```
                    (GP11 ((DET? -1 +) (DET? -3 +))
                          (CP11 NIL (N23 NIL "QUI"))
                          (GV1 NIL
                               (GV1 ((P1 1 -)) (V1 NIL "VIVENT"))
                               (GN ((P1 -1 +))
                                   (D12 NIL (DEM NIL "CETTE"))
                                   (N1 NIL ("EXPERIENCE" (NTH* ("ROBO=
TISATION")))))))))))
(C32 NIL ".")
```

8. Liste de noms-têtes de listes (dénominations d'objets)

"ROBOTISATION" "PRESSES" "COUP" "MAIN" "OUVRIER" "CONDUCTEURS" "MACINES" "HOMMES"

9. Ingrédients ("spécifiques") des objets (ou: expansion propre de chaque objet

(Procédure) Exemple 1:
(PROJEC (ROB.OBJ) (CL.OB1)'(NI(((=X(NTH* ("HOMME")))))))
(TYPE(CL.OB1))
"PRODUITS" "APPROVISIONNEMENT" "MANCHE" "PALONNIER" "YEUX" "TOUR" "CHAINES" "MACHINES" "MAITRISE" "AUTOMATES" "OUVRIER" "OUTILS"

Exemple 2:
expansion propre de "ROBOTISATION" :
"QUESTION" "AGENTS" "ENTRETIEN" "EXPERIENCE" "IDEE" "REVOLUTION"

10. Interprétation des résultats

Un nom-tête de liste correspond à la dénomination d'un objet ancré dans le discours. On le qualifie également d'"expression générique" en tant qu'il engendre une expansion. Celle-ci est faite d'"aspects de l'objet" matérialisés par des unités lexicales entretenant avec le nom-tête de liste des rapports d'anaphore ou de coréférence. En ce qui concerne l'exemple 2 ("ROBOTISATION") il sera intéressant de savoir, du point de vue de l'analyse des discours qui est le nôtre, que ROBOTISATION est anaphorisé par: QUESTION, EXPERIENCE, IDEE et REVOLUTION.
Autant de substituts contextuels au mot "ROBOTISATION" qui marquent, dans ce texte particulier les différents aspects sous lesquels est repris ce nom. On saura également qu'il existe un lien d'anaphore associative entre ROBOTISATION et AGENTS ou ENTRETIEN.
En ce qui concerne l'exemple 1, l'expansion de HOMME est beaucoup plus vaste, ce qui est assez attendu. Dans le contexte actualisé par ce discours, les aspects sous lesquels HOMME apparaît sont bien sûr liés à des activités de travail. Ainsi l'expansion de HOMME dans le texte recoupe l'expansion de HOMME dans la réalité du travail, puisqu'on y trouve:
PRODUITS, APPROVISIONNEMENT, MANCHE, PALONNIER, TOUR, CHAINES, MACHINES, AUTOMATES, OUTILS
autrement dit: des types d'activité ("approvisionnement") des résultats

de ces activitéd ("produits") et des moyens de ces activités ("outils", "machines"...).
Les relations ici implicitées sont, bien sûr, des relations d'anaphore associative. L'expansion considérée contient également: YEUX, c'est le seul ingrédient de la classe-objet HOMME qui ne soit pas directement lié au monde du travail et fasse en quelque sorte partie des aspects "naturels" de l'objet. Cette expansion très particulière "explique" alors le seul lien de coréférence apparent dans cette série, qui est le lien entre HOMME et OUVRIER.

"Notre "analyse" automatique rate malheureusement par manque de finesse dans la détermination des modèles, une autre unité coréférentielle à HOMME et qui serait: PILOTES DE LIGNES (présents dans notre expansion seulement par l'intermédiaire de deux de ses ingrédients: MANCHE et PALONNIER).
Au prix d'une légère complexification de notre algorithme, nous pourrions obtenir cependant cette relation de coréférence qui nous permettrait d'en établir une autre; celle vers laquelle le texte converge et qui lui sert de conclusion: l'équivalence entre OUVRIER et PILOTES DE LIGNES, interprétable comme substitution orientée ("aujourd'hui grâce à la ROBOTISATION, l'OUVRIER devient PILOTE DE LIGNE").

