LANGUAGE AND COGNITIVE PROCESSES
0000, 00 (00), 1–25

Psychology Press
Taylor & Francis Group

Gaze and blinking in dyadic conversation: A study in coordinated behaviour among individuals

Fred Cummins

School of Computer Science and Informatics, University College Dublin, Dublin, Ireland

Face to face conversation necessarily involves a great deal of bodily movement beyond that required for speaking. We seek to understand the systematic variation of such para-linguistic activity as a function of the ebb and flow of conversation. Gaze and blinking in dyadic conversation are examined, along with their relation to speech turn. Eight pairs provide 15 minutes of conversation each, including five participants who partake in two dyads each. This facilitates a thorough examination of the rich covariation of gaze and blinking both within an individual and as a function of the dyad. Many aspects of systematic variation are found to be relatively invariant within the individual, but individuals display large qualitative differences, one from the other.

Keywords: Blinking; Gaze; Coordinative structure; Conversation.

INTRODUCTION

It is well established that speaking is an embodied activity that involves movement and activity far beyond that required to generate speech sounds (Kendon, 1994; McNeill, 1992). Manual gestures accompany the speech stream, even when they are not visible to the listener (Goldin-Meadow, 1999), and their timing is yoked, in ways still poorly understood, to the ongoing stream of speech being produced (Wachsmuth, 1999). Facial expressions are an integral part of face-to-face communication (Argyle & Cook, 1976; Grant, 1969) and the appropriate implementation of both gesture and facial expression into the behavioural repertoire of avatars is a long-standing goal in modelling in human-computer interaction (Cassell et al., 1994). Eye gaze patterns constitute another important element in natural conversation that has been extensively studied (Argyle & Cook, 1976).

Each of these aspects to para-linguistic movement co-occurring with natural conversation has been subject to many kinds of interpretation, typically with a view to

---

Correspondence should be addressed to Fred Cummins, School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland. E-mail: fred.cummins@ucd.ie

Authors would like to thank Catharine Oertel, Donnacha Costello, Tiernan Williams, and Rob Brennan for their willingness to learn complex annotation techniques and provide annotations for methodological verification. They would also like thank three very diligent and thorough anonymous reviewers. Partial results from two dyads only were presented at the 21st National Conference on Artificial Intelligence and Cognitive Science, in Galway, Ireland, 2010.

http://www.psypress.com/lcp http://dx.doi.org/10.1080/01690965.2011.615220

identifying and even quantifying the function of the behaviour in the service of communication. We briefly review such functional approaches, before offering an alternative way of approaching such data from a coordinative perspective.

Gestures may have clearly interpretable semantic content, or they may be "content-free," such as beat gestures, that are integrated into the stream of conversation without obvious referential function (McNeill, 1992). An interpretable gesture might be obviously referential, as in tracing out a circle while saying the word; deictic gestures such as simple pointing also admit of obvious semantic interpretation. But most gestures are not of this sort. Simple beat gestures, for example, are devoid of obvious semantic content, and yet are tightly integrated into the temporal structure of a speaker's movement (Leonard & Cummins, 2010). Arguing for the irrelevance of such nonlinguistic movement, Krauss, Dushay, Chen, and Rauscher (1995) found that gesturing in a controlled experiment did not appear to facilitate the semantic effectiveness of communication.

A pragmatic standpoint that accords somewhat more weight to nonverbal contributions to conversation is provided by Quek et al. (2000). The authors emphasise the complementarity of verbal and nonverbal elements in expressive communication, and show that in multimodal communication, ambiguity or imprecision in one mode may be compensated for by information carried in another mode, even if this information is not semantically interpretable. Thus, for example, gaze and gesture can be used to help to segment an ambiguous discourse based on topic, shared ground, etc. This kind of pragmatic function is best subserved by indexical gestures, and gaze that occurs in the presence of a set of props that are referred to in the conversation.

Uniquely among para-linguistic movements, gaze behaviour has captured the attention of artists, poets, and scientists alike. Pre-theoretically, we are all acutely aware that gaze is a potent element in any visually mediated social situation. In the attribution of function to gaze, there has been a great deal of theorising, but little consensus. Gaze has been seen as critical for negotiating turn-taking (Duncan, 1972), for modulating and directing attention (Vertegaal, Slagter, van der Veer, & Nijholt, 2001), for expressing intimacy and exercising social control (Kleinke, 1986), and for highlighting the information structure of the propositional content of speech (Cassell et al., 2000).

Much less attention has been paid to blinking. Blinking is usually unconscious for both the blinker and any conversational partner. Blink rate has been found to vary systematically with specific behaviours such as reading, conversing, watching film, etc. (Drew, 1951; Fogarty & Stern, 1989; Fukuda, 1994), and this in turn is typically attributed to variation in cognitive load (Orchard & Stern, 1991). Despite evidence for some systematic variation with task, blinking has been almost universally regarded as epiphenomenal, rather than functional (with the important exception of blinking while communicating using sign language; see section headed "Blinking in conversation" below).

Coordination in conversation

Functional and epiphenomenal accounts of behaviour in conversation do not exhaust the space of theoretical options. A rather different approach to embodied conversation is provided by Shockley, Richardson, and Dale (2009), in which body movement, posture, and gaze patterning are interpreted as evidence of the establishment of an interpersonal domain of coordination. Based on the use of similar models in motor control, they call this domain a coordinative structure (Kelso, 1994). This interpretation differs fundamentally from the above functional stance, and the role of paralinguistic indices such as gesture and gaze is likewise different here (Richardson, Dale, & Kirkham, 2007). Instead of being either excluded from the business of conversation, or subserving some function in support of linguistic communication, the alignment of movement of all sort among conversational participants is taken as evidence of the establishment and maintenance of a coordinative structure. On this view, coordinated movement *is constitutive of* the conversational domain.

To further develop this view, it is necessary to approach the conversational situation with strict empirical intent. It is not known in advance which observable characteristics of a conversational participant might be found to serve in identifying the transient and emergent inter-personal coordinative domain. Shockley, Santana, and Fowler (2003) observed postural sway and found that interacting subjects engaged in a collaborative task shared postural positions, and maintained similar postural trajectories more often than when not so engaged. In related work, Richardson and Dale (2005) found coupling among the eye movements of participants who were talking while viewing a common display. The coupling of gaze was found to co-vary with performance on questions testing comprehension. Postural sway and gaze are two possible observable quantities that seem to provide evidence for the establishment of an interpersonal coordinative domain in conversation.

Gaze in conversation

Gaze has been assigned many hypothetical functions. It has been suggested that it serves to help coordinate turn-taking (Cassell, Cassell, Torres, & Prevost, 1999; Thrisson, 2002), to grab the attention of a co-speaker (Waters, Rehg, Loughlin, Kang, & Terzopoulos, 1998), to indicate an object to which attention should be directed (Lester, Towns, Callaway, Voerman, & FitzGerald, 2000; Thrisson, 2002), and to indicate the distribution of attention of the gazer (Khullar & Badler, 2001). This kind of functional interpretation of gaze lends itself well to modeling and testing in the domain of embodied conversational agents (Cassell et al., 2000; Poggi, Pelachaud, & De Rosisc, 2000; Vertegaal et al., 2001). Functional models may also be combined with statistical models of such behaviours, allowing a certain amount of personalisation of the resulting behaviour (Pelachaud & Bilvi, 2003), or differentiation in likelihood of specific behaviours as a function of role (speaking vs. listening) (Lee, Badler, & Badler, 2002).

Prior studies have shown that subjects gaze at their partners more when listening than when speaking. Kendon (1967) found ranges of 20–50% gaze at partner when speaking, compared with 30–80% when listening. Mean proportions of 38% (speaking) and 62% (listening) were found by Nielsen (1962), while Argyle and Ingham (1972) found proportions of 41% and 75%, respectively. Although the trend is clear, the ranges reported by Kendon suggest that variability within each category is larger than the variation between categories. By way of example, in the Nielsen (1962) study, the overall proportion of gaze at partner varied between 8% and 73%. Many factors may influence gaze patterns, including sex and language. Ingham (1972) recorded 12.7 glances/min, with an average glance length of 2.93 seconds for 22 pairs of English speakers (half male, half female), but a glance length of 8.2 glances/min with an average length of 5 seconds for a similar pool of Swedish subjects.

Mutual gaze is of particular interest. In nonprimates, mutual gaze is usually confrontational. In chimpanzees and humans (at least) mutual gaze appears to be integrated into natural mother–infant patterns of interaction, where it is nonaversive and comparable in many ways to touch (Bard et al., 2005). There seem to be situations in human interaction in which mutual gaze is avoided (as in embarrassing situations, or in flirting), but others in which it is actively sought (as in public speaking, and, paradoxically, flirting). Mutual gaze is also understood to be important in regulating turn taking in conversation (Argyle, 1988). Periods of mutual gaze are longer in adults than children, longer among female dyads than male, longer when participants are further apart, and vary in duration as a function of culture (Argyle & Cook, 1976).

