Facial Feedback Signals for ECAs

Elisabetta Bevacqua[1] and **Dirk Heylen**[2] and **Catherine Pelachaud**[1] and **Marion Tellier**[1]

Abstract. One of the most desirable characteristics of an intelligent interactive system is its capability of interacting with users in a natural way. An example of such a system is the embodied conversational agent (ECA) that has a humanoid aspect and the capability of communicating with users through multiple modalities such as voice, gesture, facial expressions, that are typical of human-human communication. It is important to make an ECA able to fit well in each role in a conversation: the agent should behave in a realistic and human-like way both while speaking and listening. So far most of the work on ECAs have focused on the importance of the ECA's behaviour in the role of the speaker, implementing models for the generation of verbal and non-verbal signals; but currently we are mainly interested in modelling the listening behaviour. In this paper we will describe our work in progress on this matter.

1 Introduction

In conversations participants produce behaviours that are intended to convey meaning or intentions. The producer of communicative behaviours wants the intentions he has with them to be recognized by the addressees of his message. Conversation is thus a particular, socially developed instrument to enable mindreading. Communication as we understand it here requires a Theory of Mind on the side of both producers and recipients of communicative behaviours. Producers need to design their communicative actions taking into account what they believe to be the mental state of the recipients (audience design). Recipients need to be able to recognize that behaviours were produced because of an intentional action. This is the notion of non-natural meaning as discussed by ***Grice**. ***Levinson*** paraphrases Grice's definition of non-natural meaning as follows:

> [C]ommunication consists of the 'sender' intending to cause the 'receiver' to think or do something, just by getting the 'receiver' to recognize that the 'sender'is trying to cause that thought or action. So communication is a complex kind of intention that is achieved or satisfied just by being recognized. In the process of communication, the 'sender's' communicative intention becomes mutual knowledge to 'sender' (S) and 'receiver' (H), i.e. S knows that H knows that S knows that H knows (and so ad infinitum) that S has this particular intention. Attaining this state of mutual knowledge of a communicative intention is to have successfully communicated. (Levinson, p. 16)

[1] University of Paris8, IUT de Montreuil, 140 rue de la Nouvelle France, 93100, Montreuil, France
email: e.bevacqua,c.pelachaud@iut.univ-paris8.fr

[2] University of Twente PoBOX 217, 7500 AE Enschede, The Netherlands, email: heylen@ewi.utwente.nl

During a conversation the listener is called to provide information on the successfulness of the communication. In order to ensure closure on the communicative actions, speakers will monitor listeners for cues of recognition to establish grounding which can be both natural cues as intentional signals produced by listeners to provide feedback on the speech. The term back-channel (stemming from [24]) is commonly used to denote the communicative behaviours that are produced by participants in a conversation as feedback on the reception of the communicative behaviours of the other participants. Both through linguistic and gestural signals, the listener can show his level of engagement in the conversation. According to the listener's responses the speaker can estimate how his/her interlocutor is reacting and can decide how to carry on the interaction: for example by interrupting the conversation if the listener is not interested or re-formulating a sentence if the listener showed signs of not understanding and so on.

In our research, we want to analyse not only how this behaviour is displayed, but also what kind of information it provides about the listener's reaction towards the speaker and his/her speech. Our aim is thus twofold: on the one hand we want to implement back-channel behaviour in a conversational agent in order to make it more realistic and human-like, and on the other hand we want to make sure that the user is able to interpret the agent's signals as 'intended' by the ECA, so that the user feels the ECA is displaying the appropriate level of understanding and participates actively in the conversation.

Through one or more channels like voice, head, face, gaze, posture and gesture, listeners provide back-channels signals of perception, attention, interest, understanding, attitude (belief, liking. . .) and acceptance towards what the speaker is saying [24, 1, 19]. A back-channel can be positive or negative and can have several meanings (understanding but not acceptance, believing but not agreeing and so on). Moreover, the listener can emit signals with different levels of control and intentionality: consciously deciding to emit a signal in order to show a reaction to the speaker's speech (and even deliberately choosing a specific one to provoke a particular effect on the speaker, for example: the listener decides to stare at the speaker to show disbelief or surprise expecting a confirmation by the speaker) or emitting cues without thinking, automatically reacting to the speaker's behaviour or speech, generating back-channels at a very low level of control [1].

