

Cultural Communication Idiosyncrasies in Human-Computer Interaction

Juliana Miehle¹, Koichiro Yoshino³, Louisa Pragst¹,
Stefan Ultes², Satoshi Nakamura³, Wolfgang Minker¹

¹Institute of Communications Engineering, Ulm University, Germany

²Department of Engineering, Cambridge University, UK

³Graduate School of Information Science, Nara Institute of Science and Technology, Japan

Abstract

In this work, we investigate whether the cultural idiosyncrasies found in human-human interaction may be transferred to human-computer interaction. With the aim of designing a culture-sensitive dialogue system, we designed a user study creating a dialogue in a domain that has the potential capacity to reveal cultural differences. The dialogue contains different options for the system output according to cultural differences. We conducted a survey among Germans and Japanese to investigate whether the supposed differences may be applied in human-computer interaction. Our results show that there are indeed differences, but not all results are consistent with the cultural models.

1 Introduction

Nowadays, intelligent agents are omnipresent. Furthermore, we live in a globally mobile society in which people of widely different cultural backgrounds live and work together. The number of people who leave their ancestral cultural environment and move to countries with different culture and language is increasing. This spurs the need for culturally sensitive conversation agents. Hence, our aim is to design a culture-aware dialogue system which allows a communication in accordance with the user's cultural idiosyncrasies. By adapting the system's behaviour to the user's cultural background, the conversation agent may appear more familiar and trustworthy.

However, it is unclear whether cultural idiosyncrasies found in human-human interaction (HHI) may be transferred to human-computer interaction (HCI) as it has been shown that there exist clear differences in HHI and HCI (Doran et al., 2003). To investigate this, we designed and conducted a user study with a dialogue in German and

Japanese containing cultural relevant system reactions. In every dialogue turn, the study participants had to indicate their preference concerning the system output. With the findings of the study, we demonstrate whether there are different preferences in communication style in HCI and which concepts of HHI may be applied.

The structure of the remaining paper is as follows: In Section 2, related work is presented. Subsequently, in Section 3, we present the cultural idiosyncrasies which we consider relevant for spoken dialogue systems. In Section 4, we present cultural differences between Germany and Japan supposed by the cultural models for HHI. The concept and the results of our study are presented in Section 5 before concluding in Section 6.

2 Significant Related Work

Brejcha (2015) has described patterns of language and culture in HCI and has shown why these patterns matter and how to exploit them to design a better user interface. Furthermore, Traum (2009) has outlined how cultural aspects may be included in the design of a visual human-like body and the intelligent cognition driving action of the body of a virtual human. Therefore, different cultural models have been examined and the author points out steps for a fuller model of culture. Georgila and Traum (2011) have presented how culture-specific dialogue policies of virtual humans for negotiation and in particular for argumentation and persuasion may be built. A corpus of non-culture specific dialogues is used to build simulated users which are then employed to learn negotiation dialogue policies using Reinforcement Learning. However, only negotiation specific aspects are taken into account while we aim to create an overall culture-sensitive dialogue system which takes into account cultural idiosyncrasies in every decision and adapts not only what is said, but also how it is said to the user's cultural background.

3 Integrating cultural idiosyncrasies

In a culturally aware intelligent conversation agent, the Dialogue Management (DM) sitting at the core of a dialogue system (Minker et al., 2009) has to be aware of cultural interaction idiosyncrasies to generate culturally appropriate output. Hence, the DM is not only responsible for what is said next, but also for how it is said. This is what makes the difference to generic DM where the two main tasks are to track the dialogue state and to select the next system action, i.e., what is uttered by the system (Ultes and Minker, 2014).

According to various cultural models (Hofstede, 2009; Elliott et al., 2016; Kaplan, 1966; Lewis, 2010; Qingxue, 2003), different cultures prefer different communication styles. There are four dimensions which we consider relevant for DM:

Animation/Emotion The display of emotions and the apparent involvement in a topic can be perceived very differently across cultures. While in some cultures the people are likely to express their emotions, in other cultures this is quite unusual.

Directness/Indirectness Information provided for the user has to be presented suitable so that the user is more likely to accept it. It has to be decided whether the intent is directly expressed (e.g. "Drink more water.") or if an indirect communication style is chosen (e.g. "Drinking more water may help with headaches.") whereby the listener has to deduce the intent from the context.

Identity Orientation Internalised self-perception and certain values influence the decisions of humans which depend on their culture. Hence, arguments addressing these values may be constructed based on the user's culture. In some cultures, the people are individualistically oriented which means that the peoples' personal goals take priority over their allegiance to groups or group goals and decisions are made individualistically. In other cultures, the people are collectivistically oriented which means that there is a greater emphasis on the views, needs, and goals for the group rather than oneself and decisions are often made in relation to obligations to the group (e.g. family).

