Politeness and Frustration Language in Child-Machine Interactions

Sudha Arunachalam[1], Dylan Gould[1], Elaine Andersen‡, Dani Byrd‡, Shrikanth Narayanan**

Integrated Media Systems Center, University of Southern California, Los Angeles
‡Department of Linguistics, *Department of Electrical Engineering, USC
{sarunach,dgould}@usc.edu,elaine@gizmo.usc.edu,dbyrd@usc.edu,shri@sipi.usc.edu

Abstract

Children represent a potentially crucial user segment for conversational interfaces. Computer systems interacting with children need to be tailored for these users so that they will understand child intent and so that the child will have a positive and successful experience with the system. This study focuses on discourse analysis of spoken-language child-machine interactions. In particular, politeness and frustration markers were analyzed using a database of child-machine conversations obtained from 160 children using a computer game in a wizard-of-Oz set up. Results indicate that younger children less likely to use overt politeness markers and more polite information requests compared to the older ones, with no apparent gender differences. Younger children, on the other hand, expressed frustration verbally more than the older ones; furthermore, frustration language was more predominant in male children.

1. Introduction

Enabling spoken language capability—automatic recognition, understanding, and synthesis—as a part of immersive, multimedia interfaces adds naturalness and efficiency to human-machine interactions. Children are an important user segment that will benefit from advances in multimedia technologies. Children are one of the primary potential users of computers for conversational interaction in multimedia games and computer instructional material. Children are generally comfortable and happy using spoken language interfaces. However, computer systems interacting with children need to be tailored for these users so that they will understand child intent and so that the child will have a positive and successful experience with the system [1,2,6].

1.1. Children and Language

Children are still learning linguistic rules of social and conversational interaction. Their concepts of social structure are still solidifying and are different from those of adults. This means that their behavior in interacting with a computer as an interlocutor is also different from the behavior of adults. Very little is known about these differences and their importance to speech technology. The design and implementation of ASR/NLP systems that will be provided for child users need to take the special behavior of children into account. Natural language understanding systems must accommodate the special characteristics of child-users within a developmental framework.
We feel that speech *register* is an important ingredient in this recipe for success. *Register* refers to language varieties that are specific to particular social contexts of interaction. Social factors like familiarity, relative social standing of the interlocutors, and task setting can all affect register [3]. In order to make automatic speech recognition (ASR) interfaces more enjoyable to work with, it is useful for them to be able to recognize the register of the user, and to accommodate their own registers to match or complement the user's discourse style or speech. The sociolinguistic behavior of children differs from adults in a variety of ways [4]. Very little is known about the register choices children make when talking to computer agents. Furthermore, children have different language strategies for initiating and guiding conversational exchanges. Lastly, children have less developed abilities to express complex information requests, and accordingly may have a more frequent need to reject responses to their commands (i.e., change of mind).

The current paper examines two sociolinguistic markers. The first section explores what type of vocabulary might indicate that a child is having difficulty with a task. It focuses on children's use of verbal expressions of frustration, annoyance, and rejection. In discourse between children and computers, frustration markers are expected to differ from those used in discourse between adults and computers. For example, children are expected to use less profanity, but to express frustration more often. Additionally, it is hypothesized that the use of repetition will correlate with the use of frustration markers such that when a child is experiencing difficulty with a task they will have to repeat some of their information requests. In particular, we examine differences as a function of age, sex, and task abilities.

The second section of the paper investigates how politeness and the linguistic form of information requests differ among children of different ages and sexes. Research in language acquisition shows that even six and seven year-old children have awareness and command of varying levels of politeness associated with different registers [4]. We examined the politeness of children's requests for information or action, and the register variation sensitive to the relationship between interlocutors.

The rest of the paper is organized as follows. The database and research methods are described in Section 2. The analysis of the frustration and politeness markers are given in Sections 3 and 4. Conclusions are provided in Section 5.

2. Methods

The data corpus being examined came from a 1997 study on child/machine interactions: ChIMP—*Children's Interactive Multimedia Project* database [1]. The total database included spoken interactions of 160 boys and girls, six to fourteen

years of age, with a computer. The study used a Wizard of Oz (WOZ) design, in which a human operator controlled the computer without the knowledge of the subject. The WOZ design ensured that computer language understanding and speech recognition components of the task could be performed without error. The task was to play "*Where in the USA is Carmen Sandiego*", a computer game familiar to many children in the United States. Most children played two games, but some played one or three. Text transcripts of the children's utterances were analyzed.

