

LES MARQUEURS DE LA DEMANDE D'INFORMATION DANS UN CORPUS DE DIALOGUE HOMME-MACHINE

André Meunier
Paris III-UA 1032 CNRS

Mary-Annick Morel
Paris III-UA 1032 CNRS

1. Corpus d'étude

Le corpus sur lequel nous travaillons a été recueilli dans le cadre d'une expérimentation réalisée par le GRECO Communication Parlée, Groupe Dialogue Homme-Machine (Responsable Jean-Paul Haton, CRIN, Nancy) et financée par le CNET. Il s'agit d'une recherche pluridisciplinaire (regroupant des chercheurs en informatique, reconnaissance de la parole, intelligence artificielle, ergonomie et linguistique). Les enregistrements sur lesquels nous travaillons actuellement ont été effectués en Mars 1985 au Centre d'Information et d'Orientation de l'Université de Paris V: des étudiants viennent s'informer sur les cursus en DEA et DESS d'ergonomie et de psychologie du travail. Ils obtiennent leurs informations par l'intermédiaire d'un téléphone, spécialement installé pour l'expérimentation. Ils sont prévenus qu'il s'agit d'une expérimentation.

L'enregistrement comprend trois phases, précisément définies par un protocole. Dans la phase I, le dialogue s'établit entre l'étudiant et une opératrice humaine. Dans les phases II et III, c'est toujours la même personne qui répond, mais sa voix est "mécanisée", déformée à l'aide d'un vocodeur, et son langage obéit à certaines contraintes, ce qui donne à l'interlocuteur l'impression de dialoguer avec une machine. De plus, en phase III, l'étudiant est soumis à une contrainte de compréhension, ce qui l'oblige à reformuler certains de ses énoncés à la demande de la "machine".

Les enregistrements des dialogues (15 heures au total) ont été transcrits orthographiquement sur IBM-PC, avec un système de notation des faits d'oralité. Pour rendre notre étude plus homogène, nous avons restreint ce corpus aux phases I et III et avons écarté deux communications, ce qui représente: en phase I, 7 communications et

73 pages de transcription; et en phase III, 18 communications et 72 pages de transcription. Nous n'étudions que le langage de l'étudiant, à ce stade de notre recherche.

Remarques.

1. Pour plus de détails sur le protocole d'expérimentation, voir notre article de LINX n°18, 1987.
2. Deux études (polycopiées) ont déjà été réalisées à partir de cette expérimentation (qui comprenait aussi un enregistrement effectué au Centre de Renseignements de la SNCF à Paris) et transmises au GRECO et au CNET:
 - Etude ergonomique, par J. C. Spérandio et Ch. Létang-Figeac, Université de Paris V, Février 1986.
 - Analyse linguistique du corpus SNCF, par des chercheurs de Paris III et du LIMSI, sous la responsabilité de M. A. Morel, C.I.F Paris III, Décembre 1985.

2. Objectifs

La recherche effectuée au niveau linguistique, à partir de ce corpus de dialogues finalisés (demande d'information), a pour objectif général d'observer les modifications qui se produisent dans le comportement langagier des gens face à une machine, d'une part en ce qui concerne la structure syntaxique et énonciative de leurs énoncés, et d'autre part du point de vue de la structure générale du dialogue. Les observations que nous présentons ici s'insèrent donc dans une recherche plus large. Mais elles sont centrées sur l'étude d'un certain nombre de termes, que nous appelons "Marqueurs", qui accompagnent de façon récurrente les questions posées par l'étudiant à l'opératrice (phase I) ou à la machine (phase III).

Nous commencerons donc par une présentation de ces Marqueurs, dont nous justifierons le choix. Puis nous introduirons un certain nombre de données générales, provenant de la comparaison systématique du dialogue en phase I et en phase III et nécessaires pour délimiter le cadre spécifique de notre étude et rendre compte du rôle des Marqueurs retenus. Nous soulignerons en conclusion les particularités observées dans la phase de dialogue Homme-Machine (phase III).

