

# Politeness and Summarization: an Exploratory Study\*

**Norton Trevisan Roman**

Institute of Computing  
State University of Campinas  
norton@ic.unicamp.br

**Paul Piwek**

Information Technology Research Institute  
University of Brighton  
paul.piwek@itri.brighton.ac.uk

**Ariadne Maria Brito Rizzoni Carvalho**

Institute of Computing  
State University of Campinas  
ariadne@ic.unicamp.br

## Abstract

In this paper, we describe an empirical pilot study on the role of politeness in dialogue summarization. We used a collection of four dialogues automatically generated by the NECA system. In these dialogues, the politeness of the interlocutors was systematically manipulated. We asked subjects to summarize the dialogues. Subjects were divided into groups who had to summarize the dialogues from a particular dialogue participant's point of view or the point of view of an impartial observer. Amongst other things, we found that the style of the interaction is included more often in summaries of dialogues which deviate from what would be considered normal or unmarked. We also found that the point of view of the summarizer influences which information is included in the summary and how it is presented.

## Introduction

Work on both language interpretation and generation has for a long time concentrated mainly on interpretation to and generation from truth-conditional representations of content, with some early exceptions such as (Hovy 1988). The underlying assumption has been that the content of a natural language utterance can be captured by representing the conditions under which it is true. When considering dialogue, the limitations of a strictly truth-conditional approach are apparent. Many types of dialogue act do not yield to a purely truth-conditional analysis (greetings, acknowledgments, closings, etc.) and the way a dialogue proceeds is often affected by the emotional states of the interlocutors (Craggs & Wood 2003; Fischer 1999; Schmitz & Quantz 1996; Reithinger *et al.* 2000).

Recent research on the influence of emotions on interpretation and generation includes work on recognizing the user's emotional state, e.g., anger (Huber *et al.* 2000), so that

a computer dialogue system can adapt its own behavior to prevent such feeling; the use of humor to facilitate interaction with users (Nijholt 2003); and strategies to establish a social and personal relationship with the user by means of "small talk" (Bickmore & Cassell 2000; Bickmore 2002). An overview of the literature on affect/emotion in natural language generation can be found in (Piwek 2003).

Although there is a body of work on emotion and, more specifically, politeness, in generation and interpretation in dialogue, to our knowledge there is not yet any systematic work on the role of politeness in *dialogue summarization*. We want to address the question in which circumstances and how the politeness of dialogue participants features in the summary of dialogues. In this paper we describe a pilot study carried out at the State University of Campinas – Brazil, in order to collect data to resolve (partially) the aforementioned questions. In the study, our subjects were instructed to summarize a carefully selected set of automatically generated dialogues. The resulting collection of human-authored dialogues constitutes our data. Our longer term aim is to use the insights gained from studying these data in the construction of an automatic dialogue summarizer, which will be integrated with the NECA system (Krenn *et al.* 2002).

NECA (*Net Environment for Embodied Emotional Conversational Agents*) is a conversational agent platform in which the user can create characters by specifying their roles, personalities and interests. On the basis of these settings, the system automatically generates dialogues between the characters. The result of the generation process is a script which can be performed by two or more embodied agents (Piwek *et al.* 2002; Piwek & van Deemter 2003). The system has been developed for two domains: *eShowRoom*, from which we have taken the dialogues for our study (first version of October 2002); and which concerns car sales, in which one agent is the vendor while the other is the customer. In the second domain, *Socialite*, the characters are inhabitants of a student district in Vienna. We used NECA's scripted dialogues as a test corpus because the NECA system allowed us to systematically change the agents' politeness. In this paper we focus on the results concerning the summarization of emotional/affective aspects of the dialogue. More specif-

---

\*We thank CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico - for partially sponsoring the project. Part of the research was also supported by the EC Project NECA IST-2000-28580.

Copyright © 2004, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

ically, we investigate how the politeness of both dialogue partners can influence human-authored summaries.

## Experimental setup

Our pilot study was carried out over a period of two months – July and August, 2003 – at the State University of Campinas, Brazil, with thirty subjects and four dialogues. The thirty subjects were: 1 graduate, 8 MSc students, 20 PhD students and 1 PhD, mostly male. Table 1 shows the number of subjects according to their educational degree, sex and course.

