Cooperation and flexibility in multimodal communication

Jens Allwood, Dept. of Linguistics
University of Göteborg

1.   Introduction

The design of cooperative dialog systems, new human-computer interfaces (cf. Schomacher et al 1995) and avatars in various types of virtual environments can all be improved by better knowledge of features of real human-human multimodal communication. In this paper I discuss the nature of cooperation in dialog. I will also discuss what might be called features of flexibility and conflict prevention and how they are related to cooperation. I will illustrate the features of flexibility and conflict prevention by examples drawing on videorecorded and transcribed human-human dialog. The main focus will be on nonverbal gestural means, since verbal means are somewhat better known. The paper is intended to illustrate how new ideas about the design of dialog systems (cf Pandzic et al 1996, 1997) will also lead to an interest in aspects of human-human communication that have received less attention so far.

2.   Cooperation in communication

The notion of cooperation may be defined as follows, cf Allwood 1976 and Grice 1975: Two or more agents may be said to cooperate to the extent that they

1)   Consider each other cognitively in interaction
2)   Have a joint purpose
3)   Consider each other ethically in interaction
4)   Trust each other to act according to 1-3.

If all four requirements are met, we have ideal cooperation. Normally, only some of the requirements are met, and we may therefore speak of degrees of cooperativeness. Cooperation is not an "all or nothing" phenomenon but a matter of degree. In dialog, participants can be cooperative to a greater or lesser extent, rather than just cooperative or not cooperative.

In dialog, participants cooperate through the contributions they make. These contributions in direct spoken interaction are multimodal, i.e., a *yes* can be replaced or accompanied by a nod, a *no* by a shake of the head, etc.

Let me now exemplify the four requirements mentioned above. The first requirement for cooperation - "cognitive consideration" - means that A in interacting should attempt to perceive and understand what B is doing. This is a kind of base requirement for there to be any cooperation and communication at all.

The second requirement of having a "joint purpose" means that both global and local purposes are jointly pursued. For example, A and B can cooperatively pursue the activity of teaching if they restrict their behavior in accordance with the roles of "teacher" and "student". On a local level it means that evocative intentions of preceding contributions are evaluated and responded to. So if A asks B a question, B should, if he/she is cooperative, evaluate whether he/she can answer the question and, if he/she can, do so.

The third requirement - "ethical consideration"- means that A should consider and try to make it possible for B to continue interaction as a rational motivated agent, i.e., he/she should have correct information, not be given unnecessary pain, and be allowed to act as freely as possible.

The fourth requirement for ideal cooperation - "trust" - means that A not only tries to pursue joint purposes while ethically and cognitively considering B, but also believes B to do the same, i.e., trust that B is acting in A's best interests.

Some characteristics of cooperation in dialog

Some of the main ways in which participants contribute to cooperation in dialog are therefore the following:

1. Cognitive consideration: Contributions are based on cognitive consideration if they are relevant to what other interlocutors have said or done.

2. Joint purpose: Contributions are also cooperative if they further a joint purpose. This could be a global purpose of the activity or a more local purpose. Global purposes like negotiating, card playing or planning are sometimes but not always tied to specific topics, the pursuit of which furthers the joint purpose. Global structuring of an activity can often be indicated by various sequence markers, like *O.K.* or *right.*

3. Ethical consideration: Contributions show ethical consideration if they take other interlocutors' interests into account. One of the main ways in which this can be done is through supportive feedback signals of various types.

4. Trust: Contributions show trust if they rely on the good will and ethical consideration of other participants. This can, for example, be done by admitting fault or weakness, trusting in the good will of others.

3.   Multimodal communication

Normal face-to-face communication is multimodal, cf. Duncan 1974 and Heritage 1984, employing several modalities of production and perception in order to share information.

The two primary modes of production are speech and various types of bodily gestures, perhaps primarily facial gestures, head movements and manual gestures. The two primary modes of perception are, accordingly, hearing and vision.

In this case, the spoken message will normally predominate, while bodily gestures provide additional information. The gestures are often, in turn, reinforced by prosody, resulting in a situation where utterances through words and grammatical constructions are given supplementary support by gestures and prosody.

