

Paraphrases et relatives

par

Henrik Prebensen

0. Introduction

Il y a 13 ans, j'ai publié, en collaboration avec mon collègue et ami, Ebbe Spang-Hanssen, un opuscule en danois sur la typologie des propositions relatives: *De franske relativsætningers typologi*, Romansk Institut, Københavns Universitet, 1977. Malheureusement, nous n'avons jamais trouvé le temps de rédiger une version plus définitive de ce travail. Des changements radicaux dans les conditions de travail des romanistes danois nous ont conduits l'un et l'autre vers d'autres occupations. Pour ma part, j'ai donné un résumé de nos idées dans un article sur *La proposition relative dite attributive*, *Revue Romane*, XVII, 1, 1982. Cet article a donné lieu à une discussion avec Henri Van den Bussche publiée en 1986 (*Revue Romane*, 21, 1).

Un travail de linguistique computationnelle sur la compréhension des syntagmes nominaux (désormais: *SN*) et des anaphores, m'a conduit à tenter une simulation sur ordinateur de notre théorie. Tout en confirmant la solidité de celle-ci, cette expérience a montré comment on peut soumettre certaines hypothèses linguistiques formalisées ou formalisables à des tests expérimentaux semblables à ceux qu'on rencontre dans les sciences traditionnellement qualifiées d'*expérimentales*. Elle a aussi contribué à modifier quelques détails dans la théorie primitive.

1. La problématique

Les grammairiens ont toujours eu du mal à trouver des critères concernant la typologie des propositions relatives.

Pour la distinction entre propositions relatives *explicatives* (*parenthétiques*) et *déterminatives/restrictives*, il y a des cas canoniques qui ne posent pas de problèmes, par exemple lorsque l'antécédent est un nom propre. Mais dès que le contexte se complique, les critères font défaut.

A l'origine des difficultés se trouvent surtout l'imprécision dans la délimitation du domaine d'application de la distinction et le manque de concepts sémantiques opérationnels.

Faute de critères syntaxiques, on a, par exemple, mis dans la même catégorie que les propositions relatives canoniques, des propositions qui n'en sont pas, telles les *relatives indépendantes*. Ou bien, on a voulu appliquer la distinction à des propositions qui, bien que relatives, ont une syntaxe différente des relatives canoniques.

On a essayé de définir la distinction en utilisant des concepts flous tels que «nécessaire au sens général» ou «essentiel» pour définir le concept de *déterminative*, et l'on s'est trouvé en difficulté devant des exemples comme le suivant:

- (1) Devant Marot *qui tirait son poignard*, Sagon s'enfuit (Jean Plattard, *Clément Marot*, 1938, p.76).

Il en est allé de même lorsqu'on s'est appuyé sur l'idée d'extension pour distinguer l'*explicative* (ne restreint pas l'étendue d'un concept) de la *déterminative*:

- (2) Molière a écrit une tragédie *qui est perdue*.

Certains ont douté de l'importance d'une distinction précise. Pourquoi ne pas se contenter d'un concept flou? Cependant, pour expliquer les deux interprétations possibles dans:

- (3) La hausse des matières premières *qui ont baissé depuis un an* a été déterminée par l'inflation
(= *qui ont toutes baissé/dont certaines ont baissé?*)

il faut posséder des critères sûrs.

2. La typologie

La typologie que nous avons établie dans notre étude écarte d'abord un ensemble de constructions relatives auxquelles la distinction *explicatives/déterminatives* ne s'applique pas. Il s'agit des relatives *clivées*:

- (4) a. c'est Paul *qui est venu* (cf. Paul est venu)
b. c'est à Paul *qu'il faut penser* (cf. il faut penser à Paul)

des relatives *présentatives*:

- (5) a. il y a quelqu'un *qui vous envie* (cf. quelqu'un vous envie)
b. il y a des cas *qu'on craint* (cf. on craint certains cas)

et les relatives *possessives*:

- (6) a. j'ai quelqu'un *qui m'attend* (cf. quelqu'un m'attend)
b. il a les mains *qui tremblent* (cf. ses mains tremblent).