3. Présentation des Marqueurs et justification de leur choix

Nous avons choisi d'observer un ensemble de termes de statut grammatical divers (conjonctions de coordination, "adverbes de liaison", syntagmes plus ou moins figés) qui jouent un rôle dans la structuration et la progression des demandes d'information de l'étudiant face à l'opératrice et face à la machine. Nous avons relevé, sous le nom générique de "Marqueurs", uniquement des unités figurant seules ou combinées entre elles:

1. à l'initiale absolue d'une intervention de l'étudiant:

(1)

E9 -*et où est-il mentionné?*

M10 -*je ne peux pas répondre à cette question*

E10 -*maintenant en ergonomie, quelles sont: les:| possibilités à partir de la maîtrise| les choix,,| à faire pour un étudiant (3.2)*

2. avant question, mais après un Marqueur d'enregistrement (voir lexique), lui-même simple ou combiné:

(2)

E67 - *ouais d'accord, (h) bon sinon alors maintenant| plus: terre à terre pour la maîtri: se pour être avoir des chances disons de faire le dess d'ergonomie (1.7)*

3. avant question, mais après un bref segment consistant en une reprise textuelle de la réponse antérieure ou d'une évaluation qui verbalise la satisfaction ou l'insatisfaction de l'étudiant:

(3)

O118 -*(...) vous pouvez travailler en groupe, le secteur de la recherche en ergonomie est::| est en grand développement*

E122 - *le secteur d la recherche | -| mais c'est::| c'est c'que f'ra c'est c'que vous faites comme travail non? (1.3)*

Nous avons appelé M1 les conjonctions de coordination: *et, mais, donc, ou*. Et M2 un ensemble hétérogène comportant notamment: *alors, enfin, sinon, maintenant*, représentant des classes plus ou moins fournies. Certains de ces Marqueurs ont un rôle spécifique dans la structuration du questionnement, d'autres sont polyvalents. L'analyse de ces rôles sera présentée ci-dessous (cf. 7).

4. Structuration physique du dialogue: différences globales

Pour évaluer le jeu de ces Marqueurs dans le déroulement des deux phases, nous nous appuyons sur un certain nombre d'informations tirées de l'étude ergonomique évoquée ci-dessus (cf. 1).

4.1. On retiendra notamment que la durée moyenne des communications est plus réduite en phase III (9mn 28s) qu'en phase I (15mn); qu'en I l'opératrice parle en moyenne trois fois plus que l'étudiant (O: 11mn 14s; I: 3mn 24s) tandis qu'en III l'échange est équilibré (M: 4mn 5s; E: 3mn 47s).

Le débit moyen (en mots par seconde) est plus élevé en I (O: 3,03; E: 4,05) qu'en III (M: 2,19; E: 2,58): s'il est normal que la machine, du fait des consignes données et du filtrage effectué par le vocodeur, parle plus lentement qu'une opératrice humaine, on notera l'effet de mimétisme opéré sur le débit de l'étudiant, qui est nettement ralenti (4,05 en I; 2,58 en III).

4.2. Le décompte des silences séparant les interventions fournit des indications directement utilisables dans l'étude de nos Marqueurs initiaux.

En phase I, le total moyen des silences enregistrés, par communication, est de:

- 42,13s avant intervention de l'opératrice (soit 6,4 % par rapport au temps de parole dans l'intervention);
- 48,18s avant intervention de l'étudiant (soit 23,6 %).

En phase III, les chiffres sont respectivement:

- 75,33s avant intervention de la machine (soit 30,6 %);
- 70,45s avant intervention de l'étudiant (soit 31 %).

L'augmentation globale des silences des deux interlocuteurs en phase III est caractéristique d'un dialogue d'où ont disparu pratiquement les phénomènes d'enchaînement du tac-au-tac, de chevauchement des interventions, d'interruption de la part de l'opératrice et de l'étudiant. Ce temps de latence important précédant, en phase III, la demande d'information sera à interpréter en relation avec la forte déperdition des

Marqueurs (notamment du type M2 et surtout des combinaisons de Marqueurs). Nous reviendrons plus loin sur ce point qui correspond à une des spécificités de la phase III.