Comme on le voit, les résultats obtenus par ces algorithmes programmables en DEREDEC peuvent donner lieu à des interprétations qui rappellent les essais d'analyse automatique du discours tels que AAD 69 de M. Pêcheux effectués dans la lignée des travaux de Z. Harris.

Ils s'en distinguent toutefois par une volonté de mise en relief de la structure argumentative des textes: les items mis en équivalence ne le sont plus à partir d'une identité de distribution dans des moules préétablis, mais à partir de critères de formes qui sélectionnent des syntagmes en fonction de leur rôle (ancrage, expansion, spécification,...).

ANALYSE D'INTERVIEWS ASSISTEE PAR ORDINATEUR*

Catherine PEQUEGNAT

Table des matières

1. Introduction 73
2. Description générale des objets textuels analysés 75
3. La sous-catégorisation des formes verbales 83
4. Extraction des reformulations du "changement" 89

* Ce texte a été rédigé en collaboration avec A. Lecomte. Les procédures DEREDEC qu'il évoque ont été programmées par A. Lecomte, P. Plante et moi-même. Celles-ci ont été réalisées dans l'optique de décrire des discours produits par des salariés en réponse à deux questions ouvertes, soit:

(QI) "Les nouvelles techniques récemment introduites dans votre entreprise ont-elles changé le travail?"

(QII) "Comment imaginez-vous que le développement de ces nouvelles techniques va modifier l'avenir économique des entreprises?"

ANALYSE D'INTERVIEWS ASSISTES PAR ORDINATEURS

IA IE

Catherine Péquegnat

1. INTRODUCTION

1.1 Dans le cadre d'une recherche portant sur les représentations économiques et les raisonnements qu'elles permettent, recherche commune au CNRS (Marseille), au laboratoire IRPEACS (Lyon) et au CdRS (Neuchâtel)[1], une enquête a été effectuée en France et en Suisse auprès de salariés de trois entreprises[2], au cours de l'hiver 1983-1984. Le matériel collecté lors de cette enquête comporte, outre des séries de réponses à des questionnaires de type fermé, des discours oraux, enregistrés sur bande magnétique, produits en réponse à deux questions ouvertes. Chaque enquêté est invité par ces questions à s'exprimer sur le thème des nouvelles technologies en relation avec le (son) travail et l'avenir économique des entreprises.

Ces discours ont été retranscrits de l'oral à l'écrit. Les textes qui résultent de cette opération se présentent comme des séquences dialogales complexes, dans lesquelles l'enquêteur (L_1) participe de façon active à la verbalisation des représentations. L'analyse de ces séquences pose donc de nombreux problèmes méthodologiques que nous ne prétendons pas résoudre dans les limites de ce bref papier. Nous nous contenterons ici de donner un aperçu d'une caractérisation possible des relations entre les "contenus" des questions et les "contenus" des énoncés qui composent les réponses, sans nous préoccuper pour l'instant de l'interprétation de ces relations en termes d'opérations logico-discursives[3] ou d'inférences non formelles[4], ou bien encore en termes de représentations économiques. De ce fait, les noms que nous donnerons à certaines classes d'énoncés du corpus -tels que par exemple développement ou transformation (cf. infra § 2.2)- sont purement descriptifs. Nous dirons quelques mots des hypothè-

ses sous-jacentes à cette caractérisation au paragraphe 2.

1.2 Notre analyse est basée sur l'observation d'un certain nombre de régularités sémantiques et syntaxiques repérables à la surface des textes. Autrement dit, c'est en fonction de principes de reconnaissance de forme(s) que nous induirons l'existence de relations entre contenus pour des fragments de textes. Nous tenterons de suggérer l'automatisation de cette reconnaissance au moyen du logiciel Déredec (Cf. P. Plante, ici même).

Cette reconnaissance suppose deux types de manipulations effectuées sur les énoncés du texte. D'une part, et dans un premier temps, chaque énoncé doit être décrit linguistiquement. Un ensemble de procédures tel que la Grammaire de Surface du Français (GDSF) , réalisée par P. Plante (1981) construit automatiquement une représentation structurelle des énoncés, à partir d'une première ré-écriture du texte au terme de laquelle il se présente comme une suite d'EXFAD composée chacune d'une expression atomique et de sa catégorie grammaticale. Cette représentation comporte des indications sur les relations qu'entretiennent entre elles les diverses parties d'une telle structure. Ainsi par exemple, la GDSF reconnaît un thème et son propos, différencie divers types de développement du propos et établit des relations de détermination au sein des groupes nominaux et des groupes verbaux (entre un nom et un nom, un nom et un adjectif, un nom et un groupe propositionnel, ou encore, entre un verbe et ses déterminants adverbiaux, etc.); enfin, les propositions composant un énoncé sont étiquetées relativement à leur statut au sein de la phrase (distinction entre les relatives, les complétives, les circonstancielles, etc.). Celles de ces relations qui sont utilisées dans notre analyse seront brièvement explicitées au cours de ce texte.

