Some challenges for multimodal intercultural artificial communicators

Jens Allwood, Elisabeth Ahlsén, Jia Lu
SCCIIL Interdisciplinary Center
University of Gothenburg
E-mail: jens@ling.gu.se, eliza@ling.gu.se, .uuu.edu, jia.lu@gu.se

Abstract

This paper discusses some of the challenges in constructing multimodal intercultural artificial communicators. Such challenges concern the notions of intercultural communication and culture, e.g. what aspects of culture are investigated and what methods are used. The paper also presents two example of types of investigation that are needed and concludes with a short discussion of some consequences to be considered in constructing intercultural multimodal communicative agents.

1. Purpose

The purpose of the paper is to initiate a critical discussion of some of the difficulties in constructing communicative agents that are interculturally competent. In Allwood and Ahlsén (2009), we presented a framework for the design and evaluation of a Generic, Multimodal, Intercultural Communicator (a GMIC), containing a number of features that can be adapted to different cultural patterns. The present paper discusses the basis for such adaptations, i.e. the available information on cultural patterns of interaction.

Unfortunately, studies of intercultural communication and cross-cultural comparison of patterns of communication and intercultural communication are not characterized by agreement on what cultural differences exist in communication or indeed whether there are significant cultural differences in communication at all.

A very basic reason for this is the lack of studies of actual cultural differences in communication. Most cited differences, the most famous probably being Hofstede (1980, 1997), are based on questionnaires focusing on attitudes (rather than on communication) and using averages and factor analysis to derive results. Another common source of data is what might be called studies relying on anecdotal observation. Studies based on analysis of recorded actual interactions, comparing cultures with regard to different social activities are still too few in number.

In this paper, we want to explore two ways of increasing what might be called reliable information concerning cultural differences in communication by
(i) a review of reported results concerning cultural variation in turn management
(ii) reporting on differences in smiling behavior observed in Swedish and Chinese first acquaintance conversations.

2. Introduction

2.1 What is intercultural communication?

We can define intercultural communication between A and B as follows:

> A and B are engaged in intercultural communication iff they have different cultural back-grounds and the differences in cultural background in some way influence their communication.

If the cultural difference has no effect on the interaction, we do not regard it as intercultural communication. Thus, the definition allows for multicultural competence, which would not be the case with a wider definition.

In order to construct interculturally competent agents, we need access to reliable cross-cultural comparisons and if we want them to engage in interactive communication, we need a description and understanding of how

intercultural communication actually works. We assume that the two types of study are related, so that a better understanding of how cultural difference might affect intercultural interaction usually presupposes a cross-cultural comparison of the cultural backgrounds of the intercultural communicators.

The definition of intercultural communication as well as the two needed types of study just described lead to questions like
(i) What is culture?
(ii) How can cultures be compared to capture cultural differences?
(iii) How do cultural differences influence communication? Here we should note that other factors than cultural also influence communication, e.g. aspects of personality that are not cultural (genetic, disability etc.) functional activity requirements (e.g. harvesting or programming), language (if separated from culture), human nature, or the physical environment.

2.2 What is culture?

We define culture as patterns (similarities) between people, not given by natural necessity (cf. Allwood 1982). (The ability to walk or breathe is a group similarity, but a natural rather than cultural similarity.)

The patterns can concern the following four main aspects of culture:
(i) Thoughts (factual, emotional, volitional, values etc.)
(ii) Behavior
(iii) Artifact manufacture and use
(iv) Traces in Nature
Combinations of the four aspects above.

This is a broad definition of culture, which can be contrasted with the more narrow definitions found, for example, in Hofstede or Geertz (1973), where mentalistic notions of culture, mainly focusing on "thoughts" can be found.

2.3 How can cultures be compared?

The four aspects of culture provide a starting point for cross-cultural comparisons. We can compare thoughts, (e.g. compare attitudes, beliefs, emotions or values), behavior, artifact manufacture and use (e.g. tools, housing, clothing) or traces in nature. We can also compare combinations of the aspects. Here we should note that language and communication are among the phenomena that require a combination of the four aspects.