Several different relations can hold between the messages produced in the different modalities. One possibility is that they are more or less identical - one mode of expression adding redundancy and reinforcement to the other. Another possibility is that they are partially or totally different - one mode of expression adding information to the other. A third possibility is that one modality changes the message given by another modality.

If information is added, three of the possible supplementary relations between modalities in talk are the following:

(i)   Adding emotions and attitudes to verbal messages by prosody and gesture. In this way, factual and affective information may be activated simultaneously, which is crucial to interpersonal cooperation. Another effect is that information about the identity of the speaker is given simultaneously with the message.

(ii)   Adding illustrations to verbal message by iconic or conventional gestures. In this way, communication can become livelier and more engaging. Such gestures can also be used to specify or make the verbal message more precise, thereby increasing mutual comprehension of shared information.

(iii)   Adding information pertaining to interactive communication management, i.e., giving information about who the intended addressee is, the length of one's utterance, a wish to speak etc.

When it comes to change, multimodal communication also adds to the possibilities of adapting or changing the content of what is communicated flexibly on-line. For example, it adds gesture to the means whereby a self- correction or other type of change can be made clear to an interlocutor (cf. Allwood, Nivre and Ahlsén, 1990). There are also other ways of changing the verbal content through gestural modification; a message which, on a verbal level, looks like a straightforward statement can, through gestural modification, become something very different, like a joke or an instance of sarcasm.

4.   Cooperation and Mutual Flexibility

If the analysis of cooperation given above is accepted, one of the consequences is that dialog and communication can exhibit varying degrees of cooperativeness and be cooperative in several different ways. Two concepts which are closely related to several features of cooperation are what might be called "flexibility" and "conflict prevention". They both involve cognitive and ethical consideration as well as trust. They encompass a range of phenomena in the service of the goals of keeping options open and preventing conflict or disharmony. They are therefore often a kind of prerequisite for more constructive and substantial cooperation in the pursuit of some joint activity purpose. Some of the means whereby flexibility and conflict prevention are achieved are verbal, other are nonverbal. Before discussing the various means in more detail, I will, however, briefly try to characterize the phenomena I have in mind.

Means of mutual flexibility and conflict prevention are communicative means which are used by participants to maintain one or several of the following goals.

1.   Mutual friendliness
2.   Lack of tension (tension release)
3.   Lack of need to defend a position
4.   Admitting weakness or uncertainty
5.   Lack of attempts to overtly impose opinions on others
6.   Coordination of attention and movements
7.   Giving and eliciting feedback expressing mutual support and agreement
8.   Showing consideration and interest
9.   Invoking mutual awareness and beliefs

The properties in the list are not mutually exclusive but can be related to each other in various ways. Indeed, they often condition or support each other. I will now discuss the properties one by one and relate the notion of cooperativeness to them.

1.   Mutual friendliness: Friendliness is related to cognitive and ethical consideration as well as to trust. Mostly, you don't hurt your friends, you trust them and you try to attend to their needs. Friendliness is, thus, a holistic attitude which promotes good feeling and a wish for continued interaction which is a good basis for cooperation. Even though friendliness is perhaps not a necessary feature of cooperation, cooperation is much easier if it is present and much harder if it is absent.

2.   Lack of tension: Persons who intend to cooperate can often feel some tension, especially if they are not familiar with each other. If this tension can be diminished or released, interaction can proceed with increased flexibility and smoothness. Recognizing the need for, or at least allowing for, participating in tension release can thus be related to ethical consideration and trust.

3.   Lack of need to defend a position: If we analyze the expressive and evocative functions of a statement, (cf. Allwood 1995), we find that statements, when used seriously in the most typical manner, serve, on the one hand, to express beliefs and, on the other, to evoke similar beliefs in interlocutors. If we add to

this an analysis of obligations, we find that statements when used seriously impose a requirement on the speaker of sincerity, grounding and consideration of the interlocutor. Sincerity implies stating only what is believed to be true, grounding implies stating only what one has some grounds or evidence for, and consideration of the interlocutor means that one takes the interlocutor's level of understanding and interests into account.