When in dyadic conversation, speakers often look away at the start of a speaking turn, and look toward their partner at the end of a turn (Duncan, 1972). Different authors interpret the macroscopic structure of conversation in different ways. Analyzing utterances as theme-rheme sequences (Halliday, 1967), Cassell et al. (1999) reported gazing away at the start of a theme 70% of the time, but gazing toward partner at the theme-rheme junction 73% of the time. Gaze has been used as the primary cue used in organising turn-based behaviour in embodied conversational agents (Cassell et al., 1999).

Blinking in conversation

Blinking is an activity that has stubbornly resisted clear functional interpretation. Although blinking serves a self-directed physiological need in wetting the cornea, blink patterns and blink frequency exhibit task-specific variation that demands further explanation. A common index used is the spontaneous eye blink rate (SEBR), which, as Doughty (2001) makes clear, varies greatly depending on the activity of the subject. Doughty distinguishes between three kinds of activity: reading, primary gaze, and conversation. He finds that SEBR lies between 1.4 and 14 blinks/min when reading, between 8 and 21 blinks/min during what he calls primary gaze (staring at a target), and is much increased at 10.5 to 32.5 blinks/min during conversation. Although there is clear variation with task here, the variation within each task is very substantial as well.

Several blink taxonomies have been proposed, though none has emerged as a standard suited to all purposes. Stern and Dunham (1990) distinguished between purely reflexive blinks, which are part of the startle reflex, involuntary blinks, which serve to wet the eye, but which may be inhibited when attentional needs demand, and voluntary blinks, assumed to play an overt communicative role. This taxonomy will not do in the present context, as we wish to remain agnostic about the communicative role of blinks, rather than assuming it before encountering the data. Likewise, a recent study by Herrmann (2010) provided a seven way categorisation based on the position of the blink when compared to the stream of signs. This categorisation scheme also made use of functional and etiological interpretation, such as "[to] wet eyes" or "cognitive trigger," which in the present context would anticipate some of what we hope to instead uncover. Herrmann correctly notes that there are problems with using categories based on such unobservables as "voluntary/involuntary" or "conscious/unconscious."

A further basis for categorisation is based more simply on the observable blink characteristic of blink length. Baker and Padden (1978) found average blink durations to last from 150 to 200 ms. Herrmann (2010) found an average duration of 186 ms, with a "majority" lasting between 160 and 200 ms. Longer blinks were observed, and Herrmann distinguished between long blinks and eye closure, by placing a threshold at 420 ms. Any closure of greater duration was no longer considered a blink.

Although there is little evidence of systematicity to blinking patterns in spoken interaction, blinking has often been argued to bear a relation to prosodic (or sometimes syntactic) structure in signed languages (Herrmann, 2010; Wilbur, 2009). A recent study by Bailly, Raidt, and Elisei (2010) examined gaze and speech in face-to-face interaction between pairs of human subjects, and between humans and a virtual conversational agent. In order to suitably constrain the task, the interaction was rigorously scripted, with clear roles at each point. Thus an unambiguous sequence of conversational states was available, but the complexities of negotiated interaction were absent. Based on analysis of nine subjects, they found that blink rates increased when speaking and were relatively lower when listening, although there was considerable variation among subjects.

Blinks are well known to be strongly associated with gaze change (Evinger et al., 1994), which is a physiological characteristic of vertebrates more generally, with the probability of an associated blink rising as the magnitude of the gaze change increases. Fogarty and Stern (1989) report that blinks are more likely to occur when gaze returns to the center of the visual field than when gaze moves from the center to the periphery. This seems to provide all the more reason to suspect that blink behaviour in negotiated interaction might differ from that found when reciting a monologue, or retelling a story.

This paper presents an empirical study of gaze and blinking in dyadic conversation. Data extracted from carefully annotated video recordings of dyadic conversations will allow an examination of relations between blinking, gazing, and speaking as the conversation ebbs and flows. In contrast to many previous studies, the primary unit of analysis here will be the single individual, as it will transpire that different individuals display markedly different patterns of co-variation in all variables we examine. It is hoped that this relatively straight-forward description will be of use to other researchers, irrespective of theoretical commitments.

A second goal speaks rather to theory. We will attempt to understand the observed variation in blinking and gazing as a function of both the individual and the emergent dyadic conversational domain. The inter-individual differences that are observed must be interpreted together with systematic variation in the same variables as a function of the ongoing coordination across the participants in a conversation. It will transpire that there is a great deal of systematicity to the relations among the three parallel channels of blinks, gaze, and speech, but that much of this systematicity is obscured if individual differences are glossed over, or if quantitative variables are averaged across subjects. Sensitivity to variability across subjects is nothing new in the area of movement analysis, where detailed studies of a few individuals are the norm. The incorporation of this sensitivity into accounts of higher level cognition and social interaction still represents a challenge for cognitive science.

METHODS

Corpus

Eight pairs were selected from the 20 dyadic conversations released as the IFA Dialog Video Corpus produced by the Nederlandse Taalunie (van Son, Wesseling, Sanders, & Heuvel, 2008). Each conversation was conducted by two native Dutch speaking subjects who knew each other well. They were seated on opposite sides of a table, with a video camera positioned to the left and somewhat behind the head of each, thus providing a full frontal view of both faces (Figure 1). Each dyad conversed on topics

Figure 1. Two simultaneous frames from one dyadic recording. [To view this figure in colour, please visit the online version of the journal.]

of their free choosing for 15 minutes. Video was recorded with 720×576 BGR frame resolution at 24-bits, with 25 frames/sec, thus limiting temporal granularity to the 40 ms between individual frames. Participants were entirely blind to the present analytical goals.

The pairs selected were numbers 1–7 and Pair 20 (Table 1). These were selected because they maximise the opportunity to observe the same individual in multiple dyads. Five subjects appeared in two dyads each (Subjects B, C, E, G, and H), while the remaining six took part in one pair only. Although the corpus is distributed with gaze annotation (though not blinks), gaze, speaking turn and blink were all annotated by hand from scratch. This was deemed necessary, as the tight relationship between gaze and blink was an object of particular importance, and consistent annotation of both was thus required.

All recordings were fully annotated for speech turn, gaze, and blinks. Annotation was done in Praat (Boersma & Weenink, 2010), using simultaneous video display in Elan (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Figure 2 shows a small excerpt from an annotation. The top two rows are the two speech waveforms. The next three rows are tiers for annotating blink, speech, and gaze, respectively for Subject A. The final three rows are similar for Subject B. Blinks are notated as punctate events, whereas gaze and speaking turn annotations comprise interval sequences. Principles for annotating each tier were developed and have been published to the web at http://cspeech.ucd.ie/~fred/Blinking, where they can serve both as documentation of the present project, and as a potential tutorial resource for others interested in documenting gaze and blink in conversation.

Speech turn

Speech turn appeals to the pre-theoretical notion of having the floor. In a dyad, it is usually the case that one person has the floor, in that that person is producing propositional speech or is in a position to do so without further negotiation. Speakers frequently overlap, and an overlap is to be understood as either a back channel, providing feedback to the speaker without attempt at interruption, or as turn negotiation, in an attempt to obtain the floor. In a dyad, a period of turn negotiation is usually relatively short. During such a negotiation, both speakers are marked as having turn.

Turn endings are marked in various ways. A direct question that invites a direct answer is an obvious turn end. An obviously final phrase is likewise the end of a turn,

TABLE 1
Subject details. Pair 8 is numbered 20 in the original corpus distribution and Subject K is there labeled AI

Pair(s)	*Subject*	*Sex*	*Age*
1	A	F	62
1/7	B	F	30
2/4	C	M	29
2	D	F	55
3/8	E	M	64
3	F	F	29
4/5	G	F	62
5/6	H	F	59
6	I	M	65
7	J	M	53
8	K	F	20

as is a switch from propositional content to a back channeling behaviour such as "uh-huh" or "yes...." There may be periods in which no one speaker obviously has the floor, though such periods are relatively short and infrequent. Silence does not necessarily imply surrender of turn. If one speaker picks up after a silence with a continuation of an unfinished story or a related thought, we regard turn as continuous throughout, once there is no attempt by the partner to take the floor. In general, turn onsets can be marked with more confidence than turn offsets, but there were relatively few occasions within the 2 hours of recorded conversation when turn was considered difficult to annotate or ambiguous. The training materials for annotators included several examples of turn exchange, negotiation, and back channeling.

Back channels are marked separately. We classify all feedback utterances that are devoid of novel propositional content, including filled pauses, as back channels. Unlike turn, back channels are restricted to vocalised speech. If two back channels are separated by more than 200 ms, they are regarded as two distinct instances. Laughter and respiration noises are not annotated.