5. Comparaison des différents modes d'intervention

Pour établir des statistiques comparatives sur les énoncés de question, il nous a fallu distinguer différents modes d'intervention de l'étudiant au cours du dialogue.

5.1. Interventions "avec contenu" et interventions "d'enregistrement".

L'énoncé d'enregistrement de l'information ne constitue pas à proprement parler une prise de parole, mais il est interprétable comme une marque phatique (le dialogue se déroule, il ne faut pas l'oublier, par téléphone). Il n'interrompt pas vraiment le déroulement de l'intervention en cours:

- (4)
- O56 *-(...) je ne peux pas vous l donner, étant donné que: les enseignants ne sont pas toujours les mêmes*
 - E27 *-oui*
 - O57 *-que le programme change et que (...)*
 - E58 *-mm, ouais < dans un soupir >*
 - O58 *-dans le certificat principal hein,, c't-à-dire qu'vous avez (...)*
 - E59 *-m'oui*
 - O59 *-(h) dans ce: certificat principal c'est à vous de (...)* (1.2)

Les énoncés d'enregistrement disparaissent presque totalement en phase III (1,6 % du total), alors qu'ils représentent 46 % des interventions des étudiants en phase I.

5.2. Interventions "avec contenu" qui ne constituent pas des questions.

En phase I les interventions de ce type correspondent soit à une réponse à une question ponctuelle de l'opératrice, soit à un commentaire sur l'information qui vient d'être donnée, soit encore à une digression:

- (5)
- E82 *-(...) c't-à-dire j'avais| oui ben oui c'est ça| enfin j'aurais bien aimé prendre trois matières quoi en maîtrise c'est ça qui m'embête* (1.5)

Elles représentent 17 % du total des interventions. Ce qui revient à constater que 37 % seulement du total des interventions de l'étudiant, en phase I, constituent des questions. En phase III, commentaires et digressions disparaissent et il ne reste plus que les énoncés de réponse, soit 4 % du total. Il est notable que 94,4 % du total des interventions des étudiants, en phase III, constituent des questions:

- (6)
E36 -*quel est le temps de:/ des études,/ par semaine (...)*
M37 -*en dess?*
E37 -*d'ergonomie ouais (3.9)*

5.3. Interventions comportant plusieurs énoncés interrogatifs.

On rencontre un certain nombre de cas où l'intervention de l'étudiant comporte plus d'une question (4 % en phase I; 3,5 % en phase III). Il s'agit alors soit d'une reformulation spontanée de la question initialement posée, soit de deux questions distinctes mais complémentaires:

- (7)
E130 -*enfin bon:, et:/ et c'est le/ qui décide enfin j'veux dire est-ce qu'il y a moyen de:/ de savoir (...)* (1.7)
- (8)
E31 -*(h) ah/ bon alors j'voulais vous/ je voulais savoir de quoi se compose le dess est-ce que, en dess on prend simplement/ on se spécialise simplement en psychologie du travail ou est-ce qu'il y a d'autres matières,, générales* (3.10)

5.4. Interventions comportant des questions "doubles".

5.4.1. On est frappé par la fréquence des questions doubles en phase III (12,3 % du total des questions): elles sont quatre fois plus fréquentes qu'en phase I et ont un taux de dispersion beaucoup plus important. Les deux termes de ces questions sont reliés soit par *et*, soit par *ou*. En fait, seuls les termes reliés par *et* constituent à proprement parler des questions doubles:

- (9)
E13 -*ben déjà e/ qu'est-ce que l'ergonomie,,/ et: (h) quel est/ et de la même façon quels sont les enseignements qui sont dispensés dans ce diplôme* (3.12)

Les termes reliés par *ou* sont soit en relation exclusive (pôle positif et pôle négatif ou termes antithétiques), soit en relation de complémentarité (cf. ex. (8)):

- (10)
E28 -*l'admission (...) se fait-elle après: concours après examen: après: (h) ou après une autre modalité* (3.12)

Le deuxième terme de la question double peut ne pas comporter de verbe (il est alors de type Cq, cf. ci-après 6) et peut, dans le cas d'une alternative négative, être réduit à la négation (*ou pas, ou non*).