Les représentations structurelles ainsi obtenues pourront alors être fouillées au moyen de fonctions Déredec dites exploratrices de textes. Cette fouille constitue le deuxième type de manipulations effectuées sur les textes. Nous présenterons au paragraphe 4 quelques exemples de fouilles et leurs résultats sur un exemple extrait d'une des réponses analysées.

D'autre part, il nous a semblé intéressant d'enrichir les représentations structurelles des énoncés telles qu'on peut les obtenir

par les procédures GDSF, d'un certain nombre de traits plus spécifiques comme par exemple les marques temporelles portées par les verbes, les types d'événements ou d'actions auxquels réfèrent certains verbes, ainsi que les fonctions casuelles que ces types d'actions confèrent aux groupes nominaux ou propositionnels figurant dans les énoncés où ils "occurrent". Ceci afin de pouvoir faire figurer, dans l'expression des structures permettant la fouille des textes, des indications plus sémantiques représentées par des catégories de procès et des catégories actantielles. On justifiera cet enrichissement au paragraphe 2 en présentant quelques exemples. Les procédures de catégorisations (déictiques et sémantiques) qu'il suppose seront décrites au paragraphe 3.

2. DESCRIPTION GENERALE DES OBJETS TEXTUELS ANALYSES

2.1 On peut caractériser, schématiquement, les séquences que nous analysons comme des suites réglées d'actes de langage du type "demande d'information-réponse", au cours desquelles s'élaborent et se transforment diverses représentations.

De façon plus précise, et du point de vue d'une analyse linguistique des discours produits, nous dirons que dans ces séquences, L_1 engage, sous forme d'interrogation ("est-ce que", "comment"), la reformulation d'énoncés-noyaux tels que "les nouvelles techniques récemment introduites dans votre entreprise ont changé le travail", "vous avez dû vous reformer pour ce changement", "le développement de ces nouvelles techniques va modifier l'avenir économique des entreprises", etc.

Par énoncé-noyau, on entend ici le groupe propositionnel sur lequel porte l'opérateur interrogatif. Lors de la saisie des textes, nous n'avons tenu compte des interventions de L_1 que dans la mesure où la suite du discours montrait qu'elles avaient été interprétées comme des demandes d'information. Chacune de ces interventions a été réécrites sous la forme d'une question, précédée d'un "est-ce que" pour les questions oui/non, de l'opérateur figurant dans l'énoncé pour les questions ouvertes ("pourquoi", "en quoi", etc.). Ces expressions atomiques, lors de la phase de catégorisation syntaxique, ont toutes été considérées comme des "déterminants connecteurs" (D2 dans le système des catégories pises en compte par la GDSF).

Pour simplifier, nous envisagerons ici les reformulations dans leur relation à l'énoncé-noyau figurant dans l'interrogation initiale

uniquement (cf. p. 73, QI et QII).

Les réponses des enquêtés (L_2) se présentent comme des suites ou des enchaînements d'énoncés qui, pour la plupart, reprennent partiellement ces énoncés-noyaux tout en les modifiant, de manière à satisfaire à la demande d'information impliquée par la question.

De fait, presque tous les énoncés d'une réponse sont donc susceptibles d'une double lecture. Nous excluons de cette généralisation les énoncés exprimant des "lois de passage" (cf. infra, § 2.5).

Ainsi, un énoncé e_i, en tant qu'il est "impliqué" par la question, est envisageable comme l'un des éléments de l'ensemble de ses réponses possibles; il est caractérisé alors dans sa seule relation à l'interrogation initiale. D'autre part, en tant qu'il intervient dans une suite formant une réponse globale, un énoncé e_i est une étape dans une dérivation[5]; il doit donc être analysé dans sa (ses) relation(s) aux autres énoncés de la réponse. Nous ne prendrons en considération que la première de ces lectures des énoncés produits par les enquêtés.

