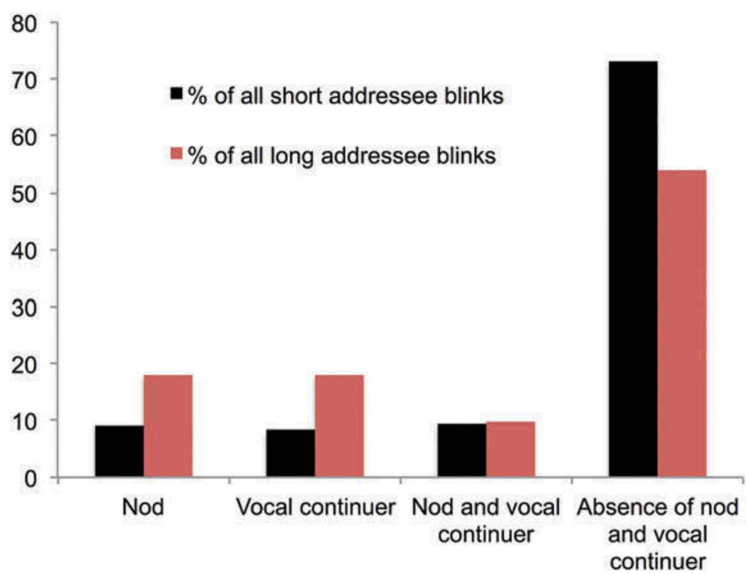

Figure 4. Co-occurrence of short blinks (in black; n = 350) versus long blinks (in red; n = 61) with nods, continuers, nods combined with continuers, and with no other addressee response.

statistically, indicating that the co-occurrence of other addressee responses reliably distinguished between short and long blinks, $\chi^2(3) = 11.6$, $p = .001$. Interestingly, long blinks were as likely to co-occur with vocal continuers *and* nods (10% [$n = 6$]) as short blinks (9% [$n = 33$]) were. One possibility is that with vocal continuers, the long blink compensated for the lack of a nod, and with nods, the long blink compensated for the lack of a vocal continuer. When nods and vocal continuers were combined into a "composite signal" (Clark, 1996), their joint "continuer effects" might have been sufficient, making the addition of a long blink more optional.

An additional analysis revealed that long blinks were also significantly more likely to be produced during the mutual "gaze window" (Bavelas, Coates, & Johnson, 2002), namely 72% of long blinks compared to 47% of short blinks.[5] Predicting blink duration (short, long) by mutual gaze (present, absent) in a binomial logistic regression analysis confirmed this observation statistically, indicating that the presence of mutual gaze reliably distinguished between short and long blinks, $\chi^2(1) = 9.7$, $p < .001$.

Overall, we would like to conclude from these patterns that (a) addressee blinks cluster at both final and nonfinal TCU ends; (b) that long blinks have a distinctive distribution regarding the concurrent production of nods and continuers; and that (c) long blinks are more likely than short blinks to be produced during mutual gaze—together suggesting that long blinks have a signaling capacity, producing feedback at critical points in the telling.

The sequential placement of long addressee blinks

The goal of this section is to further address the question of whether long blinks can serve a communicative function by examining the sequential contexts in which they were produced. To achieve this goal, we will first present a quantitative analysis of long blinks in different sequential contexts. Following this, we will provide qualitative analyses suggesting that long blinks can serve a continuer function (especially in combination with nods or vocal continuers) by displaying recipiency, passing up the opportunity to take a full turn or to initiate repair (Robinson, 2014; Schegloff, 1982) while signaling successful grounding (Clark, 1996). Finally, we will discuss potential alternative interpretations.

While short blinks occurred at a wide range of sequential locations in conversation, the use of long blinks was restricted to specific sequential contexts, namely, in response to repair solutions in self-initiated, same-turn, self-repair (47.5%), in response to disfluent speech (e.g., *uh, uhm*; 19.7%), at early recognition points (16.4%), and referring expressions (8.2%)[6,7].