5.4.2. En phase I, en revanche, on note une plus grande fréquence qu'en phase III d'énoncés de question qui se terminent par un *ou* laissé en suspens, soit que l'étudiant ne

juge pas utile d'achever la séquence (il y a alors un silence), soit que l'opératrice l'interrompt avant l'achèvement de la construction (il y a alors chevauchement):

(11)

- E5* -est-ce que c'est deux ergonomies différentes ou:
- O5* -non pas réellement (1.7)

De cette observation on peut conclure que les énoncés de question de l'étudiant sont plus achevés, plus complets et plus denses en phase III qu'en phase I.

5.5. En conclusion de cet examen des différents modes d'intervention, on peut souligner qu'en phase III le dialogue présente un caractère plus rigide, plus économique et plus dense: les marques phatiques disparaissent et les énoncés représentent de façon massive des questions.

6. Formes de questions. Marquage des questions

Si nous retenons le très fort pourcentage d'énoncés comportant une demande d'information en phase III (94,4 % contre 37 % en phase I) nous devons examiner la structure syntaxique de ces questions et la manière dont les Marqueurs y sont associés.

6.1. Nous avons distingué quatre types de questions (voir lexique):

- Q (type: *Qu'est-ce que l'ergonomie?* ou *Je voudrais savoir ce qu'est l'ergonomie*)
- Qd (type: *L'ergonomie c'est en expansion?*)
- Qm (type: *L'ergonomie c'est quoi?*)
- Cq (type: *(Et) l'ergonomie?*)

Les pourcentages sont les suivants, par rapport à l'ensemble des questions de chaque phase:

type	phase I	phase III
Q	22,7 %	67,3 %
Qd	54,5 %	15,1 %
Qm	8 %	4,7 %
Cq	14,8 %	12,9 %

Ce qui est notable est la très nette prédominance de Q (question de forme canonique) en phase III et celle de Qd en phase I.

6.2. Globalement, pour l'ensemble des questions, on constate qu'en phase I 85,7 % des interventions sont accompagnées d'un Marqueur de l'ensemble défini ci-dessus (cf. 3) contre 14,3 % sans Marqueur. En phase III, seulement 52,2 % des questions sont avec Marqueur, contre 47,8 % sans Marqueur.

Si l'on examine chaque type de question, on constate, parmi les faits les plus frappants, que:

- en phase I prédomine le type Qd accompagné d'un Marqueur (45,9%), lequel type est faiblement représenté en III (9,9 %);
- en phase III prédomine le type Q sans Marqueur (36,8 %) lequel est très peu présent en I (1,4 %).

Parmi les questions marquées on relève en phase I 47,4 % de Marqueurs simples (M1 ou M2) et 38,3 % de Marqueurs combinés: ces combinaisons de deux, trois, voire quatre M1 (ou M2) et autre(s) M2 (cf. ex.(2) ci-dessus) sont à la fois fréquentes et variées en I; parmi les plus fréquentes:

et alors, et sinon, et donc, mais alors, mais enfin, mais sinon, mais sinon enfin, mais alors donc, et alors donc sinon... (97 types de combinaisons relevés).

En phase III on ne trouve que 9,4 % de Marqueurs combinés (pour 42,8 % de Marqueurs simples). Les combinaisons sont plus rares et moins riches (32 types différents: deux Marqueurs au maximum; notamment: *et ensuite, et alors, et sans ça*).

