



DIALOGUE BEHAVIOUR INDUCED BY THE MACHINE

D. LUZZATI, F. NEEL

LIMSI-CNRS, B.P. 133 - F-91403 ORSAY CEDEX

Abstract: In order to design human-machine dialogue systems using speech input as well as written input, it is necessary to study the influence of the machine not only on the user's lexicon and syntax, but also on his dialogue behaviour. Dialogue is not handled between a man and a machine in the same way as an interpersonal dialogue.

This paper is based on the linguistic study of a corpus obtained by simulating a machine in an actual human-machine dialogue situation: it deals with a task-oriented dialogue recorded by a telephone information service of the French National Railway Company (SNCF), for train timetables (schedules).

In the paper, we focus on the basic elements of dialogue behaviour induced by the machine, namely the respect of "turns" (the way one participant gives or takes control of the communication channel to or from the other): in the experiment carried out at the SNCF, naive users spontaneously accepted this constraint. We believe that such behaviour reflects a hierarchically-structured dialogue which we formalize with a set of defined and reduced symbols.

1- INTRODUCTION

This paper presents a study of the influence of the introduction of a machine in a task-oriented dialogue. The study relies on corpora recorded in a telephone public train timetable information service at the SNCF (French National Railway Company). The experiment was carried out under the auspices of the CNRS-GRECO CP (Speech Communication) and with the financial and technical support of the CNET. The corpus was recorded in two modes: in a natural human-human interaction mode and in a human-machine interaction mode in which users were led to believe they were speaking to a machine, where, in fact, the same operator simulated a machine-like behaviour. To make the experiment more realistic, the ORSO (Computer Spontaneous Oral) protocol project (1) was used and a speech processor (SCH) fixed the operator's fundamental to a constant value. The study, described in detail in (2), is a continuation of the work presented in (3) which only focused on linguistic behaviour during requests. In this paper, we also address the dialogue behaviour induced by the machine.

Section 2 distinguishes the different types of interaction depending on whether a human thinks he is addressing another human or a machine; Section 3 concerns the modifications of the reference aspect in both interactions and, in Section 4, a model is proposed for human-machine interaction.

2- SPONTANEOUS INTERACTION AND ALTERNATE INTERACTION

The nature of the interaction was found to be quite different depending upon the mode. In human-human mode, the interaction is spontaneous, which means there is no correspondence between argumentative alternation and turns. In contrast, in human-machine mode, the interaction is alternate, which means that the correspondence is almost always systematic. The different dialogue types are shown in Examples 1 and 2.

- O1 : *SNCF bonjour (SNCF, good morning)*
C1 : *bonjour mademoiselle je voudrais avoir les heures de train de Nevers pour Nevers s'il pour Nevers s'il vous plaît (good morning, miss, I would like the time of the trains of Nevers to Nevers plea for Nevers please)*
- O2 : *oui pour quand (yes when)*
C2 : *e pour le 26 décembre (uh for December 26)*
- O3 : *pour le 26 décembre (for December 26)*
C3 : *oui (yes)*
- O4 : *dans la matinée dans l'après (in the morning in the after)*
C4 : *dans le matin la matinée (in the morning during the morning)*
- O5 : *dans la matinée oui vous ne quittez pas s'il vous plaît (in the morning, yes, hold on please)*
C5 : *oui (yes)*
C6 : *merci(thank you)*
- O6 : *allo (hello)*
C7 : *allo oui (hello yes)*
- O7 : *oui alors vous avez un train à 7h 11 (yes, then there's a train at 7:11 a.m.)*
C8 : *7h 11 (7:11 a.m.)*
- O8 : *Paris gare de Lyon (Paris Gare de Lyon)*
C9 : *oui(yes)*
- O9 : *arrivée Nevers à 10h 19 (arriving at Nevers at 10:19 a.m.)*
C10 : *10h 19 (10:19 a.m.)*
- O10 : *c'est un corail première et deuxième classe (it's a Corail, first and second class)*
C11 : *oui (yes)*
- O11 : *sinon vous en avez un à 8h 45 (otherwise, there's another at 8:45 a.m.)*
C12 : *8h 45 (8:45 a.m.)*
- O12 : *également corail première et deuxième classe (also Corail first and second class)*
C13 : *oui (yes)*
- O13 : *arrivée 10h 52 (arriving at 10:52 a.m.)*
C14 : *10h 52 (10:52 a.m.)*
- O14 : *alors je sais pas vous voulez plus tôt plus tard (then I don't know, you want one earlier, later)*
C15 : *plus tard si vous avez oui (later if there's one, yes)*
- O15 : *plus tard (later)*

