On Eye-gaze and Turn-taking

Kristiina Jokinen[1] Masafumi Nishida Seiichi Yamamoto

Doshisha University
Department of Information Systems Design
1-3 Miyakodani, Tatara, Kyoutanabe-shi,
Kyoto 610-0321, Japan
+81-774-65-6983
[kjokinen | mnishida | seyamamo]@mail.doshisha.ac.jp

ABSTRACT

In this paper we describe our eye-tracking data collection and preliminary experiments concerning the relation between eye-gazing and turn-taking in natural human-human conversations, and how these observations can be extended to multimodal human-machine interactions. We confirm the earlier findings that eye-gaze is important in coordinating turn-taking and information flow in dialogues, but note that in multiparty dialogues also head movement seems to pay a crucial role in signalling the person's intention to take, hold, or yield the turn.

1. INTRODUCTION

The role of eye-gaze in fluent communication has long since been acknowledged (Argyle and Cook 1976; Kendon 1967). Previous research has established close relations between eye-gazing and conversational feedback (Goodwin, 2000), as well as eye-gazing and focus of shared attention (Trevarthen, 1984). Bavelas (2005) has studied how eye-gazing is used in signalling turn-taking in social interactions, while Padilha and Carletta (2003) simulated non-verbal behaviour in group discussion and included eye-gaze as an important signal for turn-taking. Vertegaal et al. (2003) provides evidence that lack of eye contact decreases turn-taking efficiency in video-conferencing, and Qu and Chai (2009) showed that the coupling of speech and gaze streams in a word acquisition task can improve performance significantly.

Several computational models of eye-gaze behaviour for artificial agents have also been designed. For instance, Lee et al. (2007) describe an eye-gaze model for believable virtual humans, Sidner et al. (2005) demonstrates gaze modelling for conversational engagement, and Nakano and Nishida (2007) built an eye-gaze model to ground information in interactions with embodied conversational agents. Overviews on Embodied Conversational Agents can be found e.g. in Cassel et al. (2003) as well as in Andre and Pelachaud (2010).

Eye-tracking technology has also developed fast. Although eye-trackers have been useful tools for cognitive psychology and human-computer interaction, the technology has only recently become more robust and been successfully used for various interface applications. For instance, the eye-typing interface for disabled people uses eye-gaze as a control mechanism so that the fixation of the gaze and/or the blinking action can be used for selecting items on a virtual keyboard displayed on the screen (Majaranta and Räihä, 2002). Experiments have also dealt with interfaces where the cursor is moved towards the user's point of gaze on the screen so as to anticipate the selection of the object that the user is looking at, and this seemed to be slightly faster than the normal mouse selection. The kind of interface that takes the user's focus of attention into consideration is usually referred to as gaze-aware or Attentive User Interface (Vertegaal, 2003).

In this paper we focus on eye-gaze in natural dialogues and especially on its role as a means to coordinate and control turn-taking. The importance of eye-gaze in turn-taking has already been established by previous research (e.g. Kendon, 1967): usually the interlocutors signal their wish to give the turn by gazing up to the interlocutor, leaning back, and dropping in pitch and loudness, and the partner can, accordingly, start preparing to take the turn.

Our studies have two distinctive aspects: first we focus on three-party conversations instead of two-party dialogues, and second, we focus on free-flowing conversations rather than task related dialogues. Both these aspects have consequences on dialogue management: multiparty conversations make the participants' shared context more complex, and the interlocutors can, in fact, just observe the other partners' interaction without taking directly part in it themselves. The free-flowing conversations are expected to bring about more natural behaviour patterns than task-related conversations since the participants' interaction is not constrained by specific requirements of the task.