6.3. Enfin, s'agissant des types de Marqueurs, on peut constater que dans les deux phases ce sont les M1 qui prédominent (sur les 85,7 % questions marquées en I, 62,8 % le sont avec M1 et 22,9 % avec M2; en phase III sur les 52,2 % de questions marquées, on trouve 39,6 % de M1 et 12,6 % de M2). La tendance globale de la phase III se confirme assez nettement de ce point de vue: allègement général du marquage des questions (lesquelles tendent majoritairement à une formulation canonique, conforme à la norme académique): appauvrissement de ce marquage tant en ce qui concerne les combinaisons (comme on vient de le voir) qu'en ce qui concerne le nombre et la variété des Marqueurs simples.

Quelques sondages effectués en des points "stratégiques" du dialogue le montrent précisément:

- en début de communication: 100 % des questions avec Marqueurs en I; 25 % en III;
- au changement de thème principal (cf. lexique): 100 % de questions marquées en I; 50 % en III;
- après demandes de reformulation (en III), la nouvelle question n'est pratiquement jamais marquée (1 cas sur 42).

6.4. Corrélativement il faut signaler le fait que les Marqueurs d'enregistrement continuent d'accompagner les questions en phase III: ce qui signifie que, si l'étudiant restreint les marques de structuration de ses demandes d'information (tout comme d'ailleurs les ponctuants: *quoi, en fait, hein...* et les Marqueurs de recherche d'approbation discursive: *non? n'est-ce pas? oui? d'accord?...* entre autres signaux d'interaction, et les autres outils logiques et discursifs servant à organiser le corps de son intervention, phénomènes que nous n'étudierons pas ici faute de place), il maintient ces expressions stéréotypiques, qui ne font que signaler la prise en compte de la réponse de la machine, sans articulation réelle de la nouvelle question par rapport à l'information fournie.

Ce sont ces diverses articulations possibles que marquent justement les M1 et les M2 dont nous allons, après l'approche strictement quantitative qui précède, examiner les rôles divers en nous demandant si, de la phase I à la phase III, ils sont modifiés ou non de façon sensible.

7. Polyvalence des Marqueurs. Distribution relative de M1 et M2

7.1. Les Marqueurs M1.

Et, ou, mais, donc conservent en phase III à peu près les mêmes rôles qu'en phase I. Ils restent en tout cas polyvalents.

ET 1. Lien entre les deux termes d'une question double (cf. ex. (9))

2. Marqueur d'enchaînement dans le déroulement des questions portant sur un même thème:

- (12)
- E2 *-(...) d'autre part je voudrais savoir: e:: si, e: on peut passer un certificat de physiologie*
 - E4 *-oui: (h) et je voudrais savoir: e: si il dure pendant toute l'année*
 - E5 *-et: combien y-a-il d'heures de cours par semaine (3.14)*

3. Marqueur de reprise d'un thème antérieurement abordé:

- (13)
- E27 *-(h) et:/ e, dans quel secteur trouve-t-on le plus d'emplois < reprend le thème de E8 > (3.10)*

4. Introducteur d'un thème nouveau ou d'un sous-thème:

- (14)
- E20 *-e: le dess e d'ergonomie, se compose,, e: de/ de quels enseignements*
 - E24 *-et,, (h) quels sont les débouchés qu'il y a après (3.3)*

MAIS 1. Introducteur d'une demande de précision ou de rectification sur ce qui vient d'être dit:

- (15)
- M49 *-non pas du tout, vous pouvez vous orienter par exemple dans le secteur de la recherche*
 - E52 *-mais c'est:/ ça c'est dans l'cadre: de l'université (3.5)*

2. Introducteur d'une question qui tient à coeur à l'étudiant:

- (16)
- E61 *-(...) au niveau de la thèse quels sont:e/ (...) m des exemples de sujets (...) que l'on peut faire en ergonomie*
 - M62 *-cela dépend de vos intérêts*
 - E62 *-m: oui: (h) mais c'est: généralement cela porte: enfin sur quel type de sujet cela peut porter (3.19)*

3. Après une reprise (souvent textuelle) d'une information:

- (17)
- E11 *-(h) et: e:/ vous avez parlé d'un stage pour le dess d'ergonomie (h) mais e: doit-on préparer un s/ doit-on e faire un stage (...) au niveau de la maîtrise (3.16)*