Gaze

In marking gaze, a simple binary distinction was made, based on whether the subject was looking at the head of the partner or not. The fixed nature of gaze at a conversational partner makes this annotation relatively straightforward. Blinks often co-occur with gaze shifts (Evinger et al., 1994), and when they did so, the point at which the blink was annotated was identical to the point at which the gaze change was marked. For gaze shift that was not accompanied by a blink, gaze change was noted at the first sign in which gaze left partner, or at the first frame in which gaze was judged to be restored toward partner.

Blinks

Existing blink taxonomies were considered inappropriate for the present study on several counts. No empirical grounds are available for distinguishing between blinks that are conscious or unconscious and that are voluntary or involuntary. No distinction was made between blinks in specific structural position with respect to speech, as we wanted to see whether blinks, in fact, exhibit any systematic relationship to the speech stream. A grammatical analysis was not undertaken. We therefore

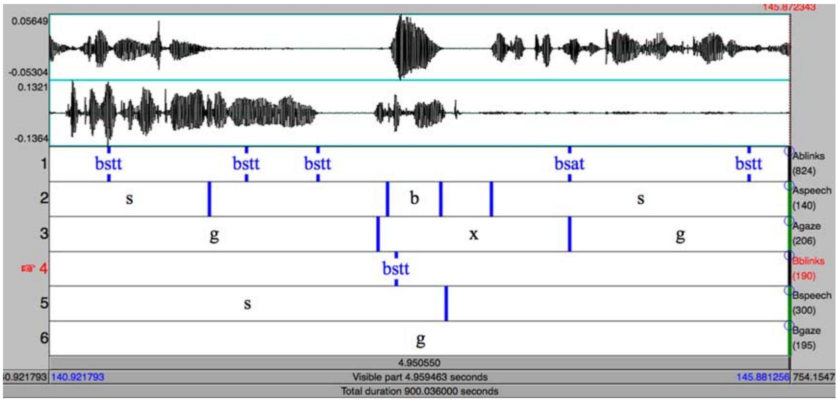

Figure 2. Sample excerpt from an annotated conversation. [To view this figure in colour, please visit the online version of the journal.]

employed a simple taxonomy based on two directly observable characteristics of each blink: the direction of gaze on either side of the blink, and the blink length.

Video materials were recorded at 25 frames/sec, yielding a 40 ms interval between successive frames. The closing phase of most blinks is approximately 50 ms, while the opening phase may be somewhat longer. The moment at which a blink was recorded was the first moment in which the visible part of the cornea was substantially occluded. This will generally be one frame later than the criterion employed by Hermann (2010), where the first movement of the eyelid was taken.

Two length classes were employed. Based on the durations provided earlier, blinks lasting 5 frames or less were considered short, while those lasting from 6 to 10 frames (240–400 ms) were considered long. Eye closure longer than 400 ms was noted at the start and end but was not considered to be a blink.

Eye gaze could be constant before and after the blink, or it could change from one side of the blink to the other. This change could be toward partner, away from partner, or from one nonpartner directed angle to another. These possibilities, crossed with the short/long distinction, yielded 10 possible blink types, as listed in Table 2. Examples of each blink type, both as real time video, and in slow motion to allow precise study, are provided in the training materials for annotators.

Some subjects sometimes displayed blinks in which the upper eyelid failed to meet the lower. For occasional blinks of this nature, the annotator was guided by the simple principle that if the visible surface of the eyeball was substantially occluded ($>50\%$), a blink was recorded.

Annotation verification

All data reported here were analyzed by the author. In order to establish whether the criteria employed were adequately defined and permitted reproduction, three annotators, all graduate students, were each given one hour's training in the annotation procedures, and they then proceeded to provide an annotation of a further 30 seconds of data (40 for one annotator), which took approximately two hours work. Each of them found the criteria to be comprehensible, but complex, and it is to be expected that facility at annotation would increase with further practice. Annotations by each were compared to the annotation originally done by the author.

TABLE 2
Blink classes used. Observations about relative frequency of occurrence are justified below

Gaze before	*Gaze after*	*Length*	*Code*	*Comments*
Toward	Toward	Short	`bstt`	By far the most common blink type
Away	Away (same direction)	Short	`bsaa`	Gazing away with no obvious change of direction
Toward	Away	Short	`bsta`	Gaze switch and blink simultaneous
Away	Toward	Short	`bsat`	Gaze switch and blink simultaneous
Away	Away (changed)	Short	`bsab`	Change of gaze direction, but not toward partner
Toward	Toward	Long	`bltt`	Often co-occurs with head nod or back channel
Away	Away	Long	`blaa`	Rare
Toward	Away	Long	`blta`	Rare
Away	Toward	Long	`blat`	Rare
Away	Away (changed)	Long	`blab`	Very rare

In what follows, annotators will be referred to as A1, A2, and A3. A3 completed 40 seconds.

Blinks

The reference annotation by the author recognised 34 blinks in the 30 second excerpt (45 in the 40 second one). A1 and A2 each recognised 32 of these 34, with no additional events. A3 recognised 43 of the 45, and annotated one additional event (a potential blink at a point of gaze change). There was agreement about the short/long distinction in 29 of 32 blinks (A1), in 31 of 32 (A2), and in 42 of 43 (A3). There was agreement in classifying blink type based on gaze in 31 of 32 blinks (A1), in 30 of 32 (A2), and 39 of 43 blinks (A3). As the inter-frame interval is 40 ms, temporal discrepancies of 20 ms or less indicate absolute agreement, discrepancies between 21 and 60 ms indicate one frame disagreement, etc. For both A1 and A2, there was no more than a single frame disagreement for all blinks. For A3, there was a 2-frame discrepancy for five blinks, and a 3-frame difference for one blink.

These results suggest that blink type, class, and temporal location are reliably annotated using the established procedures.

Gaze

For each annotator, the total time during which gaze annotation differed from the reference annotation was summed up. As annotation is done for each subject separately, this represents a 60 (A1, A2) or 80 (A3) second period. A1 disagreed about gaze for 0.23 seconds (0.4% of the total), A2 disagreed for 1.69 seconds (2.8%), and A3 disagreed for 1.7 seconds (2.1%).

Speech turn

A similar summation of disagreement about speech turn annotation provided the following results: A1 disagreed for 0.22 seconds (0.4%), A2 for 0.09 seconds (0.15%), and A3 for 2.94 seconds (3.7%).

Collectively, these results provide adequate assurance that the annotation criteria and practices are trustworthy and reproducible. Full details of the annotation procedures, together with video and audio examples of all blink types and potential ambiguities, are provided at http://cspeech.ucd.ie/~fred/Blinking.

RESULTS

In what follows, we present a series of data summaries that serve to capture both individual differences and group patterns in the co-occurrence of blinks, gaze, and speech turn. In general, we do not make heavy use of significance tests, but opt instead to let the data speak for themselves. In many cases, significance tests would, in any event, be problematical.

Blinks alone

Appendix 1 provides raw counts of the blinks observed in each category. From that it can be seen that the vast majority of blinks fall into the short category (*N* long: 296; *N* short: 7,309). The total number of blinks ranges from 812 (Subject A) or 0.9 blinks/sec to 164 (B) or 0.18 blinks/sec. Blink type `bstt` is by far the most common, making up half the observed blinks overall.

Figure 3 shows the proportion of short blinks falling into each of the five categories. Subjects B, C, E, G, and H are each represented twice, and data are arranged to facilitate intra-subject comparison. It is immediately apparent that subjects differ greatly in the proportion of each kind of blink they produce. Although fully half of all blinks are of type `bstt`, and this is the most frequent type produced for most subjects, there are notable exceptions in Subjects I and J, for whom types `bsaa` and `bsta` respectively are more common. The five subjects who provide data from two dyads each display consistency across conversational partner. It might be noted that Subject B, who produces least blinks overall (164 in Pair 1), is paired first with the most prolific blinker, A, (812 blinks), and then (Pair 7) with the subject who blinks least after herself (Subject J, 175 blinks in total). The consistency she exhibits is thus clearly robust despite the largest possible change in dyadic constitution, and so it seems that blink frequency and the distribution of blinks by type must thus be regarded primarily as a subject-specific attribute, rather than a product of the dyad.

Blink types `bstt` and `bsaa` are each characterised by unchanging gaze across the blink, and they differ in that the former occurs while looking at the partner and the latter while looking away. The ratio of `bstt` to `bsaa` varies greatly. Subjects C, G, H, A, and F all exhibit far more `bstt` than `bsaa`, while B, D, I, J, and K have `bstt/bsaa` ratios closer to unity. Only E seems to fall somewhere between the two camps, as he exhibits somewhat different ratios in the two dyads in which he partakes.

A second notable feature is the ratio of `bsta` to `bsat` blinks. These are the blinks that coincide with gaze shift away from or toward the partner respectively. Subjects E, G, H, A, F, and K all have ratios less than unity, while B, D, I, and J have ratios exceeding 1. Subject C is near unity both times.