Moreover, we explicitly base our annotations on the hypothesis that human–human interaction is cooperative activity which emerges from the speakers' capability to act in a relevant and rational manner (Allwood et al., 2007; Jokinen, 2009). Cooperation presupposes that the speakers share a goal, they consider each other cognitively and ethically, and they mutually trust the partner to follow the same communicative principles. The basic enablements of communication, i.e. Contact, Perception, and Understanding (CPU) must hold for the communication to proceed smoothly, and consequently, the agents' cooperation manifests itself to the extent in which the agents can observe and provide relevant feedback on the CPU enablements. As feedback is often expressed nonverbally, the agents must be sensitive to the relevant nonverbal signals that are used, either intentionally or unintentionally, to indicate the state of communicative activity. Such aspects as looking at the

[1] Affiliated with University of Helsinki and currently NICT Visiting Scholar at Doshisha University.

conversational partner or looking away provide indirect cues of the partner's willingness to continue interaction, whereas gazing at particular elements in the vision field tells us what the partner's focus of attention is, and thus they give guidance for appropriate presentation of information as well as suitable analysis and response to the partner's contribution. The agents thus constantly monitor each other and the communicative situation and, if some of the enablements are not fulfilled, react to the problems. For instance, if the agent has lost contact, is not interested, or does not understand the partner's contribution, he is obliged to make these CPU problems known to the partner, who, consequently, must adapt her communication strategies to the level that is appropriate in the situation.

The paper is structured as follows. We first describe the data collection setup and the data in Section 2, and then provide some initial experimental results in Section 3. Discussion concerning results and their importance for human-machine interfaces follows in Section 4, and finally, Section 5 draws conclusions and points to future research.

2. DATA COLLECTION

We collected data from speakers participating in natural, free-flowing dialogues. The collection setup is shown in Figure 1. Three participants sit in a triangle formation, and one of them has the eye-tracker in front of them to record their eye movements (the rightmost person in Figure 1). The two other participants, the left-hand speaker (LS) and the right-hand speaker (RS), are videotaped with a digital camera and they provide the reference point to what ES sees and where his gaze is focused on.

Compared with many earlier eye-tracker experiments, the special feature of the set-up is that the eye-gaze target is not a static video-screen but a moving dialogue partner, so the eye-tracker needs to be reliable concerning slight movement sideways as well as backward and forward. We also told the ES not to make overly large and exaggerated movements, and in general, they were cautions not to move too much. However, this did not seem to affect the naturalness of the dialogues: the participants chatted freely, and also used some hand gestures and head turns. The small number of hand gestures and body posture changes was also considered to be in line with the cultural characteristics of the participants rather than to stem from a special constraint of the eye-tracker setup itself.

Figure 1 Data collection setup.

We used the NAC EMR-AT VOXER eye-tracker, and its basic architecture is given in Figure 2. The desk mounted system consists of the camera optics for tracking the eye, and has a separate control monitor, and an image processing PC.

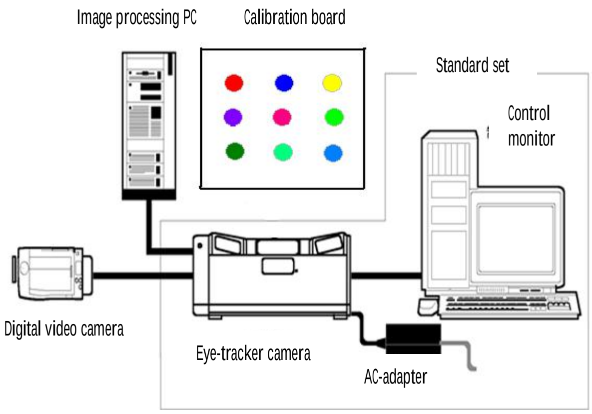

Figure 2 NAC EMR-AT VOXER eye-tracker setup.

For the data collection we recruited six participants (five male and one female) among the students in the laboratory. The participants were all Japanese, and they knew each other but not necessarily very well. In order to get a mixture of participants with minimum contact with each other in the experimental setting, the participants rotated so that the eye-tracked person was always a new participant in each triad and then moved to be the LS and the RS in the next rounds. The task of the participants was to learn more about the others and discuss issues that they were interested in. Consequently, the dialogues are natural chatting on topics that range from hobbies and weekend plans to studies and travelling.