Comparing all 61 long blinks to 61 randomly selected short blinks revealed that short blinks were significantly less likely to occur in these same contexts: in response to repair solutions in self-initiated, same-turn, self-repair (6.5%); at early recognition points (8.2%); in response to disfluent speech (3.2%); and referring expressions (1.6%). Predicting blink duration (short, long) by sequential context (self-repair, disfluency, early recognition point, referring expression) in a binomial logistic regression analysis confirmed this observation statistically, indicating that the sequential context reliably predicted blink duration, $\chi^2(4) = 45.9$, $p = .001$. The following examples focus on the two sequential contexts in which long blinks most frequently occurred, self-repairs and disfluencies.

[5]For this analysis, all 61 long blinks were compared with 61 randomly selected short blinks.

[6]Note that by "response" and "in response to" we do not intend to claim that long blinks, vocal continuers, etc., are necessarily conditionally relevant in these environments (since the present study does not demonstrate this), but see Zama and Robinson (2016) for further discussion.

[7]By "early recognition points" we refer to points at which it becomes highly predictable for the addressee what the speaker intends to say. They can occur within clauses where remaining words or syllables are highly predictable, but they can also occur at preliminary component completions within compound TCUs where the remaining units of the compound are highly predictable (e.g., in *if-then* constructions; Lerner, 1996). By "referring expressions" we refer to noun phrases identifying individual objects, events, or beings (e.g., proper names like *John*).

The long addressee blink in response to repair solutions in self-initiated, same-turn, self-repair. About half of the long blinks occurred after speakers' self-initiated, same-turn, self-repairs (Schegloff, 2013; Schegloff, Jefferson, Sacks, 1977). Addressee responses that can occur in this position are vocal continuers (Mazeland, 2007), nods (Healey, Lavelle, Howes, Battersby, & McCabe, 2013), or combinations of these. The following extracts demonstrate that long blinks also occur in this position, typically combined with nods or vocal continuers. In the transcripts, addressee blinks are bracketed between two small dots [·] and are synchronized with corresponding stretches of talk. Each large dot [●] reflects approximately 100 ms, so four [●●●●] or more large dots refer to a "long blink" (see the appendix for a full description of the transcription conventions used).

In Extract 1, A and B had talked about experiments to participate in at the university when A asks about a particularly long one B had previously mentioned (line 1).

```
Extract 1 (ETC_16_75310)

01 A: Maar wat was die andere nou? (.) tweenhalf uur nog wat
      but what was the other one now  two.and.a.half hour something
      But what was the other one now?  Two and a half hour something

02    (0.3)

03 B: +Ja en dan moest je-dan krijg je zo'n m+iddel toegediend?
      yes and then must you then get you such one agent administered
      Then you get some agent administered?
      +gaze averted                        +

04    T·MS· of zo iets,   [Ik weet niet pr]ecies,
      TMS or such something I know not precisely
      TMS or something, I don't know exactly,
   a   ·●●●·

05 A:                     [Ja:],
                           yes
                           Yeah

06 B: En dan werden je hersenen worden dan gestimuleerd .h,
      and then were your brains   are  then  stimulated
      And then your brains were then stimulated .h,

07    (0.2)

08 a   [°mhm°   ]
      ·●●●       ·

09 B: ·[Of bepa]·alde:: onderdelen van je hers·e#ne*·.hhh*hhh,
        or certain      parts      of your brains
        Or certain parts of your brain
→ a                                          ·●●●● ·
                                                 *nod  *
 fig                                           #fig.5

10 B: En+-eh: dan gaan ze ook allemaa::l (0.4) testjes met +j·e doen=
      and     then go they also all sorts of tests   with you do
      An- uh: then they go and do all sorts of tests with you
        +gaze averted                                  +
   a                                                     ·●●●->

11 B: =ik· weet niet precies.
       I   know  not precisely
      I don't know exactly.
   a   ->·
```

Figure 5. Addressee's long blink and head nod (left) following a speaker's self-initiated, same-turn, self-repair (Extract 1, lines 6–9; see also Supplementary Material 4 and 5 for video examples). Note that participants were facing each other in the laboratory and that these two images were taken from two frontal recordings (see Supplementary Material 1 for an illustration of participants' positioning in the laboratory).