4. Associé à **sinon**, introducteur d'un nouveau thème.

Ce cas ne se rencontre pas en phase III, où c'est *et* qui joue ce rôle:

- (18)
- E151 *-mm d'accord, mais sinon enfin point de vue sélection, y a plus de sélection quand même pour rentrer en dea qu'en dess (1.3)*

DONC 1. Soulignement d'une demande de confirmation

- soit sur une récapitulation
- soit sur une déduction

que E fait à partir de ce qui vient d'être dit:

- (19) E26 -après le dess e d'ergonomie, l'étudiant est donc e/ apte, à travailler (3.3)

2. Spécificateur du recentrage sur un thème antérieurement abordé:

Ce cas ne se rencontre qu'en phase I. En phase III, c'est *et* seul qui joue ce rôle ou bien il n'y a pas de Marqueur.

- (20) E26 -ouais,, et donc ce: psycho expérimentale et psycho sociale: c'est:| c'est possible < cet énoncé recentre sur le thème de O18 > (1.2)

- (21) E20 -(h) et: e:: les stages sont d'un:| d'un mois minimum il faut faire deux stages à ce moment là < cet énoncé recentre sur le thème de M8 > (3.7)

OU Lien entre les deux termes d'une question double (cf. 5.4 ci-dessus).

La polyvalence des Marqueurs subsiste donc face à la machine. Il faut toutefois souligner deux différences dans l'emploi des M1 en phase I et en phase III: d'une part la grande fréquence de *et*, au détriment de *mais* et *donc*; d'autre part l'apparition plus tardive qu'en phase I de *mais* dans le cours du dialogue (en général après le premier tiers) et celle encore plus tardive de *donc* (en général dans le dernier tiers). Ces phénomènes s'expliquent par la manière dont s'effectue la progression thématique dans les deux phases: en I elle est complexe et comporte de nombreuses imbrications par reprise de thèmes et de sous-thèmes; en III elle est très linéaire et les reprises de thème ne se font qu'en fin de communication, après énumération de tous les thèmes.

7.2. Les Marqueurs M2.

Les termes que nous avons regroupés ici le sont, comme on l'a dit plus haut, sur la base d'une distribution particulière, analogue à celle des M1. Ils peuvent avoir des rôles voisins ou différents en d'autres positions au sein des interventions. Ils ont fait l'objet d'études précises et exhaustives (notamment dans les Cahiers de Linguistique Française de Genève). Dans notre analyse nous les traitons au sein de sous-ensembles définis par tel ou tel rôle dans la structuration du questionnement de l'étudiant. Certains se retrouvent dans plus d'un sous-ensemble (les polyvalents), d'autres sont plus spécialisés. Nous distinguons:

- les Marqueurs d'introducteur de thème ou de sous-thème ou de changement d'orientation thématique
 - *alors* d'ouverture absolue (en tout début de communication):

(22)

- E1 -alors ma première question
- O1 -oui
- E2 -oui, alors j'aimerais d'abord avoir une... (1.7)

- alors d'enchaînement linéaire:

(23)

- M16 -le psychologue est seul à prendre sa décision
- E14 -/ et i(l) se base uniquement sur l'entretien et les tests
- M17 -il se base sur les techniques qu'il aura choisies
- E15 -m, bon (h) e:: m'alors quel est le rôle,, c'est-à-dire quelles sont les activités du psychologue possédant un dess d'ergonomie, par rapport au psychologue possédant un dea d'ergonomie (1.4)

- sinon, sans ça, à part ça, d'autre part (non précédé de: d'une part), autrement, je laisse tomber ma question (ce type de formule de caractère métadiscursif verbalisant le travail énonciatif est fréquent aussi bien en phase I qu'en phase III). Ces Marqueurs combinés entre eux et/ou avec *et, mais, alors*, signalent la réorientation thématique (cf. ex. (2) ci-dessus).

Ce premier type de M2 reste bien représenté en phase III, surtout en ce qui concerne le *alors* d'enchaînement linéaire.

