



HAL
open science

L'ACCESSIBILITÉ COGNITIVE DES RÉFÉRENTS, LE CENTRAGE D'ATTENTION, ET LA STRUCTURATION DU DISCOURS : UNE VUE D'ENSEMBLE

Francis Cornish

► **To cite this version:**

Francis Cornish. L'ACCESSIBILITÉ COGNITIVE DES RÉFÉRENTS, LE CENTRAGE D'ATTENTION, ET LA STRUCTURATION DU DISCOURS : UNE VUE D'ENSEMBLE. *Verbum: Analecta Neolatina*, 2000, 22 (1), pp.7-30. hal-00966837

HAL Id: hal-00966837

<https://univ-tlse2.hal.science/hal-00966837>

Submitted on 27 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'ACCESSIBILITÉ COGNITIVE DES RÉFÉRENTS, LE CENTRAGE D'ATTENTION, ET LA STRUCTURATION DU DISCOURS : UNE VUE D'ENSEMBLE

Francis Cornish
CNRS, UMR 5610
et Université de Toulouse-Le Mirail
(*Verbum* XXII, 2000, n° 1, pp. 7-30)

1. INTRODUCTION

Cet article introducteur va surtout porter (comme son titre l'indique) sur la théorie du Centrage d'attention, relativement peu connue à ce jour en France, afin de faciliter la lecture de deux des articles principaux (ceux de M. Walker et de M-P. Péry-Woodley) qui supposent ce cadre théorique. L'autre théorie principale envisagée dans ce numéro, la "Hiérarchie du Donné" de J.K. Gundel, N. Hedberg et R. Zacharski, qui figurera dans les contributions de ces trois mêmes auteurs, d'une part, et de W. de Mulder, de l'autre, y est mieux connue ; elle est d'ailleurs présentée dans le premier de ces deux articles. C'est donc pour ces raisons que je concentre l'essentiel de cette introduction sur la première de ces deux théories. Je commencerai par présenter la théorie du Centrage d'attention en tant que telle (section 2), puis j'envisagerai un certain nombre d'expérimentations destinées à vérifier la réalité psychologique du mécanisme de traitement des unités minimales de discours qu'elle postule (section 3), et terminerai par un bilan critique de ce modèle (section 4).

2. LA THEORIE DU CENTRAGE D'ATTENTION

La théorie du Centrage a pris corps à partir des travaux sur le focus d'attention dans le discours faits par C. Sidner, B. Grosz et leurs associés dans le cadre de l'Intelligence Artificielle à la fin des années 70 et au début des années 80. Les formulations les plus récentes se trouvent dans Grosz et alii (1995), dans Walker et Prince (1996), et surtout dans Walker et alii (eds.) (1998). Des validations expérimentales de ses hypothèses de travail sont présentées (entre autres publications) dans Gordon et alii (1993), Brennan (1995), Gordon et Chan (1995) et Hudson D'Zmura et Tanenhaus (1998). Des études ont été menées dans le cadre du Centrage sur des données provenant de l'anglais, de l'italien, du néerlandais, du yiddiche, de l'hébreu, du japonais, de l'espagnol, de l'allemand et du turc, mais à ma connaissance, aucun travail n'a encore été consacré au français. Pour autant que je sache, la seule présentation en français d'un travail s'apparentant (mais de loin) à cette approche est la critique radicale menée par David (1990) du système d'implémentation de l'interprétation des pronoms en termes du focus d'attention chez Sidner (1983). Mais ce système d'algorithmes, rigide et ne tenant pas pleinement compte de la dynamique du processus d'interprétation des anaphoriques en fonction de leur contexte, est largement abandonné dans le Centrage actuel¹. Voir le chapitre de Walker (ch. 19, Figs. 19.3 et 19.4, pp. 409 -10) dans le volume édité par Walker et alii (1998) pour un algorithme linguistiquement plus pertinent.

Toutefois, Marion Fossard, au laboratoire Jacques Lordat de l'Université de Toulouse-Le Mirail, est actuellement engagée dans une série d'expériences ayant le français comme langue de base. Les résultats jusqu'ici obtenus indiquent une confirmation des prédictions du Centrage quant à cette langue (voir la présentation plus loin).