In order to successfully complete the game, the child had to track, identify and arrest a (cartoon) criminal. In that process, the child has to talk to a number of cartoon characters to obtain clues about the suspect's whereabouts and appearance, and in turn use the information to track the suspect through several geographical locations. Further details about the experiment may be found in [1,2].

After the experiment, each child went through a user experience exit interview. The children gave high ratings to the speech interface (93% rated the interface 4 or 5, where 5 corresponded to the highest positive rating possible). The game also received high marks but somewhat lower than the interface (only 81% rated the game 4 or 5). It is interesting to compare the relation between the number of games won and the ratings. Losing a game had a significant negative effect on the rating of the game. However, there was no significant effect of the game outcome on the children's rating of the voice input. The 11-12 year olds gave the highest ratings for the voice input. Gender effects were negligible.

Age and speaker dependencies in interaction patterns were also analyzed. There were no noticeable differences in the dialog patterns of male and female children. However, the dialog patterns of older children (11-14 years) were different from those of younger ones (8-10 years). The older children tended to complete the game faster, did fewer database lookups, used more advanced dialog patterns, and had fewer out-of-domain utterances (about half the number as the younger group).

For the purpose of initial quantitative analysis, any speech utterance that triggered no valid game response or action was defined to be *extraneous* i.e., out of domain, primarily from the perspective of automatic speech recognition. In the data, such extraneous speech utterances corresponded to approximately 5% of all utterances spoken for the 8-10 year-olds (compared to 3.7% for all subjects), with values ranging from 0% to 25% among individual subjects (7% variance). Most extraneous speech utterances fell in one of the following categories:(i) those expressing excitement or disappointment when vital/useless information was provided by the game or success/failure was achieved in one of the game stages, (ii) those requesting game-strategy information, interpretation of game output or approval by other people in the room (an adult moderator or other children were present in the game room for about half of games played), and (iii) interacting with characters on the screen irrelevant to game goals and objectives. Overall, the extraneous speech utterances were found to be highly speaker-dependent, age-dependent, and to be preceded by a small subset of dialog states. These results motivated us to systematically investigate the importance of the linguistic patterns, initially deemed to be extraneous, to better understand child-machine interactions.

3. Frustration and Rejection Language

3.1. Frustration Markers

In order to determine the types of words children use to express frustration and politeness, we created a catalogued lexicon of the words found in the database. We identified 21 words likely to indicate frustration, difficulty, or annoyance.

Table 1: Frustration vocabulary with number of occurrences

shut up	20	bastard	1
dick	10	blah	1
(oh) man	6	don't	1
hurry (up)	5	faster	1
oops	5	geez	1
heck	4	god	1
darn	3	kill	1
jerk	3	mad	1
bad	2	nutso	1
gee	2	wrong	1
whatever	2		

"Shut up" is the most popular frustration marker, well ahead of others such as "oh man," "hurry" (or "hurry up"), "oops," and "heck." It should also be noted that there were large individual differences across children in terms of which frustration vocabulary they used. The most extreme example is "dick" which was only said by one subject (see the example dialog below).

Counts of the occurrences of the frustration markers were compared by gender, age and game outcome (win/loss). These rates are expressed as percentages because they represent the distribution across turns containing frustration markers in a game, i.e., the number of frustration markers in one game divided by the total number of turns in that game. Males used frustration markers four times more often than females (0.16% & 0.4% respectively). Additionally, the youngest children used more frustration markers than the older children. The small sample of adults also recorded as part of the experiment also showed a large frequency of frustration markers (.29%), such than young children and adults were comparable on this measure. Finally, verbal expressions of frustration occurred more than twice as often in games that ended up in a loss than in those which were won (0.13% & 0.06% respectively). These results are quantified in the tables below.

Table 2: Frustration marker use grouped by sex and age

female	male	8-9 y/o	10-11 y/o	12-14 y/o
0.04%	0.16%	0.22%	0.04%	0.07%

Table3: Frustration marker use grouped by game data

game 1	game 2	won	lost
0.08%	0.15%	0.06%	0.13%

Below we present a sample dialog which exemplifies the use of frustration language (just the user portion of a sub-dialog):

- I don't know how to spell Albuqer- Albuquerque
- don't you know how to spell
- I'll just exit
- geez they should make better computers
- hey mister talk
- where did the suspect go

- quit the chit chat

3.2. Rejection Language—"no"

Recall that we predicted that since young children have less developed abilities to express complex information requests, they may have a more frequent need to reject responses to their commands. Since the game scenario did not involve asking the children any explicit yes/no questions in solving the game, the text transcripts could be examined for occurrences of the word "no" from the children as this word systematically indicated a rejection of a system response or action. Rejection using the word "no" usually happened when the child said she wanted something and then changed her mind after the computer had begun that action. Table 4 below shows averages over the first and second game played by the child and whether the game was won or lost; both occurrences (average number across all subjects) and percent occurrences per turn (average value of per subject values) are given. Losing games included more use of "no" than winning games, and second games in a series had more occurrences than first games.