The six dialogues are about 10 minutes long. We chose about 4-6 minutes stretches from the beginning of each dialogue for the annotation so that the selected part forms a cohesive topic. The excerpts were annotated by three naïve annotators who got instructions in English and in Japanese. They annotated the behaviour of all the three dialogue participants (ES, LS, and RS) using labels for dialogue acts, gaze, face, head, turn-taking, feedback and emotion/attitude.

The annotated features are listed in Table 1. The dialogue act annotation followed the guidelines of the AMI corpus (www.amiproject.org), while non-verbal communication was annotated according to a modified MUMIN scheme (Allwood et al., 2007). The features in the original MUMIN scheme also include hand gestures and body posture, but they were not included in the present study.

It must be noticed that the eye-tracked participants were not videotaped, so their face and head movement could not be annotated either. The features that were used only for LS and RS are italicized in Table 1. However, these features can be interpreted as describing observations of the head movement that is taken into account by the eye-tracker when it processes ES's gaze points, although of course any correspondence is only conceptual. On the other hand, the features *GazeObject* (ES) and *GazeToInterlocutor* (LR/RS) roughly correspond to each other.

Annotation features	Feature values
Dialogue Act	Backchannel, Stall, Fragment, BePositive, AskUnderstand, GiveUnderstand, AskAssesment, GiveAssesment, Suggest-offer, Inform, Ask, Other
GazeObject (only for ES)	RS, LS, Other
GazeToInterlocutor	*ES, Partner, Other*
GazeDirection	*Up, Down, Side, Otherway*
HeadMovement	*Nod, Jerk, Backward, Forward, Tilt, TurnToPartner, TurnSide, Waggle, Other*
HeadRepetition	*Single, Repeated*
Face	*Basic, Smile, Laughter, Scowl, Other*
FeedbackDirection	Give, Elicit
Feedback	CPU, Agree NonAgree
Turn	Give, Take, Hold, No_turn
Emotion/Attitude	Happy, Sad, Interested, Uninterested, Surprised, Disgusted, Angry, Frightened, Certain, Uncertain, Disappointed, Satisfied, Other

Table 1 Annotation features. The five features in italics in the middle are used only for LS/RS. The features GazeObject (ES) and GazeToInterlocutor (LS/RS) roughly correspond to each other.

Annotation was done with the Anvil annotation tool (Kipp, 2001). The annotation board has separate tracks for each of the three partners, and it groups the annotation features into the levels of utterance (words, dialogue act) and gaze/face. A view of the annotation board is shown in Figure 3.

The annotations were further checked pairwise. The annotator agreement was measured by Cohen's kappa-coefficient, and we reached the kappa value 0.46, which corresponds to moderate agreement.

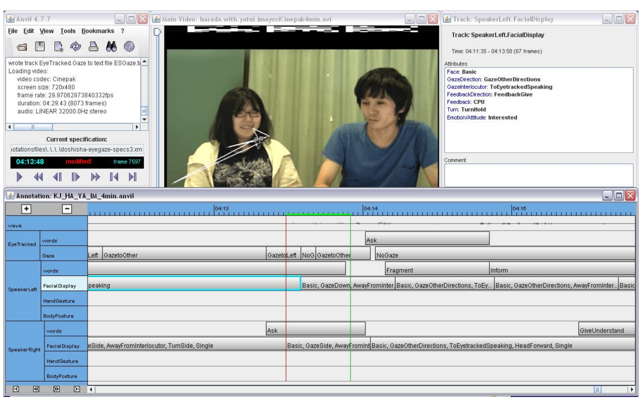

Figure 3 Anvil annotation board.

Concerning eye-gaze, this was manually annotated on the basis of the video data, while the eye-tracked person's gaze was based on the eye-tracker analysis. The eye-tracker data is not always clear as the gaze trace sometimes breaks. If the break (= no gaze) is shorter than 0.2 seconds, the gaze elements were regarded as part of the same gaze event (unless there was a shift), otherwise they were considered separate events (and no shift could be recorded).