Although all these features of statements are desirable in some contexts, they do not necessarily allow for flexibility. Speakers have therefore adopted a number of strategies which allow statements to be used in a more flexible way. One of the ways is to make what is claimed by the statement so vague that it almost has to be true. This can be achieved by so-called hedges or other qualifiers that make statements vaguer. Two examples in English are *sort of* and *kind of.*

It is sort of ugly
It is kind of ugly

Use of these expressions has the effect of making a statement both more like a platitude and easier for an interlocutor to accept. It also keeps your options open for a flexible specification if you were to be challenged. Another way to weaken a statement is to remove or soften its truth claim. This can be done by indicating that it is not serious, or that it least might not be serious.

Both the move of making statements vaguer and the move of making them less serious have the effect of diminishing their clarity and weakening their truth claims. In the case of a pure joke, the truth claim completely disappears. The effect of this, in turn, reduces the need to defend the statements, since it is not clear whether anything specific really has been claimed, which, in turn, removes or weakens the obligations of sincerity and grounding. Both moves, thus, represent ethically acceptable ways of lifting ethical restrictions which, if broken, might otherwise lead to rigidity and disharmony.

4. Admitting own weakness or uncertainty: Another way to increase flexibility is to admit weakness and uncertainty. This makes any statement or opinion expressed open to revision and removes the need to defend. If not exaggerated, this type of move can, by showing trust, help to create further trust.

5. Lack of attempts to overtly impose opinions on others: If we return to the analysis of the communicative functions of a statement given above, we see that its evocative function is that the listener shares the belief expressed in the statement; i.e. an attempt to influence the listener. If the listener is cooperative, he/she has to evaluate whether he/she is willing and able to share the belief and in some way indicate the result of this evaluation to the speaker. If beliefs are expressed which the listener either does not want to share or is unable to share, this means that there is a risk of overt non-mutuality of beliefs. Since this is not desirable in discourse which is supposed to be smooth and flexible, there will be attempts to reduce the evocative force of the statements which are made. This can be achieved by the means just discussed, i.e. making statements vague, uncertain or non-serious.

An alternative strategy is not to reduce the evocative force of the statement but to accompany it by clear indications of friendliness, in this way dampening the potential irritation that might result from nonagreement.

6. Coordination of attention and movements: Flexibility and cooperation are also aided by coordination of attention and movements among speakers. Means of this are perhaps primarily nonverbal, such as attending to speakers by gazing at them, or moving in synchrony with other interlocutors. Coordination of attention and movements is closely related to cognitive consideration, where coordination of movement might even be on a more basic level than cognitive consideration.

7. Giving and eliciting mutual support and agreement: An important ingredient in creating an atmosphere of flexibility, trust and consideration can be achieved by giving positive and supportive feedback to other speakers. This promotes activity and reduces insecurity and inhibitions. Some interlocutors use this strategy even when they do not in fact share the beliefs expressed. From a short- term flexibility point of view, this can be effective but, from a long- term perspective, it may of course raise ethical problems.

One of the most important ways of giving support and signalling agreement is through linguistic and more generally communicative feedback, (cf. Allwood, Nivre and Ahlsén 1992 and Heritage 1984). Feedback signals are one of the main ways in which cooperation is pursued in dialog. Through communicative feedback, interlocutors can inform each other whether and in what way the following basic requirements of communication, but also of cooperation, are met:

(i) continuation of contact
(ii) perception
(iii) understanding
(iv) evaluation and response to evocative intentions

Feedback signals can be related to all requirements of cooperation. Basically, they indicate cognitive and ethical consideration of a speaker's attempt to bring about shared understanding by directly signalling information as to whether this is successful. Their use is therefore a prerequisite for trust and the pursuit of further goals which require shared understanding.