Ces types obéissent à des contraintes sémantiques spécifiques et font ressortir des rapports morpho-syntaxiques entre l'antécédent et le pronom relatif qu'on ne retrouve pas dans les relatives canoniques. Ils comportent *être* et *avoir* dans des acceptions sans contenu sémantique spécifique. Grammaticalement, on peut les décrire comme des transformations de phrases simples. Cette description correspond bien à leur emploi sur le plan pragmatique.

Quant aux relatives pour lesquelles la distinction *explicatif/déterminatif* est pertinente, on peut utiliser le rapport anaphorique entre l'antécédent et le pronom relatif. Pour cela, on a recours à des transformations paraphrastiques («paraphrasages»).

En explicitant la relation anaphorique, on peut constater si l'antécédent constitue une base anaphorique par rapport au pronom relatif ou non, par exemple, si l'ensemble des référents du discours couvert par l'antécédent et celui couvert par le pronom sont co-extensifs, comme dans le rapport normal entre un SN et son terme anaphorique.

Si la relation: antécédent-pronom relatif est un rapport anaphorique ordinaire, la proposition relative est *explicative* (ou *parenthétique*). Sinon, elle est ou *déterminative* ou *restrictive*.

Si l'antécédent est un nom propre ou un SN entièrement déterminé par le contexte ou par la situation, le pronom relatif s'y rapporte comme une anaphore ordinaire et les deux syntagmes sont co-extensifs:

- (7) a. je n'ai pas vu Paul *qui a dû rater son train*
 b. ses fils *qui avaient 3 et 5 ans* nous ont fait des sourires
 c. le roi du pays *qui fume sans arrêt* fera une apparition brève.

Un paraphrasage au moyen d'une proposition incise ou postposée rend explicite le rapport anaphorique:

- (8) a. je n'ai pas vu Paul – *il a dû rater son train*
 b. ses fils – *ils avaient 3 et 5 ans* – nous ont fait des sourires
 c. le roi du pays fera une apparition brève – *il fume sans arrêt*.

On peut généraliser et formaliser (7) et (8) dans le schéma suivant (\emptyset = l'élément vide):

- (9) a. $\emptyset/\text{un}/\text{le N qui } V_i V_j <=> \emptyset/\text{un}/\text{le N } V_j - \text{il } V_i$
 b. $\emptyset/\text{un}/\text{le N qui } V_i V_j <=> \text{le N} - \text{il } V_i - V_j$

et définir le concept de *relative explicative* (*parenthétique*) ainsi:

- (10) Une proposition relative est *explicative* (ou *parenthétique*) si elle est paraphrasable à l'aide d'un des schémas (9)a. ou b.

S'il n'y a pas de rapport anaphorique entre l'antécédent et le pronom, d'où il s'ensuit que les deux ne sont pas co-extensifs, on a affaire, au contraire, soit à une *relative déterminative*, soit à une *relative restrictive*. La déterminative entre

dans la détermination sémantique de l'extension de l'antécédent, qui est toujours défini:

- (11) a. la question *qui se pose maintenant* est la suivante
 b. les taux *qu'on se donne au départ* déterminent le profit

ont comme paraphrases:

- (12) a. une question se pose maintenant – *elle est la suivante*
 b. on se donne au départ certains taux – *ils déterminent le profit*

et sont formalisables par le schéma:

- (13) le N qui $V_i V_j$ \Leftrightarrow un N V_i – il V_j

En comparant (9)a. et (13), on voit que les deux définitions prédisent l'existence d'un rapport entre les relatives explicatives et déterminatives:

- (14) a. le N qui $V_i V_j$ \Leftrightarrow un N V_i – il V_j (13)
 b. un N qui $V_j V_i$ \Leftrightarrow un N V_i – il V_j (9a.)
 c. le N qui $V_i V_j$ \Leftrightarrow un N qui $V_j V_i$ (13+9a.)