Degree		Sex		Course	
Graduated	1	Masc.	18	Physics	1
MSc Student	8	Fem.	12	Biology	1
PhD Student	20			Civil Eng.	1
PhD	1			Computer Sci.	27

Table 1: Subjects.

The study was conducted as follows. First, four dialogues were generated using the NECA *eShowroom* system. Each dialogue involves a customer and a vendor: a character named Ritchie (who we will refer to as “he” from now on) tries to sell a car to another character named Tina (who we will refer to as “she”). The NECA system allows us to control a number of dialogue parameters (interests and politeness of dialogue partners, for example). For the purpose of our study we varied the politeness of both interlocutors. Currently, the system only distinguishes between polite and impolite behaviour.

The four dialogues that were generated can be found in Appendix I to this paper. They were printed, and presented to the subjects in a fixed order. The first dialogue is more or less neutral, in the sense that both customer and vendor are polite and do not use any language that is likely to cause offence. The second dialogue is characterized by the impoliteness of the customer, while the vendor remains polite throughout the dialogue. Conversely, in the third dialogue, the vendor is impolite while the customer is always polite. Finally, in the fourth dialogue, both characters are again polite. This last dialogue differs from the first primarily in its length, being much shorter than the first. The shortness was a result of limiting the number of topics the customer is interested in and of changing the agreeableness attribute for the customer: this led to the customer being easily persuaded to buy the car.

Note that our use of the NECA generator for generating dialogue gives us greater control over the properties of such dialogues than would have been possible if, for instance, we had collected a corpus of real car sales dialogues. In particular, it would have been difficult to find a corpus in which politeness was systematically varied, while other parameters were held constant. The NECA system can provide us with dialogues of approximately the same length, but varying in

the politeness of the interlocutors. This was achieved by instructing the system to hold interests and other personality traits constant, but vary politeness. The scripted dialogues generated by the NECA system are, of course, only an approximation of naturally occurring dialogues. This need not be a deficiency, many dialogues that we see and hear every day are not naturally occurring dialogues (dialogues in film, soaps, etc. on radio, TV, etc.).

All subjects were told to summarize the dialogues as if they had been present during the dialogue and were telling a friend what happened. In particular, this included 10 subjects who were told to summarize the dialogue from the point of view of the customer, a further 10 who were told to summarize the dialogues as if they were the vendor and, finally, 10 who were told to summarize the dialogues as if they had overheard the dialogue without directly participating. The summarizers were also instructed to summarize the dialogues either in Portuguese or in English. In the next section we will discuss some preliminary results from the study.

## Results and Analysis

The dialogues were presented in the following order (see Appendix I for full transcripts of the dialogues):

- Dialogue D<sub>1</sub> shows an interaction in which both vendor and customer were polite. The dialogue begins with the vendor attending to the customer, and ends with the customer buying the car;
- D<sub>2</sub> shows an interaction in which the vendor is polite and the customer is impolite. The interaction begins with the customer not finding the vendor in the showroom (i.e., the vendor was apparently absent for a moment) and ends with the customer not buying the car;
- D<sub>3</sub> shows an interaction in which the vendor is impolite and the customer is polite. Again, the dialogue begins with the vendor being apparently absent for a moment, and ends with the customer not buying the car; and
- D<sub>4</sub> shows a very short interaction in which, again, both vendor and customer are polite. Like the first dialogue, it begins with the vendor introducing himself to the customer, and ends with the customer buying the car.

The produced summaries were manually annotated, sentence by sentence, for politeness content. A summary is classified, for example, as + *interaction* if it has some remark concerning the politeness, which may be an entire phrase, such as “Ritchie is too self-assured and a horrible vendor” or may be part of a phrase, such as “I asked a very disgusting vendor about a car...”. Two examples of these summaries can be found in Appendix II.

Tables 2 and 3 summarize the results. In these tables, the rows differentiate among the users who have taken the politeness degree of the interaction into account, as well as technical information (+ *interaction* in the tables), when producing the summary; the users who have not taken the

politeness degree into account at all (- *interaction*), i.e., those who have only included the exchanged technical information in the summary; and those who have only taken the interaction's politeness into account (+ *interaction (only)*, in the table), ignoring technical information. The main columns indicate the number of subjects for each dialogue. Each dialogue column, in turn, is subdivided according to the three points of view present in the research, i.e., the customer's (C), the vendor's (V) and an observer's (O) point of view.