In this paper we present our first experiments along these lines. We start with a characterisation of back-channels. Then, we present a perceptual test we have conducted and preliminary results. Finally we explain how we aim at introducing the evaluated signals in a computational model for a conversational virtual **listener**.

2 Back-channels

Several research traditions have studied the behaviours that listeners display in conversations. Back-channels, or similar phenomena with a different name such as response tokens, have been studied in the conversational analysis literature, for instance, with the purpose of understanding what role the various contributions of all of the participants play in shaping the conversation. Most relevant in this respect are papers such as [21], [22], [15] but there are many others. The literature on turn-taking, both from the CA and other perspectives, also provides useful notes on the behaviours of participants that assume the primary speaker role and the auditors. In the series of papers by Duncan and co-authors[3], for instance, auditor back-channel signal are one of three classes of signals, besides speaker within-turn and speaker continuation signals, that serve to mark units of interaction during speaking turns.

A general assumption behind the concept of back-channel is that all the participants in a face-to-face conversation are both producers and recipients of communicative signals, but that there are different levels on which this occurs. Communicative signals on the primary track, to use the term by [5], are by the participants that have the floor and the secondary track, ‘in the back’, is constituted by the feedback on the behaviours in the primary track. As [24] points out there may be cases of iteration where speakers provide feedback on the back-channels of listeners.

Several studies of nonverbal behaviours have paid attention to the behaviours displayed by listeners. One kind of phenomenon that has received some attention is the way in which behaviours of participants are synchronized and in particular how body movements of listeners are coordinated with the verbal utterances of the speaker. [14] showed that about a quarter of the head movements by listeners are in sync with the speaker’s speech. Interactional synchrony in this sense has been studied, amongst others by [16], [20], [6]. Mirroring is a particular type that has often been commented upon. Scheflen suggests that this often reflects a shared viewpoint. Also [16] hypothesized that the level to which behaviours are synchronized may signal the degree of understanding, agreement or support. These kinds of phenomena show that the behaviours of listeners arise not only from ‘structural concerns’ (e.g. turn-taking signals) but also from ‘ritual concerns’. We take these terms from [12] who points out that it is sheer impossible to assign to behaviours a function of only one of these types of concerns (see also [3]).

Besides these synchrony behaviours, listeners display various other nonverbal behaviours as feedback. [4], looking in particular at facial expressions, classifies these behaviours in a small set of semantic categories of listener comment displays. These are, besides displays for agreement:

- Back-channel: Displays that were produced by listeners while the speaker was talking or at the end of the speaker’s turn. They take the form of brow raises, mouth corners turned down, eyes closed, lips pressed. In Chovil’s corpus the displays could be accompanied by typical back-channel vocalizations such as “uhuh”, “mhmm”, “yeah”, etc.
- Personal reaction displays: A reaction in response to what the speaker had said rather than just acknowledging the content.
- Motor mimicry displays: displays that might occur in the actual situation that the speaker is talking about (e.g. wincing after hitting ones’ thumb with a hammer, eyes widened and an open mouth in response to a frightening situation). These are interpreted as messages that indicated a sincere appreciation of the situation being described.

In the discussion so far, we have mentioned several functions that are served by the behaviours of listeners. They provide feedback to the speaker, acknowledging reception of the signal, possibly its understanding or some kind of comment expressing a particular attitude towards what is being expressed. From its nature as a kind of joint communicative action, conversations require that participants come to react to each other’s actions to ground the actions and provide closure. Feedback is an important part of establishing grounding in the interactional achievement of having a conversation. The variety of functions that feedback serves is partly explained by the various levels on which grounding needs to take place: i.e. levels at which the participants need to have a mutual understanding of each other’s intentions. [5] suggests that grounding needs to occur on at least four levels with each step a kind of joint action.