For each speaker, a vector was constructed of blink counts in the four most prevalent categories (`bstt`, `bsaa`, `bsat`, `bsta`), on which a principal component analysis was then performed. The first two principal components accounted for 89.4% of the variance in the data set, and individual subjects are shown in Figure 4. The relative invariance of blink type distribution within an individual is readily apparent.

Looking only at the long blinks (240–400 ms), we see that these blinks are far less frequent (Appendix 1). The only types of long blinks that occurred with any regularity were of type `bltt` and `blta`, and both were largely restricted to two subjects, G and H. It seems therefore that a first pass understanding of blinks in conversation can safely restrict its attention to blinks of 200 ms duration or less.

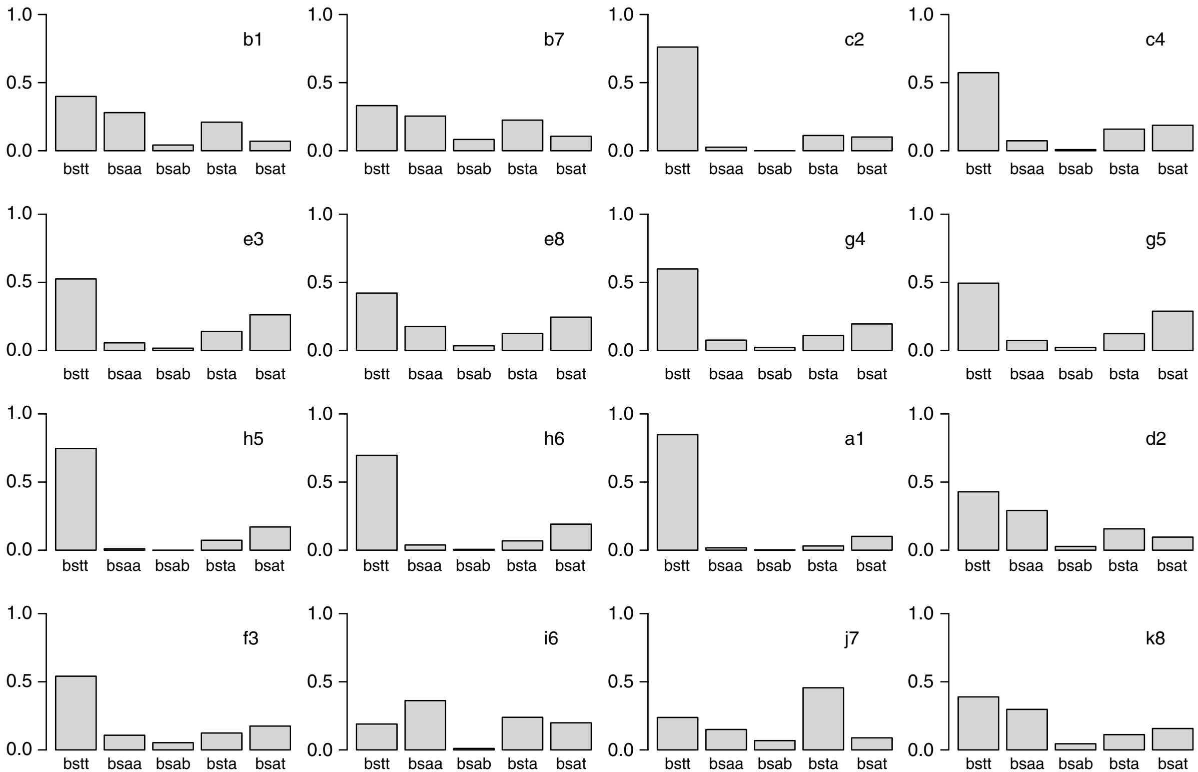

Figure 3. Proportion of short blinks falling into each type. Legends on the individual panels give subject ID and pair number. Blink codes are as in Table 2.

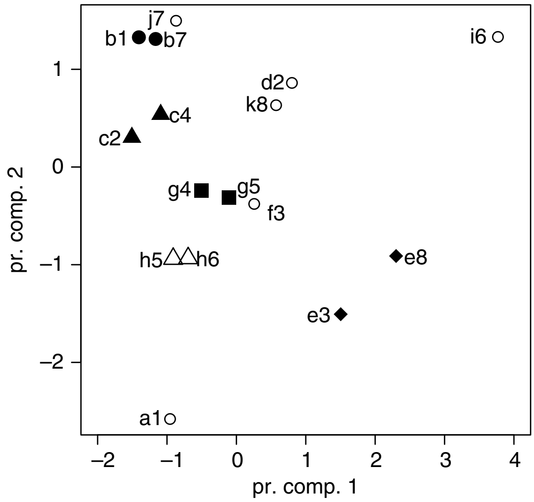

Figure 4. First two principal components of vectors based on counts of the four blink types `bstt`, `bsaa`, `bsat`, and `bsta`.

This preliminary presentation of blink frequency by type serves to illustrate several themes that will be developed here. Subjects differ greatly, but we can look more closely at this variation, with a particular focus on consistency within subject across dyads, and its counterpart, which is dyadic specific effects on both participants. Variation across subjects might be unbounded, but it might also be that distinct patterns, representative of interactional styles, emerge.

Gaze alone

The proportion of time each subject gazed at partner was calculated, along with the proportion spent in mutual gaze. Time gazing at partner ranged from 59% to 95%, while the time spent in mutual gaze ranged from 36% to 71%. With only four male subjects, a direct male–female comparison is not meaningful here, but we note in passing that all four extreme values just quoted came from male subjects. For subjects who were recorded in two dyads, variation in proportion of mutual gaze across dyad was found for Subject B (0.68 in Pair 1, 0.39 in Pair 7), while the other four were relatively consistent across different partners. Proportions for all subjects are provided in Table 3.

The distribution of the duration of episodes of mutual gaze were examined. All distributions were skewed heavily right, as is typical for duration distributions, and so they were log-transformed. Figure 5 shows the distributions. It should be borne in mind that because mutual gaze is identically defined for each subject in a specific dyad, each distribution appears twice here. In several dyads (most notably pairs 3, 4, and 7), in addition to the well-defined central mode, there may be a second mode at the shortest intervals. It might be interesting to look for evidence of a preponderance of fleeting moments of mutual gaze in future data sets, but the present data do not warrant strong conclusions here.

Blinks and gaze

In examining the possible relation between blink behaviour and direction of gaze, we exclude the substantial proportion of blinks that occur at points of gaze change.

TABLE 3
Proportion of time spent gazing at partner and proportion of time spend in mutual gaze

Subject	*Pair*	*p (Gaze)*	*p (Mutual)*
B	1	0.82	0.68
B	7	0.68	0.39
C	2	0.95	0.71
C	4	0.88	0.71
E	3	0.65	0.44
E	8	0.60	0.36
G	4	0.80	0.71
G	5	0.77	0.67
H	5	0.87	0.67
H	6	0.81	0.55
A	1	0.86	0.68
D	2	0.76	0.71
F	3	0.74	0.44
I	6	0.66	0.55
J	7	0.59	0.39
K	8	0.59	0.36

Because gaze and blink were annotated together, this is readily done, and does not require the setting of any arbitrary threshold. In what follows, we further restrict our analysis to blinks of type `bstt` and `bsaa` only. The analysis is unchanged in essence if the few long blinks and the even fewer blinks of type `bsab`/`blab` are included.

For each subject, we calculated blink rate as a function of gaze by counting blinks in each of the four joint gaze states and dividing by the amount of time spent in that state. We notate gaze using G to indicate subject gazing at partner, and X for gaze directed elsewhere, so GG represents mutual gaze, etc. Figure 6 shows blink rates for each of the four states separately. Figures above each bar give the number of seconds spent in that state, because, for example, far more time is spent in state GG than XX.

Firstly, it is abundantly clear that for many subjects, blink rate varies systematically with gaze, both their own, and that of their partner. Subjects A, H, and E (in Pair 3) blink almost only when they are looking at partner. Subjects B, D, I, and K, in contrast, blink more when they are not looking at partner. A clear distinction between blink rates in states GG and GX is evident for Subjects E, H, A, D, F, and I: In every case, the subject blinks more when the partner is not looking at them. Interpreting blink rates for state XX is problematic, as comparatively little time is spent in this state for some subjects. The five subjects for whom two data sets each are present exhibit considerable consistency across dyads for the most prevalent categories of GG, GX, and XG.

We can also selectively examine blinks that occur together with gaze changes. We restrict our analysis to short blinks, and ignore blinks of type `bsab`, both because they are numerically far fewer than types `bsat` and `bsta`, and also because blinks that guide gaze toward or away from the partner (`bsat`, `bsta`) may be presumed to have a special significance in conversation. From prior work (Fogarty & Stern, 1989), more blinks of type `bsat` are expected overall than blinks of type `bsta`. Figure 7 shows the probability of a blink accompanying a gaze shift both away from and toward partner.