Nous entendrons alors par reformulations certains types d'énoncés de L_2 pour lesquels on peut envisager une relation à l'énoncé-noyau exprimable dans les termes d'éléments marquant d'une part une reprise, d'autre part une modification de cet énoncé. Nous tenterons d'établir, pour chacun de ces types, les relations qui les définissent selon les transformations structurelles -syntaxiques ou sémantiques- supportées dans chaque cas par l'énoncé-noyau.

2.2 Parmi les réponses qui composent notre corpus, on remarque donc plusieurs types de reformulations. On distingue par exemple dans les énoncés produits par L_2[6], (la liste qui suit n'étant pas exhaustive!)

- des développements

(1) <les nouvelles techniques récemment introduites dans votre entreprise ont changé le travail>

——>(1') <elles ont changé le travail pour plusieurs personnes que je connais>

- des commentaires

(2) <les nouvelles techniques vont modifier l'avenir économique des entreprises>

——> (2') <je ne pense pas que ça va tellement modifier>

- des transformations[7])

(3) <les nouvelles techniques vont modifier l'avenir économique des entreprises>

⟶ (3') <le patron va faire des bénéfices aussi plus grands>

- des évaluations-appréciations

(4) <les nouvelles techniques (...) ont changé le travail>

⟶ (4') <c'est de la folie de pousser les gens comme ça>

⟶ (4") <c'est moins monotone>

⟶ (4"') <c'est beaucoup plus simple>

ou encore des énoncés dans lesquels figurent des indices signalant une combinaison de plusieurs types de reformulations, comme par exemple dans (5'), qui ajoute à l'aspect transformation l'aspect commentaire:

(5) <les nouvelles techniques vont modifier l'avenir économique des entreprises>

⟶ (5') <à mon avis, il va se créer un nouveau climat de travail>

ou dans (6'), qui combine les aspects développement, transformation et évaluation-appréciation :

(6) <les nouvelles techniques vont modifier l'avenir économique des entreprises>

⟶ (6') <les nouvelles techniques, c'est qu'ils veulent faire bien plus de montres et puis aussi précises qu'avant, avec beaucoup moins de monde>

2.3 Pour chacun des types de reformulations dégagés ci-dessus on peut exprimer la relation qu'il entretient avec l'énoncé-noyau qui l'"implique" sous la forme d'un ensemble de traits syntaxico-sémantiques marquant d'une part qu'il y a eu reprise, d'autre part qu'il y a eu modification. Ainsi, dans le passage (1)→(1'), la reprise est marquée par la paire anaphorisé-anaphorisant ("les nouvelles techniques" - "elles"), la répétition lexicale ("changé" ~ "travail") et la conservation des marques temporelles portées par le verbe. L'aspect développement de la reformulation est donné par l'ajout d'un complément prépositionnel ("pour plusieurs personnes que je connais").

Dans le passage (2)→(2'), la reprise est également marquée par une paire anaphorisé-anaphorisant ("ça"), une répétition lexicale

("modifier"), et la conservation de la marque temporelle portée par le verbe; l'aspect commentaire résulte dans l'enchâssement de l'énoncé reformulé dans un énoncé où figure un verbe délocutif ("penser") ainsi que dans l'insertion d'un déterminant verbal ("tellement").

La transformation effectuée dans (3)→(3'), passage dont le seul élément de reprise est la conservation de la marque temporelle, résulte dans l'introduction d'un verbe dénotant un passage d'un état à un autre ("modifier") par un verbe désignant un processus causatif-résultatif ("faire") et l'un de ses compléments possibles ("des bénéfices aussi plus grands").

Dans (4)→(4')(4")(4"') , le statut appréciation-évaluation de l'énoncé second est donné à notre avis par la mise en équivalence interprétationnelle de l'énoncé-noyau avec une forme pleine (nominale: "jolie", ou adjectivale: "monotone", "simple"). La reprise n'est plus ici signifiée par la récurrence d'une marque temporelle portée par le verbe, mais par un déictique ("c"), qui opère une référenciation de l'énoncé-noyau relativement à l'énoncé reformulant.