Thought Patterns and Rhetorical Style Different cultures use different argumentation styles (e.g. linear, parallel, circular and digressive). In a discussion, the way arguments are presented

helps to provide necessary information to the user in an appropriate way. Additionally, some cultures have low-context communication whereas other cultures have high-context communication. In low-context communication, there is a low use of non-verbal communication. Therefore, the people need background information and expect messages to be detailed. In contrast, in high-context communication, there is a high use of non-verbal communication and the people do not require, nor do they expect much in-depth background information. Taking these facts into account means that the DM has to make a very detailed decision about how to present the information to the user.

4 Cultural differences

According to the aforementioned cultural models, various cultural differences are expected to exist between Germany and Japan. However, concerning *Animation/Emotion*, both Germans and Japanese are not expected to be emotionally expressive. According to (Elliott et al., 2016), both cultures avoid intensely emotional interactions as they may lead to a loss of self-control. Lewis (2010) affirms the fact that both Germans and Japanese don't like losing their face. Hence, emotionally expressive communication is not a preferred mode and the people try to preserve a friendly appearance.

Regarding *Directness/Indirectness*, Elliot et al. (2016) and Lewis (2010) indeed supposes differences between Germany and Japan in their cultural model. While Germans tend to speak very direct about certain things, Japanese prefer an implicit and indirect communication.

According to (Hofstede, 2009; Elliott et al., 2016; Lewis, 2010; Qingxue, 2003), the *Identity Orientation* is also expected to be different for Germans and Japanese. Germans are supposed to be rather individualistically oriented and the personal goals take priority over the allegiance to groups or group goals. In contrast, Japanese are more collectivistically oriented and often make their decisions in relation to obligations to their family or other groups. They tend to be people-oriented and the self is often subordinated in the interests of harmony.

In terms of *Thought Patterns and Rhetorical Style*, the cultural models also suppose various differences between Germans and Japanese. First of all, Qingxue (2003) states that Germans have a

low-context communication while Japanese have a high-context communication. Therefore, Germans need background information and expect messages to be detailed. In contrast, Japanese provide a lot of information through gestures, the use of space, and even silence. Most of the information is not explicitly transmitted in the verbal part of the message. Furthermore, according to (Elliott et al., 2016), the two cultures are expected to use different argumentation styles. For Germans, directness in stating the point, purpose, or conclusion of a communication is the preferred style while for Japanese this is not considered appropriate.

5 Concept and Evaluation

Based on the cultural differences in the dimensions *Directness/Indirectness*, *Identity Orientation* and *Thought Patterns and Rhetorical Style* which have been presented in Section 4, we have designed a study to investigate if these differences may be transferred to HCI. We formulated four hypotheses:

1. Germans choose options with direct communication more often than Japanese do.
2. Japanese choose options with motivation using group oriented arguments more often than Germans do.
3. Germans choose options with background information more often than Japanese do.
4. There are differences in the selection of argumentation styles.

Experimental Setting For the study, a dialogue in the healthcare domain has been created. This domain has the potential capacity to reveal such differences as very sensitive subjects are covered. For every system output, different variations have been formulated. Each of them has been adapted according to the supposed cultural differences. The participants assumed the role of a caregiver who is caring for their father.

In the beginning of the dialogue, the agent greets the user. The user also greets him and tells him that their father doesn't drink enough. The agent asks how much he usually drinks and the answer is that he drinks only one cup of tea after breakfast. Afterwards, different possibilities for the agent's output are presented. The first one doesn't contain any background information: "You're right, that's not enough. Do

you know why your father doesn't drink enough?" In contrast, the other four options include some background information why it is important for an adult to drink at least 1.5 litres of water per day. However, they differ in the argumentation style (parallel, linear, circular, digressive). The user answers that he doesn't know why their father doesn't drink enough. Then, the agent has different proposals how the water-intake may be increased and there are four different options for each proposal how it is presented to the user. The first option contains background information and expresses the content directly. The second option is also direct but doesn't give any background information. For the third and the fourth options an indirect communication style is chosen, whereby one option contains background information and the other doesn't. An example for the different options can be found in Table 1.

Option	Formulation
1	Offer him tea instead of water. It tastes good and is not as bad as soft drinks.
2	Offer him tea instead of water.
3	Offering tea instead of water can help. It tastes good and is not as bad as soft drinks.
4	Offering tea instead of water can help.

Table 1: There are four different options for each proposal how it is presented to the user: (1) direct, background information, (2) direct, no background information, (3) indirect, background information, (4) indirect, no background information.