- Les Marqueurs-balises indiquant la progression du questionnement, son déploiement chronologique. On trouve dans ce groupe, très rarement en phase I, beaucoup plus en phase III: *en premier lieu, (tout) d'abord, ensuite, finalement* (fin d'énumération), *maintenant* (cf. ex. (1) ci-dessus), *deuxième question, en premier lieu déjà...*
Alors que globalement les Marqueurs sont plus fréquents en I qu'en III, ceux de ce type sont deux fois plus fréquents en III qu'en I: l'étudiant, face à la machine, "planifie" sa demande d'information, l'ordonne en catalogue: l'impression de non-dialogue perçue à l'écoute des communications provient souvent de ces types d'enchaînements (assez systématiques dans certaines communications).
- Les Marqueurs d'association de thèmes, *également, aussi, de la même façon, à propos, au fait, justement, ainsi que...*, ont une valeur anaphorique et indiquent une liaison thématique avec la réponse de l'opératrice ou de la machine. Ils sont présents en I et III, mais nettement plus nombreux en I où l'étudiant, par ces termes, justifie l'enchaînement des questions, ce qu'il est peu tenté de faire face à la machine.
- Les Marqueurs de connexion logique: spécification, inférence à la suite d'une réponse antérieure. C'est le rôle que jouent les deux *alors* de cet exemple:

(24)

- E38 -*mm,, (h) je voudrais savoir si tous les étudiants qui souhaitent passer c/ un/ ou: ces deux dess est-c' qu'i(ls): |est-c' qu'ils peuvent| le: passer tous: e*
M39 -*non tous les étudiants ne sont pas admis en dess*
E39 -*alors selon quels critères se/ sont-ils admis?*
M40 -*les critères d'admission en dess sont: un dossier est ouvert pour chaque étudiant, le dossier comprend, les certificats qu'il aura choisis au niveau de la maîtrise, des résultats aux examens de la maîtrise et enfin deux appréciations explicites d'enseignants de son choix*
E40 -*m: m:*
M41 -*cette réponse vous convient-elle?*
E41 -*(h) m: oui/ m d'accord,, m,, alors si j'ai bien compris si les résultats sont:| ne sont pas très bons on n'est pas admis en dess: (3.14)*

Ce qui n'était que relation de contiguïté dans le lien associatif est ici marque d'un raisonnement guidant le questionnement. Jouent ce rôle, essentiellement: *alors, mais aussi, à ce moment-là, dans ce cas*, (voir aussi les combinaisons: *mais alors, donc alors, mais alors donc*). Cette valeur logique de *alors* est très rare en phase I, fréquente en phase III (autant que la valeur d'enchaînement linéaire).

Ce soulignement du travail d'inférence en III est significatif de l'effort de l'étudiant pour rationaliser sa stratégie d'interrogation face à un interlocuteur peu coopératif: destinée moins à la machine qu'à lui-même, cette verbalisation du raisonnement signale la tension observable par ailleurs dans les échanges Homme-Machine.

- Les Marqueurs de reformulation de question (paraphrase ou spécification) ou de demande de confirmation (souvent associée à une question de type Qd en phase I). Des termes comme: *c'est-à-dire (que), je veux dire (que), disons (que), enfin, par exemple...* apparaissent souvent à l'intérieur des interventions comme Marqueurs d'auto-correction. Ceux que nous avons relevés précèdent une question. La reformulation peut intervenir soit après une demande de la machine:

(25)

- E25 -*oui|- | c'qui signifie que ça fait plus de:,| de débouchés e:| que prévu*
M26 -*qu'entendez-vous par là?*
E26 -*je veux dire que finalement: e:,, y a un éventail de possibilités e:| assez énorme (3.2)*

soit spontanément, comme présentation d'une interprétation du propos antérieur à laquelle on veut associer l'interlocuteur:

(26)

- E50 -*d'accord (h) pour ce genre de poste quelle est | les possibilités d'évolution possible ensuite dans l'entreprise*
M48 -*cela dépend de vos intérêts*
E51 -*| disons qu'a priori e: une formation mettons en ergonomie (h) est-ce que ça n'limite pas à faire de l'ergonomie toute sa vie en fait e*
M49 -*non pas du tout (...)* (3.5)

Ce type de Marqueur (notamment *enfin*) est moins fréquent en III qu'en I.