Ainsi:

- (15) a. la question *qu'on se pose maintenant* est la suivante
 b. il ne peut pas le faire dans l'état où il se trouve

\Leftrightarrow

- c. on se pose maintenant une question *qui est la suivante*
 d. il se trouve dans un état où il ne peut pas le faire

On peut donc ajouter la définition:

- (16) Une proposition relative est *déterminative* si elle est paraphrasable à l'aide des schémas (13) ou (14c.)

On remarque que la définition (14) marche dans les deux sens:

- (10) bis Une proposition relative dont l'antécédent est indéfini est explicative si elle est paraphrasable à l'aide du schéma (14c)

Une *relative restrictive* restreint le domaine d'extension de l'antécédent:

- (17) Molière a écrit des comédies *qui sont perdues*

équivalent à:

- (18) a. Molière a écrit des comédies – *certaines d'entre elles sont perdues*
 b. Certaines des comédies qu'a écrites Molière sont perdues

En formalisant, on obtient les deux schémas suivants:

- (19) un/des N qui $V_i V_j$ \Leftrightarrow des N V_j – l'un/certains d'entre eux V_i

- (20) un/des N qui $V_i V_j$ \Leftrightarrow un/certains des N qui $V_j V_i$

Il s'ensuit la définition (21):

- (21) Une proposition relative est *restrictive* si elle est paraphrasable à l'aide des schémas (19) ou (20).

Dans tous les cas ci-dessus le SN enchâssant la relative est *spécifique*. Lorsqu'il s'agit de SN *non-spécifiques*, l'application des critères de définition se complique du fait que les SN non-spécifiques sont sous le contrôle d'un opérateur: verbe d'attitude, expression modale, négation, quantificateur, conjonction, etc. Or, il est difficile, généralement, d'extraire un SN non-spécifique ou une partie d'un tel SN du contrôle de l'opérateur, comme l'exigeraient les paraphrases de type (10),(13) et (19):

- (22) a. Si Molière a écrit des comédies *qui sont perdues*, on les retrouvera un jour
 b. Je n'ai pas d'amis *qui puissent m'aider*
 c. Elle veut un mari *qui lui soit inférieur sur le plan intellectuel*
 d. Qu'un pays mène une politique étrangère *qui soit égoïste*, est considéré comme normal.

On peut, dans ces cas, avoir recours aux formes variantes (10bis) (14) et (20) pour montrer que ces relatives ne sont ni restrictives ni déterminatives, mais *explicatives* (le terme traditionnel, est-il bien adéquat?):

- (23) a. Si certaines des comédies *qu'a écrites Molière* sont perdues, on les retrouvera un jour
 b. Les amis *que j'ai* ne peuvent pas m'aider
 c. Le mari *qu'elle veut* doit lui être inférieur sur le plan intellectuel
 d. Que la politique étrangère *que mène un pays* soit égoïste, est considéré comme normal.

Restent les propositions relatives dites *attributives*. Elles obéissent à un ensemble de contraintes spécifiques, mais appartiennent aux relatives *explicatives*, comme le montre la co-extension entre l'antécédent, qui peut être un nom propre ou un pronom personnel, et le relatif:

- (24) a. Je vois Pierre *qui joue dans le jardin*
 b. Je le vois *qui joue dans le jardin*.

Je n'entrerai pas dans plus de détails, mais terminerai par le tableau d'ensemble suivant:

(25)	A: Relatives transformées:	1. clivées 2. présentatives 3. possessives
	B: Relatives non-transformées:	1. explicatives 2. non-explicatives 2.1. déterminatives 2.2. restrictives

3. Le paraphrasage, outil de méthode

La paraphrase a toujours joué un rôle important dans les raisonnements linguistiques. Sous la forme de principe d'heuristique, par exemple: *épreuve de commutation* chère à Hjelmslev. Comme dispositif formel permettant d'identifier et d'effectuer des opérations syntaxiques correctrices: *transformations* chomskyennes.