In responding to this question, B produces two self-initiated, same-turn, self-repairs—each specifying some aspect of the respective prior unit. His first self-repair targets the type of agent (*middle*, line 3) he thinks is being administered by specifying it in line 4 (*TMS of zo iets*). The addressee registers this specification with a vocal continuer in line 5 (*Ja*). When B continues his response, resulting in a possible completion of a TCU in line 6, the addressee produces a vocal continuer (*mhm*; line 8) following a short gap (line 7). Note that the possible delay (line 7) may have influenced the speaker's production of the self-repair in line 9 and that the addressee's vocal continuer (line 8) occurs simultaneously with the onset of that self-repair. This second self-repair of the turn (line 9) again targets a referent mentioned in the prior unit as the repairable. In this case, the speaker provides a repair solution, specifying that it is not necessarily the brain as a whole that is being stimulated (as might be inferred from line 6) but only certain parts of the brain (*of bepaalde onderdelen van je hersenen*, line 9). The *or*-preface in this self-repair may already project that the trouble source formulation (line 6) is not about to be discarded as a whole but that merely a specification is coming instead (Lerner & Kitzinger, 2015, on *or*-prefaced self-repair in English). The addressee registers the repair solution with a long blink and a nod in line 9, before the speaker averts his gaze from the addressee and continues his telling in line 10.

The long addressee blink in response to disfluencies. In Extract 1, the long blink was combined with a nod. On the one hand, the co-occurrence of long blinks with nods supports the argument that long blinks serve a similar function. On the other hand, it raises the question whether long blinks per se can have a continuer function. Do long blinks *alone* also occur in positions where a continuer is relevant? They do, indeed. Long blinks alone were especially observed in response to intra-TCU troubles in speaking, that is, in response to disfluencies (e.g., sound stretches, cut-offs, *uh, uhm*; Clark & Tree, 2002; Schegloff, 2010; Lickley, 2015) or to what Levelt (1989) called the "editing phase" in self-repairs—the interval between the end of a repairable and the beginning of a repair solution.[8] Previous research has already demonstrated that addressees responses (nods in this case) are especially relevant at points where the speaker has difficulties in producing a turn (Healey et al., 2013). In line with this research, Extract 2 demonstrates that continuers *in general* (in this case a vocal continuer and a long blink) are relevant in response to troubles in speaking, while Extract 3 exemplifies a case where a long blink *alone* is produced in this position. Prior to Extract 2, A and B talked about an upcoming pool party.

[8]Four cases of long blinks in isolation also occurred at points of possible but not actual sequence completion—which would further support the argument that they can serve a continuer function (Schegloff, 1982, p. 84). However, since most of these were accompanied by laughter or larger gaze shifts—which could alternatively explain the eye closures (Evinger et al., 1994; see also Rossano, 2012; on sequence-final gaze withdrawal)—we did not include them in this analysis.

```
Extract 2 (ETC_19_582999)

01 A: Wat is nou-wat is nou het thema van-van-van de pool party
      what is now what is now the theme of of of the pool party
      What is the theme of the pool party now?

      ((Lines omitted))

02 B: Wat ik had bedacht,
      what I had invented
      What I had invented

03    en dat vonden ze allemaal leuk,
      and that found they all lovely
      and they all liked the idea

04    is een Finding Nemo thema.
      is a   Finding Nemo theme
      is a   Finding Nemo theme

      ((Lines omitted))

05 B  en dan uh: (0.2)·[(0.2)+dumpen we]· ergen+s een visje,
      and then uh             dump  we somewhere a little fish
      and then uh              we dump a little fish somewhere
                             +averted gaze    +
   a                   [mhm             ]
→                      ·●●●●●            ·
```

In response to A's question regarding the theme of the upcoming pool party (line 1), B answers by mentioning the "Finding Nemo" theme he had come up with (lines 2–4). When B has trouble producing an elaboration on the details of this theme as reflected by disfluencies (*en dan* followed by *uh* and a pause; line 5), the addressee produces a vocal continuer (*mhm*) and a long blink. This illustrates that continuers in general are not only relevant at TCU ends but that they can also be relevant within a TCU-in-progress, in response to disfluency-related disruptions of progressivity.