We should also note that a cross-cultural comparison and taxonomy based on one of the aspects does not automatically give access to the other aspects. So a taxonomy of attitudes (based on questionnaires which have been analyzed using factor analysis and averages, like the one presented in Hofstede (1980, 1997), does not necessarily say anything about behavior or artifacts. To acquire reliable data on the combination of aspects

present in human communication, human communication itself must be studied. This will often require a combination of methods like analysis of audio and video recorded interaction, interviews, questionnaires and available reported research. Relying on only indirect, reported data provided through questionnaires will usually not give reliable data on communication.

2.4 Cultural influence on communication?

What needs to be studied is human communication in a number of culturally significant activities, like education, negotiation, quarreling, preaching, discussing and counseling, where these activities are held constant when a comparison is made between two cultures, i.e. class-room teaching in China has to be compared with class-room teaching in France, so that we can say something about cultural differences in class-room teaching between China and France. In many studies of cultures, the level of abstraction is too high (e.g. Chinese versus French national culture) to allow a more precise understanding of why and how differences in communication occur.

In studying communication in different social activities, the following dimensions need to be described:
1. Dimensions of production of communicative behavior, e.g. vocal verbal and gesture.
2. Dimensions of interpretation (perception and understanding)
3. Dimensions of communicative interaction (turn management, feedback, sequencing, rhythm, spatial configuration).
4. Dimensions of context (especially the social activity for which the communication serves as an instrument.

Below we will now study twp parameters from the list in Allwood and Ahlsén (2009), namely turn management (an aspect of communicative interaction), and smiling (anaspect of communicative behavior). We will employ two different methods (i) a review of existing research reports and (ii) analysis of audio/video recordings of interaction. The two studies are meant to discuss and exemplify how cultural features that could be used in constructing culturally competent virtual agents can be obtained. In this way hoping to contribute to what we can claim about cultural differences in communication.

3.   Some observations on studies of culturally relevant turn management

Turn management (or turntaking) as an aspect of interaction has been studied quite extensively, not least from the point of view of cultural variation. The area was made popular by Sacks, Schegloff and Jefferson (1974), who outlined a set of rules for turntaking, basically stating that we strive for one speaker at a time with no long pauses between speakers and that the current speaker generally selects the next speaker. If no speaker is selected, the turn goes back to the last speaker and if there is a longer pause anyone can take the turn. Turn Turn transition relevance points i.e. points were a

speaker shift can be made, were also discussed.

3.1 Collectivistic (high-context) vs. Individualistic (low-context)

In cross-cultural comparisons of turn management patterns, cultures are often divided into two types, individualistic and collectivistic, based on the attitude based classification of national cultures made by Hofstede (1980), or high context and low context cultures, based on work by E. T. Hall (REF). Sometimes the two classifications are combined, so that collectivistic and high context and individual and low context are merged. Two of the cultures that are classified as collectivistic/high context are Japanese and Chinese culture, while U.S., Canadadian, U.K, French, "European" or in some studies "Western" cultures have been classified as individualistic/low context (cf., for example, studies by Gudykunst 1983, Ng et al. 2000, Yamada 1992). The studies are mainly based on data from business meetings, but also from family conversations and usually compare an individualistic/low context culture with a collectivistic/high context culture.

There are not only cross-cultural comparisons, but also studies of intercultural communication and comparisons between intra- and intercultural interactions, especially for business meetings in multinational companies. Here we find the same division into the two types of cultures (individualistic and collectivistic) (cf. Du-Babcock and Tanaka 2010, Bilbow 1992). Some authors, for example Tanaka (2006a, 2006b) emphasize also other factors, such as linguistic and power differences, but the collectivistic-individualistic division is still dominant in these studies.

Claims from the above comparisons are that a culture claimed to be collectivistic/high context shows the following features of turn management: an even distribution of turns between the speakers (also including the initiator of the topic), shorter turns and fewer turns directed at a single addressee in group interactions. In intercultural interactions with individualistic/low context cultures they also show less verbal activity/more silence and they produce more speech in intracultural than in intercultural interactions.