7.3. On peut remarquer, en conclusion sur cette distribution, que la diminution globale des Marqueurs en phase III s'accompagne de quelques phénomènes spécifiques: si la polyvalence est dans l'ensemble maintenue, on observe cependant certaines tendances à privilégier les rôles de structuration "objective" du questionnement.

La prédominance très forte en phase III des Marqueurs simples (par rapport à la phase I, cinq fois plus de *et* simples, quatre fois plus de *mais* simples en phase III que de *et* et de *mais* combinés) et la diminution très sensible de tous les types de combinaisons et de certains M2 sont à rapprocher de l'augmentation très sensible du nombre d'occurrences de *et*, qui tend à devenir le Marqueur universel en raison même de sa "transparence". Donc, à la raréfaction du marquage des questions il faut ajouter un phénomène de neutralisation, constatable uniformément sur l'ensemble des communications avec la machine.

8. *Éléments de conclusion*

Cette étude des Marqueurs qui accompagnent les questions dans une demande d'information a montré leur rôle spécifique et leurs valeurs particulières dans le déroulement du dialogue entre humains. La comparaison avec la phase de dialogue Homme-Machine a certes révélé la disparition de certains faits, en particulier la diminution des questions avec Marqueurs et la disparition des Marqueurs combinés complexes. Mais elle a aussi révélé le maintien des Marqueurs simples (en particulier *et*, mais aussi *mais*, *donc*, *alors...*), des Marqueurs d'enregistrement et l'augmentation des Marqueurs-balises. Ils peuvent, nous semble-t-il, être considérés comme des Marqueurs fondamentaux de ce type de dialogue finalisé, appartenant à la norme structurelle de ces dialogues, vu la normativisation générale qui se produit dans le dialogue Homme-Machine, tant dans la structure syntaxique des énoncés (formes des questions) que dans le déroulement du dialogue (succession des thèmes).

Lexique des termes employés

- Enregistrement:** énoncé qui ne comporte qu'un mot ou une suite de mots qui signifie la prise en compte de ce qui vient d'être dit (*oui, d'accord, mm*, etc.): cf. ex. (4).
- Marqueur:** terme jouant un rôle dans la spécification du type d'intervention.
- Marqueur simple:** un seul mot (*et, alors, maintenant*, etc.): cf. ex. (1), ex. (3).
- Marqueur combiné:** deux ou plusieurs Marqueurs simples co-présents dans un énoncé. Ils peuvent être conjoints (*et sinon*, etc.), discontinus (*mais... c'est-à-dire...*, etc.), ou à la fois conjoints et discontinus (*et alors... enfin...*, etc.): cf. ex. (2).
- Question:** énoncé comportant une demande d'information. Il peut revêtir quatre formes (Q, Qd, Qm, Cq).
- Q =** question syntaxiquement ou lexicalement marquée (*est-ce que*, inversion du sujet, *je voudrais savoir*, etc.): cf. ex. (1), ex. (12).
- Qd =** question formulée dans un énoncé de type déclaratif: cf. ex. (3).
- Qm =** question qui comporte un mot interrogatif (*quoi, comment*, etc.), mais celui-ci n'est pas placé en tête de l'énoncé: cf. ex. (14), E20.
- Cq =** énoncé interrogatif sans verbe: cf. ex. (6), M26.
- Reformulation:** il convient de distinguer trois types:
- a. spontanée (dans la même intervention): cf. ex. (7);
- b. dans le but de spécifier (dans une autre intervention): cf. ex. (25);
- c. imposée par la machine (pas d'exemple dans l'article).
- Thème:** domaine dans lequel l'étudiant s'informe et qui peut s'étendre sur plusieurs interventions: cf. ex. (20), ex (21).