In the end of the dialogue, the agent tries to motivate the user. Two different kinds of motivation are formulated and presented by the agent. The first one uses individualistically oriented arguments ("You're really doing a great job! It's impressive that you are able to handle all of this.") whereas the second one uses group oriented arguments ("You're really a big help for your family!"). Afterwards, the agent and the user say goodbye and the dialogue ends.

The survey has been conducted on-line. A video for each possible system output has been created using a Spoken Dialogue System with an animated agent. For all recordings, the same system and the same agent have been used. In each dialogue turn, the participants had to watch videos representing the different variants of the system output and decide which one they prefer. An example



Figure 1: In each dialogue turn, the participants had to watch different videos and decide which one they prefer.

of this web page is shown in Figure 1. During the survey, all descriptions have been provided in English, German and Japanese. The videos have been recorded in English and subtitled in German and Japanese. The translations have been made by German and Japanese native speakers who were instructed to be aware of the linguistic features and details of the cultural differences to assure equivalence in the translations.

The survey Altogether, 65 Germans and 46 Japanese participated in the study. They have been recruited using mailing lists and social networks. The participants are aged between 15 and 62 years. The average age of the Germans is 25.7 years while the average age of the Japanese participants is 27.9 years. The gender distribution of the participants is shown in Table 2. It can be seen that 65 % of the German and only 17 % of the Japanese participants are female.

	German	Japanese
male / female	23 / 42	38 / 8

Table 2: The participants' gender distribution.

Evaluation results The evaluation of the survey confirms our main hypothesis that Germans and Japanese have different preferences in communication style in HCI.

Our first hypotheses says that Germans choose options with direct communication more often than Japanese do. The study contains four questions where the participants have to choose be-

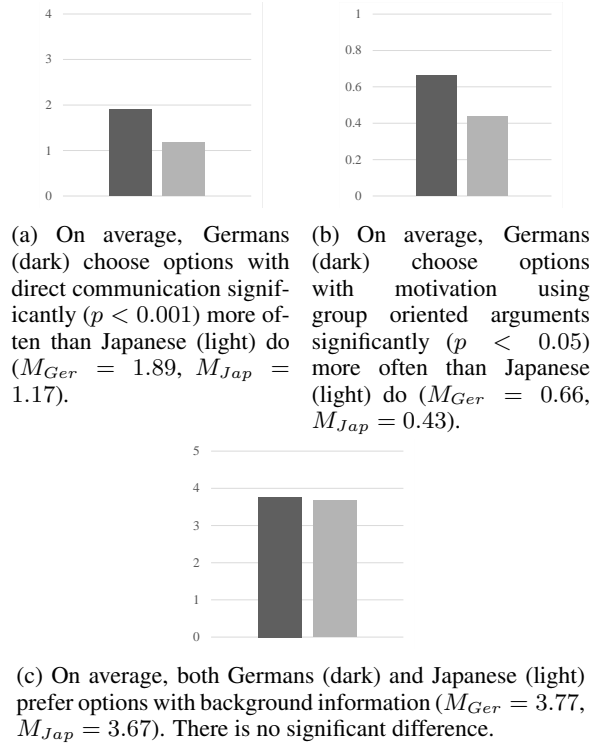


Figure 2: Differences between Germans/Japanese.

two direct and indirect options. Figure 2a shows the mean of how often Germans (dark grey) and Japanese (light grey) selected the direct option. The German mean is with 1.89 significantly higher than the Japanese mean ($p < 0.001$ using the T-Test) thus confirming our hypothesis.

Our second hypotheses says that Japanese choose options with motivation using group oriented arguments more often than Germans do. The survey includes one system action where the agent motivates the user. Figure 2b shows the mean of how often Germans (dark grey) and Japanese (light grey) selected the motivation with group oriented arguments. It can be seen that the opposite of the hypothesised effect occurred. On average, the Germans chose the option with group oriented arguments more often than the Japanese ($p < 0.05$ using the T-Test). An explanation for this result might be that motivation may be dependent on the topic of the dialogue. In our case, the dialogue is in the healthcare domain and caring for a family member is inherently group oriented. Therefore, it is most likely that motivating using group oriented arguments is more preferred for individualistically oriented people. However, if for someone it is natural to care for a family member because he is group oriented, then motivation using group oriented arguments is not needed and individualis-

tically oriented arguments seem to be favoured.