1. Joint[A executes behavior t for B to perceive; B attends perceptually to behavior t from A]
2. Joint[A presents signal s to B, B identifies signal s from A]
3. Joint[A signals to B that p, B recognizes that A means that p]
4. Joint[A proposes a joint project to B, B takes up the joint project]

As speakers make their utterances, they are usually also monitoring the interlocutors behaviours to find signs of their participatory involvedness on all of these levels.

1. A monitors B for signs of perception activity / B’s behaviour provides cues of perception activity
2. A monitors B for signs that B has identified the signal / B indicates that he has identified the signal...

The utterance of speakers and the accompanying behaviours will often be designed to invoke behaviours of interlocutors to ensure this. A typical case of this behaviour is analysed by [13], consisting of hesitations and repetitions of speakers at the beginning of their utterance to evoke gaze behaviours in interlocutors.

In a similar vein, [1] distinguishes four basic communicative functions on which the speaker may require feedback:

1. Contact: is the interlocutor willing and able to continue the interaction
2. Perception: is the interlocutor willing and able to perceive the message
3. Understanding: is the interlocutor willing and able to understand the message
4. Attitude: is the interlocutor willing and able to react and respond to the message (specifically accepting or rejecting it).

The various feedback behaviours are thus not only varied in their form but also in their function. In one of the experiments that we are carrying out and report on below, we are looking at back-channel behaviours in which facial expressions, gaze, and head movements are controlled. We look at various classes of expressivity. The general classes that we consider are the following:

- Performatives such as agree, disagree, criticize, refuse, accept, approve, confirm, question
- Affectives: liking, disliking, disgust, sorry-for, surprise, fear, anger, reproach, gratitude
- Epistemics: believe, disbelieve, scepticism, certainty, doubt

[3] See [7], [8], [9], [10], [11],.

- Meta-cognitives: thinking, planning, remembering

These functions and behaviours go beyond the usual back-channel behaviours such as nodding that are mostly discussed in the computational literature. An important issue to consider is the degree to which people agree on the interpretation of the behaviours. The experiment described next is supposed to shed some light on this.

3 Recognition Test

3.1 Hypotheses

Our first hypothesis is that most back-channel signals either convey a positive or a negative connotation. Therefore, we are trying to find out the general meaning for each signal when there is one. We can assume that signals containing nods and smiles will be interpreted as positive feedback signals such as agree, accept, like, understanding and believe whereas signals containing shakes, frown and tension will rather be associated with negative meanings such as disagree, refuse, dislike, do not understand and disbelieve. Our second hypothesis is that back-channel signals are polysemic: the same signal can have different meanings and a single meaning can be expressed with different signals or a combination of signals. We are thus assuming that a single signal can be interpreted by subjects in different ways. To test these hypotheses we have conducted recognition tests on subjects who were asked to judge a set of 14 different signals displayed by a 3D agent Greta [18].

3.2 Participants

Twelve French students have been tested so far. They are students in computer science, age range 18-20.

3.3 Material

The test was done with our 3D agent Greta. The graphic interface of the test application can be seen in Figure 1. In the little window on the left Greta's videos are shown once at a time. Two buttons under the window, *play* and *next movie*, allow the user respectively to play the movie (in this way the movie is shown only when the user is paying attention) and to move on to the next movie. For a more controlled procedure, we decided that participants could not rewind the video. On the right a list of possible meanings is proposed to the participant who, after each movie and before moving on, can select one meaning according to his/her opinion about which meaning fits that particular back-channel signal best. It is possible to select several meanings for one signal and when none of the meanings seems to fit, partecipants can just click on *next movie*. In this case the absence of answers will be treated as "no answer" in the data.

In this test we decided to consider the meanings belonging to the class of expressivity *performative*:

- agree (AG)
- disagree (DA)
- accept (AC)
- refuse (RE)
- interested (IN)
- bored (BO)

From the class *epistemics* we selected the following meanings:

- believe (BE)