Feedback signals giving these kinds of information in direct face-to-face communication are multimodal. They involve prosodic, lexical and syntactic features of spoken utterances as well as bodily gestures. However, feedback is not only given in dialog, it is also actively elicited. By use of verbal means such as *right,* or tag questions, and nonverbal means such as raising a hand or an eyebrow, speakers actively seek feedback from other interlocutors in order to ascertain whether communication continues to be coordinated and cooperative.

8. Showing consideration and interest: Consideration and interest can be shown by attending to other speakers. It can also be shown by reacting clearly and by sympathetically showing, for example, surprise or pleasure at what others say. Finally, it is indicated through an interest in the reactions of others, something which can be overtly expressed by attempting to elicit feedback. Show of consideration and interest is, thus, a fairly direct manifestation of the cognitive and ethical consideration mentioned above as two of the features of cooperation.

9. Invoking mutual awareness and belief: Another function which is related to the ones already discussed is that of invoking rather than expressing consensus. In English, this can, for example, be achieved by use of the phrase *you know* which when added to a statement (cf. Aijmer 1996) indicates that what is stated is already shared by the interlocutor, thus obviating any need for controversy.

5.   Means of achieving mutual flexibility and preventing conflict

Let me now turn to a discussion of some of the means of achieving mutual flexibility and preventing conflict. As we have already noted, the means can be both verbal and nonverbal. My focus will mainly be on nonverbal gestural means.

5.1   Verbal means

Verbal means of expressing friendliness include various ways of showing liking and appreciation of other interlocutors, as well as refraining from showing dislike or non-appreciation.

As we have seen, verbal means also include use of humor, e.g. jokes, and making claims and evocative functions vaguer. They include expressions of uncertainty like *I don't know* and all the various means of giving and eliciting linguistic feedback by phrases such as *m, yes, yeah, sure, great, precisely*, *eh* and tag questions (cf. Allwood, Nivre & Ahlsén 1992). Finally, they include means of invoking consensus such as *you know*.

5.2   Nonverbal gestural means

The nonverbal gestural means of achieving mutual flexibility are perhaps even richer than the verbal ones. Perhaps they are also more basic, since interlocutors often believe or feel that nonverbal communicative expressions are more directly causally linked to the true feelings and attitudes of a speaker and thus more genuine. A classification of nonverbal means of communication can be very detailed, which is what is needed, for example, for an analysis of deaf sign language.

However, since my main purpose at present is to capture some of the main nonverbal means of mutual flexibility and conflict prevention, I will here use a simpler scheme of classification, including only the following gestural articulators:

1.   Head   -   head/movements e.g. nods, shakes, tilts
2.   facial gestures (other than smiles)
3.   gazing
4.   smiles
5.   Laughter
6.   Body posture
7.   Movements of arms and hands

For each of these gestural articulators, I will now mention some of their main functions, indicating by the use of "bullet signs" which of these serve as a means of achieving mutual flexibility or preventing conflict. In the tables below, the terms *indicate*, *express* and *show* are used more or less synonymously to refer to the manifestation of inner states through gestures. The list is based on an analysis of videorecorded and transcribed human-human conversations and is meant to be illustrative rather than exhaustive.

Gesture	**Function**
1. Head movements	
Shaking the head	- rejection, denial both as a proper turn and as overlapping with another person's turn
	• agreement with negated statement both as a proper turn and as overlapping support of another person's statement
	- support of own negated statement
	• indicating bewilderment over content in own or other person's statement
	• non-insistence on point made
Nodding (the head)	• agreement both as supportive overlap and as own turn
	- reinforcing own turn
	- rejection of negative statement
Tilting head to one side	• accompanying own objection to soften it
	• eliciting feedback
	• indicating insecurity, shyness
Rocking head	• indicating lack of knowledge,
	• indicating humor
	• indicating bewilderment
Jerking head backwards	• "what do I know"
Raising head	• indicating surprise
Forward rocking	• eliciting feedback
Pushing head forward	• indicating surprise
	• indicating that someone else's statement is noteworthy

2. ***Facial gestures***

Wrinkling eyebrows	• indicating lack of understanding of facts related
	• indicating something unpleasant in own utterance
	• indicating difficulty of finding the right word
	• indicating surprise at other's utterance
Raising eyebrows	• indicating surprise at other's utterance