	<i>D</i> <sub>1</sub>			<i>D</i> <sub>2</sub>		
	<i>C</i>	<i>V</i>	<i>O</i>	<i>C</i>	<i>V</i>	<i>O</i>
+ <i>interaction (only)</i>	0	0	0	0	0	0
+ <i>interaction</i>	2	0	0	6	10	9
- <i>interaction</i>	8	10	10	4	0	1

Table 2: Data for the first 2 dialogues.

	<i>D</i> <sub>3</sub>			<i>D</i> <sub>4</sub>		
	<i>C</i>	<i>V</i>	<i>O</i>	<i>C</i>	<i>V</i>	<i>O</i>
+ <i>interaction (only)</i>	5	2	1	0	0	0
+ <i>interaction</i>	4	7	9	3	1	1
- <i>interaction</i>	1	1	0	7	9	9

Table 3: Data for the last 2 dialogues.

From these tables, we can see that although for *D*<sub>1</sub> and *D*<sub>4</sub> there was a small percentage of subjects who included the interaction characteristics in the summary (respectively, 7% for *D*<sub>1</sub> and 17% for *D*<sub>4</sub>), this percentage increases remarkably when we consider *D*<sub>2</sub> and *D*<sub>3</sub> (respectively, 83% for *D*<sub>2</sub> and 93% for *D*<sub>4</sub><sup>1</sup>).

The difference can be explained by looking at the politeness in the dialogues. In *D*<sub>1</sub> and *D*<sub>4</sub>, both the customer and the vendor were polite, which may be characterized as a “normal”, or “neutral”, situation. In *D*<sub>2</sub> and *D*<sub>3</sub>, first the customer and then the vendor displayed impolite behaviors. For run-of-the-mill car sales dialogues, the resulting interactions come across as quite unusual and, perhaps, abnormal. This may explain why a high percentage of subjects included a description of these unusual traits of the interaction in the summary.

It also seems that because in a customer-vendor interaction the vendor is expected to be polite and not the customer, more people noticed the vendor's impoliteness. However, this cannot be a final conclusion, owing to the small amount of data. Even so, the results suggest that it could be fruitful to investigate further how pre-conceived expectations about a particular genre of conversation can affect the summaries of such conversations. Such expectations could, for example, explain the number of subjects who have taken only the politeness into account for *D*<sub>3</sub>, where the vendor was impolite. In this group we have subjects who explicitly blamed

the vendor for dissuading the customer from buying the car. Another possible explanation for the vendor's behaviour to be noticed by more subjects may be that, owing to cultural issues, *D*<sub>2</sub> is actually less impolite than *D*<sub>3</sub>. Again, this is a matter that deserves further investigation.

Now we take a closer look at each individual dialogue. For *D*<sub>1</sub>, those subjects who did take the politeness into account did so by praising the vendor's behavior. We have summaries in which the summarizer mentioned that the vendor was “a great vendor”, or was “nice and solicitous”. Note that the only subjects who took the politeness into account were those who summarized *D*<sub>1</sub> according to the customer's point of view.

*D*<sub>2</sub> is marked by the customer's impoliteness, a fact considered important enough to be mentioned in the summary by 83% of the subjects. The way in which the interaction was taken into account varied according to the point of view of the summarizer. For this dialogue, although the customer was rude, all summarizers who took the customer's point of view and commented on the style of the interaction, blamed the vendor for his behaviour, accusing him of “not being in the mood” and “not being convincing”. Sometimes, they even accused him of not treating the customer satisfactorily.

Conversely, all 10 subjects who summarized this dialogue from the vendor's point of view classified the customer, among other things, as “impolite”, “rude”, “irritated” and “impatient”. Sometimes they used irony, such as classifying the customer as a “nice woman”. One subject observed that the customer behaved rudely even than the vendor was solicitous. Interestingly, two “vendor” subjects volunteered a “*mea culpa*” for the customer's impoliteness, saying that the customer was nervous because the vendor had kept her waiting at the beginning of the dialogue.