Avant tout, elle sert d'outil de vérification au grammairien ordinaire. Celui qui veut raisonner sur le rôle du syntagme nominal (sujet ou attribut?) dans:

(26) a. son plus grand désir est de voyager

peut avoir recours à ces paraphrases:

(26) b. son plus grand désir est-il de voyager?

c. quel est son plus grand désir?

d. son plus grand désir qui est de voyager ...

et utiliser dans son raisonnement des schémas qui souvent restent implicites, mais sans lesquels l'argument n'a pas de force démonstrative:

(27) a. SN est de V_{inf} \Leftrightarrow

b. SN est-il de V_{inf} ?

c. quel est SN?

d. SN qui est de V_{inf} ...

Le signe d'équivalence, « \Leftrightarrow », pose un problème. Les paraphrases ne sont pas des identités, ni ici, ni en général. Quel que soit le changement opéré au niveau de la forme, on peut l'investir d'une signification. Même dans la paraphrase entre l'actif et le passif, l'équivalence sémantique n'est pas évidente.

Néanmoins, le pivot du raisonnement grammatical est souvent l'identité postulée entre des paraphrases. Le grammairien est amené à distinguer entre différence accidentelle et identité essentielle dans le maniement des paraphrases.

Le grammairien, souvent, aura l'impression qu'un certain «bon sens», un «sentiment linguistique», des intuitions le guident. Sandfeld, à qui l'on doit le raisonnement (26)-(27), n'insiste pas sur une évidence: les faits (les exemples) sont parlants (*L'Infinitif*, 1943, §§44 ss.)

Malheureusement, lorsque le «bon sens» s'avère réparti avec parcimonie, il faut convaincre. C'est ici qu'intervient la *formalisation*.

Formaliser un raisonnement, une théorie, une méthode, revient à choisir une représentation symbolique de ses concepts fondamentaux, donner des règles syntaxiques pour en dériver des expressions complexes (formules), trouver des règles de transformation ou de calcul permettant de relier les formules entre elles de façon incontestable et finalement spécifier des règles

sémantiques (ou d'interprétation) pour établir le rapport entre les formules et le domaine auquel on les applique.

Le but d'une formalisation est d'introduire la calculabilité et la décidabilité dans la recherche pour voir si une controverse est d'ordre technique (dans l'application de concepts sur lesquels on est en principe d'accord, par exemple), ou d'ordre théorique (sur la validité d'un concept quel que soit son degré de précision, par exemple.)

L'idée directrice des définitions à la base de notre typologie est qu'elles délimitent des catégories décidables, c.-à-d. reposent sur des tests opérationnels. Etant donné:

- (28) Pour maintenir leurs recettes, les constructeurs d'automobiles qui doivent faire face à un renchérissement de leur coût ont, ces derniers mois, relevé leurs tarifs

et l'ensemble des schémas ci-dessus, il faut (i) pouvoir identifier les schémas pertinents, à savoir:

(9) a. $\emptyset/\text{un/le } N \text{ qui } V_i V_j \iff \emptyset/\text{un/le } N V_j - \text{il } V_i$

(13) $\text{le } N \text{ qui } V_i V_j \iff \text{un } N V_i - \text{il } V_j$

(ii) pouvoir identifier les syntagmes de (28) qui correspondent aux «variables» des deux schémas:

(29) le N: = *les constructeurs d'automobiles*

qui := *qui*

V_i := *doivent faire face à un renchérissement de leur coût*

V_j := *ont ... relevé leurs tarifs*

(iii) effectuer la transformation paraphrastique prescrite:

(30) a. les constructeurs d'automobiles ont relevé leurs tarifs – ils doivent faire face à un renchérissement de leur coût

b. certains constructeurs d'automobiles doivent faire face à un renchérissement de leur coût – ils ont relevé leurs tarifs

Le test donne deux solutions différentes: (30)a. – sens d'une relative explicative, (30)b. – sens d'une relative déterminative. Pour interpréter ce résultat, il faut trouver à (30)a. et b. des contextes appropriés, soit le contexte original, soit des contextes inventés.