3. ***Gaze***

Gazing around	• to elicit confirmation
at other interlocutors	- to announce new information
	• observing reactions of other
Gazing at own hands gesturing	- directing attention
Gazing down	• breaking contact when silence occurs, avoiding confrontation
	- indicating insecurity
Gazing at handling of artefacts	• allowing for break and decrease of concentration
Gazing at speaker	• attending to speaker
Seeking eye contact	• elicitation of feedback

4. ***Smiles***

Smiles	• indicating insecurity, uncertainty
	• giving confirmation
	• indicating friendliness
	• eliciting confirmation
	• removing seriousness
	• removing effects of own statement
	• apologizing
	• indicating self-irony
	• removing danger
	• indicating humour
	• indicating that something is daring or controversial
	• weakening opposition
	• indicating that something is unpleasant

5. ***Laughter***

Laughter	• releasing tension, collective and individual
	• showing agreement, consensus, collective and individual
	• expressing surprise
Laughter + smile	• expressing uncertainty
	• indicating that something is a joke
	• indicating insecurity
	• giving confirmation
	• eliciting confirmation

6. ***Body posture***

Leaning forward	• showing interest
Body contracted	• indicating insecurity
Moving shoulders	• indicating that something is to be taken as a rough estimate

7. ***Movements of arms and hands***

Fidgeting with clothes, hair	• expressing insecurity
	• tension release
Striking out	• tension release
Iconic illustrations	- supplementing content
Baton gesture	- emphasizing
Arms crossed	- indicating negative attitude
Moving artifacts	- obviating need for talk
Moving finger	- indicating waiting, expectation
Pointing	- symbolic or concrete deixis

If we try to summarize the observations on gestures in relation to the various ways to achieve flexibility and non-conflict mentioned above, we can see the following.

1. Mutual friendliness: Friendliness is above all expressed through smiles.

2. Lack of tension: Lack of tension is primarily achieved through laughter and hand movements, but also probably by downward gazing allowing for breaks of concentration.

3. Lack of need to defend a position: This is above all achieved by smiles softening the content, removing seriousness and apologizing, as well as by head movements indicating non-insistence and non-seriousness. There are also shoulder shrugs indicating vagueness.

4. Admitting weakness or uncertainty: This is achieved by head movements expressing lack of knowledge, lack of understanding, uncertainty, insecurity, bewilderment.

5. Lack of attempts to overtly impose opinions on others: The means mentioned in (3) and (4) above also serve to diminish the evocative force of the statements made.

6. Coordination of attention and movements: This is primarily indicated through coordinated hand movements and body postures.

7. Giving and eliciting mutual support and agreement: Support and agreement are primarily given through head nods, head shakes and smiles. Elicitation of feedback is done through gaze, head raising and smiles.

8. Showing consideration and interest: This can be seen in a gaze directed at the speaker or gazing around to see the reactions of other interlocutors. It can also be seen in clear, overt, friendly bodily reactions to other speakers.

9. Invoking mutual awareness and beliefs: Perhaps this is chiefly done by attempting to create an atmosphere of mutual non-seriousness.

The most important of the bodily means used seem to be head movements, gazing, smiles and laughter, and the most important functions of flexibility and conflict prevention seem to be giving support, showing friendliness, releasing tension, indicating non-seriousness and admission of one's own weakness or uncertainty. All of the means and functions are frequent in normal face-to-face communication. In interchanges that are characterized by cooperation, flexibility and conflict prevention, one or other of the means seems to accompany almost every utterance. This implies that multimodality is a crucial means of achieving these goals in normal human-human dialog.

We have also seen that all three of the relations between verbal and nonverbal communciation discussed above – support, supplementary information and change – occur. Gestures can be used to support both your own contributions and those of others in different respects. They can be used for supplementary purposes, to give iconic or indexical illustrations, as well as for attitudinal information or tension release, or to change the message, for example, by weakening the expressive and evocative functions of different communicative acts.