All nine subjects who reported the dialogue from the observer's point of view and who took the politeness degree of the interaction into account noticed the customer's behavior, classifying her as “irritated”, “impatient”, “stressed” and “impolite”. Only one subject remarked that the vendor was “polite and patient”, and one did not take the interaction into account at all. Three subjects explicitly linked the customer's behavior to the vendor's delay at the beginning of the dialogue, i.e., they mentioned that the customer was impatient due to the vendor's delay (two subjects) and that the customer was in a hurry and the vendor took too long to help her (one subject).

*D*<sub>3</sub> yielded similar results. Although the vendor was responsible for the impolite behavior, some subjects, summarizing from the vendor's point of view, blamed the customer for the vendor's behavior, complaining that the customer was “too inquisitive”, “complaining”, “unpleasant” and “impatient”. Two subjects blamed themselves for their behavior (in the dialogue, as the vendor), saying “it was not my day”, “I was not in the mood” etc. It is important to notice that,

<sup>1</sup>The percentage is actually the sum of the users who have somehow taken the interaction into account and those who have only taken characteristics of the interaction into account.

from the vendor's point of view, only two subjects limited the summary to a description of the interaction, although a lot of technical information about the car was exchanged in the dialogue.

From the customer's point of view, on the other hand, the number of subjects who only described the interaction increases to five. This increase suggests that people pay more attention to an interaction when it is unusual from their perspective. The unmarked situation is the vendor being polite, for he wants the customer to buy the car. The customer expects to be treated well and, when it does not happen, the customer (or, in this case, the summarizer with the customer's point of view) thinks it is worth noticing it. Again, the customer's remarks varied from "it seemed that it was not his – the vendor's – day" to "how can a vendor be so rude?"

Subjects with the observer's point of view were almost unanimous in claiming that the vendor was "rude", "not respectable" and that the customer was not "well treated". From the nine subjects who have included the interaction in the summary, two have explicitly blamed the vendor for the customer not buying the car.

D<sub>4</sub> is very much like D<sub>1</sub>, except that it is much shorter. Even so, for this dialogue, the percentage of subjects who included politeness in the summary increased from 7% to 17%, when compared to D<sub>1</sub>. The reason for this increase may be the location of this dialogue in the sequence of dialogues which were presented to the subjects. It came after two more "charged" dialogues. This may have influenced the subjects' behaviour, in the sense that they were paying more attention to the politeness of the interaction when they read the fourth dialogue. This is a hypothesis that needs to be investigated in future experiments.

Evidence for the above conclusion is that the subjects that included the interaction in the D<sub>4</sub> summary – five in total – mainly noticed that the vendor was "nice", "helpful" and "kind", while the customer was "polite", exactly the opposite of the behavior shown in D<sub>2</sub> and D<sub>3</sub>. Of those five subjects, the ones who were summarizing according to the customer's and observer's point of view – four in total – congratulated the vendor for his behaviour, as if the vendor's behavior had improved over time; the subjects who summarized with the vendor's point of view, on the other hand, applauded the customer's behaviour.

The study suggests that the politeness degree of the interaction becomes more important as the dialogues become less neutral and more marked. The politeness degree of the interaction is included in the summary in different ways, according to the point of view assumed by the summarizer. Also, the summaries of apparently neutral dialogues can be affected by the immediate previous experience of the summarizer, in the sense that, despite the unmarked politeness of interaction, it is sometimes included in the summary.

## Conclusion

In this paper we have described a pilot study. The objective was to uncover some of the factors which contribute to the inclusion of interaction characteristics (in particular, those related to politeness) in dialogue summaries. The study revealed that the point of view of the summarizer influences which interactional characteristics are included and how they are presented.