(30)a. et b. sont tous les deux appropriés dans un contexte, s'il y a un ensemble de référents pour l'expression *les constructeurs d'automobiles*.

(30)a. décrit une situation où l'ensemble (tous les constructeurs) ont relevé leurs prix, prévoyant une augmentation de leurs coûts. Cette interprétation laisse deviner un rapport de cause à effet entre les deux.

(30)b., par contre, décrit une situation où seul un sous-ensemble est menacé par la perspective d'une hausse des coûts, et on affirme que les membres de ce sous-ensemble ont déjà relevé les prix.

Le test est opérationnel en ce sens qu'il précise comment faire pour obtenir une réponse à la question concernant le type de la relative, et que les opérations qu'il prescrit sont mécaniquement applicables. Ce qui n'est pas mécanique et qui ne pourra jamais le devenir dans ce genre de test, c'est l'évaluation du résultat. Dans le cas d'un test linguistique, il faut comprendre la langue objet.

Je reviens à la question de l'équivalence entre paraphrases. Si (30)a. et b. sont tous les deux des paraphrases de (28) et différents l'un de l'autre, ces paraphrases ne sont pas des équivalences. Néanmoins, l'évaluation du test a comme prémisses qu'elles le sont: *les constructeurs...qui* et soit *les constructeurs...ils* soit *certain constructeurs..ils* doivent avoir le même sens pour être interchangeables. Pour expliquer cette contradiction apparente, nous avons recours à la logique et aux techniques de la mesure.

Le fondement de la mesure est la possibilité de décider, avec la précision adéquate, de l'équivalence entre deux objets. Ce jugement est d'ordre empirique et non formalisable. Par contre, *l'étalon*, c.-à-d. l'ensemble d'objets canoniques qu'on propose pour effectuer les mesures, est, lui, un objet formel.

L'étalon réduit les possibilités de celui qui mesure. Pour décider de la longueur d'un objet, on force le mesureur à choisir dans une collection d'objets bien délimités, représentés par les petits traits sur le mètre, l'objet qui lui semble le plus exactement représenter l'objet mesuré sous le rapport pertinent.

Avoir une longueur (un poids ou un volume) dans un système de mesure veut donc dire avoir *la même longueur* (le même poids ou volume) qu'un objet bien déterminé de l'étalon avec le degré de précision requis par la tâche. L'étalon définit un ensemble de classes d'équivalence dans un univers mesuré. Chaque classe a un représentant: l'objet canonique. Ma table mesure 1,74m. Cela signifie que l'objet de l'étalon métrique désigné par l'expression «1,74m» la représente. Cette mesure n'est certes pas exacte au micron près. L'équivalence absolue n'existe pas en dehors des mathématiques. Il suffit que la précision soit adéquate pour mon but.

L'étalon permet de prendre des décisions rationnelles dans des cas litigieux. Notre pratique concernant les poids et mesures le montre.

La typologie proposée ici repose sur un étalon qui comprend les trois schémas canoniques primaires: (9), (13), (19), et les deux schémas secondaires (14) et (20). Elle requiert l'identification d'un objet linguistique avec le représentant d'une classe à l'aide d'un schéma canonique. L'enquê-

teur est forcé de faire un choix dans un ensemble restreint. Il doit par la suite évaluer le résultat en faisant entrer en jeu sa compétence linguistique.

Parfois, la quantité d'informations présentes est insuffisante pour effectuer ce choix. C'est le cas de (28) où le choix entre les deux sens (30)a. et b. n'est possible, en pratique, que si le contexte contient les informations nécessaires sur la situation des constructeurs d'automobiles.