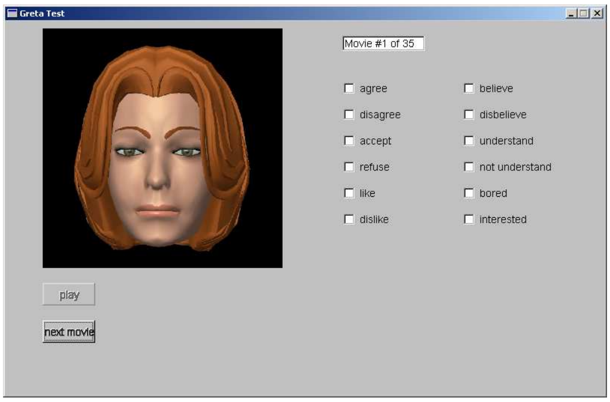

Figure 1. Graphic interface of the test application

- disbelieve (DB)
- understand (UN)
- not understand (NU)

Finally, from the class *affective*, we considered:

- like (LI)
- dislike (DL)

To make the videos we have selected 14 signals chosen among the back-channel signals which were analysed and proposed by [2, 19]. Some signals are simple, containing just a single action (like a nod or a shake), while others are obtained by combining several actions (like a nod and a raise eyebrows or a head tilt and a frown). The 14 signals are:

1. a single head nod (N)
2. a head nod with a smile (NS)
3. a head nod and a raise of the eyebrows (NRE)
4. a head shake (S)
5. a head shake and a frown (SF)
6. a head shake, a frown and a tension in the lips that tighten getting thinner (SFT)
7. a frown and a tension in the lips that tighten getting thinner (FT)
8. a raise of the left eyebrow (RLE)
9. the eyes roll up in the head (ER)
10. a head tilt on the left and sad eyebrows (TSE)
11. a head tilt on the left and a frown (TF)
12. a head tilt on the right and raise eyebrows (TRE)
13. a head tilt on the right and the gaze turns on the down right (TG)
14. the eyes wide open (EWO)

3.4 Procedure

Participants were given instructions for the test through a written text. They were told that Greta would display back-channel signals as if she was talking to an imaginary speaker. They were asked to evaluate these signals by choosing among the available list of meanings. This way we made sure that participants were aware that they were evaluating back-channel signals. The signals were shown randomly at least twice and at most three times so that the participants had to watch 35 movies in all: we wanted to find out whether people gave always the same answer, or if they tended to remember the signals and associate them to more possible meanings.

Signals	Positive answers	Negative answers	No answer	Total of answers
N	26	2	0	28
NS	38	1	0	39
NRE	57	1	0	60
RLE	17	22	3	42
TSE	5	27	0	32
ER	0	21	4	25
TG	6	29	8	43
SFT	2	35	0	37
FT	3	27	0	30
S	0	38	2	40
SF	1	31	0	32
TF	8	33	1	42
TRE	18	20	2	40
EWO	11	13	13	37

Table 1. Results positive and negative for each signal.

3.5 Results and Discussion

Signals	Significant	Meaning
N	Yes $p<0.0001$	positive
NS	Yes $p<0.0001$	positive
NRE	Yes $p<0.0001$	positive
RLE	No $p=0.5224$	No distinct meaning
TSE	Yes $p<0.0001$	positive
ER	Yes $p<0.0001$	positive
TG	Yes $p<0.0001$	positive
SFT	Yes $p<0.0001$	positive
FT	Yes $p<0.0001$	positive
S	Yes $p<0.0001$	positive
SF	Yes $p<0.0001$	positive
TF	Yes $p<0.0001$	positive
TRE	No $p=0.8714$	No distinct meaning
EWO	No $p=0.8388$	No distinct meaning

Table 2. Results of the binomial tests.

One of our main concerns in this experiment was to find out whether certain signals are globally considered positive or negative. We also expected to find meaningless signals that is to say signals that do not convey positive nor negative meaning on their own and need to be matched with other signals to be meaningful. To analyse the data, we coded the answers given by the subjects as positive or negative, according to the following principles: agree, accept, like, believe, understand and interested were considered as positive answers and disagree, refuse, dislike, not understand and bored were considered as negative answers. Table 1 shows the results for each signal. For the treatment of the data we have left out the cases in which subjects have not answered ("no answer") but it will be taken into account during the analysis of the results when relevant. The differences in the total number of answers is explained by the fact that some signals have been presented twice to subjects and others three times and by the fact that subjects could give several answers for each signal. The null hypothesis is that a signal has no distinct meaning (positive or negative) so that there is no significant difference between the amount of positive and negative answers for that signal. The alternative hypothesis is that the amount of answers for one signal is so high that it proves that subjects detected a distinct meaning. To test the hypothesis, binomial tests have been performed for each signal. Signals for which the p value is less than the 0.05 level of significance, reject the null hypothesis. Table 2 shows the results.