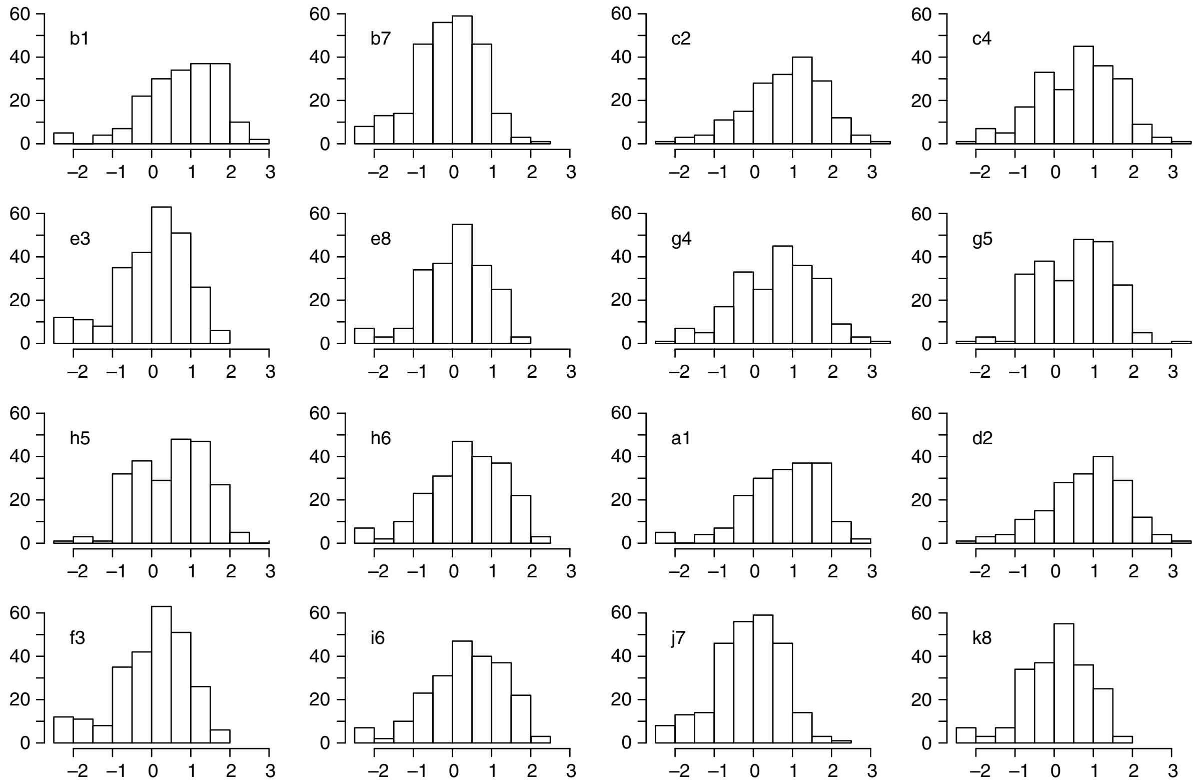

Figure 5. Distribution of log transformed durations of mutual gaze.

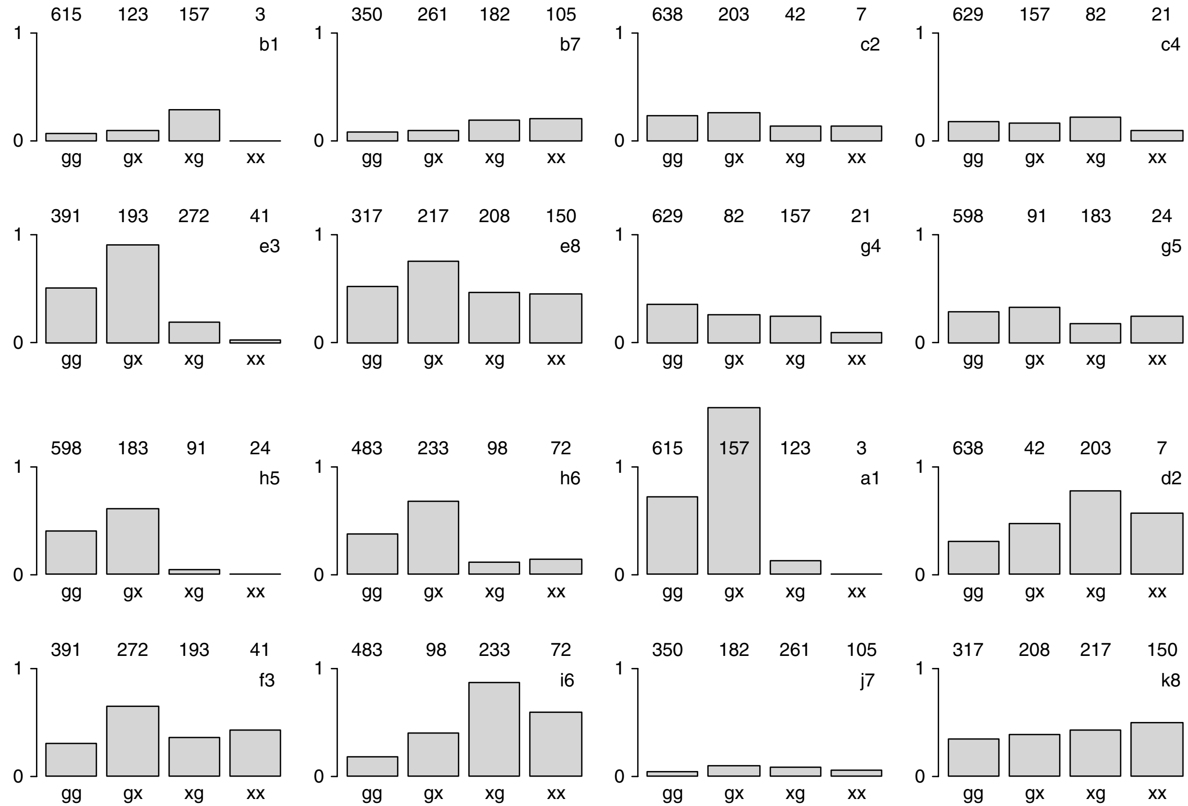

Figure 6. Blink rates (blinks/sec) as a function of joint gaze state. G = gaze toward partner, X = gaze away from partner. Thus GX for Subject K means K is looking at partner, but partner is looking elsewhere. Numbers above each bar give the total amount of time in seconds spent in that state. Only blink types `bstt` and `bsaa` are included in this analysis.

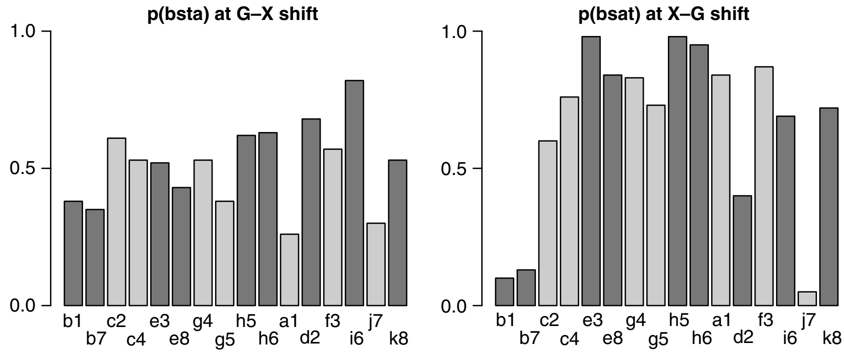

Figure 7. Probability of a blink accompanying a gaze shift away from the partner (left) or toward partner (right).

The simple prediction of a greater probability for shifts toward partner (blink type `bsat`) is confirmed, with significant individual differences. Subjects B (twice), D, I, and J display an effect in the opposite direction. This is particularly interesting in the case of B, who we have seen takes part in two rather different conversations. Nevertheless, the pattern of blink–gaze association remains constant across the change in partner.

Some of the variation appears almost categorical. Subjects B and J almost never blink on returning their gaze to the partner, while Subjects E and H almost always do so. Variation of this sort is important, as it should caution against any kind of averaging across subjects. Statements linking blink rate to specific behaviours are rife in the literature, but rarely acknowledge the degree of inter-individual variation seen here.

Speech turn and gaze

There is little to say about speech turn alone, except to note that the conversations are all reasonably well balanced, with the proportion of the entire time during which a given subject has turn ranging from 0.37 to 0.71, $M = 0.52$.