Our preliminary results indicate that politeness is largely included in the summary by human summarizers, when this interaction is more unusual, i.e., when it is different from what the summarizers may consider to be an unmarked situation. If we compare the results for the 120 summaries generated from two "normal" and two "impolite" dialogues, in comparison to what we could expect to have by chance, we obtain a  $\chi^2$  value of 70.53, which is highly significant ( $p < 0.1\%$ ), and a  $\Phi$  value of 0.766. Actually, some summarizers go even so far as to only discuss the characteristics of the interaction in the summary, ignoring the remaining information (information about the topic of the conversation, i.e., the car that is under discussion). In what follows we summarize some preliminary conclusions which we might draw from the study:

- People include the politeness of the dialogue interaction in the summary, in particular when it deviates from the kind of interaction that would have been expected, i.e., from the genre of conversation that was previously expected for the dialogue;
- Once people have seen an extreme occurrence of such a deviation, some of them tend, in the remaining interactions, to pay more attention to the points that were considered problematic in the previous dialogues. For example, despite the fact that D<sub>4</sub> consists of a relatively unremarkable interaction, some people pointed out that it was a good interaction, perhaps on the basis of the problematic interactions presented in D<sub>2</sub> and D<sub>3</sub>; and
- Different points of view generate different summaries. This difference is greater when the politeness of the interaction is included in the summary. In these cases, we have noticed a high number of subjects trying to defend their positions and, therefore, producing a biased summary.

An automatic summarization system that complies with these insights needs to determine whether a dialogue is unusual. Alternatively, such a system could allow the user to indicate whether s/he would like to see the politeness degree of the interaction in the summary. Additionally, a user might be given the option of defining under which point of view the summary should be generated. We are currently working on such a system, which will be embedded into the NECA conversational agents platform, taking as input a NECA scripted dialogue, generating its summary according to a point of view defined by the user, and including the politeness of the interaction, if the user so chooses.

We are also carrying out another study concerning the in-

fluence of summary length on the information included. We would like to know whether subjects, when limited to a maximum summary length, still include the politeness of the interaction in the summary. We would also like to investigate the influence of the viewpoint assumed by the summarizer in dialogues and, more specifically, in more complex dialogues, like those produced by human script writers for TV and cinema.

## References

- Bickmore, T., and Cassell, J. 2000. How about this weather? social dialogue with embodied conversational agents. In *Proceedings of the AAAI Fall Symposium on Socially Intelligent Agents*.
- Bickmore, T. 2002. Social dialogue is serious business. In *Proceedings of the CHI 2002 Workshop on Socially Adept Technologies*.
- Craggs, R., and Wood, M. M. 2003. Annotating emotion in dialogue. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*.
- Fischer, K. 1999. Annotating emotional language data. Technical Report 236, VerbMobil Project.
- Hovy, E. 1988. *Generating Natural Language Under Pragmatic Constraints*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Huber, R.; Batliner, A.; Buckow, J.; Nöth, E.; Warnke, V.; and Niemann, H. 2000. Recognition of emotion in a realistic dialogue scenario. In *Proceedings of the International Conference on Spoken Language Processing*, volume 1, 665–668.
- Krenn, B.; Pirker, H.; Grice, M.; Baumann, S.; Piwek, P.; van Deemter, K.; Schroeder, M.; Klesen, M.; and Gstrein, E. 2002. Generation of multimodal dialogue for net environments. In Busemann, S., ed., *KONVENS 2002, Deutsches Forschungszentrum fuer Kuenstliche Intelligenz (DFKI)*. 91–98.
- Nijholt, A. 2003. Humor and embodied conversational agents. CTIT Technical Report series 03-03, University of Twente.
- Piwek, P., and van Deemter, K. 2003. Dialogue as discourse: Controlling global properties of scripted dialogue. In *AAAI Spring Symposium on Natural Language Generation in Spoken and Written Dialogue*. Menlo Park, California: AAAI Press. Technical Report SS-03-06.
- Piwek, P.; Krenn, B.; Schröder, M.; Grice, M.; Baumann, S.; and Pirker, H. 2002. RRL: A rich representation language for the description of agent behaviour in neca. In *Proceedings of the Workshop "Embodied conversational agents - let's specify and evaluate them!"*, held in conjunction with AAMAS-02.
- Piwek, P. 2003. An annotated bibliography of affective natural language generation. Technical Report ITRI-02-02, ITRI – University of Brighton.
- Reithinger, N.; Kipp, M.; Engel, R.; and Alexandersson, J. 2000. Summarizing multilingual spoken negotiation dialogues. In *Proceedings of the 38th Meeting of the Association for Computational Linguistics (ACL – 2000)*.
- Schmitz, B., and Quantz, J. 1996. Dialogue acts in automatic dialogue interpreting. Technical Report 173, VerbMobil Project.