Pour conclure provisoirement,

- le paraphrasage est un outil de méthode qui peut être formalisé à l'aide de schémas canoniques qui servent d'étalon dans une typologie, c.-à-d. une classification par classes d'équivalence
- le but de la formalisation est de prescrire un procédé quasi mécanique pour conduire les opérations de classification, c.-à-d. faire de la question de l'appartenance à une classe un problème décidable
- la formalisation contribue ainsi à une clarification théorique et méthodique en séparant clairement les questions de technique des questions de validité; une classification opérationnelle n'est pas nécessairement valide ni l'inverse
- la question de savoir si la typologie ici proposée est opérationnelle peut être discutée rationnellement; le travail que nous avons publié en 1977 s'efforce de montrer qu'elle l'est; pour réfuter cette preuve, il suffit d'alléguer un matériel empirique où les tests sont effectivement inapplicables
- d'un tout autre caractère est la question de savoir si la typologie est valide; nous avons essayé de la justifier en démontrant qu'elle est conforme à l'intuition et qu'elle est fructueuse: elle permet par exemple de formuler des règles grammaticales claires sur l'emploi du subjonctif, du passé simple, etc., et d'expliquer les phénomènes sémantiques.

Dans la section suivante, je voudrais montrer comment on peut utiliser l'ordinateur pour justifier la validité de la typologie.

4. Typologie des relatives, paraphrases et simulation

Un procédé formalisé, quasi mécanique est déjà un algorithme. Les opérations de paraphrasage décrites ci-dessus peuvent donc être implantées dans un programme et effectuées par un ordinateur.

Puisqu'il est possible de réaliser une évaluation du sens sur ordinateur, il est également possible de simuler une compréhension – très limitée – du sens de différentes phrases et d'autres expressions.

Si une machine peut simuler la compréhension d'expressions d'un certain type, on peut aussi en principe l'utiliser pour confirmer expérimentalement des hypothèses sur les propriétés de ces expressions.

On peut monter une expérimentation selon les lignes suivantes:

(i) Un ordinateur peut effectuer des actions. On peut donc le programmer pour l'amener à exécuter des procès, par exemple activer un robot, réel ou simulé sur l'écran, jouer aux échecs ou à d'autres jeux, dessiner des figures, analyser un texte, consulter une base de données, etc. Appelons *monde* l'ensemble des états et des processus qu'on peut obtenir dans une telle application. Ainsi l'échiquier, les pièces, les différentes situations d'un jeu, les mouvements, les stratégies, etc. constituent *le monde* contrôlé par un programme d'échec.

(ii) On peut ensuite implémenter une interface, basée sur un dictionnaire, une grammaire syntagmatique augmentée et un moteur d'inférence, le tout capable d'engendrer des représentations sémantiques à partir de phrases et d'expressions entrées en langue naturelle et de les interpréter dans le monde d'un programme.

(iii) L'interface n'a pas besoin d'être excessivement grande, puissante et compliquée. Au contraire, élément d'un montage expérimental, on a intérêt à ne pas la surcharger de mécanismes encombrants qui nuiraient à la transparence. Il suffit qu'elle donne à l'utilisateur le sentiment d'avoir un accès linguistique *au monde*, par exemple, de pouvoir poser certaines questions, donner des ordres, etc.

(iv) Dictionnaire, grammaire et moteur d'inférence doivent incorporer dans leur formalisme les hypothèses sur le phénomène de langue à vérifier expérimentalement. Un utilisateur qui tape une phrase au clavier, a une intention et, partant, des idées sur des réponses appropriées. Si la machine réagit en accord avec son attente, il conclura que le message a été «compris» (ou que le programme donne l'illusion de comprendre.) S'il accepte la réaction de la machine comme naturelle, son jugement, qui n'a rien de spécifiquement linguistique, corrobore les hypothèses incorporées.

Ici, il s'agit de réaliser une confirmation expérimentale d'hypothèses sur l'interprétation des propositions relatives explicatives, déterminatives et restrictives.