Similarly, Extract 3 shows an example of a case where a long blink *alone* is produced in response to a disfluency-related disruptions of progressivity within a TCU-in-progress. Prior to Extract 4, A and B have talked about possible romantic matches among friends.

```
Extract 3 (ETC_10_72939)

01 A: Maar wat zou je doen als je zelf in di-die positie zit=
      but what would you do if you self in thi this position sit
      But what would you do if you were in that position

02    =dan ga je dat ook wel doen
      then go you that  also  do
      then you would also do it

      ((Lines omitted))

03 B: En die zei ja ik v- ik denk da-z-prec-uh:-·geschikt ty·pe
      and he said yes I f I think tha sh prec uh suitable type
      And he said well I f I think tha sh prec uh is the right type
→ a                                                ·●●●●      ·

04    voor Helen is
      for  Helen is
      for Helen
```

In response to A's question (line 1), B reports that one friend had stated that another friend would be a suitable type for Helen (lines 3–4). After B produces several disfluencies in succession (multiple cutoffs plus *uh: ik v- ik denk da-z- prec-uh*, line 3), A produces a long blink *alone*. The succession of disfluencies in line 3 starts with a cut-off *ik v-* (recognizable as the beginning of the word *vind* 'find'), which is replaced with *ik denk*. Subsequently, the cutoff *da-z-* (potentially recognizable as the beginning of the words *dat* 'that' and *ze* 'she' respectively) is followed by another cut-off *prec-uh* (potentially recognizable as the beginning of the word *precies* 'exactly'), which is then replaced with *geschikt* 'suitable.' Note that the long addressee blink in this case is not produced in response to a repair *solution* (as in Extract 1), but it is produced in response to the disfluent editing phase (immediately after *uh*; line 3), and it only partially overlaps with the beginning of a repair solution (*geschikt type*, line 3) in this "replacing" type of self-initiated, self-repair (Schegloff, 2013).

Summary and discussion. First, we have seen that long addressee blinks were especially produced in response to repair solutions in speakers' self-initiated, same-turn, self-repairs. The long addressee blink in these positions (combined with a nod in 83% of cases) appears to pass up the opportunity to initiate repair or to take a full turn (Schegloff, 1982). By orienting to the continuation of a turn-in-progress, the addressee's long blink also seems to signal that by having received a repair solution the speaker has reached sufficient informativeness for current purposes. It may signal to that there is no need to clarify or specify further by producing more self-repairs, thereby helping the speaker in avoiding underinforming and overinforming (Grice, 1975; Mazeland, 2007; Sacks & Schegloff, 1979). If this is true, then speakers may strategically produce self-repairs in order to mobilize addressee responses (see Goodwin, 1980; Jefferson, 1974; Stivers & Rossano, 2010). Furthermore, we have also seen that long blinks often co-occurred with other continuers, which raises the question whether long blinks contribute anything at all in these cases and whether, for example, blinks are not a side effect of nodding. Since head movements often co-occur with gaze shifts and gaze shifts often co-occur with blinks (Evinger et al., 1994), blinks may simply be a way to optimize vision while nodding. Evinger et al. (1994) stated that, "since saccadic suppression impairs vision during a saccadic gaze shift, a gaze-evoked blink can lubricate the cornea without interfering with vision" (p. 342). However, there are at least two reasons to doubt such an account. First, in our data most blinks did not co-occur with nods (see Figure 4), and 36% (n = 30) of all nods (n = 83) were not combined with blinks at all. Secondly, addressees typically do not shift gaze while nodding; they tend to keep looking at the speaker's face. Yet, while blinks may not be caused by nods, there may be motoric synergies between producing a nod and producing a blink (e.g., the downward movement of the head facilitating the downward movement of the upper eyelid).

Secondly, we have seen that long blinks alone (i.e., combined with neither a nod nor a vocal continuer in 75% of cases) were also used in positions where a continuer is relevant, namely, in response to intra-TCU disfluency-related disruptions of progressivity. What might be the function of a long blink produced in this position? First, producing a long blink in this position may display continued recipiency. Like TCU ends, intra-TCU disfluencies may be vulnerable places at which addressees might steal the turn (Jefferson, 1983; Maclay & Osgood, 1959). By producing a long blink, the addressee may display attentiveness through responsiveness, passing up the opportunity to take a turn and thereby aligning with the telling in progress. Secondly, it has been argued that speakers prosodically distract their addressee's attention from the repairable in their self-repairs by increasing the pitch and loudness of the following repair proper (Noteboom & Quené, 2014). Similarly, long blinks (as the ones in Extracts 2–3) may contribute to camouflaging or rendering invisible the presumably unintended reduction in progressivity, thereby minimizing the potentially face-threatening act (Brown & Levinson, 1987) of "watching" the speaker having difficulties speaking. On a related note, research has shown that perceived direct gaze is cognitively demanding (Conty, Gimmig, Belletier, George, & Huguet, 2010). If this is true, then addressees may produce long blinks during disfluencies to momentarily reduce the speaker's cognitive load. By blinking, the addressee disrupts eye contact, which might help the speaker to make her speech fluent again—orienting to progressivity while minimizing "joint effort" (Clark, 1996). In short, in the face of disfluencies, then, long addressee blinks are presumably used to display continued recipiency, to camouflage the disfluencies, to reduce the speaker's cognitive load, or are due to a combination of these potential functions.