Some basic statistics of the data is given in Table 2.

	eye-tracked speaker	left + right speaker
Dialogue acts	359	557
Total speaking time	487,81s	950,74
Av length of dact /s	1,36	1,70
Max /s	8,74	11,98
Min /s	0,17	0,07
Face elements	567	725
Av length of gaze /s	2,82	2,85
Max /s	51,92	38,17
Min /s	0,1	0,17

Table 2 Basic statistics of the annotated dialogue samples.

3. SOME EXPERIMENTAL RESULTS

As discussed above, gaze has many functions in face-to-face interaction. We investigated the relation between eye-gaze and turn-taking, especially how eye-gaze shows the speaker's intention to prepare to talk (turn-requesting), to continue to talk (turn-holding) and to give up the turn (turn-yielding).

3.1 Turn-taking and eye-gaze

We used the Support Vector Machine technique in the Weka software package (Witten and Frank, 2005) to classify the gaze instances with respect to turn management (*give, take, hold, and no_turn*). We combined the data from LS and RS together, but kept the data from ES separate. We first used all the seven features annotated for the LS/RS gaze instances and the five features annotated for ES gaze instances, and in the second test run we left out the specification for LS/RS gaze (we did not run an analogous classification for ES gaze).

The results are given in Table 3 and Table 4. They concern the %-correct and weighted average f-measure, respectively. As a baseline classification, we used Weka's ZeroR algorithm which selects the most common category. The significance differences are calculated by Weka's Paired T-tester.

On the confidence level 0.01, the classifications obtained by the Support Vector Machine (SVM) are significantly better than the baseline classification. Comparing the f-measures obtained by SVM for the "with gaze" and "without gaze" conditions, we notice that gaze features improve the classification performance. However, it turns out that the difference is not statistically significant on the confidence level 0.01.

Confusion matrix for the LS/RS speaker with gaze information is depicted in Table 5. We notice that when the interlocutor is actually listening (no_turn), only few confusions (9%, n=49) occur considering whether the interlocutor holds the turn (i.e.

would be speaking instead), but when the interlocutor is actually holding the turn, a good number of confusions (38%, n=51) occur considering the interlocutor not having the turn (i.e. would be listening instead). However, the actual turn-changing events, turn-give and turn-take, were not very well recognized. Almost 42% of the turn-giving events (n=17) were confused with no_turn (listening), and 60% of turn-taking events (n=12) were confused with no_turn (listening), too.

We can interpret this to mean that although eye-gaze is necessary for the interaction coordination, there are many eye-gaze events which occur during the interlocutors either speaking themselves or listening to their partners speaking, and the selected features do not distinguish these events clearly enough.

Dataset	RulesZ	SVM
SVM LSRS with gaze	73.11 +/- 0.51	80.00 +/- 4.08
SVM LSRS without gaze	73.11 +/- 0.51	78.14 +/- 3.57
SVM ES with gaze	51.68 +/- 0.52	92.68 +/- 1.32
SVM ES without gaze	51.68 +/- 0.52	90.56 +/- 1.07

Table 3 Percent correct classification results. RulesZ is the majority baseline algorithm, SVM is Support Vector Machine. The difference between RulesZ and SVM is statistically significant, but the difference between with gaze and without gaze conditions is not.

Dataset	RulesZ	SVM
SVM LSRS with gaze	0.62 +/- 0.01	0.79 +/- 0.04
SVM LSRS without gaze	0.62 +/- 0.01	0.74 +/- 0.04
SVM ES with gaze	0.35 +/- 0.01	0.90 +/- 0.01
SVM ES without gaze	0.35 +/- 0.01	0.87 +/- 0.01

Table 4 F-measure for the classification results. RulesZ is the majority baseline algorithm, SVM is Support Vector Machine. The difference between RulesZ and SVM is statistically significant, but the difference between with gaze and without gaze conditions is not.

Total	a	b	c	d	← classified as
530	474	7	0	49	a = no turn (listening)
41	17	23	1	0	b = give
20	12	1	5	2	c = take
134	51	0	2	81	d = hold (speaking)

Table 5 Confusion matrix for the LS/RS speaker with gaze information.

