Linguistic features of Human-Machine oral Dialogue

DOMINIQUE DELOMIER, ANDRE HEUNIER, MARY-ANNICK MOREL

Université de Paris III- Sorbonne Nouvelle
Centre de recherche en Morphosyntaxe du Français contemporain

A thorough analysis of two corpuses made of oral dialogues gathered in authentic situations and comprising two phases of a dialogue with a machine has allowed the identification of three structuration levels in the structure of a dialogue: the informational one, the interactional one and the syntactic one (concerning the global syntax as well as the local one). These three structuration levels can be described in terms of specific features of pragmatic and, above all, linguistic nature, which appears as absolutely essential to the existence of the dialogue as such.

This linguistic study focuses on two very different corpuses which were collected in real situations in Paris: one in a telephone information center of the French railways, in December 1984 (corpus named: SNCF); the other in an information center for students, Ph. D. Level, at the University of Paris V (corpus named: C.I.O.) in April 1985. In both cases the language used is spontaneous oral French. The two corpuses have been completely transcribed into writing. The first situation is characterized by a large public and a limited task (schedule and prices), whereas in the second one the public is restricted, the task is more diversified and the goals less precise (university syllabus and courses, professional prospects and interests in those professions). The experiment consisted in three phases. In phase I, the human speaker communicated with a human operator. In phases II and III, the operator's voice was mechanized in order to give the user the impression he was talking with a machine. In phase III in order to make himself better understood, the speaker was asked to re-phrase some of his utterances.

Introduction

The study, presented in an extremely summarized way, is a result of a thorough analysis of linguistic facts encountered in the three phases of the experiment and aims at isolating the experiment and aims at isolating the irreducible constants about these facts. We would like to demonstrate the following:

- in the structure of a finalized dialogue, it is necessary to distinguish three structuration levels (i.e. the structure of information, the one of interaction and the one of syntactic patterns);
- there are specific rules that apply at these three levels;
- it is impossible not to comply with these rules without destroying the conditions of existence of an authentic dialogue.

We shall, in our study, consider mainly two examples from corpus C.I.O., one at the beginning and the other in the course of the dialogue:

(1) E1--alors ben j'voudrais savoir d'abord la: /la différence qu'il y a entre les deux
O2--la différence entre la psychologie du travail et l'ergonomie
E2--et l'ergonomie ben je croyais qu'il était la même chose
O2--bien, alors on va définir ce que sont ces deux: enseignements, ces deux professions
E3--oui
O3--la psychologie du travail, prépare à l'exercice des fonctions de psychologue, dans les entreprises industrielles et commerciales
E4--dans les entreprises industrielles et commerciales, d'accord [...]
(28 exchanges between the speakers on this theme)
E32--d'accord d'accord, et pour e /pour l'ergonomie alors
E30--pour l'ergonomie, bon l'ergonomie, e: /est:, l'application, du de la psychologie du travail, et de la physiologie (CIO 1.3)

(2) E67--[...] le: c trois par exemple ça serait psychologie sociale et le c quatre psychologie experimentale
O65--davantage, en psychologie du travail, il est vrai qu'il: le tronc principal au niveau d la maitrise c'est la psychologie sociale
E68--"mm" d'accord <h>... bon ça c'est pour la maîtrise alors,
The codification of the oral specific phenomena (silence, pause, hesitation, intonation,...) is explained in Leroy (1985). The reference of the examples in the corpora are noted at the end of the quotation with the name of the corpus (SNCF or CIO), the number of the phase (1, 2 or 3), and the number of the dialogue. The speakers are noted O = operatrice (human expert), M = machine, E = etudiant (CIO), C = correspondant (SNCF). The number O and M end "caller" E and C. The number which follows the designation of the speaker corresponds to the place of the utterance in the dialogue.

1. Structure of information

In order to provide an adequate informative reply, the expert has to know all the thematic data (or informative units) necessary for his task to be performed. Three series of linguistic operations accomplished during the dialogue can be identified.

1.1 Negotiation of the thematic data

To fill the missing items which correspond to the essential informative units, the expert has to ask additional questions. In (1) E1 for instance the anaphoric item "les deux" is ambiguous and the expert has to ask for a specification of the referent: O1 "la différence entre la psychologie du travail et l'ergonomie".

1.2 Storage of the thematic data

The general thematic data provided at the beginning and in the course of the negotiation must be stored until the end of the dialogue. In fact, a number of questions which imply recalling of a previously mentioned theme (e.g. SNCF: return schedule of a train) can only be interpreted as a function of that previous question as well as thematic data already stored at that time.

We observe that phenomena in (1) E32: there has been 28 exchanges between the expert and the caller. These two types which we have analysed in our corpora.