We noted before that gaze while listening is likely to be greater than gaze while speaking, with significant individual variation (Argyle & Cook, 1976). Listening is not a recorded state in our data, but we can look at gaze as a function of speaking turn. Figure 8 shows the proportion of the time during which a speaker gazes at partner when he/she has turn or when partner has turn, along with the proportion spent gazing at partner overall. Gaze is, indeed, more likely to be partner-directed when partner has turn, and once more, there is great variation. Speaker I (and to a lesser extent, C) is unaffected by speaking turn, whereas E (and perhaps B) displays huge sensitivity to who has turn.

The literature is unclear about the degree to which gaze switch is tied to specific points in the ebb and flow of speech. In our data set it is possible to look at the probability of a gaze switch lying within some fixed interval around the onset or offset of a turn. Turn onsets are recorded with considerably more certainty than turn offsets, so any systematicity is more likely to be found at onsets. We employed a threshold of 100 ms either side of the onset of a speech turn, and counted the number of G–X and X–G gaze switches (i.e., gaze change away from or toward partner respectively) that lay within that window. In order to test whether there was any evidence of an association between turn onset and gaze switch, a Monte Carlo sampling method was

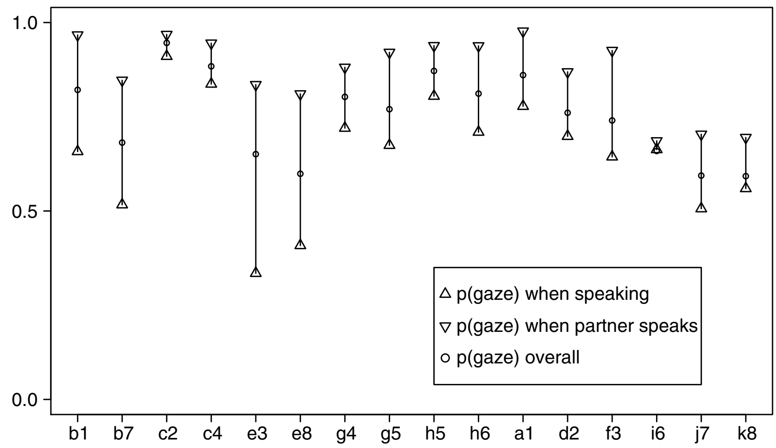

Figure 8. Proportion of time gazing at partner as a function of speaking turn, and overall.

used, whereby a number of points equal to the number of turn onsets was randomly selected from a uniform distribution over the 15 minutes of conversation, and the number of gaze switches that lay within similarly defined windows centered at these random points was calculated. This was repeated 10,000 times, and the proportion of such cases in which the number of gaze switches was greater than or equal to that actually observed at turn onsets was noted.

Table 4 provides the likelihood that the observed number of gaze changes occurring around turn onsets arose by chance. There is clear evidence here, despite some variation, that the start of a turn is associated with a heightened probability of a gaze change away from partner (G–X). This accords well with loosely comparable data

TABLE 4
Association of gaze change and blinks with speech turn onsets. Estimated likelihood of finding this many or more gaze changes or blinks at this many events is provided

Subject	*Pair*	*n Onsets*	*n(G–X)*	*p(G–X)*	n*(X–G)*	*p(X–G)*	*n Blinks*	*p Blink*
B	1	53	11	***	0	–	5	*
B	7	58	8	***	2	–	6	*
C	2	89	9	***	5	**	19	***
C	4	93	4	–	1	–	12	**
E	3	33	5	*	0	–	4	–
E	8	65	14	***	0	–	8	–
G	4	107	9	**	2	–	11	–
G	5	74	6	–	6	–	15	**
H	5	76	2	–	3	–	13	–
H	6	80	5	*	2	–	13	–
A	1	54	1	–	2	–	5	–
D	2	81	10	***	5	–	13	–
F	3	40	6	**	0	–	6	–
I	6	83	9	**	11	***	28	***
J	7	65	8	*	2	–	2	–
K	8	66	1	–	5	*	17	**

Note: X–G = gaze changes toward partner; G–X = gaze changes away from partner. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

from Duncan (1972) and Cassell et al. (1999). Furthermore, this does not seem to be a stable and reliable characteristic of an individual speaker, as three of the five subjects that are represented twice in the data set are inconsistent in this regard. Gaze change at speech turn onset may thus be a dynamic feature of a specific conversational situation.

Speech turn and blinks

A similar analysis of the co-occurrence of blinks with speech turn onset did not yield a clear pattern of results. Subjects C, I, and K displayed a strong association between turn onset and blinking. Subject B showed a weak association, while Subject G was inconsistent in this regard across dyads. These data are also provided in Table 4.

We can also examine the variation in blink rate as a function of speaking turn (Figure 9). As with gaze, we can identify four qualitatively different states: SS (both have turn), SO (subject has turn), OS (partner has turn), and OO (no turn evident). In this instance, most time is spent in the two states SO and OS. A pattern whereby subjects blink substantially more when partner has turn than when they themselves do is evident for Subjects E, H, A, and F, but is not universally present. There is considerable consistency in patterning within subject and across dyads, suggesting that the association of blinking and speaking turn is an individual, rather than a dyadic, characteristic.

Back channels were annotated separately from speaking turn. A total of 1,206 back channels were observed. In order to examine the possibility of a co-dependence of back channel behaviour and gaze, the proportion of back channels in each of the four gaze states (GG, GX, XG, and XX) was computed for each subject in each dyad, and this proportion was scaled by dividing it by the proportion of time the subject spent in that joint gaze state. In this way, if back channels bear no relation to gaze state, the expected value of the ratio is 1.0. Higher values are indicative of a tendency for back channels to preferentially occur in that state. Table 5 provides figures for each subject. Overall, for all subjects, back channels occur preferentially during periods of mutual gaze, as all ratios for all subjects exceed 1.0, while very few ratios in the other three joint gaze states exceed unity.

As with speech turn onset, we can ask whether blinks are more likely than chance to occur near a back channel event. Once more, we count the number of blinks that occur within 0.1 second on either side of a back channel onset. This is compared to the number of times out of 10,000 that an equal or greater number of blinks is found within a similar window centered around random events that are equal in number to the number of back channels. In Table 6 it can be seen that a strong association of blinks to back channels is occasionally found, but that this is not a stable characteristic of a specific speaker, as Subjects C and E each display a strong association once and no such association in their other dyad.

In a highly redundant signaling system such as spoken communication, it is possible that blinking and back channeling serve similar roles in providing feedback from listener to speaker. If so, we might expect to find a trade-off between the number of blinks produced (or the blink rate) and the number of back channels. No such correlation was found, however. Overall, the correlation between number of back channels and total number of blinks in 900 seconds was 0.01, suggesting that these two variables are entirely independent of one another.

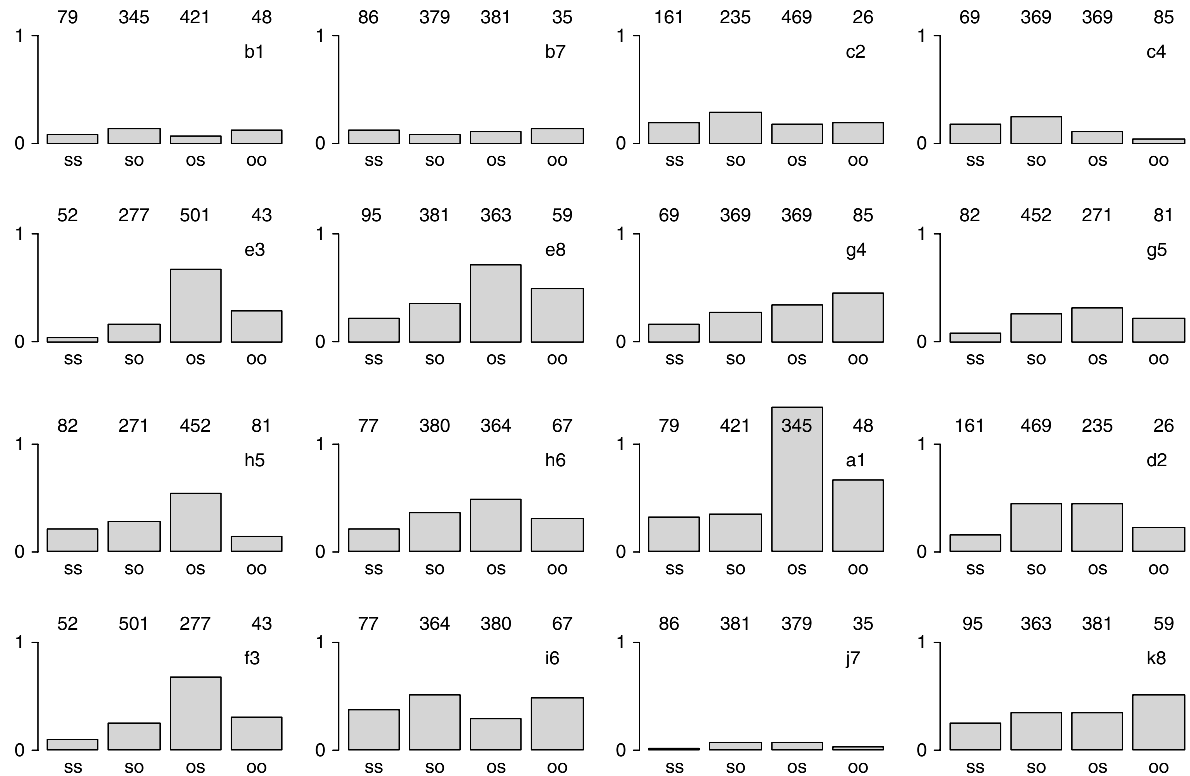

Figure 9. Blink rate (blinks/sec) as a function of speaking turn. SS, both have turn; SO, Subject only has turn; OS, partner only has turn; OO, no turn evident. Numbers above bars give the total amount of time spent in each state.