Enfin, dans (5)→(5'), l'antéposition d'un complément prépositionnel où figure une forme nominale construite à partir d'un verbe délocutif ("à mon avis") marque l'aspect commentaire. L'aspect transformation, quant à lui, réside à nouveau dans le remplacement du verbe initial par un processus causatif-résultatif ("créer"), et l'un de ses compléments possibles. L'agent, par contre, est élidé. Dans cet exemple encore, la seule marque de reprise est la conservation du temps du verbe.

Le dernier exemple (6)→(6') peut se décomposer en une transformation (introduction d'un agent: "ils", substitution verbale: "faire" et introduction d'un complément impliqué par le verbe substitué); en un développement (ajout d'un complément prépositionnel: "avec beaucoup moins de monde"), et en une évaluation-appréciation (présence dans le second énoncé d'une copule au présent et d'un déictique).

2.4 Des descriptions d'exemples qui précèdent, nous retiendrons les points suivants:

A. 1/ Parmi les éléments formels signalant la reprise partielle de l'énoncé-noyau, la récurrence d'une marque temporelle (le passé composé, le futur proche[7], le futur) d'une part, l'occurrence conjointe d'un

déictique en position thématique et d'une copule au présent d'autre part, nous tiendrons lieu, dans un premier temps, de critères de reconnaissance (ou d'extraction) pour l'ensemble des énoncés reformulants d'un texte ou d'un groupe de textes.

Aussi triviaux soient-ils, ces deux critères nous permettrons alors d'isoler, de façon négative en quelque sorte, un certain nombre d'énoncés plus généraux exprimant des "lois de passage" ou des "raisons" restées implicites[8] dans la plupart des relations énoncés-noyaux-reformulations, soit par exemple:

(7) <...parce que je pense toujours que l'être humain est créatif>

(8) <...pour moi, il y a une nécessité que ça évolue du côté technologique>

(9) <...parce que ça fait déjà longtemps qu'on travaille là-dessus>

Dans (7) figure une copule "être" au présent; mais le thème ("l'être humain") n'est pas un pronom. Dans (8) et (9), on trouve des verbes au présent qui ne sont pas des copules[9] ("évolue", "travaille").

2/ Exprimés sous forme de modèles d'exploration (cf. Plante ici même) qui seront projetés sur les descriptions structurelles des énoncés, ces critères permettront l'extraction automatique de l'ensemble des énoncés qui devront à leur tour être fouillés de manière plus sélective, afin de distinguer les divers types de reformulations en fonction des marques de modifications par développement, commentaire, évaluation-appréciation, etc.

Cette première extraction suppose deux types d'indications figurant dans les représentations structurelles des énoncés, et distinctes des indications de catégorio-relations effectuées par la GDSF:

a) Une catégorisation des formes verbales selon les marques temporelles qu'elles portent. Elle est réalisée par un automate Déredec (AUTOTPS) qui intervient après la catégorisation syntaxique, et avant l'application de la GDSF. Nous décrirons partiellement cet automate au paragraphe 3.

b) Une distinction entre copule, ou prédicats d'état et autres verbes. Cette distinction est effectuée par la projection d'un dictionnaire sur les textes, au moyen de la fonction TRADES (cf. § 3). Cette projection a lieu avant les descriptions de textes par AUTOTPS; celui-ci devra en effet "reconnaître" les formes verbales utilisées à la

voix passive et les distinguer des formes verbales utilisées à la voix active: la distinction entre auxiliaires (ETRE ou AVOIR) et l'identification des verbes de déplacement sont deux conditions préalables à cette reconnaissance. Cette catégorisation sémantique des verbes nous permettra également de simplifier considérablement les modèles d'exploration correspondant à l'expression des structures pertinentes pour l'extraction des différents types de modifications évoquées ci-dessus.

B. 1/ Afin de distinguer ces types de modifications, nous avons donc procédé à une sous-catégorisation sémantique des verbes du corpus; en plus d'une indication temporelle, chaque forme verbale sera ainsi dotée automatiquement d'une catégorie indiquant le type de procès auquel elle réfère.

Nous avons repris la définition que donnent B. Fradin et J.-M. Marandin (1979, 71-72) du sens d'un verbe, lequel consiste en une spécification "des relations sémantiques qu'entretiennent tous les actants (les N des constructions syntaxiques) qu'il est susceptible de prendre dans les phrases où il occurre". Cette spécification permet d'exprimer "la construction la plus vaste dans laquelle un verbe s'insère, ou encore, sa suite maximale"[11].