As a conclusion we can say that the data with the annotated features confirm the earlier findings, namely that eye-gazing is important in turn management. However, it seems that in our data, with the annotation used, eye-gaze has not such a significant but also other features, especially feedback behaviour is crucial in signalling turn taking. As the data set is still rather small, it may be that the data contains some idiosyncrasies which affect the classification results, so we refrain from drawing definite conclusions on the topic.

3.2 Head movement, eye-gaze and turn-taking

Head movement as a turn taking signal has not been paid so much attention compared with eye-gaze, and e.g. Argyle and Cook (1976) talk about gaze as a social signal which opens a channel to receive non-verbal messages and synchronize speech. At the end of the utterance, the speaker usually has a long eye-contact with the partner, who then shifts his gaze away before starting to speak.

In our previous work (Jokinen et al., 2009), we found out that in multiparty conversations, the speaker's turn-taking signalling seemed to be based more on head movement than eye-gazing, and that eye-gaze seemed to follow head movement. Gaze events occur usually just before turn-taking, as the next speaker tend to lift their head and gaze before starting to speak or, as in some cases, to laugh. In the laughing context, the speaker often gives verbal agreement signals to the speaker and nods his/her head so as to make the positive effect stronger.

As we had collected more data and had more than twice the amount of annotated data, we investigated the relation further. We used Weka to investigate the relative contribution of the different attributes for turn management. This gives us some indication of the effects of the communicative signals upon turn management, i.e. how the head movement, perceived emotion, and gazing, seem to contribute to the classification, and thus to the correct interpretation of the turn management. The method evaluates each attribute with respect to the information gained by selecting it for turn management, and ranks the results with the Ranker method in Weka, using 10-fold cross-validation. The most highly ranked attributes are shown in Table 6 and Table 7.

As can be seen, for ES, gaze object information does not seem very useful. The best indicators for ES turn management are related to feedback giving process, and the perceived emotional state rather than eye-gaze. With the LS RS partners, feedback seems to play a crucial role and head movement and the perceived emotional state. However, now also gaze has some important effect: gazing to the partner is an indication of turn taking.

average merit	average rank	attribute
0.985 +- 0.001	1.2 +- 0.6	Feedback
0.983 +- 0.002	2.3 +- 0.46	Emotion/Attitude
0.982 +- 0.001	2.8 +- 0.6	FeedbackDirection
0.98 +- 0.003	3.7 +- 0.9	GazeObject
0.964 +- 0	5 +- 0	DialogueAct

Table 6 Ranking of attributes for ES turn taking

However, in multiparty conversations, the interlocutors' head movements may be more noticeable signals to the partners than eye-gazing, and thus they can also function as more reliable cues for turn management. The interlocutors effectively turn their head to the speaker if they want to take the turn, and if they do not want to take the turn, they move their head towards the other partner who will talk next. In multiparty conversations, head movement may thus be one of the most important cues to signal one's intention to speak or to yield turn to the partner.

It will be interesting to see if the hypothesis also works for the ES. In this data set we did not have a video of the ES but in the new data set collected, also ES is videotaped from the front so the hypothesis can be checked. On the other hand, as already mentioned, the head movement of ES may be somewhat restricted by the fact that ES is aware of the eye-tracker and that they should not be move their head so freely.

average merit	average rank	attribute
0.27 +- 0.007	1 +- 0	Speaker
0.179 +- 0.003	2 +- 0	Feedback
0.111 +- 0.006	3 +- 0	GazeInterlocutor
0.096 +- 0.006	4.3 +- 0.46	HeadMovement
0.088 +- 0.004	4.8 +- 0.6	Emotion/Attitude
0.08 +- 0.005	5.9 +- 0.3	GazeDirection
0.052 +- 0.004	7 +- 0	HeadRepetition
0.037 +- 0.003	8 +- 0	Face
0.029 +- 0.001	9 +- 0	FeedbackDirection

Table 7 Ranking of attributes for LSRS turn taking

4. HUMAN-MACHINE INTERFACE

In human machine interactions, one of the main goals has been to design systems that are natural and easy to use, yet robust and efficient in their task completion. In particular, in the prevailing ubiquitous computing paradigm (Weiser, 1990), it is assumed that our environment will be populated with several context-aware and networked devices that communicate with each other, and also with users. Applications range from various home appliances to robotic companions and social media, and from the users' point of view, it is expected that input modalities are available such that they best suit to the users' needs and to the task in hand. This requires that the applications have appropriate communicative skills so that the users can interact with them in a natural and smooth manner. Moreover, as the devices can also request the user's attention to a particular task or item in the environment, such skills must also include simultaneous coordination of action and communication.