C16 : plus tard jusqu'au début de l'après de l'après-midi quoi (later until the beginning of the after, of the afternoon yes)

O16 : oui (yes)

C17 : hein (what)

O17 : alors j'en ai un à 10h 53 (then there's one at 10:53 a.m.)

C18 : 10h 53 (10:53 a.m.)

O18 : arrivée 13h 38 (arriving at 1:38 p.m.)

C19 : 13h 38 ... très bien bon ben ça ira comme ça allez c'est parfait bon ben je vous remercie mademoiselle (1:38 p.m... very well ok well it's fine well it's perfect well ok thank you miss)

Example 1 : Representative example of a human-human interaction mode.

M1 : SNCF bonjour quels renseignements désirez-vous obtenir (SNCF good morning what information would you like to obtain)

C1 : je voudrais connaître les trains vendredi 28 décembre Paris Clermont-Ferrand le matin. (I'd like to know the trains on Friday 28 December Paris Clermont-Ferrand in the morning)

M2 : vers quelle heure désirez-vous partir (about what time would you like to leave)

C2 : ben le matin (well in the morning)

M3 : ne quittez pas (hold on please)

M4 : un corail première et deuxième classe part de Paris gare de Lyon à 7h 02 mn arrive à la gare de Clermont-Ferrand à 10h 47 mn ce train vous convient-il (a first and second class Corail leaves Paris gare de Lyon at 7:02 a.m. arrives at Clermont-Ferrand at 10:47 a.m., does this train suit you)

C3 : non plus tard (no, later)

M5 : un corail première et deuxième classe part de Paris gare de Lyon à 8h 45mn arrive à la gare de Clermont-Ferrand à 12h 40 mn ce train vous convient-il (a first and second class Corail leaves Paris gare de Lyon at 8:45 a.m. arrives at Clermont-Ferrand station at 12:40 p.m., does this train suit you)

C4 : parfait merci (fine thank you)

M6 : désirez-vous d'autres renseignements (would you like some more information)

C5 : non merci (no thank you)

Example 2: Representative example of a human-machine interaction mode.

2.1 Spontaneous Interaction

During spontaneous interaction, the correspondence between speech-acts (4) and turns may not be respected. Several speech acts can be found with one turn, as in the example: *je voudrais un train pour Châlon vous allez peut être me dire quel Châlon (I'd like a train to Châlon, you will perhaps ask me which Châlon I mean)*. Alternatively, a speech-act may also expand on several turns, as in Example 1.

One particular explanation is that nearly one third of control-taking has a phatic status and consists of *oui (yes)* or "echoes" (5) also called "diaphonie" by Roulet (6) (echoes consist of repeating one element of the preceding exchange). Two main roles can be distinguished:

- they may appear as simple approbatory punctuation, (C9 and C13 for yes; C8, C10, C12, C18 for echoes) a sort of control-taking which is spoken aside only to interrupt the development of one of the operator's turns. In Example 1, there is only one speech-act expanding over 7 turns, starting with message O7 to O10, (4 for the operator and 3 for the caller): *oui alors vous avez un train à 7h 11 (7h 11) Paris gare de Lyon (oui) arrivée Nevers à 10h 19 (10h 19) c'est un corail 1ère et 2ème classe (yes, then*

there's a train at 7:11 a.m. (7:11 a.m.) Paris gare de Lyon (yes) arriving at Nevers at 10:19 a.m. (10:19 a.m.) it's a Corail, first and second class). It seems here that this type of echoing also allows a better memorization of the information, as pointed out by Beun (7).