Table 5 shows occurrences of "no" as a function of child age and sex (averaged across subjects). Results indicate that males reject more than females and that the youngest children make more frequent rejections than the older children.

There are at least two possible interpretations of the sex difference. Female children might reject system actions/responses less often because they are more patient or because their information requests have been created more successfully. Alternatively, female children might in fact show comparable rejection rates to the male children but simply be using some other verbal form to do so. The age effect is also of interest. It provides initial support for our hypothesis that the information-request format of the game might create a situation in which the less developed cognitive skills of the youngest children put them in the position of having to more frequently reject system responses to the requests they've formulated.

Table 4: Occurrences of the rejection word "no"

	game 1	game 2		Won	Lost
occurrences	1.05	1.28		1.00	1.28
turns	191	154		184	162
occurrences per turn	0.47%	0.70%		0.55%	0.62%

Table5: Occurrences of the rejection word "no"

	8-9 y/o	10-11 y/o	12-14 y/o	female	male
occurrences	1.70	0.77	0.64	0.67	1.36
turns	196	166	144	142	195
occurrences per turn	0.74%	0.40%	0.47%	0.42%	0.60%

3.3. Insistent Requests

We expect insistent requests for information to include a high degree of repetition of words from one user-turn to the next. In a final analysis, each line of each dialog was compared to the line following it by a program that returns percentage similarity score. The scores are based on the ratio of words in the first sentence which also appear in the second, with some penalties for differing word orders. These similarity scores were calculated for three cases: A) no substitutions, B) substitution of words commonly used synonymously, and C)

substitution of common and less common synonyms. Since the positions of the frustration words in the dialogs are known, the similarity scores of the lines containing and surrounding these words will be compared to the average similarity of utterances. Thus it can be determined whether correlations exist between turn-to-turn word repetition and frustration and rejection markers. This work remains in progress.

4. Politeness and Form of Information Requests

Different varieties of language are warranted by different social situations; such varieties are known as registers. Linguists find that different types of information requests (i.e. different registers) occur depending on the relative social standing of interlocutors. Most of the children's speech in the game consisted of requests to the computer for information or action. We examined these requests to determine the level of politeness children used when interacting with the computer. Analysis of the information requests used by the children can inform as to what social standing children assign to their computer-interlocutor.

4.1. Politeness Markers

The transcripts of the child-machine dialogues were searched for *please*, *thank you/thanks*, and *excuse me*. The number of occurrences was divided by the number of turns in each dialog to indicate frequency of usage.

For the first analysis the children were divided into three groups: the youngest aged 6-8 years, the middle group aged 10-11, and the oldest group aged 13-14. A two-factor ANOVA with the independent variables of age group and sex indicates no significant effects on frequency of these terms. The means and standard deviation data are shown in the table below. Note that the small mean values are due to the fact that normalization was done using the total number of user turns in the dialog (typically over 150); a better way would have been to normalize on turns belonging to specific dialog states.

Table 6: Politeness markers (average per turn per subject)

	Younger (6-8 yrs)		Middle (10-11 yrs)		Older (13-14 yrs)	
	M	SD	M	SD	M	SD
Thanks/Thank you	.08	.067	.093	.066	.068	.075
Excuse me	.004	.019	.022	.048	.025	.047
Please	.039	.086	.039	.072	.055	.123

However a more fine-grained examination indicates that these age grouping may be too broad. The two six-year olds were highly variable in their behavior. When they are excluded, the following patterns of variation emerges across the ages *(age 7, n=2; age 8, n=36; age 9, n=47; age 10, n=42; age11, n=37; age 12, n=28; age 13, n=24; age 14, n=13).* While little age-related variation is seen for *excuse me*, the use of *please* and *thanks/thank you* seems to vary with age. *Thanks/thank you* is most common among the youngest and the middle ages, with the oldest children showing more variability in their usage. This pattern is repeated for *please* except that its occurrence is lacking for the seven year olds. When one looks at the overall usage of these terms in the figure below, a pattern of increasing use with age *and* increasing variability with age is apparent.