Thus, only three signals do not reject the null hypothesis: "raise of the left eyebrow", "head tilt on the right" and "raise eyebrows and eyes wide open". This means that these three signals do not convey enough meaning when displayed alone. Looking at the distribution of the answers, we can notice that subjects' answers are almost equally shared between positive and negative items and for the last signal, "eyes wide open", the amount of "no answer" is extremely high (13 out of 37) which confirms the meaningless aspect of this particular signal. Every other signal reject the null hypothesis which proves that they either convey a positive or negative connotation. Our data shows that the positive meaning of "head nod", "head nod and smile" as well as "head nod and raise of the eyebrows" is significant. It also shows that the negative aspect of "head tilt on the left and sad eyebrows", "eyes roll up in the head", "head tilt and gaze", "head shake, frown and tension", "frown and tension", "head shake", "head shake and frown" and "head tilt on the left and frown" is significant.

Table 3 shows the statistical results en percentage, signals were played two or three times and the table contains the results of the all the repetitions. On the rows there are the signals, while on the column (from the second to the fourteenth one) there are the meanings. The first column (#Ans) contains the number of answers given for the corresponding signal.

In general we have seen that subjects tend to give more and more answers for each signal as the test goes on, probably because they become accustomed to the movies and to the aim of the test. Moreover the more complex is the signal the more answers the subjects gave. For example "head nod and smile" obtained 39 answers while "head nod" 28. "head nod and raise of the eyebrows" had 60 answers, but it is important to notice that it was displayed three times while "head nod and smile" and "head nod" just twice. However in the first two

repetitions the signal “head nod and raise of the eyebrows” obtained more answers that “head nod and smile” and “head nod”.
We have the same result for the signals “head shake”, “head shake and frown” and “head shake, frown and tension”; results in the table 3 show that “head shake” obtained more answers than “head shake and frown” and “head shake, frown and tension” which are more complex signals, however “head shake” was displayed three times while “head shake and frown” and “head shake, frown and tension” just twice. During the first two repetitions “head shake and frown” and “head shake, frown and tension” still obtained more answers than “head shake”.
As expected, participants associated positive meanings to signals containing nods and smiles and in particular they related the smile to the meaning of liking (39.90%). Negative meanings were linked to shakes and frowns; for example the signal “head shake and frown” was associated above all to refuse and disagree (37.5%). The other signals were less easily associated to a constant set of meanings, as we assumed head tilts and rolling of the eyes were seen as signals of disbelief, not understanding and boredom, but they also suffered the more evident dispersion of answers and sometimes the percentage are not so relevant. For example, the signal “head tilt and gaze” (TG) was interpreted above all as a back-channel of boredom (37.20%), but all the other meanings were also selected at least once and the 18.60% of answers was “no answer”. The signal “eyes wide open” was the hardest to interpret, most answers were “no answer” (35.14%) and even if the second highest percentage classify this signal as a back-channel of disbelief, such percentage is not very relevant (16.23%). Perhaps these signals were hardest to interpret because they can convey more meanings according to the context and to the listener’s personality.
Some of the signals we took into account in this test are complex signals, composed by several single actions which have not been all tested individually. Thus, in further experiments we will analyse some actions separately, for example “smile”, “frown” and “sad eyebrows”.