3.2 Studies of turn transitions

Another difference that has been studied concerns the timing of turn transitions, where cultures with a strong preference for "one speaker at a time" and a high tolerance for silent pauses between turns, e.g. Finnish, have been contrasted with cultures with frequent interrupting overlaps and a high tolerance for interruptions (sometimes called "high involvement" style), e.g. New York Jewish or Mediterranean interaction patterns (Tannen 1982, 1984 Fant 1989). Fant suggests that the latter style in Spanish business meetings either can be said to provide counter-evidence to the rules for turntaking proposed by Sacks et al., since the Spanish interactions contain frequent overlaps. Fant therefore suggests that the rules of Sacks et al. rather than empirically descriptive should be seen as an

"idealized type" of turn taking.

A possibly different result was obtained in a fairly recent study by Stivers et al (2009), who compare turn management data from naturalistic interactions in 10 different languages/cultures and conclude that differences are only "quantitative" and that all the cultures seem to strive for "one speaker at a time" with no pauses between turns. The patterns are similar, only differing in time within a certain range, and the authors claim that humans have a specific sensitivity within this range that makes small differences in timing be perceived as big. They argue in favor of a universal turn taking pattern with an evolutionary basis. The relative homogeneity of the results in this study, can, however, be affected by the fact that only the subset of turn transitions occurring in "question-answer" pairs were extracted and used for the study.

3.3 Comments - different types of overlaps, different activities, different subsequences

Two main points are worth considering in relation to the type of studies mentioned above.

1) Overlaps can be cooperative. Much overlap is often supportive feedback during another person's turn (Allwood 2001b) or high involvement simultaneous talk, which is quite different from intrusive interruptions in a power struggle for the floor. It therefore seems necessary to differentiate overlaps qualitatively according to different features of content and function, in order to be able to make any more substantial claims about cultural patterns of overlap in turntaking.

2) Interaction patterns are determined by many factors, such as the *background of the speakers* involved, the *social activity* being pursued, the *roles* of the participants in this activity, the *power relations* in terms of roles and *position* and in terms of *command of the language spoken*, the *communication medium*, the *group size* etc. (cf. Allwood 2000, 2001a), In the light of the complex relations between these factors, the quite abstract dimension collectivistic vs. individualistic (or high context vs. low context) in order to explain differences in patterns of turn management could very well be overused or exaggerated in importance in relation to other factors, which are not always taken into account as competing or contributing explanations.

3) Different exchange types and subsequences are likely to lead to differences in turn taking patterns. (For example, only question-answer sequences are not necessarily representative for turn management in all other exchange types.)

We conclude that, so far, mostly divisions of cultures based on attitudinal questionnaires leading to fairly abstract dimensions have been used in studies of turn management cultural patterns. Different types of overlaps and pauses have not generally been considered and used systematically for a more fine grained analysis. Nor have the different determining factors behind the actual interaction patterns found in a specific activity (e.g. a business meeting) always been taken into account. Even if some available studies focus on a specific activity in two cultures, this provides only partial information about cultural variation in turn management and more activities need to be studied.

4.   A study of smiles in Chinese and Swedish first encounters

4.1 Introduction and purpose

Smiles are a common type of facial expression that plays an important role in human communication, often communicating emotions and attitudes and managing the interaction flow. A main function of smiles is to facilitate friendly communication. However, the number of empirical studies on cultural (or gender) variation in smiles is relatively small. This study attempts to investigate smiling in four Chinese-Chinese and four Swedish-Swedish video-recordings of university students' first encounters.

The paper primarily investigates three questions. What are some of the features of how and why people smile to strangers in human communications? Is there a cultural difference variation between Chinese and Swedish subjects? and Is there any gender variation?

4.2 Materials and method

The study is based on four Chinese-Chinese and four Swedish-Swedish video-recordings of face-to-face dyadic dialogs between university students studying in Sweden. In order to make a comparative study of smiling with respect to differences in culture and gender, each subject was studied in two mono-cultural dialogs varying in the gender of the other interlocutor (see Table 1). For instance, Chinese female Cf2 was communicating with both a Chinese female Cf1 in dialog 1 and a Chinese male Cm1 in dialog 2. The communicative languages are Chinese and Swedish respectively.