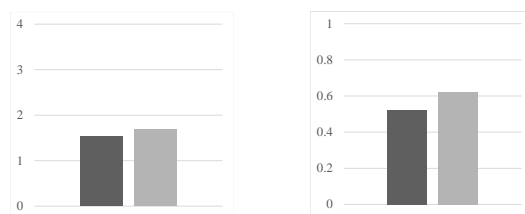
Our third hypotheses says that Germans choose options with background information more often than Japanese do. The survey comprises five questions where the participants could select between system outputs with and without background information. Figure 2c shows the mean of how often Germans (dark grey) and Japanese (light grey) selected the option with background information. On average, both Germans and Japanese preferred the options with background information. This suggests that there is no non-verbal communication in this kind of HCI which is only based on speech and does not include other modalities (the agent in the videos does not produce any output but the speech). In this case, Japanese tend to miss the non-verbal communication which they use to have in HHI and therefore need verbal background information.

Our last hypotheses says that there are differences in the selection of argumentation styles. The survey contains one system output where the participants have to choose between different argumentation styles. However, no significant difference could be found.

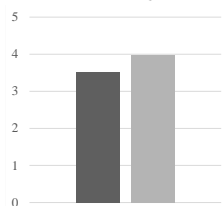
Due to the difference in the gender distribution, it is important to investigate whether this has an effect on the overall results. As can be seen in Figure 3, only for *Thought Patterns and Rhetorical Style*, a significant difference has been found: on average, women chose options with background information more often than men. However, as the majority of both genders and both cultures chose the options with background information ($M_m > 2.5$, $M_w > 2.5$, $M_{Ger} > 2.5$, $M_{Jap} > 2.5$), the difference between the genders is not supposed to effect the result based on the culture.

6 Conclusion and Future Directions

In this work, we presented a study investigating whether cultural communication idiosyncrasies found in HHI may also be observed during HCI in a Spoken Dialogue System context. Therefore, we have created a dialogue with different options for the system output according to the supposed differences. In an on-line survey on the user's preference concerning the different options we have shown that there are indeed differences between Germany and Japan. However, not all results are consistent with the existing cultural models for HHI. This suggests that the communication pat-



(a) On average, both men (dark) and women (light) prefer options with indirect communication ($M_m = 1.52$, $M_w = 1.68$). There is no significant difference. (b) On average, both men (dark) and women (light) prefer options with motivation using group oriented arguments ($M_m = 0.52$, $M_w = 0.62$). There is no significant difference.



(c) On average, women (light) choose options with background information significantly ($p < 0.05$) more often than men (dark) do ($M_m = 3.52$, $M_w = 3.98$).

Figure 3: Differences between men/women.

terns are not only influenced by the culture, but also by the dialogue domain and the user emotion. Moreover, it is shown that not all cultural idiosyncrasies that occur in HHI may be applied for HCI.

In this work, only one specific dialogue has been considered. To get a more general view and exclude effects which may depend rather on the domain than on the culture, in future work other dialogues from different domains should be examined. Furthermore, we have to identify how the defined cultural idiosyncrasies may be implemented in the Dialogue Management to design a culture-sensitive spoken dialogue system.

Acknowledgements

This work is part of a project that has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 645012. This research and development work was also supported by the MIC/SCOPE #152307004.

References

- Jan Brejcha. 2015. *Cross-Cultural Human-Computer Interaction and User Experience Design: A Semi-otic Perspective*. CRC Press.
- Christine Doran, John Aberdeen, Laurie Damianos, and Lynette Hirschman. 2003. Comparing several

- aspects of human-computer and human-human dialogues. In *Current and new directions in discourse and dialogue*, pages 133–159. Springer.
- Candia Elliott, R. Jerry Adams, and Suganya Sockalingam. 2016. Multicultural toolkit: Toolkit for cross-cultural collaboration. Awesome Library. <http://www.awesomelibrary.org/multiculturaltoolkit.html>. Accessed: 2016-05-01.
- Kallirroi Georgila and David Traum. 2011. Learning culture-specific dialogue models from non culture-specific data. In *Universal Access in Human-Computer Interaction. Users Diversity*, pages 440–449. Springer.
- Geert Hofstede. 2009. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Sage.
- Robert B. Kaplan. 1966. Cultural thought patterns in inter-cultural education. *Language learning*, 16(1-2):1–20.
- Richard D. Lewis. 2010. *When Cultures Collide: Leading Across Cultures*. Brealey.
- Wolfgang Minker, Ramón López-Cózar, and Michael F. McTear. 2009. The role of spoken language dialogue interaction in intelligent environments. *Journal of Ambient Intelligence and Smart Environments*, 1(1):31–36.
- Liu Qingxue. 2003. Understanding different cultural patterns or orientations between east and west. *Investigationes Linguisticae*, 9:21–30.
- David Traum. 2009. Models of culture for virtual human conversation. *Universal Access in Human-Computer Interaction. Applications and Services*, pages 434–440.
- Stefan Ultes and Wolfgang Minker. 2014. Managing adaptive spoken dialogue for intelligent environments. *Journal of Ambient Intelligence and Smart Environments*, 6(5):523–539.