6. Concluding remarks

This paper has explored some of the ways cooperativeness is multimodally manifested in dialog. Claiming that cooperativeness is a matter of degree, it is suggested that it can therefore be related to phenomena like coordination, flexibility and conflict prevention. An attempt to analyze this relation is made by relating cooperativeness to nine subgoals for flexibility and conflict prevention. A further attempt to clarify the role of flexibility and conflict prevention is made by examining some of the nonverbal (and verbal) means to achieve these goals. Hopefully, the types of communicative flexibility and conflict prevention that have been discussed are of a fairly generic nature, even though they are, in fact, based on particular empirical data from Swedish face-to-face conversation.

Hopefully, they can therefore serve both to throw light on human dialog and as an addition to knowledge about communicative functions which could be incorporated in the design of human-computer interfaces, cooperative dialog systems, or avatars in virtual environments. We might, for example, pose questions such as the following: Should systems be friendly? Should they allow for release of tension? Should they sometimes be non-serious or vague? Should they be non-imposing? Should they be coordinated with the user? Should they give and elicit supportive or other types of feedback? Should they show consideration and interest, and should they be able to invoke mutual awareness and belief?

If the answer to any of these questions is yes, and there are already a number of systems which have some of these features, the next question is, of course, what means to use. With the advent of multimodal cooperative systems, it is more than

likely that a lasting source of inspiration will be the way cooperation, flexibility and conflict prevention is achieved multimodally in human-human dialog.

References

Aijmer, K. 1996. *Conversational Routines in English - Convention and Creativity.* London: Longman.

Allwood, J. 1976. *Linguistic Communication as Action and Cooperation.* Gothenburg Monographs in Linguistics 2, University of Göteborg, Department of Linguistics.

Allwood, J., Nivre, J. and Ahlsén, E. 1990. Speech Management – On the nonwritten Life of Speech. *Nordic Journal of Linguistics* 13, 3-48.

Allwood, J., Nivre, J. and Ahlsén, E. 1992. On the Semantics and Pragmatics of Linguistic Feedback. *Journal of Semantics,* 1992, also in *Gothenburg Papers in Theoretical Linguistics,* 64, University of Göteborg, Department of Linguistics.

Cassell, J. 1995. Speech, Action and Gestures as Context for Ongoing Task-oriented Talk. In *AAAI Fall Symposium Working Notes: Embodied Language.*

Duncan, S. 1974. Some Signals and Rules for Taking Speaker Turns in Conversations. in Weitz, S. (Ed.) *Nonverbal Communication.* New York: Oxford University Press.

Grice, H.P. 1975. Logic and Conversation. In Cole, P. and Morgan, J.L. (Eds.) *Syntax and Semantics* Vol 3: Speech acts. New York: Seminar Press, 41-58.

Heritage, J. 1984. A Change of State Roken and Aspects of its Sequential Placement. In Atkinson, M. and Heritage, J. *Structures of Social Action: Studies in Conversation Analysis.* Cambridge: Cambridge University Press.

Mehrabian, A. 1971. *Silent Messages.* Belmont, Calif. : Wadsworth publ. comp.

Pandzic, I., Capin, T., Thalmann, N. and Thalmann, D. 1996. Towards Natural Communication in Networked Collaborative Virutal Environments. http://miralabwww.unige.ch/ARTICLES/five96B.html

Pandzic, I., Capin, T., Lee, E., Magnenat, N. and Thalmann, D. 1997. A Flexible Architecture for Virtual Humans in Networked Collaborative Virtual Environments. http://miralabwww.unige.ch/ARTICLES/eurographics97B.html

Schomacher, L, Nijtmans, J., Camurri, A., Lavagetto, F., Morasso, P., Benoît, C., Guiard-Marigny, T, Le Goff, B., Robert-Ribes, J., Adjoudani, A., Defée, E., Münch, S., Hartung, K. and Blauert, J. 1995. A taxonomy of Multimodal Interaction in the Human Information System. *A Report of the ESPRIT PROJECT 8579*, WP1.