Finally, while the long addressee blink cases presented here already provide suggestive evidence, future research will be required to provide conclusive evidence. More conversation analytic research would be desirable to further demonstrate that participants orient to long blinks as consequential social actions, while experimental research will be required to confirm the hypothesis that speakers are *causally* influenced by addressee blink behavior. Based on the present study, however, we would like to suggest that since long addressee blinks appeared in the same sequential positions as continuers, they are likely to serve similar functions: orienting to progressivity by passing up the opportunity to take a turn or to initiate repair (Robinson, 2014; Schegloff, 1982) and aligning with the telling-in-progress (Stivers, 2008) while signaling successful grounding (Clark, 1996).

General discussion

Does addressee blinking serve a feedback function in conversation? First, the present results are incompatible with a purely physiological interpretation of blinking in conversation. While all blinks also lubricate the cornea, addressees blinked too frequently (approximately 30 times per minute) to serve solely this physiological function (Doane, 1980). More strikingly, these results are consistent with a social-communicative interpretation of blinking because addressee blinks were timed around TCU ends, the location where addressee responses are typically produced and where speakers tend to visually monitor addressees for feedback (Bavelas et al., 2002).

Second, short and long addressee blinks appear to fulfill partially different functions. While there was no clear difference in timing between short and long blinks relative to TCU ends, long blinks—in contrast to short blinks—were (a) less frequent; (b) more likely to co-occur with nods or vocal continuers; (c) more likely to occur during mutual gaze; and (d) their production was restricted to sequential contexts in which it was structurally relevant to display recipiency and understanding, together suggesting that long blinks are particularly likely to serve a social-communicative function—especially when combined with other feedback responses.

Can a cognitive interpretation of the timing of addressee blinks be ruled out? The fact that the majority of addressee blinks were timed to TCU ends is also consistent with a cognitive interpretation of blinking (Nakano et al., 2009, 2013; Siegle et al., 2008). Nonblinking while listening to the core of a TCU may be a symptom of addressees' high cognitive load, while blinking at the end of a TCU may reflect addressees' relative decrease in cognitive load. While this cognitive interpretation may hold for short blinks, it seems less plausible for long blinks. Short blinks occurred throughout whole conversations. Long blinks, however, were placed in specific sequential positions and, although this is an open issue requiring further investigation, it seems unlikely that in these positions addressees' cognitive load was particularly low relative to other sequential positions.

Cognitive, perceptual, and social functions of addressee blinks are not mutually exclusive. It is possible that the cognitive and perceptual functions underlie and evolutionarily preceded any potential social-communicative function. Perhaps blinking as a symptom of low cognitive load or attentional disengagement and the need to control blinking to minimize audiovisual information loss during speech comprehension (McGurk & MacDonald, 1976; Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007) have been co-opted for communicative purposes, such that they are now used as a semiotic signal. As squinting the eyes—as if to see more clearly—seems to signal lack of understanding, closing the eyes by blinking seems to convey "no need to see anymore" because understanding has been established (Lakoff & Johnson, 1999, on the Understanding-Is-Seeing metaphor).

Interestingly, our results suggesting a continuer function for addressee blinking in Dutch are in line with studies on blinking in Yélî Dnye (Levinson, 2015; Levinson & Brown, 2004) and American Sign

Language (Sultan, 2004). At least based on this limited number of studies, addressee blinks as signals of successful grounding seem to be independent from language modality—since it is used in spoken as well as signed language—as well as from language history—since it has also been described in Yélî Dnye, a Papuan language. If addressee blinking as a signal of recipiency and successful grounding is shared across a wider range of unrelated languages, it may have evolved due to common pressures of a shared conversational infrastructure (Dingemanse, Torreira, & Enfield, 2013; Stivers et al., 2009).