On simule pour cela le monde classique des objets (blocs, pyramides, etc.) du fameux programme SHRDLU de Winograd, dans lequel opère un robot capable de déplacer les objets et de répondre à des questions sur leurs positions.

Une partie de l'interface linguistique, qui nous intéresse ici, est capable d'interpréter des SN comprenant des propositions relatives comme:

- (31) a. Regarde l'objet blanc *qui se trouve sur le bloc qui vient d'être déplacé et place-le sur un bloc que porte le grand bloc noir!*
 b. Mets un bloc sur la table *qui soutient la pyramide!*
 c. Quel objet *qui supporte un autre objet* se trouve sur la table?

Pour interpréter les SN et les propositions relatives, la grammaire sépare d'abord les SN définis: *le bloc (qui), celui (qui), le (pronom)*, des SN indéfinis: *un bloc (qui), quel objet (qui)*. Le déterminant est considéré comme contenant des informations sur l'unicité du référent.

Si le déterminant est défini, le moteur d'inférence essaiera de l'interpréter en posant des questions à deux bases de données dynamiques qui représentent le monde à son état actuel et le discours précédent. En délimitant les syntagmes minimaux de gauche à droite: *le bloc noir, – le bloc noir sur la table, – le bloc noir sur la table qui supporte une pyramide*, il vérifiera s'il existe un *bloc noir* unique soit dans le monde, soit dans le discours précédent (une base anaphorique.)

Si un SN défini, antécédent d'une proposition relative, a un référent *unique*, il constitue une base anaphorique complète, et le pronom relatif est un terme anaphorique. Le schéma de paraphrase valable est le schéma (9)a. Pour constater cela, le moteur d'inférence transforme la relative en interrogative totale: *est-ce que le bloc noir supporte une pyramide?* et cherche à trouver une réponse à cette question.

Si la réponse est *oui*, il aura validé la paraphrase correspondant à (9)a. Si la réponse est *non*, il aura décelé une anomalie présuppositionnelle. L'interlocuteur a une vue du monde différente de la «réalité».

Si l'antécédent est défini, mais sans avoir un référent *unique*, la relative doit être analysée selon le schéma (13). Le moteur d'inférence transforme la relative en interrogative partielle: *quel bloc noir sur la table supporte une pyramide?*

Si la réponse à cette question donne la référence à un bloc précis: *le petit bloc noir*, par exemple, le moteur d'inférence aura validé la paraphrase correspondant à (13). Si la réponse est *aucun* ou *plusieurs*, il aura de nouveau décelé un désaccord entre les présuppositions du discours (où la relation *supporter une pyramide* prétend définir un être unique) et la réalité.

Si l'antécédent est indéfini: *un bloc qui se trouve sur la table*, la question du référent unique ne se pose pas. Le SN doit être analysé selon l'un des schémas (9)a. ou (19)/(20). Le moteur d'inférence transformera la relative en interrogative totale contenant un indéfini à la place du pronom relatif: *est-ce qu'un bloc se trouve sur la table?*

Si la réponse trouvée par le système est *oui* – accompagnée du nom de l'objet ayant servi à la vérification: *le bloc blanc*, par exemple, la relative est conforme au schéma (9)a., donc explicative.

Si la réponse est *oui, plusieurs*, c'est soit le schéma (9)a., soit le schéma (19)/(20) qui entre en jeu. Si le moteur d'inférence peut achever l'interprétation de la phrase entière en choisissant un objet quelconque parmi les blocs désignés, l'interprétation valide le schéma (9)a.

Si ce choix se limite à un seul bloc déterminé, elle valide le schéma (19)/(20). La raison en est que l'ensemble des objets désignés par la relation *se trouver sur la table* n'est pas inclus dans l'ensemble déterminé par la relation que définit le prédicat de la proposition supérieure (cf. *Molière a écrit une comédie qui est perdue* – dans l'interprétation selon (19)/(20), il y a des comédies écrites par Molière qui ne sont pas perdues et des comédies perdues qui ne sont pas écrites par Molière).