## Acknowledgments

We thank all subjects who have participated in the study for their patience, friendship and active performance. We also thank Richard Power, Kees van Deemter, and Catalina Hallett for their valuable remarks on a draft of this paper.

## Appendix I

In this section we show the four dialogues used in the experiment, in the same order they were presented to the subjects.

### Dialogue 1

- R: Hello! I am Ritchie.  
 T: Can you tell me something about this car?  
 R: It is very safe.  
 R: It has anti-lock brakes.  
 R: It has airbags.  
 T: How much does it consume?  
 R: It consumes 8 liters per 60 miles.  
 T: Interesting.  
 T: What kind of luggage compartment does this car have?  
 R: It has a spacious luggage compartment.  
 T: Excellent!  
 T: What kind of interior does this car have?  
 R: It has a spacious interior.  
 T: Excellent! (Tina)  
 T: How much does this car cost?  
 R: It costs 25 thousand Euros.  
 T: Well, well.  
 T: All in all this is a perfect car. It is a deal!  
 R: Really? I am sure you wont regret it.

### Dialogue 2

- T: Anybody here? I haven't got all day.  
 R: I am sorry to have kept you waiting. I will be with you in a moment.  
 R: How can I help you?  
 T: I want you to tell me about this car.  
 R: It is very friendly for the environment.  
 R: It is made of recyclable materials.  
 R: It has a catalytic converter.  
 T: How much does it consume?  
 R: It consumes 8 liters per 60 miles.

T: It is not exactly very friendly for the environment.  
 R: Just consider this.  
 R: This car is very prestigious.  
 R: It has leather seats.  
 R: It has broad tires.  
 R: It drives 100 and 10 miles per hour.  
 R: It has 80 horse power.  
 T: What kind of luggage compartment does it have?  
 R: It has a spacious luggage compartment.  
 T: Well, well.  
 T: What kind of interior does this car have?  
 R: It has a spacious interior.  
 T: Well, well.  
 T: How much does this car cost?  
 R: It costs 25 thousand Euros.  
 T: It is not exactly very cheap.  
 R: Just consider this.  
 R: This car is very safe.  
 R: It has anti-lock brakes.  
 R: It has airbags.  
 T: Enough is enough! You keep your car!  
 R: Well, sorry that I cannot help you.

### Dialogue 3

T: Anybody in today?  
 R: Yeah yeah, keep your calm.  
 T: Can you tell me something about this car?  
 R: It is very safe.  
 R: It has anti-lock brakes.  
 R: It has airbags.  
 T: Does it have power windows?  
 R: What do you think? Of course not!  
 T: This car is not exactly very prestigious.  
 R: Just consider this.  
 R: This car has leather seats.  
 R: It has broad tires.  
 R: It drives 100 and 10 miles per hour.  
 R: It has 80 horse power.  
 T: What kind of luggage compartment does it have?  
 R: It has a spacious luggage compartment.  
 T: Well, well.  
 T: What kind of interior does this car have?  
 R: For god's sake, how should I know?  
 T: No problem.  
 T: Thank you for your help. I have to think a bit more about this.  
 R: Great! Thanks for wasting my time.

### Dialogue 4

R: Hello! My name is Richie.  
 R: What can i do for you?  
 T: Could you tell me something about this car?  
 R: This is a very sporty car.  
 R: It can drive 100 miles per hour.  
 T: How much does it consume?

R: It consumes 8 liters per 60 miles.  
 T: Not bad.  
 T: Ok! I'll take it.  
 R: Congratulations! Exquisite taste.

## Appendix II

In this section we present two example summaries for Dialogue 3 produced by two different subjects<sup>2</sup>. Both summaries were written according to the customer's point of view. The sentences in *italic* are the ones we have considered as referring to the politeness of the interaction.

### Summary 1

*I asked a very disgusting vendor about a car. He gave me some information about it, but it seemed that he was not interested in selling anything.*

### Summary 2

Tina didn't make up her mind about buying the car. It has a good luggage compartment, but it doesn't have power windows and it isn't prestigious.

---

<sup>2</sup>These two summaries were originally written in English.