We pursue a multilevel analysis of the data so as to integrate both signal processing and human communication studies. Thus our method is hybrid in that it combines the top-down human annotation approach with the signal analysis (Jokinen et al., 2009). By studying human-human interactions we investigate intuitive human behaviour patterns and dependencies among different modalities, while signal level analysis gives us an estimate of how a machine can recognize the same activity. The models can then be deployed in human-computer interactions to enhance interaction possibilities for context-aware and intelligent devices.

In human-machine interfaces, cooperation manifests itself in the system properties that allow users to interact in a natural manner, that is, in the ways in which the system affords cooperative interaction. Recent developments in speech, eye-tracking, and gesture recognition technologies form enabling technologies which can greatly enhance future human-technology interfaces by allowing natural, intuitive, easy, and friendly interactions in ubiquitous environment. The eye-gaze experiments reported in this paper also aim to contribute to the theories of human-human interaction and to the modelling of gaze behaviour for the design and development of robust and natural interactions with smart objects, services, and environments.

5. CONCLUSIONS

Eye-gaze gives the interlocutors guidance as to when to take the turn, or if the partner wants to continue talking or is preparing to talk, and thus helps in smooth and effective interaction without explicit spoken expressions. In this paper we confirmed earlier observations concerning eye-gaze and turn taking, but also discussed some interesting results related to turn-signalling. In particular, we noticed that in multiparty conversations, the turn management seem to be signalled with head turns rather than eye-gazing, while eye-gaze is important as an initial signal of who could be next speaker.

In multiparty conversations, head movement may function as a more visible signal of the speaker's focus of attention and willingness to take turn. In two-party dialogues, eye-gaze may be enough to signal the partner's intention to take the turn or to give the turn, but in multi-party dialogues, the participants may not share the context completely and the partner's focus of attention needs to be expressed in a more visible manner.

There are further challenges that concern the correlations between eye-gaze and turn-management and we will investigate these further. On the basis of the gaze elements and the features that were annotated in our data, we noticed that turn-taking could be predicted fairly accurately. However, aspects such how long and how often the partner is looked at, are also relevant from the communicative point of view, as well as the whole eye region (eyebrows, eyelids, wrinkles), and the upper body and gesture movements.

We have also collected more data, so that the dialogues now comprise 20 dialogues. Ten of the conversations are among participants who do not know each other in advance. This allows us to compare eye-gazing behaviour with familiar and unfamiliar contexts. We also intend to do thorough speech analysis on the annotated dialogue acts so as to study the relation between eye-gaze, prosody, and turn-taking. The new collection also contains video of the eye-tracked persons so as to allow more detailed analysis of their face and body. As a future work, we plan to include an analysis of gestures in the research.

6. ACKNOWLEDGMENTS

We would like to thank all the students who helped us with the experiments and data annotation, and our subjects for taking part in the experiments. The first author would also like to thank NiCT for the grant that allows her to conduct research in Japan.

7. REFERENCES

[1] Allwood, J., L. Cerrato, K. Jokinen, C. Navarretta and P. Paggio 2007. The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management, and Sequencing Phenomena. In Martin, J.C., P. Paggio, P. Kuenlein, R. Stiefelhagen and F. Pianesi (Eds.) Multimodal Corpora for Modelling Human Multimodal Behaviour.

Special issue of the International Journal of Language Resources and Evaluation, 41(3-4), 273-287. Springer.

[2] André, E. and C. Pelachaud (2010). Interacting with Embodied Conversational Agents. In K. Jokinen and F. Cheng (Eds.), *New trends in speech-based interactive systems*. New York: Springer.