4 Future Work

In the future we will submit the test to a more relevant number of subjects in order to obtain more accurate and significant results. Moreover we aim at proposing this test to subjects of different cultures in order to see if back-channel signals are perceived differently in other countries or if they are interpreted in the same way.
With this test we also want to define a set of recognizable signals to be used in the implementation of a listener model for our conversational agent. As we said in the Introduction, the listener can emit signals with different levels of control and intentionality, he can provide a back-channel consciously or unconsciously. Consequently a single listener model is not enough; two computational models are needed, respectively a cognitive model (to generate intentional back-channel signals) and a reactive model (to generate non-intentional back-channel signals). Since the instinctive listener’s back-channel is often elicited by the speaker’s behaviour, a set of rules can be defined to implement a reactive model [17]. For example, from a corpus of data, Maatman derived a list of rules useful to predict when back-channel can occur according to the speaker’s actions. Back-channel continuers (like head nods, verbal responses) appear at a pitch variation in the speaker’s voice; frowns, body movements and gaze shifts are produced when the speaker shows uncertainty; facial expressions, postural and gaze shifts are provided to reflect those made by the speaker (mimicry). Even variation in the speaker’s pitch of voice usually elicits a back-channel signal from the listener [23].
As for the cognitive model, it is complex to implement. To elaborate reasoned reactions from a listener, one must have access to not only the extrapolated speech content, but also information about the listener’s personality. For this reason, we will begin by implementing a Wizard of Oz system to provide consciously back-channel. The intentional listener behaviour is driven by a wizard while our virtual agent interacts with a user.

REFERENCES

[1] J. Allwood, J. Nivre, and E. Ahlsén, ‘On the semantics and pragmatics of linguistic feedback’, *Semantics*, **9**(1), (1993).
[2] Jens Allwood and L. Cerrato, ‘A study of gestural feedback expressions’, in *First Nordic Symposium on Multimodal Communication*, eds., P. Paggio, K. Jokinen, and A. Jönsson, pp. 7–22, Copenhagen, (2003).
[3] J.E.G.F.J. Bernieri, ‘The importance of nonverbal cues in judging rapport’, *Journal of Nonverbal behavior*, **23**(4), 253–269, (1999).
[4] Nicole Chovil, ‘Social determinants of facial displays’, *Journal of Nonverbal Behavior*, **15**(3), 141–154, (1991).
[5] Herbert Clark, *Using Language*, Cambridge University Press, Cambridge, 1996.
[6] W.S. Condon and W.D. Ogston, ‘A segmentation of behavior’, *Journal of Psychiatry*, **5**, 221–235, (1967).
[7] S.D. Duncan, ‘Some signals and rules for taking speaking turns in conversations’, *Journal of Personality and Social Psychology*, **23**, 283–92, (1972).
[8] S.D. Duncan, ‘Towards a grammar for dyadic conversations’, *Semiotica*, 29–46, (1973).
[9] S.D. Duncan, ‘On the structure of speaker-auditor interaction during speaking turns’, *Language in Society*, **2**, 161–180, (1974).
[10] S.D. Duncan, ‘Language, paralanguage, and body motion in the structure of conversations’, in *Language and Man. Anthropological Issues*, eds., W.C. McCormack and S.A. Wurm, 239–268, Mouton, The Hague, (1976).
[11] S.D. Duncan and G. Niederehe, ‘On signalling that its your turn to speak’, *Journal of Experimental Social Psychology*, **10**, 234–47, (1974).
[12] E. Goffman, *Forms of Talk*, Oxford University Press, Oxford, 1981.
[13] C. Goodwin, *Conversational Organization: Interaction between Speakers and Hearers*, Academic Press, New York, 1981.
[14] U. Hadar, T.J. Steiner, and Clifford F. Rose, ‘Head movement during listening turns in conversation’, *Journal of Nonverbal Behavior*, **9**(4), 214–228, (1985).
[15] John Heritage, ‘A change-of-state token and aspects of its sequential placement’, in *Structures of Social Action*, eds., J. M. Atkingson and J. Heritage, Cambridge University Press, Cambridge, (1984).
[16] Adam Kendon, ‘Movement coordination in social interaction: some examples described’, *Acta Psychologica*, **32**, 100–125, (1970).
[17] R.M. Maatman, J. Gratch, and S. Marsella, ‘Natural behavior of a listening agent’, in *5th International Conference on Interactive Virtual Agents*, Kos, Greece, (2005).
[18] Catherine Pelachaud and Massimo Bilvi, ‘Computational model of believable conversational agents’, in *Communication in Multiagent Systems*, ed., Marc-Philippe Huget, volume 2650 of *Lecture Notes in Computer Science*, 300–317, Springer-Verlag, (2003).
[19] I. Poggi, ‘Backchannel: from humans to embodied agents’, in *Conversational Informatics for Supporting Social Intelligence and Interaction - Situational and Environmental Information Enforcing Involvement in Conversation workshop in AISB’05*, University of Hertfordshire, Hatfield, England, (2005).
[20] A.E. Scheflen, ‘The significance of posture in communication systems’, *Psychiatry*, **27**, 316–331, (1964).
[21] Emanuel A. Schegloff, ‘Discourse as interactional achievement: Some uses of ”uh huh” and other things that come between sentences’, in *Analyzing discourse, text, and talk*, ed., D. Tannen, 71–93, Georgetown University Press, Washington, DC, (1982).
[22] Emanuel A. Schegloff, ‘Issues of relevance for discourse analysis: Contingency in action, interaction and co-participant context’, in *Computational and Conversational Discourse. Burning issues - An interdisciplinary account*, eds., Edward H. Hovy and Donia R. Scott, 3–35, Springer, (1996).