Funding

This research was funded by the European Research Council (Advanced Grant #269484 INTERACT awarded to Prof. Levinson) and the Max Planck Gesellschaft.

References

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes*. Retrieved from http://CRAN.R-project.org/package=lme4

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*(6), 941–952. doi:10.1037/0022-3514.79.6.941

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, *52*(3), 566–580. doi:10.1111/jcom.2002.52.issue-3

Beach, W. (1993). Transitional regularities for 'casual' "Okay" usages. *Journal of Pragmatics*, *19*, 325–352. doi:10.1016/0378-2166(93)90092-4

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, *53*, 301–344.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge, England: Cambridge University Press.

Brunner, L. J. (1979). Smiles can be back-channels. *Journal of Personality and Social Psychology*, *37*(5), 728–734. doi:10.1037/0022-3514.37.5.728

Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81. doi:10.1016/j.jml.2003.08.004

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*, 73–111.

Conty, L., Gimmig, D., Belletier, C., George, N., & Huguet, P. (2010). The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition*, *115*(1), 133–139. doi:10.1016/j.cognition.2009.12.005

Cummins, F. (2012). Gaze and blinking in dyadic conversation: A study in coordinated behaviour among individuals. *Language and Cognitive Processes*, *27*(10), 1525–1549. doi:10.1080/01690965.2011.615220

Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is "Huh?" a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *Plos ONE*, *8*, e78273. doi:10.1371/journal.pone.0078273

Doane, M. G. (1980). Interaction of eyelids and tears in corneal wetting and the dynamics of the normal human eyeblink. *American Journal of Ophthalmology*, *89*(4), 507–516. doi:10.1016/0002-9394(80)90058-6

Doughty, M. J. (2001). Consideration of three types of spontaneous eyeblink activity in normal humans: During reading and video display terminal use, in primary gaze, and while in conversation. *Optometry and Vision Science*, *78*(10), 712–725. doi:10.1097/00006324-200110000-00011

Dunbar, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*(6), 469–493. doi:10.1016/0047-2484(92)90081-J

Evinger, C., Manning, K. A., Pellegrini, J. J., Basso, M. A., Powers, A. S., & Sibony, P. A. (1994). Not looking while leaping: The linkage of blinking and saccadic gaze shifts. *Experimental Brain Research*, *100*(2), 337–344. doi:10.1007/BF00227203

Ford, C. E., Fox, B. A., & Thompson, S. A. (1996). Practices in the construction of turns: The "TCU" revisited. *Pragmatics*, *6*, 427–454. doi:10.1075/prag.6.3.07for

Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*. Amsterdam, The Netherlands: John Benjamins.

Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry*, *50*(3–4), 272–302. doi:10.1111/j.1475-682X.1980.tb00023.x

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York, NY: Academic Press.

Goodwin, C. (1986). Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies*, *9*(2–3), 205–217. doi:10.1007/BF00148127

Grice, H. P. (1975). Logic and conversation. *Syntax and Semantics*, *3*, 41–58.

Healey, P. G. T., Lavelle, M., Howes, C., Battersby, S., & McCabe, R. (2013). How listeners respond to speaker's troubles. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the Cognitive Science Society* (pp. 2506–2511). Berlin, Germany: Cognitive Science Society.

Heath, C. (1992). Gesture's discreet tasks: Multiple relevancies in visual conduct and in the contextualization of language. In P. Auer & A. D. Luzio (Eds.), *The contextualization of language* (pp. 101–127). Amsterdam, The Netherlands: John Benjamins.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 299–347). Cambridge, England: Cambridge University Press.

Herrmann, A. (2010). The interaction of eye blinks and other prosodic cues in German Sign Language. *Sign Language and Linguistics*, *13*(1), 3–39. doi:10.1075/sll.13.1

Jefferson, G. (1974). Error correction as an interactional resource. *Language in Society*, *3*, 181–199. doi:10.1017/S0047404500004334

Jefferson, G. (1983). *Two explorations of the organization of overlapping talk in conversation: Notes on some orderliness in overlap onset*. Tilburg papers in language and literature, No. 28. Tilburg, The Netherlands: Tilburg University, 11–38.