¹ Pourtant, ceci ne veut nullement dire que la théorie du Centrage actuelle soit exempte de critiques - bien au contraire, comme nous le verrons surtout dans la section 4 du présent article.

La théorie du Centrage vise à modéliser “le centre d’attention des interlocuteurs au sein d’un discours [en termes de] rapport entre l’état attentionnel, de complexité inférentielle et de forme des expressions référentielles employées” au sein d’un segment de discours donné (Walker et alii, éds., 1998 : 1 ; je traduis). En un mot, cette théorie essaie de prédire le degré de cohérence d’un segment de discours (unité minimale de discours), et donc la difficulté relative de son traitement par un usager, en prévoyant les enchaînements/développements les plus probables à partir de référents évoqués au début du segment : ces référents sont classés en fonction de leur niveau de topicalité. Le type des marqueurs référentiels (forme zéro, pronom clitique, SN défini ou démonstratif, nom propre..) employés pour assurer cet enchaînement signalera s’il s’agit d’une continuation, d’un maintien temporaire, ou d’un déplacement de la focalisation établie via l’énoncé initial dans le segment.

D’une part, la structure globale ainsi que locale d’un segment de discours donné affecte le choix de marqueur référentiel employé pour reprendre tel ou tel référent de discours qui y aura été évoqué, et de l’autre, le choix de marqueur influence la nature et la cohérence des relations qui peuvent exister entre les différents segments de discours, ainsi qu’entre énoncés au sein d’un seul segment. Cependant, à part les chapitres dans le volume édité par Walker et alii (1998), par Grosz et Ziv, Passonneau, Roberts, et par Walker elle-même, qui traitent de l’interaction entre les relations de focus global et local, la préoccupation principale du Centrage porte sur les relations de focus d’attention *local* au sein d’un segment de discours donné.

2.1 Le segment de discours

L’installation du topique local (Cr) d’un énoncé donné au sein d’un segment de discours détermine la manière dont cet énoncé va devoir être intégré dans ce segment, dont le but à l’intérieur du discours dans son ensemble (sans doute en tant qu’il est conçu par l’allocutaire) est appelé *but de segment de discours*. Le segment de discours peut être réalisé par un seul énoncé ou bien par un groupe d’énoncés qui, pris ensemble², mettent en œuvre une sous-intention de la part de l’énonciateur dans le discours à l’œuvre. Par exemple, camper la scène dans le récit d’une histoire, raconter un épisode dans un récit, ou établir et justifier une prémisse au sein d’une argumentation (voir à cet égard Grosz et Sidner, 1986). En corrélation avec la Structure Intentionnelle d’un discours donné, nous trouvons un focus d’attention global (qui correspond au but macro-discursif) et des focus locaux (qui sont en relation avec des buts internes au segment de discours).

La structure Intentionnelle d’un discours donné est associée à un focus d’attention global (qui correspond au but macro-discursif) et à des focus locaux (qui sont en relation avec des buts internes au segment de discours). Le segment de discours est déterminé par trois facteurs : le fait qu’il permet au locuteur d’accomplir un **but local de discours homogène**, contribuant, d’une façon ou d’une autre, à un but de discours plus global ; le fait qu’un « **espace de focus** » lui est associé, domaine qui réunit les entités, propriétés et relations sur lesquelles le locuteur veut focaliser l’attention de son allocutaire à ce moment-là du discours ; et le fait qu’il délimite une unité minimale de discours signalée par des marqueurs linguistiques (expressions « signaux » - *cue phrases* - , choix des temps, des modes, de l’aspect, ainsi que de certains types de marqueurs référentiels plutôt que d’autres). Pour bien comprendre la notion d’*espace de focus* et notamment sa pertinence pour le fonctionnement de l’anaphore, examinons l’exemple anglais attesté suivant :

² Tel que l’allocutaire le conçoit.