- they may occur as real participation in the interaction, as an effective approbatory sign received as such by the other participant (C3 and C5 for yes, O3 and C4 for echoes). Whereas phatic yes and echoes, in the first case, are nearly exclusively uttered by callers (and the number depends on each user's linguistic behaviour), in this case, they are spoken by both participants and are no more merely approbatory punctuation but may be considered as approbatory reaction. They are included in a subordinate exchange and behave as a request for implicite confirmation requiring either a yes (which has the same reaction value) or a disconfirmation (in case of bad communication or interpretation).

2.2 Alternate Interaction

In alternate interaction, as observed in Example 2, the entire communication consists of very rigid and robust exchanges. Each turn almost always corresponds to one speech-act and each exchange consists of two turns.

The operator simulating the machine is led to add an evaluation request at the end of each of her responses: *ce train vous convient-il (does this train suit you)* (M4 and M5). This request evokes a corresponding evaluation from the user, sometimes the only speech-act (C4) and sometimes the first element of the exchange (C3). The phatic message announcing the search for information *ne quittez pas (hold on)* (M3) is not a part of the exchange system: it is a ritual way informing the user that the operator is searching for information in such an enquiry-task dialogue. In a real time, machine-based query system, such messages will no more be necessary.

The existence of such rigid exchanges is, in fact, the consequence of the voluntary avoidance of psychology in the communication. In an enquiry task, the aim is often to attract the sympathy of the person who possesses the information, in order to involve him or her in the problem to ensure a good response. In human-machine communication, users do not need to argue for their requests and do not attempt to anticipate counter arguments. Therefore, they do not feel obliged to give any causal, consecutive or adversative observations (8).

This situation bears a resemblance with face theory (9) which stipulates that humans are guided, in their dialogue behaviour, by the will of not losing face either positive (the image which they want to give of themselves) or negative (the integrity they want to keep of their own territory). In human-machine interaction, humans know the machine is incapable of feeling and has no face to save. Therefore, they are not inclined to attract the sympathy of the machine or defend their own face against it. They will not take the precautions they typically take with other humans.

In the same way yes and echoes, either mere approbatory punctuations or approbatory reactions tend to disappear almost completely. The artificial tone of the simulated machine (in fact, the mechanized operator's voice) also discourages users from interrupting the machine. This explains that, for a negotiation of equal importance, three to four times more exchanges are necessary in human-human communication.

3- DIALOGUE AND REFERENCES

Spontaneous and alternate interactions correspond to different types of dialogue. Both also imply different linguistic behaviour, especially as far as reference is concerned. A great number of phenomena observed in human-human interaction and constituting a real difficulty for automatic processing, tend to disappear in the presence of the machine.

3.1. Contextual references

It is not easy to automatically process frequent contextual references or anaphoras: in Example 3, it is difficult to interpret *first* (C5) (even by the human operator) as referring to O3 and having the meaning of *is the train you are proposing to me the first with regard to the time limit I specified before (7:30 p.m.)*.

O1 : SNCF bonjour (French National Railway Company (SNCF), good morning)

C1 : ou bonjour madame est-ce-que je pourrais avoir e les numéros de de Paris les heures de trains directs Montargis à partir de 7h 30 jeudi oui de 19h 30 jeudi (ye good morning, ma'am, could you give me uh the numbers of of Paris the schedule of trains direct to Montargis from 7:30 on Thursday yes of 7:30 p.m. on Thursday)

O2 : et pour quel jour (and for which day)

C2 : ben pour aujourd'hui (well for today)

O3 : pour aujourd'hui à partir de 19h 30 oui vous ne quittez pas (for today, after 7:30 p.m., yes please hold on)

C3 : merci(thank you)

O4 : allo (hello)

C4 : oui (yes)

O5 : oui alors vous avez un train e départ Paris gare de Lyon à 20h 09 (yes well there's a train uh from Paris gare de Lyon at 8:09 p.m.)