Recording	Participants	Analyzed Time	Language
Dial.1	Cf1—Cf2	7:00 min.	Chinese
Dial.2	Cf1—Cm2	7:00 min.	Chinese
Dial.3	Cm1—Cf2	7:00 min.	Chinese
Dial.4	Cm1—Cm2	7:00 min.	Chinese
Dial.5	Sf1—Sf2	7:00 min.	Swedish
Dial.6	Sf1—Sm2	7:00 min.	Swedish
Dial.7	Sm1—Sm2	7:00 min.	Swedish

Dial.8	Sm1—Sf2	7:00 min.	Swedish

Table 1: Video-recordings
(Note: C=Chinese, S=Swedish, f=female, and m=male.)

The study focuses on how two strangers who have no earlier acquaintance go about the task of getting to know each other. Each interaction was video-filmed by three video cameras (left-, center-, and right-position) with interlocutors face to face in a standing position. Each video-recording lasted approximately seven to ten minutes, and the first seven minutes were analyzed in detail in this study. The video-recorded data was transcribed and checked according to the GTS (Göteborg Transcription Standard) version 6.2 (Nivre, 1999). They were manually annotated by means of the MUMIN multimodal coding scheme for feedback (Allwood, Cerrato, Dybkjær, Jokinen, Navaretta & Paggio, 2005; Navarretta, Allwood, Cerrato, Jokinen & Paggio, 2006).

4.3 Smiles with and without vocal verbal or other gestural accompaniment

In the next sections, smiles wihout and with other vocal verbal or gestural accompaniment and the interactive communication management (ICM) functions of smile are analyzed.

4.3.1 Analysis and results – smiles with and without vocal verbal or other gestural accompaniment
In the data, smiles are used on their own or accompanying words or other gestures.
 In total, Swedish subjects smiled more than Chinese, with a frequency of 109 compared to 78 (see table2). Chinese females smiled more than males (with 44 compared to 34), but Swedish females and males (55 and 54) have roughly the same frequency.

Culture	Gender	Smile only	Smile accompanying words or gestures	Total
Chinese	Female	4	40	44
	Male	5	29	34
	Total	9	69	78
Swedish	Female	13	42	55
	Male	13	41	54
	Total	26	83	109
Total		35	152	187

Table 2: Chinese and Swedish female and male differences in using smile.

Table 2 shows that Chinese female and male subjects smile without accompaniment (vocal verbal or gestural) with roughly the same frequency (4 and 5 respectively); whereas, Chinese females smiled almost 1.5 times as

much with other accompaniment as males (40 compared to 29). The frequency of smiles for Swedish females and males is roughly the same, both without (13 and 13) and with accompanying words or other gestures (43 and 41). The accompaniments of smiles fall into three categories: other gesture (83), vocal verbal (56) and gesture + vocal verbal (13).

4.3.2 Analysis and results - gestural accompaniment of smiles
If we consider smiles that accompany other gestures, the most common gestures accompanying smile are presented in Table 3. Both Chinese and Swedish subjects combine smiles most often with single or repeated down-nods (with a frequency of 21) and hand movement (7).

Accompanied Gestures	Chinese			Swedish			Tot
	F	M	Tot	F	M	Tot	
Down-nod	2	0	2	3	4	7	9
Down-nods	8	0	8	2	2	4	12
Head complex	1	1	2	0	0	0	2
Gaze at	0	3	3	0	0	0	3
Gaze sideways	0	3	3	0	0	0	3
Gaze around	0	2	2	0	0	0	2
Hands	1	2	3	2	2	4	7
Eyebrow raise	1	1	2	0	0	0	2
Up-nod	0	0	0	1	1	2	2
Up-nods	0	0	0	2	0	2	2
Gaze down	0	0	0	1	1	2	2
Chuckle	0	0	0	3	1	4	4
Laughter	0	0	0	1	1	2	2
Total	13	12	25	15	12	27	52

Table 3: The most common gestures accompanying a smile (with a frequency of more than 1 for the 'Total')

We can note that with smiles repeated down nods are more frequent in the Chinese data, while single down nod is more frequent in the Swedish data. Gaze and eyebrow movement is used are only used with smiles in the Chinese data.

4.4 Analysis and results – Interactive Communication Management (ICM) functions of smile

Smiles also play a role in interactive communication management (ICM), which includes feedback, turn giving, and feedback+turn giving.