- (1) « ...*He [Kenny Rogers] grew up with four brothers and three sisters, the son of a labourer and a cleaning lady, in a poor area of Houston, Texas. « My father was an alcoholic, but it wasn't disruptive because he was a wonderful man with a great sense of humour. The worst he did for our family was use money for alcohol rather than food or clothes. But he earned it, and had the right to get something out of life. He didn't drink for the last four years. » His parents were not keen on him being a musician, and the early years were tough. »* (Entretien avec Kenny Rogers, *The Radio Times*, 7-13/08/99, p. 18)
- '...Il [Kenny Rogers] a été élevé avec quatre frères et trois sœurs, il était le fils d'un travailleur manuel et d'une femme de ménage, dans un quartier pauvre de Houston, Texas. « Mon père était alcoolique, mais ça ne nous dérangeait pas car c'était un homme merveilleux qui avait un magnifique sens de l'humour. La pire chose qu'il a faite pour notre famille a été d'utiliser de l'argent pour de l'alcool plutôt que pour de la nourriture ou des vêtements. Mais il l'avait bien gagné, et avait bien le droit d'attendre quelque chose de la vie. » *Ses parents n'étaient pas favorables à ce qu'il devienne musicien, et les premières années étaient dures...*'

Le segment de discours du début de cet extrait a pour énonciateur le journaliste qui a recueilli l'entretien dont il est question : la référence à l'interviewé (*He...*) est à la troisième personne. Cependant, cet espace de focus est interrompu par un extrait du discours direct, où l'énonciateur est dorénavant l'interviewé lui-même (voir l'ouverture des guillemets, ainsi que les auto-références à la première personne, ex. *My, our*). L'interruption est clairement délimitée au plan graphique par les guillemets. Au début de l'interruption, nous pouvons dire que l'espace de focus initial est « empilé » ('pushed down') dans la pile ('stack') des espaces de focus évoqués dans ce discours, et remplacé dans la position privilégiée par l'espace ouvert via l'interruption (le passage en discours direct). Une fois cette interruption terminée (par l'entremise de la fermeture des guillemets), l'espace de focus qui lui correspond est « dépilé » ('popped') par rapport à la pile. Les entités saillantes évoquées (surtout le père de l'interviewé, le topique local de ce segment) ne sont donc plus accessibles pour une reprise au sein de l'espace de focus repris (le segment initial), et les pronoms de 3^{ème} personne qui apparaissent dans la phrase immédiatement suivante (*His parents, him*) ne sont de ce fait pas ambigus, aucune concurrence entre référents potentiels n'ayant lieu³.

2.2 La notion de “centre”

Le terme de *centre* a été introduit afin d'éviter l'ambiguïté potentielle qu'il peut y avoir dans l'emploi du mot *focus* (en anglais, tout au moins) entre les sens “référent de discours psychologiquement prééminent”, et “élément d'information relativement le plus important au sein de la phrase” (dans le cadre de l'opposition *Topique/Focus*, telle qu'elle est formulée par Dik (1997,1 : ch. 13) ou par Lambrecht (1994)⁴. Néanmoins, il existe une relation évidente entre ces deux notions : le ‘focus’⁵ conçu en terme de la structure du message, ou structure informationnelle de la phrase (en tant qu'énoncé potentiel) est assigné aux parties de l'énoncé auxquelles l'énonciateur souhaite attirer l'attention de son allocutaire, donc les *rendre* saillantes de ce fait. En tant que tels, les dénotata en question sont supposés par l'énonciateur ne pas déjà se trouver au devant de la mémoire à court terme de son allocutaire. Par contre, un référent ou dénotatum cognitivement en focus⁶ est supposé déjà revêtir un tel statut. Ces deux statuts sont exprimés par des moyens phonologiquement

³ Voir la contribution à ce numéro de M. Walker pour de plus amples considérations à propos du *segment de discours*, notion introduite par Grosz & Sidner (1986).

⁴ En français, toutefois, il existe la possibilité d'employer le terme *focus* dans la première acception, et de réserver celui de *foyer* pour la seconde.

⁵ Appelé “focus sémantique” par Gundel (1999).

⁶ Appelé “focus psychologique” par Gundel (1999).

opposés : accentuation forte, hauteur élevée et plénitude phonologique dans le premier cas, et absence d'accentuation, hauteur basse et atténuation phonologique dans le second⁷.