Comme dans les autres cas, une réponse négative révèle un désaccord entre présuppositions et réalités.

Les opérations effectuées en interprétant (31)a.-c. peuvent être représentées ainsi:

(32) a. Regarde l'objet blanc qui se trouve sur le bloc qui vient d'être déplacé et place-le sur un bloc que porte le grand bloc noir!

=>

Quel bloc vient d'être déplacé? schéma (13)

Quel objet blanc se trouve sur lui? (13)

Regarde cet objet!

Est-ce que le grand bloc noir porte un bloc? (9)a.

Place l'objet sur ce bloc!

b. Mets un bloc sur la table qui soutient la pyramide!

=>

Est-ce que la table soutient la pyramide? schéma (9)a.

Mets un bloc sur elle!

c. Quel objet qui supporte un autre objet se trouve sur la table?

=>

Est-ce qu'un objet supporte un autre objet?

Est-ce qu'il se trouve sur la table? schéma (19)

On voit que le moteur d'inférence transforme les relatives en interrogatives et non en principales comme dans les schémas paraphrastiques que nous avons proposés dans le paragraphe 2. La raison en est un changement de perspective locutionnelle. Les schémas originaux étaient conçus dans la perspective du locuteur. Lui connaît déjà les référents du discours qu'il est en train d'encoder. Sa tâche est de trouver les formules qui permettent à son interlocuteur de les identifier. Il peut choisir pour cela la construction relative, mais il pourrait aussi choisir d'autres types d'expressions. En effet, on pourrait imaginer une société dirigée par un fou, ayant décidé pour des raisons insondables de punir de la peine de mort tout emploi de propositions

relatives. Y aurait-il une seule pensée que les habitants de cette société ne pourraient plus exprimer? La réponse est certainement *non*.

Le moteur d'inférence voit un message linguistique dans la perspective du *décodage*. Pour interpréter le discours, identifier les objets auxquels réfère le message et ajuster sa version de l'univers du discours à la version du locuteur, il doit poser des questions et y répondre. Lorsqu'il s'agit de propositions relatives, ces questions prennent la forme des interrogatives ci-dessus.

On peut donc mettre en parallèle les mouvements d'encodage et de décodage:

(33) encodage	message	décodage
(9) Ø/un/le N V _i – il V _j	Ø/un/le N V _i qui V _j	est-ce qu'Ø/un/le N V _j – il V _i
(13) un N V _i – il V _j	le N qui V _i V _j	quel N V _i – il V _j
(19) des N V _i – certains d'entre eux V _j	un N qui V _j V _i	est-ce qu'un N V _j – un d'entre eux V _i

6. Conclusion

Un paraphrasage formalisé à l'aide de schémas canoniques a servi à établir une typologie des relatives.

La formalisation aboutit à un procédé quasimécanique, donc décidable, pour les opérations de classification.

La validité de la typologie peut être justifiée en démontrant qu'elle est conforme à l'intuition et qu'elle est fructueuse. Cependant, on peut aussi profiter de son caractère opérationnel pour monter une expérimentation par simulation sur ordinateur.

L'ordinateur peut contrôler un monde restreint. Une interface peut simuler la compréhension d'une (partie limitée d'une) langue naturelle en engendrant des représentations sémantiques et en les interprétant dans ce monde.

Cette interface peut incorporer dans son formalisme les hypothèses sur lesquelles repose la typologie.

Lorsque la machine réagit en accord avec l'attente d'un utilisateur, ce montage a contribué à la corroboration des hypothèses et a confirmé expérimentalement la validité de la typologie.

Henrik Prebensen
Université de Copenhague

Résumé

Une hypothèse sur la typologie des propositions relatives, conforme à l'intuition linguistique et établie au moyen de paraphrases applicables mécaniquement, est proposée. Sa validité a été expérimentalement corroborée sur ordinateur.