C5 : c'est le premier (it's the first one)?

O6 : pardon (I beg your pardon)

C6 : c'est le premier (it's the first one)?

O7 : c'est le premier que je puisse vous donner oui ben vous m'avez dit après 19h 30 (it's the first one I can give you, yes, well you told me after 7:30 p.m.)

Example 3: Example of contextual allusion, of anaphoric expression which would not easily be processed with the dialogue model and which is not immediately perceived by the operator.

A great number of interchanges can only be understood with the help of knowledge far beyond the task domain: in the request *allo bonjour madame je voudrais avoir des horaires e je sais pas on m'a donné des horaires e suivants pour Paris Le Creusot le 14 décembre à 16h:59 et moi je les trouve pas on moi je les trouve pas (hello, good morning, ma'am, I would like the schedule, uh I don't know, someone gave me the following uh time for Paris Le Creusot on the 14th of December at 4:59 p.m. and I can't find it some I can't find it), on and les (someone and it) must be correctly interpreted: the system should infer that someone has already given the caller information, that the caller possesses a schedule and that he is asking for confirmation of information which he is unable to find in his schedule.*

In the entire sequence of Example 4, it is necessary to infer the profession of the caller and of the future user, in order to be able to correctly identify the different references of *il(s)* (*he (they)*) (C5, C10 and C11) as being the train and the persons for whom the caller is asking the information: from the dialogue, we can infer that the caller makes the request for someone else for whom she works and who is supposed to attend a meeting with other persons. It is a constant phenomenon, in human-human interaction, to refer to the other person's world of knowledge.

O5 : bon non alors j'ai un train mais le enfin le problème c'est que c'est un train à supplément (well no well there's a train but the well the problem is that it's a train that requires a supplement)

C5 : oui et il part à quelle heure (yes and it leaves at what time)

O6 : alors j'en ai un qui part à 19 18h 33 (well there's one which leaves at 7, 6:33 p.m.)

C6 : oui (yes)

O7 : et il arrive à Angers à 20h 46 (and it arrives at Angers at 8:46 p.m.)

C7 : 20h 46 (8:46 p.m.)

O8 : c'est bien pour aujourd'hui hein (it's for today, isn't it?)

C8 : oui oui (yes yes)

O9 : bon alors sinon e (well then otherwise uh)

C9 : après (after)

O10 : sinon j'ai des problèmes plus tard (otherwise I've problems later)

C10 : ben il me il m'a dit entre 18h et 19h (well he me he told me between 6 and 7 p.m.)

O11 : 18 et 19h et j'en ai un à 19h mais il est pas pour e il marche que le vendredi (6 and 7 p.m. and there's one at 7 p.m. but it's not for uh it only runs on Friday)

C11 : parce que comme ils savent pas quelle heure qu'ils vont sortir de la réunion (because since they don't know the time that they're going to leave the meeting)

Example 4: Example of complex anaphora *il(s)* (it/he/they).

3.2. Adaptation to the dialogue partner

Such general world knowledge appears to be less necessary in human-machine interaction: just as there are attempts to adapt machine behaviour to the user, we found that the callers also spontaneously adapted themselves to what they assumed the machine's capabilities to be. The behaviour of the operator (simulating the machine in the way she imagined the machine would act), also helped condition the callers. It is obvious, however, that callers do not attribute the same function to the communication when they speak to a machine.

As was pointed out before, they have no face to defend against the machine: if, in Example 4, the caller makes all these references, it is a way of justifying his behaviour as he thinks he will be considered as a stickler for details by the operator. In Example 3, the caller feels obliged to justify why he is asking for confirmation of information he already knows.