Les 'centres' d'un énoncé ont un statut sémantico-pragmatique, et ne sont pas des objets formels : en tant que tels, ils sont dits être *réalisés* (autrement dit, "représentés", ou "évoqués") au sein d'un énoncé donné à l'intérieur d'un discours. La définition qu'en donnent Walker et alii (1998 : 3) est comme suit : "Un énoncé E *réalise* un centre c si c est un élément de la situation décrite par E, ou si c est l'interprétation sémantique de quelque constituant de E." La définition se veut neutre quant au moyen par lequel le centre en question aura été installé dans le discours. Ainsi, un centre donné peut avoir été réalisé explicitement, ou **directement**, auquel cas il sera "l'interprétation sémantique d'un syntagme de l'énoncé" (Walker et alii, 1998 : 4) ; ou bien il peut avoir été réalisé implicitement au moyen d'une inférence à partir de la situation de discours. Grosz et alii (1995) ainsi que la plupart des chapitres dans Walker et alii (éds.) (1998) limitent leur étude à des centres explicites, évoqués textuellement, laissant ainsi de côté des référents de discours en focus qui auront été introduits dans le discours par d'autres moyens⁸ (voir aussi à cet égard l'attitude de Sanford et Garrod, 1981 : 154 envers ce qu'ils appellent le "focus explicite"). Cette lacune est également relevée par Brennan (1995 : 164).

Le Centrage implique donc une relation entre un ensemble d'entités de discours évoquées *via* un énoncé donné - les *centres anticipateurs*, comme on les appelle, ou Ca - et un seul *centre rétroactif*, le Cr, récupéré dans l'énoncé immédiatement suivant, en même temps, potentiellement, que dans des énoncés subséquents au sein du segment de discours à l'œuvre. Le Ca est un ensemble de centres (**anticipateurs**) réalisés ou évoqués dans un énoncé, et dont l'allocutaire pourra anticiper, avec divers degrés de probabilité, qu'ils vont devenir le **centre rétroactif** (Cr, ou topique de discours local) de l'énoncé suivant. Le Cp évoqué dans un énoncé est le "centre préféré", autrement dit, le centre qui est classé premier dans le Ca de cet énoncé. Le Cr est l'objet de discours psychologiquement le plus saillant à la fois pour l'énonciateur et l'allocutaire au moment où l'expression qui le réalise est employée : c'est l'entité topique de discours local.

2.3 Les paramètres qui déterminent la hiérarchisation des membres de l'ensemble des centres anticipateurs (Ca)

Les membres de l'ensemble du Ca sont des référents dont on prédit, à des degrés de probabilité divers, qu'ils vont devenir le topique d'énoncés ultérieurs. Ces référents sont partiellement hiérarchisés en terme de prééminence - autrement dit, en terme de leur probabilité relative de figurer comme Cr de l'énoncé suivant. Les paramètres retenus pour déterminer le niveau de saillance relatif des membres du Ca sont d'ordre formel - cependant, il y a une discussion au sein de la communauté des chercheurs "centragistes" sur la nature exacte de ces paramètres (voir à ce sujet les contributions au recueil édité par Walker et alii, 1998). Certains (Cote, Turan) sont d'avis qu'ils sont d'abord d'ordre sémantique (équivalents aux rôles sémantiques associés à telle ou telle position argumentale à l'intérieur de la structure lexico-sémantique qui correspond à une signification d'un verbe ou autre

⁷ Voir Gundel (1999) pour une discussion de la distinction entre ces deux acceptions du terme *focus*, ainsi que l'évocation d'une troisième : le "focus contrastif".

⁸ Par exemple, par inférence à partir d'un énoncé (ou d'une énonciation), ou de la situation perceptible par les participants à la communication, ou par la perception directe d'une situation. Mais aucun algorithme n'est prévu par la théorie pour établir le rang dans la hiérarchie qu'occuperaient ces référents non-explicites, soit par rapport aux autres référents introduits textuellement par le biais d'un SN, soit par rapport à d'autres référents du même type. Dans ce dernier cas, la seule méthode envisageable pour les classer semblerait être selon l'ordre linéaire de leur évocation dans le discours.

prédicateur) ; d'autres (Turan également, Kameyama, di Eugenio) font valoir la nécessité de faire intervenir la notion de 'point de vue', de sujet de conscience dans l'établissement du Ca d'un énoncé donné. La notion d'empathie avec le point de vue du référent de tel ou tel SN dans une phrase, ainsi que celle de topique grammatical (codé par la particule *wa*), est postulée comme facteur pertinent à cet effet dans le cas du japonais, par Iida (1998). Il ressort d'une lecture attentive de Walker et alii (éds.) (1998) que les paramètres qui permettent d'établir la hiérarchie des centres au sein d'un Ca donné sont spécifiques à la langue décrite, et ne sont pas universels.