ICM function	Chinese			Swedish			Tot
	F	M	Tot	F	M	Tot	
Feed-back	26	18	44	33	20	53	97
turn giving	6	3	9	1	2	3	12
Feed-back+ turn giving	4	4	8	1	0	1	9
no ICM func-tion	8	9	17	19	33	52	69
Total	44	34	78	54	55	109	187

Table 4: The ICM functions of smile.

69 (37%) of the smiles do not have an ICM function, while 118 (63%) do, Among the smiles that are used for interactive communication management, the most common function is feedback; 57% of all smiles are used for feedback. Chinese females have slightly more feedback smiles than Chinese males (68% vs 65%) and Swedish females used smile as feedback and feedback+turn giving much more often than Swedish males (63% versus 36%).

4.5 Concluding remarks - smiles

Thus, Swedish subjects smile more than Chinese. Females smile more than males, especially with other gestural or vocal-verbal accompaniment. Both Chinese and Swedish subjects smile more with interlocutors of the other gender than with the same gender.

At the moment, it is unclear what Hofstede inspired attitudinal factor could explain the fact that there are more smiles in Swedish first acquaintance conversations than Chinese. Does it show that Swedes are more individualistic, collectivistic, uncertainty avoiding, feminine, egalitarian, short term directed?

In general, in this type of data on cultural differences, it is important to consider the specific context and possible influencing factors, such as, for example, gender differences (which are reported here), age differences and social role differences (the participants are all students) and the possible influence of living in Sweden for the Chinese participants.

The safest policy seems to be to note the observed difference and wait for confirmation from further observation. If it turns out to be tenable to claim that Swedes smile more in conversation than Chinese, perhaps this is a fact about conventional communicative behavior that cannot readily be explained by attitudinal factors. In this it could be similar, for example, to the absence or presence of the definite article in a language

5. Concluding remarks

In this paper, we have pointed to some challenges that need to be faced in constructing culturally savvy and sensitive communicative agents.

In addition to the issues raised above, we may now ask how information about cultural differences in turn management should be used in constructing communicative agents. In cultural awareness training programs, the culturally competent agents can act according to what is known about the culture they represent, in this way providing training for persons who want to learn how to interact with members of the culture. But what happens outside of the educational/training context? Should Chinese and Japanese agents strive for an even distribution of turns and produce more silence or New York Jewish or Mediterranean agents interrupt their interlocutors or should they try to adapt to the style of the user? This, in turn, raises the question of how the style of the user can be determined.

The basic problem is that we do not have enough certain information about behavioral cultural variation. We do not know how it varies with contextual circumstances like social activity. We do not know how to safely move from collective generalizations about a communicative pattern to the actual behavior of a particular individual in a particular situation. Other factors than cultural, like personality and gender, also play a role. One of the problems here is that we have no good way of linking claimed abstract national attitudinal differences to actual behavior.

To remedy the problem, we suggest that a combination of methods is needed. However, we probably need less of the studies that hitherto have dominated the field, i.e. interviews and questionnaires probing attitudinal differences and we need more of the kind of studies we have exemplified in this paper, i.e. (i) studies and reviews of the claims that have been made in the literature and (ii) direct analysis of recorded communicative interaction in different social activities.

6. References

Allwood, J. (1982). Finns det svenska kommunikationsmönster? In "*Vad är svensk kultur*", *Papers in Anthropological Linguistics* 9, University of Göteborg, Dept of Linguistics.

Allwood, J. (2000). An Activity Based Approach to Pragmatics In Bunt, H., & Black, B. (Eds.) *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*. Amsterdam, John Benjamins, pp. 47-80.

Allwood, J. (2001a). Capturing Differences between Social Activities in Spoken Language. In Kenesei, I., & Harnish, R.M. (Eds.) *Perspectives on Semantics,*

Pragmatics and Discourse. Amsterdam: John Benjamins, pp. 301-319.

Allwood, J.(2001b) Cooperation and Flexibility in Multimodal Communication. In *Cooperative Multimodal Communication* Harry Bunt and Robbert-Jan Beun (Eds.). Lecture Notes in Computer Science 2155 Springer Verlag, Berlin/Heidelbert, Also in *SALSA - Symposium About Language and Society - Austin*. Texas Linguistic forum. Vol. 44, Nos. 1 & 2, 2002, pp. 21-17.