	#Ans	AG	DA	AC	RE	LI	DL	BE	DB	UN	NU	BO	IN	NONE
N	28	21.43		39.28		7.15	3.75	3.75		17.86		3.75	3.75	
NS	39	17.95		15.38		35.90		5.13		5.13		2.56	17.95	
NRE	60	38.33		20	1.67	6.67		6.67		13.33		3.33	10	
RLE	42	14.28	2.38	4.77	4.77	2.38		7.14	23.80	4.77	9.53	11.90	7.14	7.14
TSE	32		12.5		6.25		9.38	3.12	18.75	3.12	31.25	6.25	9.38	
ER	25		4		4		12		20		12	32		16
TG	43	2.33	4.65	2.33	4.65	2.33	6.97	2.33	9.30	2.33	4.65	37.20	2.33	18.60
SFT	37	2.70	24.33		24.33		18.91		8.10		16.22	2.70	2.70	
FT	30		36.67	3.33	20	3.33	3.33		13.34		13.34	3.33	3.33	
S	40		25		40		20		7.5		2.5			5
SF	32		37.5		37.5		6.25		6.25	3.13	9.37			
TF	42	4.76	19.04	2.39			9.52	2.39	19.04	7.14	21.42	9.52	2.39	2.39
TRE	40	2	10	10		5	5	17.5	12.5		7.5	15	7.5	5
EWO	37	5.40	5.40	5.40			2.70		16.23	5.40	5.40	5.40	13.53	35.14

Table 3. Statistical results en percentage. On the rows there are the signals, while on the column there are the meanings. In Table 4 a reminder of the meaning of the abbrevations.

Signals		Meanings	
N	a single head nod	**AG**	agree
NS	a head nod with a smile	**DA**	disagree
NRE	a head nod and a raise of the eyebrows	**AC**	accept
RLE	a raise of the left eyebrows	**RE**	refuse
TSE	a head tilt on the left and sad eyebrows	**IN**	interested
ER	the eyes roll up in the head	**BO**	bored
TG	a head tilt on the right and the gaze turns on the right and down	**BE**	believe
SFT	a head shake, a frown and a tension in the lips that tighten getting thinner	**DB**	disbelieve
FT	a frown and a tension in the lips that tighten getting thinner	**UN**	understand
S	a head shake	**NU**	not understand
SF	a head shake and a frown	**LI**	like
TF	a head tilt on the left and a frown	**DL**	dislike
TRE	a head tilt on the right and a raise eyebrows		
EWO	the eyes wide open		
#Ans	number of answers given for the corresponding signal		

Table 4. Meanings of the abbrevations.

[23] N. Ward and W. Tsukahara, ‘Prosodic features which cue back-channel responses in english and japanes’, *Journal of Pragmatics*, **23**, 1177–1207, (2000).

[24] V.H. Yngve, ‘On getting a word in edgewise’, in *Papers from the sixth regional meeting of the Chicago Linguistic Society*, pp. 567–77, Chicago: Chicago Linguistic Society, (1970).