The following example never occurred before the simulated machine: *bonjour madame je vais vous demander un renseignement s'il vous plaît ... e en ce qui concerne les billets de congés payés ... pour l'année alors voilà alors moi je travaille mais je travaille je ne travaille que par intermittence ... donc je n'ai pas toute l'année e je suis jamais au même endroit et je n'ai pas toute l'année des fiches de paye bon d'un autre côté je suis mariée est-ce que ça peut marcher sur le sur parce que moi je suis sur la sécurité sociale de mon mari est-ce que ça peut ma comment ça se passe est-ce que vous (good morning ma'am I'm going to ask you for some information please ... uh concerning the paid vacation tickets ... for the year then well then I work but I work I only work some periods of time... then I haven't the whole year uh I'm never at the same place and I don't have pay slips for the whole year well on the other hand I'm married will it work on the on because I'm declared on my husband's medical coverage and social security will it be all how is it do you...).* The operator cried for mercy after 8 exchanges and switched the caller on to another service!

One may wonder what usefulness is borne by such psychological processes. They are used frequently enough that they must be thought to be necessary in interpersonal interaction, but it is also certain that, for such a task, their absence does not hinder the successfulness of the communication: despite the absence of any justification and independently of the length of the interaction, the callers obtained a satisfactory response from the simulated machine. However, it must be noted that, in the satisfaction dialogue which the operator had with the users after the experiment, several callers indirectly asked her for confirmation of the information they had just obtained from the machine! That observation would tend to prove

that, even if the users obtained the good information, they probably missed something and somehow distrusted the machine because of lack of confirmation.

4- DIALOGUE AUTOMATIC PROCESSING

It is obvious a dialogue devoid of psychological aspects can be more easily implemented than more natural ones. We tried to substitute automatic processing (DIALORS) (2) for the operator simulating the machine.

4.1. Dialogue model

The simulation is rendered possible because one of the main characteristics of alternate interaction is to respect turns. In the dialogue model proposed here, all the utterances of the corpus can be represented by parenthesized formulas. The relatively limited nature of the task led us to choose a hierarchical frame structure to represent the task knowledge: different frames represent different types of requests (concerning train schedules, ticket prices, bookings, etc.) (10). We distinguish among different levels of questions and associated answers:

QP/RP are Principal questions/answers which concern the principal slot of a frame and require the instantiation of this frame.

QS/RS are Secondary questions/answers which lead to modification of one of the secondary slots of the frame (change of day or time for a time request, for instance).

QI/RI are Incident questions/answers for confirmation, explanation or reformulation and do not necessitate the modification of any slot.

Example 2 may be represented by the following formula:

(QP/C1(QI/M2 RI/C2)RP/M4(QS/C3 RS/M5))

Such a descriptive model takes the machine/caller alternation into account, as well as the inevitable sequences of questions which are difficult to handle automatically.

C1 : quelle heure est-il ? (what time is it?)

M1 : *pardon ? (I beg your pardon?)*

C2 : quelle heure est-il ? (what time is it?)

M2 : *vous me demandez quelle heure il est ? (you're asking me what's the time?)*

C3 : oui quelle heure est-il ? (yes, what time is it?)

M3 : *voulez-vous savoir l'heure à New-York ou à San Francisco ? (do you want to know the time at New-York or at San Francisco?)*

C4 : quelle heure est-il à San Francisco ? (what time is it at San Francisco?)

M4 : *midi (twelve o'clock)*

Example 5 : Imaginary dialogue composed of 7 consecutive messages in an interrogative mode.

The dialogue shown in Example 5, invented for the circumstance, may easily be represented by the following formula, because the dialogue model is based on speech-acts and not only on the utterance surface forms:

(QP/C1(QI/M1 RI/C2)(QI/M2 RI/C3)(QI/M3 RI/C4)RP/M4)

In the whole corpus (125 communications), only 6 communications failed to be processed with this model, and in these communications, the callers adopted a contestation attitude, which means that they did not respect the principle of cooperation described by Grice, at least in the context of human-machine dialogue (11).