Kaneko, K., & Sakamoto, K. (1999). Evaluation of three types of blinks with the use of electro-oculogram and electromyogram. *Perceptual and Motor Skills*, *88*(3), 1037–1052. doi:10.2466/pms.1999.88.3.1037

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York, NY: Basic Books.

Lerner, G. H. (1996). On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. *Studies in Interactional Sociolinguistics*, *13*, 238–276.

Lerner, G. H., & Kitzinger, C. (2015). *Or*-prefacing in the organization of self-initiated repair. *Research on Language and Social Interaction*, *48*(1), 58–78. doi:10.1080/08351813.2015.993844

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levinson, S. C. (2015). Other-initiated repair in Yélî Dnye: Seeing eye-to-eye in the language of Rossel Island. *Open Linguistics*, *1*(1), 386–410. doi:10.1515/opli-2015-0009

Levinson, S. C., & Brown, P. (2004, February). *Comparative feedback: Cultural shaping of response systems in interaction*. Talk presented at the workshop on Feedback in Interaction, Nijmegen, The Netherlands.

Lickley, R. J. (2015). Fluency and disfluency. In M. A. Redford (Ed.), *The handbook of speech production* (pp. 445–474). Hoboken, NJ: John Wiley.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, *15*(1), 19–44. doi:10.1080/00437956.1959.11659682

Malisz, Z., Włodarczak, M., Buschmeier, H., Skubisz, J., Kopp, S., & Wagner, P. (2016). The ALICO Corpus: Analysing the active listener. *Language Resources and Evaluation*, *50*(2), 411–442.

Mandel, A., Helokunnas, S., Pihko, E., & Hari, R. (2014). Neuromagnetic brain responses to other person's eye blinks seen on video. *European Journal of Neuroscience*, *40*(3), 2576–2580. doi:10.1111/ejn.12611

Mandelbaum, J. (2013). Storytelling in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 492–508). Chichester, England: Wiley-Blackwell.

Mazeland, H. (2007). Parenthetical sequences. *Journal of Pragmatics*, *39*(10), 1816–1869. doi:10.1016/j.pragma.2007.05.005

McClave, E. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, *32*(7), 855–878. doi:10.1016/S0378-2166(99)00079-X

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. doi:10.1038/264746a0

Mondada, L. (2011). Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, *43*(2), 542–552. doi:10.1016/j.pragma.2010.08.019

Mondada, L. (2014). The temporal orders of multiactivity. In P. Haddington, T. Keisanen, L. Mondada, & M. Nevile (Eds.), *Multiactivity in social interaction* (pp. 3–32). Amsterdam, The Netherlands: John Benjamins.

Nakano, T., Kato, M., Morito, Y., Itoi, S., & Kitazawa, S. (2013). Blink-related momentary activation of the default mode network while viewing videos. *Proceedings of the National Academy of Sciences*, *110*(2), 702–706. doi:10.1073/pnas.1214804110

Nakano, T., & Kitazawa, S. (2010). Eyeblink entrainment at breakpoints of speech. *Experimental Brain Research*, *205*(4), 577–581. doi:10.1007/s00221-010-2387-z

Nakano, T., Yamamoto, Y., Kitajo, K., Takahashi, T., & Kitazawa, S. (2009). Synchronization of spontaneous eyeblinks while viewing video stories. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 3635–3644. doi:10.1098/rspb.2009.0828

Nooteboom, S., & Quené, H. (2014). Do speakers try to distract attention from their speech errors? In J. Caspers, Y. Chen, W. Heeren, J. Pacilly, N. Schiller, & E. Van Zanten (Eds.), *Above and beyond the segments: Experimental linguistics and phonetics* (pp. 203–217). Amsterdam, The Netherlands: John Benjamins.

Ponder, E., & Kennedy, W. P. (1927). On the act of blinking. *Quarterly Journal of Experimental Physiology*, *18*(2), 89–110. doi:10.1113/expphysiol.1927.sp000433

R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robinson, J. D. (2014). What "what?" tells us about how conversationalists manage intersubjectivity. *Research on Language and Social Interaction*, *47*(2), 109–129. doi:10.1080/08351813.2014.900214

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, *17*(5), 1147–1153. doi:10.1093/cercor/bhl024

Rossano, F. (2012). *Gaze behavior in face-to-face interaction* (Unpublished doctoral dissertation). Radboud University Nijmegen, Nijmegen, The Netherlands.

Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. In J. Sidnell (Ed.), *Conversation analysis: Comparative perspectives* (pp. 187–249). Cambridge, England: Cambridge University Press.

Sacks, H., & Schegloff, E. A. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 15–21). New York, NY: Irvington.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735. doi:10.1353/lan.1974.0010

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 71–93). Washington, DC: Georgetown University Press.

Schegloff, E. A. (2010). Some other "uh(m)"s. *Discourse Processes*, *47*, 130–174.

Schegloff, E. A. (2013). Ten operations in self-initiated, same-turn repair. In M. Hayashi, G. Raymond, & J. Sidnell (Eds.), *Conversational repair and human understanding* (pp. 41–70). Cambridge, England: 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*, 361–382. doi:10.1353/lan.1977.0041

Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, *29*(4), 477–517. doi:10.1017/S0047404500004012

Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, *45*(5), 679–687. doi:10.1111/psyp.2008.45.issue-5

Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction*, *41*(1), 31–57. doi:10.1080/08351810701691123

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592. doi:10.1073/pnas.0903616106

Stivers, T., & Rossano, F. (2010). Mobilizing response. *Research on Language & Social Interaction*, *43*(1), 3–31. doi:10.1080/08351810903471258

Sultan, V. D. (2004). *Blinking as a recipient resource in American Sign Language discourse* (Unpublished doctoral dissertation). University of California, Santa Barbara, CA.

Tada, H., Omori, Y., Hirokawa, K., Ohira, H., Tomonaga, M., & Sueur, C. (2013). Eye-blink behaviors in 71 species of primates. *Plos One*, *8*(5), e66018. doi:10.1371/journal.pone.0066018

VanderWerf, F., Brassinga, P., Reits, D., Aramideh, M., & de Visser, B. O. (2003). Eyelid movements: Behavioral studies of blinking in humans under different stimulus conditions. *Journal of Neurophysiology*, *89*(5), 2784–2796. doi:10.1152/jn.00557.2002

Whitehead, K. A. (2011). Some uses of head nods in "third position" in talk-in-interaction. *Gesture*, *11*(2), 103–122. doi:10.1075/gest.11.2

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)* (pp. 1556–1559). Retrieved from http://www.mpi.nl/publications/escidoc-60436/@@popup

Xiong, X., & De la Torre, F. (2013). *Supervised descent method and its applications to face alignment*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 532–539). Los Alamitos, CA: IEEE Computer Science Conference Publishing Services.

Yngve, V. H. (1970). On getting a word in edgewise. In M. A. Campbell (Ed.), *Papers from the sixth regional meeting, Chicago Linguistics Society* (pp. 567–578). Chicago, IL: Department of Linguistics, University of Chicago.

Zama, A., & Robinson, J. D. (2016). A relevance rule organizing responsive behavior during one type of institutional extended telling. *Research on Language and Social Interaction*, *49*(3), 220–237. doi:10.1080/08351813.2016.1196551

Zametkin, A. J., Stevens, J. R., & Pittman, R. (1979). Ontogeny of spontaneous blinking and of habituation of the blink reflex. *Annals of Neurology*, *5*(5), 453–457. doi:10.1002/(ISSN)1531-8249

Appendix

Transcription conventions (Based on Mondada, 2014)

Gestures and descriptions of embodied actions are delimited between two identical symbols and are synchronized with correspondent stretches of talk (one symbol per embodied action: · for addressee blinks [each ● reflects approximately 100 ms], * for nods, + for speaker gaze aversion during the telling, ∞ for addressee gaze aversion during the telling, ♦ for other bodily conduct). The action described continues until after the excerpt's end (*—≫) or across subsequent lines (*—>) until the same symbol is reached (—->*). Participants are identified in the margins. Capital letters (e.g., A) indicate speakers, small letters (e.g., a) indicate addressees. The moment at which a still image has been taken is indicated with a # showing its position within the turn. The corresponding figure number is shown in a separate line (e.g., #Figure 5).