Ainsi, deux verbes appartiendront-ils à la même catégorie sémantique si leurs suites maximales sont identiques. Soit par exemple les deux verbes figurant dans les énoncés-noyaux dont nous analysons les reformulations: "changer" et "modifier". Ces deux verbes dénotent des passages d'un état à un autre supportés par deux actants, éventuellement sous le contrôle d'un agent. De tels procès présupposent de plus des "localisations" (de départ ou d'arrivée), assimilées d'une part à l'état initial, d'autre part à l'état résultant, localisations qui peuvent par ailleurs être absentes de la surface de l'énoncé.

2/ Nous avons catégorisé de tels verbes TRAN (pour Transformation). Cette catégorie regroupe également certains verbes de déplacement (ex: "déplacer", "reclasser",...) qui se distinguent des "vrais déplacements (DEPL, ex: "passer", "quitter",...) en ce qu'ils admettent la construction passive. Par ailleurs, les verbes catégorisés TRAN se distinguent d'autres verbes qui dénotent également un passage d'un état

à un autre mais n'admettent pas la construction passive (ex: "évoluer", "disparaître"...) et qui seront catégorisés CHGM (pour Changement), etc.

Afin de la rendre la plus générale possible, nous avons défini la suite maximale des verbes catégorisés TRAN de la manière suivante: lorsqu'un tel verbe "occurre" à la voix active, son THEME (élément positif de la relation TP construite pas le GDSF) sera vu comme AGENT, son complément P_1 (premier développement de propos, caractérisé par l'absence de préposition introductive) sera vu comme AGI ou comme OBJET, et ses compléments P2 seront vus, selon le type de préposition introductive, comme des LOCALISATIONS: LOC_1(ORIGINE) et LOC2(BUT).

Ainsi, les modifications par développement seront-elles extraites en fonction de la présence, dans un énoncé, d'un verbe catégorisé TRAN; la répétition lexicale observée pour les exemples (1)→(1') ci-dessus constituait un cas particulier (et idéal) de maintien du type de prédicat caractéristique des développements. On distinguera ensuite différents développements selon la nature des P2 qui figurent dans les énoncés retenus: soit qu'il s'agit des bornes où des limites présupposées par ce type de prédicat, et repérables aux prépositions introduisant le complément ("à", et "de"), soit qu'il s'agit d'ajout comme dans (1'), où la préposition introductive est différente ("pour").

N.B. On peut légitimement supposer que la représentation des bornes ou des limites présupposées par les prédicats de type TRAN figurant dans les deux énoncés-noyaux de notre corpus, notamment en ce qui concerne les notions d'état initial/état résultant, n'est pas liée exclusivement à la répétition d'un prédicat de ce type, et dépasse le cadre strict de l'énoncé. Ainsi par exemple, et pour les développements de la "modification", tous les énoncés du type <il y aura + GN>, ou <ce sera + GN> peuvent être considérés comme des expressions de cet état résultant. Pour les bornes impliquées par le "changement", on peut admettre qu'elles sont non seulement représentées par les énoncés copulatifs ou les prédicats d'état dont le verbe porte une marque d'itération (en général l'imparfait: ex: <la vie d'une entreprise était (...) une grande vie de famille>; <il y en a qui étaient plus détendus avant>, <Longines était majoritaire en actions>), mais, par extension, par tous les énoncés dont le prédicat porte une telle marque, que ce prédicat soit une copule ou non (ex: <avant on travaillait>). Cependant, les énoncés dont le verbe est non copulatif et porte une marque d'itération sont relativement rares dans notre corpus; la combinaison la plus fréquente est ETRE/AVOIR + imparfait. D'autre part, comme on le verra plus loin (§ 4) dans l'analyse d'une réponse, presque toutes les occurrences du verbe "travailler" tendent à représenter des états.