4.2. Oral Phenomenon

Even with a machine, communication is made through an oral channel, and, therefore, oral phenomena linked to spontaneous interaction still remain, especially "echo" phenomena. As was mentioned before, echoes appear to play the role of iterative approbation signs in human-human communication.

When occurring in human-machine communication, we may wonder if these repetitions are mere approbations or implicit requests for confirmation or precision. If a caller utters *midi (12 o'clock)* when the system has just proposed a train leaving at 12 o'clock, it means either that the caller has memorized the information, or that he fears he has misunderstood it and is asking for a repetition of the information.

It appeared to be necessary to integrate a rule in the system for handling "echoes": if the repeated segment is immediately at the beginning of the exchange, it is considered as a mere echo and is ignored. If the segment is slightly shifted in the message, it has to be taken into account. A *midi oui (12 o'clock, yes)* answer is therefore processed differently from *oui midi (yes, 12 o'clock)*. This extra rule allows the system to correctly process 16 interventions in a corpus (of 750) without introducing any errors.

5- CONCLUSION

In conclusion, we stress two points. First, the observations reported here are task-dependent. The simplicity of the task, the knowledge of which can be represented by a limited set of frames, allows for the implementation of a hierarchical-structured dialogue; more complex tasks such as the one concerning a student Information and Orientation Center (CIO) (2) appear to need a more sophisticated structure interweaving hierarchical and linear elements.

Second, oral characteristics even if they are still present tend to occur less frequently in front of the simulated machine, making automatic processing possible. Moreover, the speaker dialogue attitude is influenced by the machine behaviour. The system capabilities as well as the way the dialogue is handled have an effect on the speaker behaviour, even if there seems to be a general paradigm of what people think the machine is. In particular, the way the messages are delivered by the machine largely contributes to modify the image people have of the machine. This observation emphasizes the importance that must be placed on linguistic studies and message generation.

REFERENCES

- (1) LUZZATI, D. "ORSO: projet pour la constitution et l'étude de dialogues homme-machine" n° 1, note scientifique, LIMSI, 1984.
- (2) LUZZATI, D. "Recherches sur le Dialogue Homme-Machine: Modèles Linguistiques et Traitements Automatiques", thèse d'Etat Paris III, 9 Juin 1989.
- (3) LUZZATI, D. & NEEL, F. "Linguistic Behaviour brought about by the machine", Proceedings of the 1st European Conference on Speech Technology, Edinburgh, Sept. 1987.
- (4) ALLEN, J.F. & PERRAULT, C.R. "Analyzing Intention in Utterances", Artificial Intelligence, n° 15-3, p. 143-178, 1980.
- (5) TAYLOR, M.M., NEEL, F. & BOUWHUIS, D. (eds) "The Structure of Multimodal Dialogues", Amsterdam North Holland, 1988 (in press).
- (6) ROULET, E. & al., "L'Articulation du Discours en Français Contemporain", Berne, Peter Lang, 1985.
- (7) BEUN, R.J. "The Function of Repetitions in Information Dialogues", IPO Annual Progress Report, 20, 1985.
- (8) MOESCHLER, J. "Argumentation et Conversation, éléments pour une Analyse Pragmatique du Discours", Paris, Hatier. 1985. (coll. LAL)
- (9) GOFFMAN, E. "Forms of Talk", Basil & Blackwell, Oxford, 1981.
- (10) BOBROW, D. & al., "GUS, a frame-driven Dialog System", Artificial Intelligence, n° 8, p. 155-174, 1977.
- (11) GRICE, H.P. "Logic and Conversation", in Syntax and Semantics 3: Speech-Acts, Coles & Morgan, Academic Press, New York, p.41-58, 1975.