Allwood, J. ; Ahlsén, E. (2009). Multimodal Intercultural Interaction and Communication Technology - A conceptual framework for designing and evaluating Multimodal Intercultural Communicators. In: Kipp, M., Martin, J.-C., Paggio, P., & Heylen, D. (eds) (2009). *Multimodal Corpora*. Berlin:Springer. (Lecture notes in computer science, 5509). pp. 160-175.

Allwood, J., Cerrato, L., Dybjær, L., Jokinen, K., Navaretta, C. & Paggio, P. (2005). The MUMIN Multimodal Coding Scheme. *NorFA Yearbook* 2005.

Bilbow, G (1992) Pragmatic failure in cross-cultural business meetings. In T. Boswood, R. Hoffman & S. Hsia (Eds) *Perspectives on English for professional communication*. Hong Kong: City Polytechnic of Hong Kong, pp. 169-181.

Du-Babcock, B. & V Tanaka, H. (2010). Turn-taking behavior and topic Management Strategies of Chinese and Japanese Business Professionals. *Proceedings of the 75th Annual Convention of the Association for Business Communication*, Oct 27-30, Chicago, Illinois.

Fant, Lars M. (1989), Cultural mismatch in conversation: Spanish and Scandinavian communicative behaviour in negotiation settings. *Hermes Journal of Linguistics* 3. Århus: Aarhus School of Economics. 247-265.

Geertz, C.(1973). *The Interpretation of Culture*. New York: Basic Books.

Gudykunst, W. (1983). Uncertainty reduction and predictability of behavior in low and high context cultures. *Management Communication Quarterly* 31, 49-55.

Hall, E. T. (1959). *The Silent Language*. Garden City, N. Y.: Doubleday.

Hall, E. T. (1966). *The Hidden Dimension*. Garden City, N. Y. Doubleday.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverley Hills, CA: Sage.

Hofstede, G. (1997). *Cultures and Organizations: Software of the Mind*. McGraw-Hill, New York.

Li, H. Z. (2001) Cooperative and intrusive interruptions in inter- and intracultural dyadic discourse. *Journal of Language and Social Psychology* 20: 1-24.

Navarretta, C., Allwood, J., Cerrato, L., Jokinen, K., Paggio, P. (2006). The MUMIN Anotation Scheme for Feedback, Turn Management and Sequencing. In Proceedings from the Second Nordic Conference on Multimodal Communication. *Gothenburg Papers in Theoretical Linguistics 92*. Göteborg University.

Ng, S. H. Ng, Loong, C. S. F., An Ping, H., Liu, J. & Weatherall, A. (2000). Communication correlates of individualism and collectivism. *Journal of Language and Social Psychology*, 19(2) p. 26-46.

Nivre, J. (1999). *Göteborg Transcription Standard. Version 6*.2, pp. 38. Göteborg: Göteborg University,

Department of Linguistics.

Sacks, H., Schegloff, E.A., Jefferson, G. (1974). A simplest systematics for the organization of Turn-Taking for conversation. *Language* 50(4), 696–735.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, E-Y. & Levinson, S. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS* 196(26):10587-10592.

Tanaka, H. (2006a). Emerging English speaking business discourse in Japan. *Journal of Asian Pacific Communication, 16*(1), 25-50.

Tanaka, H. (2006b). Corporate language policy change: the trajectory of management discourse in Japan, the oppressed or the oppressor? In R. Dolón, M. Labarta & J. Todolí (Eds.), *Critical Discourse Analysis* (pp.279-288). Spain: Universitat de Valencia.

Tannen, Deborah (1981). "New York Jewish Conversational Style." *International Journal of the Sociology of Language* 30, *The Sociology of Jewish Languages*: 133-149.

Tannen, D. (1984). Conversational style: Analysing talk among friends. Norwood, N.J.: Ablex.

Yamada, Haru (1992). *American and Japanese business discourse: A comparison of interactional styles*. Advances in Discourse Processes XLV. Norwood, N.J.: Ablex.