TABLE 5
Back channels as a function of gaze. The number of back channels are reported, along with the proportion of back channels in each of the four gaze states, divided by the proportion of the total time spent in that state

Subject	*Pair*	*No. back channels*	*p(BC|GG)/ p(GG)*	*p(BC|GX)/ p(GX)*	*p(BC|XG)/ p(XG)*	*p(BC|XX)/ p(XX)*
B	1	96	1.32	0.43	0.22	–
B	7	75	2.03	0.31	0.40	0.33
C	2	106	1.06	0.87	0.6	2.0
C	4	53	1.11	0.53	1.0	1.0
E	3	157	1.43	0.29	0.90	0.60
E	8	95	1.50	0.54	0.78	0.94
G	4	45	1.18	0.22	0.76	0
G	5	57	1.15	0.90	0.70	0
H	5	114	1.13	1.10	0.10	0.33
H	6	86	1.35	0.77	0.09	0.63
A	1	13	1.13	0.83	0.57	–
D	2	49	1.24	0.40	0.43	0
F	3	40	1.36	0.93	0.38	1.0
I	6	42	1.25	0.45	0.81	0.63
J	7	36	1.28	0.85	1.07	0.25
K	8	85	1.22	0.78	0.63	1.41

TABLE 6
Association of back channels and blinks. *p*-values refer the likelihood of finding this many or more blinks within 0.1 second on either side of the onset of a back channel

Subject	*Pair*	*n Back channels*	*n Blinks*	*p*
B	1	96	9	**
B	7	75	4	*ns*
C	2	106	20	***
C	4	53	3	*ns*
E	3	157	25	*ns*
E	8	95	35	***
G	4	45	6	*ns*
G	5	57	6	*ns*
H	5	114	5	*ns*
H	6	86	10	*ns*
A	1	13	3	*ns*
D	2	49	10	*ns*
F	3	40	10	*
I	6	42	16	***
J	7	36	1	*ns*
K	8	85	17	*

*$p < .05$; **$p < .01$; ***$p < .001$.

DISCUSSION

The series of empirical results presented here make it abundantly clear that blinks and gaze are indeed integrated into the structure of conversational interaction. Individuals display systematic modulation of blink and gaze behaviour as a function of the joint state of the conversation, but this systematicity varies greatly between subjects. A brief review of the principal findings seems to be in order, before a discussion of the consequences for our understanding of the nature of conversational interaction.

A novel blink taxonomy was proposed, based on two objective characteristics of the blink: its length and the direction of gaze on either side of the eye closure. The length distinction is of interest only in establishing that long blinks are relatively rare, and so infrequent as to play no significant part in understanding how blinks co-vary with gaze and speech. The distinction based on the direction of gaze revealed a great deal of information about individual characteristics of coordinative behaviour. Although five blink classes based on gaze were identified, only those four that make reference to the person-directed nature of conversational gaze were important. The fifth class, `bsab/blab`, was very rare and otherwise uninformative.

The four remaining blink classes of note divide into holding blinks (`bstt`, `bsaa`) in which the direction of gaze is constant, and blinks that co-occur with gaze changes (`bsat`, `bsta`). These classes seem to be well motivated, as blink rates in each class differ greatly across subjects. Furthermore, the relative frequency of each of these four classes appears to be a stable characteristic of an individual, and relatively unaffected by the dyad in which the subject participates. Individual-specific patterns of covariation with joint gaze state were found for types `bstt` and `bsaa` (Figure 6). Likewise, blinks at gaze turns (types `bsat/bsta`) vary systematically within individuals, but with wide variation between individuals (Figure 7). Blinking behaviour is thus inseparable from gaze, while the way in which the two forms of behaviour are linked serves to characterise an individual communicative style.

Mutual gaze is necessarily a feature of a given dyad, although the propensity of the individual to engage in mutual gaze will greatly affect the result. The distribution of mutual gaze durations hinted at the possibility of a distinct behaviour of fleeting glances, resulting in elevated counts at the shortest intervals, but further observations will be required to investigate this rigorously.

Further hallmarks of individual style in the mutual relations among gaze, blink, and speech turn were found in the differential probability of gaze toward partner as a function of who is speaking (Figure 8) and in the dependence of blink rate on speaking turn (Figure 9). In both these cases, individuals seemed to be relatively unaffected by the specific dyad. In contrast to these features that serve to characterise an individual, there was less stability in the association between turn onsets and gaze shifts or blinks (Table 4) and in the degree to which back channels were associated with blinks (Table 6). These associations were sometimes very strong but were found to vary within an individual across dyads.

We thus have an emerging picture of gaze and blinking behaviour in which individuals each bring a set of propensities to a given conversational situation, but the manifestation of those propensities is then modulated by the ebb and flow of the conversation. Individuals may be frequent or infrequent blinkers, but that base rate will still be influenced by the joint speaking and gaze state obtaining among conversational participants. The relative frequency of blinks when looking at the partner, when changing gaze direction, or when looking away characterises a given individual, but the dynamics of gaze that makes this apparent emerges from the joint gazing behaviour of the participants. Finally, the degree to which blinks become associated with specific speech-based events (turn onsets, back channels) varies among dyads, and is not a stable characteristic of a single individual.

Collectively, these findings help to understand why functional interpretations of blink and gaze behaviour have been relatively tentative and have consistently revealed large amounts of variation across individuals. Even when tasks have been highly constrained, as in the structured interactions found in Bailly et al. (2010), claimed associations between blink rate and "speaking" or "listening," or between gaze patterns and interaction state, need to be tempered with the acknowledgment of a great deal of individual variability, even to the point at which two subjects may display effects of comparable magnitude but opposite sign. There were several examples of this in our data, for example, when we compare the association between blink and gaze shift for Subjects A and B (Figure 7).

If we find that we cannot assign specific functions to particular kinds of behaviour, and yet that behaviour is found to vary with great systematicity within an individual as she partakes in conversational interaction, an alternative account is required that can help to understand the systematic variation in behaviour as an integral part of the conversational situation. The interpretation of a conversation as an interpersonal coordinative domain, or coordinative structure, as suggested by Shockley et al. (2003) represents a potential starting point. If we take this approach to understanding a conversation between two people, we might hypothesise that the heart of the conversation lies in the establishment and maintenance of a novel inter-personal domain: the conversation. Now real, fluid and engaged human conversation is far too complex that we might ever hope to capture every detail of its ebb and flow using a few differential equations. However, we can still look for the hallmarks of coordination by seeking out observable quantities that exhibit co-varying relationships as the conversation progresses. We find these, not only in the words exchanged, but in the co-varying characteristics of gaze, and even blinking. Gaze and blinking go on even when the conversation is not taking place, but during conversational interaction, they become sucked into the interpersonal coordinative domain. Because neither is strictly necessary in order for a conversation to take place (we can talk just fine on the telephone), their involvement in the domain is not obligatory, and so we find highly individual patterns of recruitment. A cursory glance at Figure 6 is enough to suggest that blinking has been captured by the ongoing interaction for Subject A, whose blink rate is exquisitely dependent on joint gaze state, but blinking has hardly been affected at all for Subject J, whose blink rate is largely independent of gaze.

Of course this account of the nature of conversation is an unfinished sketch. It is of potential interest, however, as it suggests a way of coming at coordinated behaviour of multiple individuals that acknowledges their individuality while simultaneously recognising that some degree of autonomy is surrendered when we enter into situations of joint coordination. These degrees of freedom are not lost, rather they are now located within a transient system that extends across two individuals. This accords well with the relatively recent enactive treatment of social interaction:

> Social interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter's scope can be augmented or reduced). (De Jaegher & Di Paolo, 2007, p. 493)

Evidence for the establishment, maintenance, and ultimately dissolution of a transient, yet autonomous, domain of organisation must take the form of measurable characteristics of participating systems that exhibit co-variation in time that is best understood (i.e., that behaves lawfully) with respect to the joint domain itself. The dependence of blinking and gaze on the joint coordinative state of the participants in the present study may contribute to this emerging picture.