The use of politeness markers in interacting with the computer is least common in the youngest age group, suggesting that perhaps these children preserve their cognitive resources for negotiating the game. The middle age group seems to productively use politeness markers in interacting with their interlocutor, attributing it a higher social standing than their own. The older children are particularly variable in whether they choose to use overt politeness markers. Some older children may not view their interlocutor as 'animate,' therefore not requiring extremely polite language, or they may view it as more of a peer than an interlocutor with higher social standing.
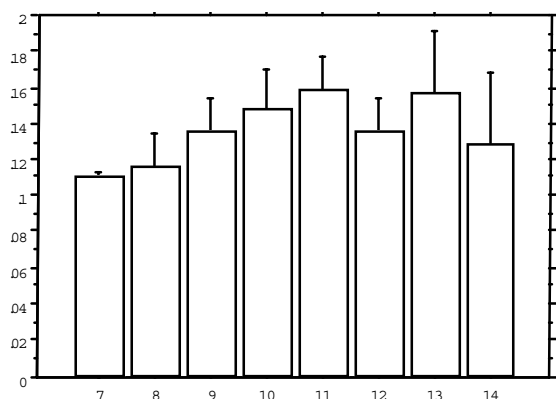

Figure1: Cummulative use of politeness terms as a function of age.

4.2. Information Requests Analysis

Another way of indicating politeness is to use different forms of questions for information or action requests. The following five forms range in politeness, from least to most polite:

- *Can you go to the map?*
- *Will you go to the map?*
- *Could you go to the map?*
- *Would you go to the map?*
- *May I see the map?*

These question forms are called modals. The transcripts were analyzed to see if use of these modal types varied by gender and/or age. The modal types were separated into the less polite (*can* & *will*) and the more polite (*could*, *would*, & *may*). The older children use significantly more of the more polite forms ($M = .06$, $SD = .13$) than the younger group ($M = .01$, $SD = .03$), with the middle group in between ($M = .03$, $SD = .05$). There were no significant differences across age groups for the less polite forms. Thus modal politeness increases with age. There were no significant differences between males and females in modal politeness. The following is a typical example from an 8-year-old male's speech:

Child: **turn left**
System: {character in focus turns left}
Child: **hi there you**
System: Hello. Nice to see you around these parts.
Child: **Can you tell me what she was wearing?**
System:.....
Child: **okay thank you**

The child uses the modal *can*, as well as overt politeness markers. Below is an example from a 14 year-old female, who uses overt politeness markers, more interrogatives and more polite modal types.
Child: **can I talk to you please?**
System: hi there.
Child: **Do you know where the suspect went?**
System:....
Child: **could you put that in my notebook?**
System:....
Child: **could I look at the book?**

4.3. Politeness Language Conclusions

The results suggest that the younger children don't yet have the language development and cognitive resources for politeness markers and complex forms of information requests. The preadolescent children uses overt politeness markers but don't yet fully employ polite request forms. The older children express politeness by information request forms as well as overt politeness markers, but they don't always 'deign' to use overt markers with the computer.

5. Conclusions

System designers for conversational interfaces can expect different vocabulary and information request forms from children and adults. Spoken language systems should accommodate and be responsive to the different language of child users of varying ages and sex. Children will overtly express frustration, annoyance, and rejection in a computer-child dialogue, and some common 'warning words' can indicate such difficulty on the part of the child user. Furthermore, child age and sex affect children's choices about politeness and register. Future research will illuminate child preferences regarding the linguistic register used by systems in generating spoken responses to children's requests.The data of the present study came from a WOZ experiment that assumed "perfect" speech recognition and understanding errors. We believe that future work on the analysis of child-machine conversations under errorful scenarios is crucial to facilitate better design of spoken language systems for children.

6. References

[1] A. Potamianos and S. Narayanan., "Spoken dialog systems for children", *Proc. of ICASSP*, p 197-200, 1998.
[2] S. Narayanan and A. Potamianos. [under review]. Creating conversational interfaces for children. *IEEE Trans. Speech and Audio Processing*.
[3] E. Andersen., *Speaking with style: The Sociolinguistics Skills of Children*. London: Routledge, 1991.
[4] E. Andersen et al., "Cross-linguistic evidence for the early acquisition of discourse markers as register variable", *Journal of Pragmatics, 31*, p. 1339-1351, 1999.
[5] D. Byrd. 1994. Relations of sex and dialect to reduction. *Speech Communication*, 15:39-54.
[6] S. Oviatt, "Talking to thimble jellies: Children's conversational speech with animated characters", Proc. ICSLP (Beijing, China), pp. 67—70, 2000.