[3] Argyle, M., and M. Cook (1976). Gaze and Mutual Gaze. Cambridge University Press, Cambridge

[4] Bavelas, J. B. 2005. Appreciating face-to-face dialogue. In AVSP-2005, 1.

[5] Cassell, J. Sullivan J., Prevost, S., Churchill E. (Eds.) 2003. *Embodied Conversational Agents*. Cambridge University Press

[6] Goodwin, C. 2000. Action and embodiment within Situated Human Interction. Journal of Pragmatics, 32:1489-1522

[7] Jokinen, K. 2009. Constructive Dialogue Management. Rational Agents and Speech Interfaces. John Wiley.

[8] Jokinen, K., Nishida, M., and Yamamoto, S. 2009. Eye-gaze Experiments for Conversation Monitoring. Proceedings of the IUCS'09 conference, ACM, Tokyo

[9] Kendon, A. 1967. Some Functions of Gaze Direction in Social Interaction. Acta Psychologica 26, 22–63

[10] Kipp, M. (2001). Anvil – A generic annotation tool for multimodal dialogue. In *Proceedings of the Seventh European Conference on Speech Communication and Technology* (pp. 1367–1370).

[11] Lee, J., Marsella, T., Traum, D., Gratch, J., and Lance, B. 2007. The Rickel Gaze Model: A Window on the Mind of a Virtual Human. In C. Pelachaud et al. (Eds.) Proceedings of the 7th International Conference on Intelligent Virtual Agents (IVA-2007). Springer Lecture Notes in Artificial Intelligence 4722, pp. 296-303. Springer-Verlag Berlin Heidelberg.

[12] Majaranta, P. and Räihä, K. 2002. Twenty years of eye typing: systems and design issues. In Proceedings of the 2002 Symposium on Eye Tracking Research & Applications (New Orleans, Louisiana, March 25 - 27, 2002). ETRA '02. ACM, New York, NY, 15-22.

[13] Nakano, Y., Nishida, T. 2007. Attentional Behaviours as Nonverbal Communicative Signals in Situated Interactions with Conversational Agents. In Nishida, T. (Ed.) Engineering Approaches to Conversational Informatics, John Wiley.

[14] Padilha, E. G. and J.C. Carletta (2003) Nonverbal behaviours improving a simulation of small group discussion. Proceedings of the First International Nordic Symposium of Multi-modal Communication, Copenhagen, Denmark, 25-26 September, 2003.

[15] Qu, S. and J. Chai (2009). The Role of Interactivity in Human-Machine Conversation for Automatic Word Acquisition. SIGDIAL Conference 2009, 188-195

[16] Rietveld, T. and R. van Hout 1993. Statistical Techniques for the Study of Language and Language Behaviour. Mouton de Gruyter, Berlin.

[17] Sidner, C.L., Lee, C., Kidd, C., Lesh, N., Rich, C. 2005. Explorations in engagement for humans and robots. Artificial Intelligence 166(1-2), 140–164

[18] Trevarthen, C. 1984. Emotions in Infancy: Regulators of Contact and Relationships with Persons. In K. Scherer and P. Ekman (eds.) Approaches to Emotion. Hillsdale: Erlbaum

[19] Umata, I., Ito, S., Iwasawa, S., Suzuki, N. and Inoue, N. 2009. Non-verbal Information in Communication: A Preliminary Study in a Tourist-Information Setting. Proceedings of the HCII Conference, pp. 653-657.

[20] Vertegaal, R., Weevers, I., Sohn, C. and Cheung, C. 2003. GAZE-2: conveying eye contact in group videoconferencing using eye-controlled camera direction. Proceedings of CHI 2003. Fort Lauderdale, TX: ACM Press.

[21] Vertegaal, R. 2003. Attentive User Interfaces. Editorial, Special Issue on Attentive User Interfaces, Communications of ACM 46(3), March 2003.

[22] Weiser, M. (1991). The computer for the twenty-first century. *Scientific American*, 94–10.

[23] Witten, I.H. and Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. 2nd Edition, Morgan Kaufmann, San Francisco.