3/ Nous admettrons qu'une reformulation est une transformation dès lors que le type du verbe figurant dans l'énoncé-noyau n'est pas conservé dans le second énoncé. Par exemple, la sous-catégorisation des verbes "faire" et "créer" par CAUS, ou CAUS-RES (pour CAUSATIF-RESULTATIF), permettra d'extraire des énoncés tels que (3'), (5') et (6'). On observera d'autres transformations liées à l'occurrence de verbes (ou des formes nominales résultant d'une nominalisation) appartenant à d'autres types, comme par exemple des actions transitives (ACTR). Exemples:

(10') <on va tellement bien robotiser que...>

(11') <on saisissait les commandes pour les clients>

(12') <ça va supprimer du monde>

ou encore, des verbes dénotant des déplacements (DEPL):

(13') <ça a été un passage d'une machine vraiment manuelle datant de peut-être cinquante ans pour les premiers modèles à la machine CMC sortie l'année dernière pour le dernier modèle>

4/ Les commentaires seront extraits en fonction de la présence dans l'énoncé, d'un délocutif (DELO). Traditionnellement, un verbe est dit délocutif si l'on peut établir qu'il est construit à partir d'une locution (ex: "saluer" "dire": "salut"). Nous avons élargi considérablement cette définition et admis dans la classe des délocutifs des verbes tels que "justifier", "ordonner", "parler", ou encore "connaître", etc[12]). La suite maximale des délocutifs sera brièvement décrite au paragraphe 3. On ne dira rien de plus ici de l'insertion des modalités de re évoquées au paragraphe 2.2, et caractéristiques des commentaires, ni de l'enchâssement signalé à propos de l'exemple (2').

5/ Les évaluations-appréciations seront extraites en fonction des critères suivants: présence dans l'énoncé d'une copule ETRE au présent et d'un THEME où figure un pronom déictique (N22 dans les catégories primitives de la GDSF). Le "corps" ou le "contenu" de ces reformulations (soit, pour les exemples (4'), (4"), (4"') cités au paragraphe 2.2: "Folie", "moins de monotonie", "plus de simplicité") correspondra soit à l'élément de plus haut niveau catégorisé P1 + par la GDSF, soit à l'élément (du plus haut niveau également) catégorisé P1P2 + par la GDSF.

N.B. les éléments positifs d'une relation P1P2 correspondent aux développements d'un propos introduits par une préposition "de", ou "des", etc.

6/ Enfin, un ordre d'applicatior sur les modèles d'exploration d'une part, un jeu sur les variables HAUTEUR (cf. Plante ici même) et POSITIF d'autre part, nous permettrons de repérer certaines des combinaisons de reformulations présentes dans les réponses.

3. LA SOUS-CATEGORISATION DES FORMES VERBALES[13)]

3.1 La première étape consiste donc à définir des sous-catégories verbales et à les indexer aux verbes ou à certains noms grâce à un dictionnaire projeté sur le texte. (fonction TRADES). Ce dictionnaire se présente comme une suite d'EXFAD composée chacune :

1/ du nom d'une catégorie; nous avons retenu 12 catégories, soit ACTR, CHGM, ETRE, AVOIR, DEPL, CAUS, ABEN (pour "action bénéficiaire", ex: "apporter", "abandonner"), DELO, ASPC (pour "aspectuel", être: "commencer", "finir",...) TRAN, CONS (pour "consécutif", ex: "découler", résulter"), DIT

2/ d'une expression atomique: celle-ci est soit une racine: ex: "ECRAS", "ENERV",...; soit une expression atomique complète: ex: "SUIS", "VENIEZ"

3/ d'une valeur pour la variable LEMMA (T ou NIL), qui indique si la sous-catégorisation doit s'effectuer pour chaque occurrence d'une racine ou pour chaque occurrence d'une expression atomique complète (ex: son évaluation à NIL pour l'E.A "PERIT" a pour but de me pas catégoriser CHGM une "PERITONITE")

4/ de l'indication d'une catégorie morpho-syntaxique[14)] (N, N1, V, V21 par exemple...). Il s'agit d'éviter, par exemple, la sous-catégorisation AVOI pour l'E.A : "AVIONS", lorsque celle-ci est déjà catégorisée N1 ![15)]

Exemples d'EXFAD extraites du dictionnaire :

```
:
(ACTR "ADAPT" T)
:
(DELO "COMPAR" T  V)
(DELO "COMPAR" T  N)
:
(ACTR "PRODU" T)
:
(ETRE "SERA" NIL)
:
```