On a more practical note, it suggests that modeling of human behaviour will demand that we model individuals, rather than average types. An embodied conversational agent built on the basis of aggregate data from multiple individuals will look like nobody at all, but one built with an idiosyncratic individual style may ultimately leap across the uncanny valley and into the domain of apparent naturalness.

Manuscript received 7 December 2010
Revised manuscript received 11 August 2011
First published online month/year

REFERENCES

Argyle, M. (1988). *Bodily communication* (2nd ed.). London, UK: Taylor & Francis.

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

Argyle, M., & Ingham, R. (1972). Gaze, mutual gaze, and proximity. *Semiotica*, *6*(1), 32–49.

Bailly, G., Raidt, S., & Elisei, F. (2010). Gaze, conversational agents and face-to-face communication. *Speech Communication*, *52*(6), 598–612.

Baker, C., & Padden, C. (1978). Focusing on the nonmanual components of American Sign Language. In *Understanding language through sign language research* (pp. 27–57). New York, NY: Academic Press.

Bard, K., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A., & Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees (*Pan troglodytes*). *Developmental Psychology*, *41*(4), 616–624.

Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer (version 5.2)* [computer program]. Retrieved November 1, 2010, from http://www.praat.org/

Cassell, J., et al. (2000). Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, & E. F. Churchill (Eds.), *Embodied conversational agents* (pp. 1–27). Cambridge, MA: MIT Press.

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., et al.. (1994). Animated conversation: Rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In D. Schweizer, A. Glassner, & M. Keeler (Eds.), *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques* (pp. 413–420). ACM, Orlando, FL, USA.

Cassell, J., Torres, O., & Prevost, S. (1999). Turn taking vs. discourse structure: How best to model multimodal conversation. In Y. Wilks (Ed.), *Machine conversations* (pp. 143–154). The Hague, The Netherlands: Kluwer.

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, *6*(4), 485–507.

Doughty, M. (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 & Vision Science*, *78*(10), 712.

Drew, G. C. (1951). Variations in reflex blink-rate during visual-motor tasks. *Quarterly Journal of Experimental Psychology*, *3*(2), 73–88.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, *23*(2), 283–292.

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

Fogarty, C., & Stern, J. A. (1989). Eye movements and blinks: Their relationship to higher cognitive processes. *International Journal of Psychophysiology*, *8*(1), 35–42.

Fukuda, K. (1994). Analysis of eyeblink activity during discriminative tasks. *Perceptual and Motor Skills*, *79*(3 Pt 2), 1599.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, *3*(11), 419–429.

Grant, E. (1969). Human facial expression. *Man*, *4*(4), 525–692.

Halliday, M. (1967). *Intonation and grammar in British English*. Berlin, Germany: De Gruyter Mouton.

Herrmann, A. (2010). The interaction of eye blinks and other prosodic cues in German Sign Language. *Sign Language & Linguistics*, *13*(1), 3–39.

Ingham, R. (1972). *Cross-cultural differences in social behaviour*. Oxford, UK: Oxford University Press.

Kelso, J. A. S. (1994). Elementary coordination dynamics. In S. Swinnen, H. Heuer, J. Massion, & P. Casaer (Eds.), *Interlimb coordination: Neural, dynamical and cognitive constraints* (Ch. 15, pp. 301–318). New York: Academic Press.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*(1), 22.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, *27*, 175–200.

Khullar, S., & Badler, N. (2001). Where to look? Automating attending behaviours of virtual human characters. *Autonomous Agents and Multi-Agent Systems*, *4*(1), 9–23.

Kleinke, C. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*(1), 78–100.

Krauss, R., Dushay, R., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gesture. *Journal of Experimental Social Psychology*, *31*(6), 533–552.

Lee, S., Badler, J., & Badler, N. (2002). Eyes alive. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 637–644). ACM, San Diego, CA, USA.

Leonard, T., & Cummins, F. (2010). The temporal relation between beat gestures and speech. *Language and Cognitive Processes, 26,* 1295–1309.

Lester, J., Towns, S., Callaway, C., Voerman, J., & FitzGerald, P. (2000). Deictic and emotive communication in animated pedagogical agents. J. Cassell, J. Sullivan, S. Prevost, & E. F. Churchill (Eds.), *Embodied conversational agents* (pp. 123–154). Cambridge, MA: MIT Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Nielsen, G. (1962). *Studies in self confrontation*. Copenhagen, Denmark: Munksgaard.

Orchard, L. N., & Stern, J. A. (1991). Blinks as an index of cognitive activity during reading. *Integrative Physiological and Behavioral Science*, *26*(2), 108–116.

Pelachaud, C., & Bilvi, M. 2003. Modelling gaze behaviour for conversational agents. In *IVA'03 International Working Conference on Intelligent Virtual Agents* (LNAI 2792), Germany, September 15–17.

Poggi, I., Pelachaud, C., & De Rosisc, F. (2000). Eye communication in a conversational 3D synthetic agent. *AI Communications*, *13*(3), 169–181.

Quek, F., McNeill, D., Bryll, R., Kirbas, C., Arslan, H., McCullough, K., et al. (2000). Gesture, speech, and gaze cues for discourse segmentation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)* (vol. 2, pp. 247–254). Hilton Head Island, South Carolina, June 13–15.

Richardson, D., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045–1060.

Richardson, D., Dale, R., & Kirkham, N. (2007). The art of conversation is coordination. *Psychological Science*, *18*(5), 407.

Shockley, K., Richardson, D., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*(2), 305–319.

Shockley, K., Santana, M., & Fowler, C. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology, Human Perception and Performance*, *29*(2), 326.

Stern, J.A. & Dunham, D.N. (1990). The ocular system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology: Physical, social, and inferential elements* (pp. 513–553). Berkeley, CA: University of California Press.

Thórisson, K. (2002). Natural turn-taking needs no manual: Computational theory and model, from perception to action. In B. Granström, D. House, & I. Karlsson (Eds.), *Multimodality in language and speech systems* (pp. 173–207). Dordrecht, The Netherlands: Kluwer Academic Publishers.

van Son, R., Wesseling, W., Sanders, E., & Heuvel, H. (2008). The IFADV corpus: A free dialog video corpus. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis, & D. Tapias, (Eds.), *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)* (pp. 1–8), Marrakech, Morocco.

Vertegaal, R., Slagter, R., van der Veer, G., & Nijholt, A. (2001). Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (p. 308). ACM, Seattle, WA, USA.

Wachsmuth, I. (1999). Communicative rhythm in gesture and speech. In A. Braffort, R. Gherbi, S. Gibet, D. Teil & J. Richardson (Eds.), *Gesture-Based Communication in Human-Computer Interaction. Vol. 1739 of Lecture Notes in Computer Science* (Ch. 25, pp. 277–289). Berlin, Germany: Springer.

Waters, K., Rehg, J., Loughlin, M., Kang, S., & Terzopoulos, D. (1998). Visual sensing of humans for active public interfaces. In R. Cipolla & A. P. Pentland (Eds.), *Computer vision for human–machine interaction* (pp. 83–96). Cambridge, UK: Cambridge University Press..

Wilbur, R. (2009). Effects of varying rate of signing on ASL manual signs and nonmanual markers. *Language and Speech*, *52*(2–3), 245.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). Elan: A professional framework for multimodality research. In *Proceedings of Language Resources and Evaluation Conference (LREC)*, Genoa, Italy.

APPENDIX 1
Blink counts

Subject	*Pair*	bstt	bltt	bsaa	blaa	bsab	blab	bsta	blta	bsat	blat	*n (short)*	*n (long)*	*Total*
B	1	57	14	40	1	6	0	30	4	10	2	143	21	164
B	7	56	6	43	0	14	0	38	9	18	1	169	16	185
C	2	204	18	7	0	0	0	30	0	27	2	268	20	288
C	4	141	8	18	0	2	0	39	0	46	5	246	13	259
E	3	375	0	40	0	12	0	100	0	187	0	714	0	714
E	8	329	0	137	0	27	0	97	0	191	0	781	0	781
G	4	252	28	32	0	9	1	46	10	82	10	421	49	470
G	5	204	8	30	4	9	0	51	12	119	6	413	30	443
H	5	359	10	5	0	0	0	35	13	82	0	481	23	504
H	6	346	27	19	0	3	0	34	22	95	2	497	51	548
A	1	686	2	14	1	2	0	25	0	82	0	809	3	812
D	2	219	11	149	2	14	0	80	0	49	0	511	13	524
F	3	297	2	59	0	29	0	68	2	96	2	549	6	555
I	6	127	1	242	4	7	2	160	0	133	1	669	8	677
J	7	35	3	22	8	10	4	67	13	13	0	147	28	175
K	8	191	0	146	1	22	2	55	8	77	4	491	15	506
Totals		3,878	138	1,003	21	166	9	955	93	1,307	35	7,309	296	7,605