1.3 The expert's re-clarification of certain data at particular moments of dialogue.

It is natural that thematic data are not systematically re-clarified in replies or in additional questions. However such a re-clarification seems necessary on occasions, either in order to restrict the area of information validity (cf (2) O65="avantage en psychologie du travail") or to avoid referential ambiguity (cf (2) O66=trois certificats conseillés pour le dess de psychologie du travail"). It corresponds to authentic rules in a Human-Human dialogue.

2. Structure of interaction

2.1 In an authentic finalized dialogue, the speakers devote part of their exchanges to make sure that the information has been properly conveyed. Therefore, we have

a) on behalf of the caller:
- questions which constitute a request for confirmation, to make sure that he understands the information properly (cf (2) E65-"bon donc y a trois/trois couples conseillés")
- requests for additional information
(3) M26--le nombre d'heures est nettement inférieur en ergonomie
E27--autant c'est-à-dire
M27--de dix à douze environ (CIO 3.5)

b) on behalf of the expert:
- request for confirmation of thematic data: (cf (1) O1 already mentioned
  confirmation statements: cf (2)
O65--"trois couples conseillés pour le dess de psychologie du travail"

- statements which rectify an information not properly understood: (cf (2) O65--"il est vrai que le: tronc principal au niveau d la maîtrise c'est la psychologie sociale".

The utterances at that level of interactional structure are either more elliptic, even in the experimental dialogue with the machine (phases II and III).

2.2 Two main types of interactional scheme can therefore be identified (we do not present here the subdivisions of these two types which we have analysed in our corpora).

a) First type of interactional scheme

We found the scheme on the pattern of example (1)

E1- [Main initiating question]

O1- [Additional question (to confirm a proper understanding of the question)]

E2- [Main informative reply]

b) Second type of interactional scheme

Example (2) is concerned in this scheme

E67- [Initiating question]

O65- [Informative reply]

E68- [Recording process + question to confirm the information]

O66- [Reply to confirm]

E69- [Recording process + New initiating question]

Should this interactional component be intentionally eliminated (as it was the case during the recording of pha-
ses II and III of our corpus), the interaction tends to be obstructed and the dialogue is then set in a predictable form in which the caller is constantly forced to manage himself the exchanges.

c) Predictable type of non interactional scheme

(4) E27—et: ou m: /* est-ce qu'il est préférable qu'il e: /qu'il complète sa formation par autre chose, après le dess

M28—après le dess cela est préférable

E28—et qu'est-il e préférable /après le dess: e de: /de faire

M29—après le dess vous pouvez vous engager dans diverses formations, cela depend de vous

E29—et quel /: dans quel type de formation peut-on s'engager, après le dess: (CIO 3.17)

Scheme:

S27—[Initiating scheme

M28—[Informative reply (deceptive one)

E28—[Question to obtain more precise information

M29—[Informative reply (deceptive once more)

E29—[Question to obtain more precision (once more)

3. Structure of syntactic patterns

In a Human-Human dialogue, the form of statements produced is not predictable, since it is a function of a dialogue project common to both participants and it involves permanent integration and adjustment of linguistic productions of both speakers.

Now, in a Human-Machine dialogue, the syntactic patterns, inherent in a finalized dialogue, tend to be more regular in global syntax (which cannot be eliminated) as well as in local syntax.

3.1 Global syntax of dialogue exchanges

The regularities appear at very special points.

Specific form of question statements

In order to make himself better understood (maximum intelligibility concern), the caller tends, in both corpuses, to first provide the thematic data (which we call topic, abbreviated TOP) and, only after that, the wording of the question itself (abbreviated Q), following the general scheme: [TOP], [Q]

(5) C4—[Jusqu'à 19 heures 30] [je voudrais les heures des trains] (SNCF 2.73)

which can be refined as follow:

a) Both components of the statement are most often accompanied by specific marks: topic introducers (I) (for instance: "au niveau de, du point de vue, pour, en ce qui concerne...") and question markers (Mq) ("je voudrais savoir, est-ce que, pouvez-vous me dire...").

The scheme is: [I + TOP], [Mq + Q]

(6) E44—e: m (h) et: (h) [au niveau du /deal], m: [pouvez-vous, pouvez-vous vous m'donner une idée des; matières enseignées en psychologie du travail]

(CIO 3.19)

It has been noted that these markers (I and Mq) appear much more frequently and are more explicit and more various, when the caller is talking to the machine.

b) The topic itself can often be analyzed in sub-topics (TOP), thus allowing maximum refining of thematic data

[I + TOP = (I1 + top1) + (I2 + top2) + ... ], [Mq + Q]

et: j'ai entendu dire, (h) que

[(I1+top1) pour faire e: le dess de psychologie du travail

(I2+top2) au niveau de la maitrise

(o+top3) il faut choisir ses valeurs e

alors quelles se / quelles seraient ces /les valeurs e: recommandées (CIO 3.11 El6)

Eliminated this type of dislocated statements, to obtain — in a intentionally induced re-phrasing (as it was experimented in our corpuses phase III) — a question statement complying with the standard of written language appears, not only useless, but even prejudicial to the proper transmission of the information.