Une fois qu'on a appliqué ce dictionnaire à un texte catégorisé morpho-syntaxiquement, celui-ci se présente comme suit:

```
(N211 NIL "JE)
(V1 NIL (DELO NIL "CROIS"))
(D2 NIL "QUE")
(N22 NIL "CA")
(V1 NIL (ETRE NIL "SERA"))
(D11 NIL "PAS")
(D11 NIL "TRES")
(D13 NIL "BRILLANT")
.
.
(V23 NIL          (ACTR NIL "ADAPTE"))
.
.
(N1 NIL "ATELIERS")
(C22 NIL "DE")
(CN1 NIL (ACTR NIL "PRODUCTION"))
.
.
```

. 3.2 Il s'agit ensuite d'indexer les groupes verbaux d'une catégorie temporelle, et, pour ce faire, de repérer tout d'abord des marques des temps des verbes. Ce repérage se fait en projetant certaines modèles sur le texte.

Exemple

```
(= V1 (("ERAI"))(("IRAI"))(("RRAI"))(("DRAI"))(("URAI"))
      (("IRAS"))(("RRAS"))(("DRAS"))(("URAS"))(("ERAS"))
      (("ERA"))(("RRA"))(("IRA"))(("DRA"))((URA"))
      (("ERONS"))(("IRONS"))(("RRONS"))(("DRONS"))(("URONS"))
      (("IREZ"))(("RREZ"))(("DREZ"))(("EREZ"))(("UREZ"))
      (("ERONT"))((RRONT"))(("IRONT"))(("DRONT"))(("URONT")))
```

détectera, à condition que les variables LEMMA et REVER soient à T, certaines formes du futur. Du fait que plusieurs terminaisons sont communes à plusieurs temps (celles du conditionnel et de l'imparfait par exemple, ou bien celles de l'imparfait et du présent de certains verbes -"faire", "savoir", "aller",...) on doit hiérarchiser l'application des différents modèles et considérer de nombreux cas particuliers. D'où le recours à un automate (AUTOTPS), qui fonctionne (dans les grandes lignes!) de la manière suivante:

Dans son état S2, il distingue les cas:
- verbe infinitif (V21)
- verbe participe passé (V23)
- verbe conjugué (V1)

Pour un verbe conjugué:
Dans son état S3, il dépiste d'abord les cas particuliers au présent, et catégorise PRES ces cas (opération CAT). Puis, pour tous les autres cas, il donne aux variables LEMMA et REVER la valeur T, avant de pouvoir, dans l'état S4: projeter hiérarchiquement les modèles de terminaison correspondant au conditionnel, au futur, à l'imparfait et au présent.

D'autres états serviront à repérer les formes: futur proche (construites avec le verbe "aller", ou certains modaux, comme auxiliaire), et passé composé (participes passés avec "avoir" comme auxiliaire, ou avec "être" dans le cas des verbes sous-catégorisés DEPL).
Les participes passés non catégorisés comme "passé composé" seront dès lors considérés comme des marques de la voix PASSIVE.

Après application d'AUTOTPS, un texte se présente comme suit:

```
.
(N211 NIL "JE")
(PRES NIL (V1 NIL (DELO NIL "PENSE")))
(D2   NIL "QUE")
(N22  NIL "C")
(PRES NIL (V1 NIL (ETRE NIL "EST")))
(T13  NIL "UN")
(D11  NIL "PEU")
(D13  NIL "NORMAL")
.
(N211 NIL "J")
(V1   NIL (AVOI NIL "AI")
(PC NIL (V23 NIL (ETRE NIL "ETE")))
(V23  NIL (TRAN NIL "DEPLACEE"))
.
```

Ce qui résume les informations suivantes:
- le locuteur introduit un délocutif au présent ("pense"), puis, dans le contexte de ce délocutif une occurrence du verbe "être" au présent suivie par une qualification (D13: "normal"). Nous pourrons ensuite (après passage des automates GDSF) récupérer toute cette information par deux modèles d'exploration correspondant aux structures des commentaires et des évaluations-appréciations. Ensuite, le locuteur introduit une TRANSFORMATION ("déplacée") au passé composé et à la voix passive (un V23 n'est pas catégorisé PC dans une EXFAD). Lorsque les groupes propositionnels seront formés, "j'ai été déplacée" pourra être extrait comme un développement. Les catégories actantielles présupposées par l'emploi d'un verbe sous-catégorisé TRAN ne sont pas encore représentées à ce stade de la descrip-