(7) E30—[...] oui vous m'avez dit qu'en /qu'on /que pour le dess il y avait différents secteurs, mais quand on passe le dess il y a /autre chose que la psychologie du travail. Il qu'est-ce que /qu'est-ce qu'on nous propose

[N asks for a re-phrasing]

E31—alors je voudrais savoir de quoi se compose le dess

(CIO 3.10)

Specific form of reply statements

Very frequently, the human expert re-calls, at the beginning of his reply, one or several thematic data of the question, without integrating this component syntactically in the reply itself which follows. Here is another linguistic specificity which seems necessary to the proper functioning of a finalized dialogue. Scheme is:

[Hanging Topic] + Mark + [Reply = Achieved Sentence]

See exemple (1) O30

MT: pour l'ergonomie,

Mark: bon

Reply: l'ergonomie, e: /est:...

Elliptic form of certain utterances

The ellipsis appears to be a linguistic fact which is inherent in a finalized dialogue and does not seem, all the contrary, to be an obstacle to the intercomprehension between the

EUROSPEECH '89, Paris, France, September 1989 2238
speakers. In our corpuses, all the strategies used by the machine which tend to avoid the ellipsis (cf the verb "existent" in (8) to avoid an Hanging Topic) end in producing bizarre and at least unnatural statements.

(8) E6--[... et e comment est-on admis dans chaque dessus ?] M7--les conditions d'accès existent au niveau de l'entree en dessus, un dossier est ouvert pour chaque étudiant [... ] (CIO 2.7)

The ellipsis is compulsory in the following cases:

a) Replies. The syntactic form of the reply is a function of the form of the corresponding question and the couple Q-R constitutes an indissociable syntactic unit.

(9) E27--[... le dessus d'année combien d'années ?] M27--une année (CIO 2.2)

b) Additional question: cf (2) E59--"et puis ensuite ergonomie c'est pareil !"

c) Stating that the information is being processed: cf (1) E4--"dans les entreprises industrielles et commerciales, d'accord", or (2) E29--"un dess de psycho".

3.2 Local syntax within a statement

The local syntactic rules affect mainly the marking of thematic data (cf above I + TOP) and components of the question, which are more easily noticeable in the dialogue with the machine. In fact, it appears that in a dialogue with more constraints (phase III), the caller tends to use strong marks, meaning full content, rather than polysemic grammatical words, such as "a" and "de". There is a large variety of marks, very frequently used in the Human-Machine dialogue, such as prepositional phrases ("jusqu'à, en direction de, en partance de, ..."), prepositional groups ("pour aller à, pour aller vers, ..."), gerunds or present participles ("en allant à, partant de, ...") and even relative clauses:

(10) C1--e: les trains pour brussel, e: les deux prochains

[M asks for a re-phrasing]

C2--e: les trains e: qui partent pour brussel de la gare du nord, les deux prochains trains

[M asks once more for are-phrasing]

C3--ce que vous pouvez m'dire, les /les deux prochains trains qui partent pour brussel, de la gare du nord (SNCF 3.38)

These strong marking procedures of data, necessary to the information processing, express in our opinion, the maximum intelligibility concern in a situation where the human being is facing a "restricted partner", the machine.

Conclusion

In conclusion, we would emphasize that the specificities of the finalized dialogue presented here do not disappear in the less constraint phase of the dialogue with the machine (phase II), where only a more controlled language and the mechanized voice give the illusion of a dialogue with a computer. However, a certain influence of the machine has been noticed in this phase, namely: slowing down of the speech, disappearance of some phonic marks and restriction in the variety of those markers which accompany the question ("alors, et donc sinon, ... ").

At the same time, when the constraints are enforced (phase III), when the machine is constantly asking the caller to re-phrase his question or to give additional precisions on the referential field of the terms he is using, the re-phrased statements show a tendency to regular forms and a densification of the utterances which could be seen as helping the machine in understanding the statements. However, as we have attempted to demonstrate, this use of regular forms and this densification do not respect the natural intelligibility law, which has a negative effect on the proper functioning of the dialogue: the silences, the "ehu" (hesitation marks) tend to replace requests for confirmation and statements confirming the processing of the information. Moreover, the caller's tone, the metadiscursive comments he produces ("je laisse tomber ma question" for instance), and even the nature of some markers he uses ("ah bon, tant pis, ..."), express his disappointment and irritation which eventually lead him to abandon the dialogue. While these remarks apply essentially to the CIO corpus, it is most interesting to notice that, in the two corpuses (SNCF and CIO) dialogues with the machine are three to four times shorter than those with a human operator, which means that it is unnatural for the caller to manage himself the whole dialogue. Therefore, in order to elaborate an operational pattern of oral Human-Machine dialogue, it seems necessary to take into account all the specific languages which we have identified as pertaining to the three structuration levels of the finalized dialogue.

References