Le paramètre "standard" est donc, d'abord, celui de la fonction syntaxique du SN qui les réalise :

- (2) *Echelle des Ca* : Sujet > objet indirect animé > objet direct > objet indirect inanimé > objet oblique

et ensuite, l'ordre linéaire d'introduction, une évocation précédente l'emportant sur une évocation subséquente (mais ce dernier paramètre n'est pas déterminant dans le cas du turc, selon Turan, 1998 et Hoffman, 1998). Cette échelle est justement très proche de l'une de celles proposées par Schnedecker (1997 : 65), à savoir : *critère syntaxique* :...< génitif/attribution < accusatif-objet < datif-objet < sujet ; cependant, cette échelle est "orientée" dans la direction opposée par rapport à celle du Centrage (elle en est l'image "miroir", en quelque sorte).

Il faut noter qu'il y a des liens très évidents entre l'échelle de fonctions syntaxiques dans (2) et la relation Topique-Focus au sein de l'énoncé : voir la notion de "dynamisme communicatif" de l'Ecole de Prague, ainsi que Lambrecht (1994). Le sujet grammatical code en effet très fréquemment le Topique, et les autres FG font partie, à des degrés divers, du Focus (au sens de la Structure informationnelle)⁹. Grosz et alii (1995 : 214) suggèrent comme autres facteurs déterminants la subordination propositionnelle et la sémantique lexicale, ainsi que l'interaction qui existe entre elles. Brennan (1995 : 141) en ajoute encore un autre, celui des fonctions sous-catégorisées du verbe principal de l'énoncé par rapport aux adjoints (les référents d'expressions ayant le premier statut sont classés plus haut dans la hiérarchie par rapport à ceux revêtant le second). A ces paramètres, il faut ajouter, me semble-t-il, le niveau d'enchâssement syntaxique au sein d'un SN : par exemple, le référent d'un SN complément ou modifieur au sein d'un SN "matrice" est classé moins haut que celui de ce dernier (cf. *les saucisses de Toulouse*, où il sera plus difficile de reprendre le référent 'Toulouse' dans la suite au moyen d'un pronom, que celui du SN global, 'les saucisses de Toulouse'). Plus un SN sera enchâssé à l'intérieur d'un autre SN, plus il sera difficile de récupérer son référent par rapport à celui qui correspond à un SN moins enchâssé que lui.

2.4 Les quatre types de transition entre énoncés au sein d'un segment

Le centre préféré (Cp) est le membre classé premier dans l'ensemble de centres réalisés par un énoncé donné au sein d'un segment de discours. En tant que tel, on prédit que cette entité de discours deviendra le Cr de l'énoncé qui le suit immédiatement (dans le cadre du type de transition appelé 'Continuation' entre énoncés - voir ci-dessous).

Le Cr d'un énoncé donné permet au contenu de la proposition qui le contient d'être attaché à l'interprétation du membre du Ca de l'énoncé immédiatement précédent en fonction du choix particulier d'expression indexicale utilisée pour le l'exprimer (forme nulle, pronom

⁹ A mon avis, c'est en tant que l'échelle des fonctions grammaticales reflète la structure informationnelle de la phrase en terme de topicalité et de focalité qu'elle est pertinente pour classer les Ca.

de 3^{ème} personne, SN défini, SN ou pronom démonstratif), conjointement à d'autres aspects pertinents de l'énoncé contenant le Cr. Quatre types de relations de transition entre énoncés adjacents sont retenus (voir aussi Walker et alii, 1998 : 5) (les exemples - traduits en français - sont tous tirés de Grosz et alii, 1995) :

I Continuation du Cr antérieur. Dans la transition du type Continuation, le Cr du second énoncé ((3b) ci-dessous) dans un segment est le membre classé premier du Ca de l'énoncé initial (autrement dit, son Cp), ainsi que de celui de ce second énoncé. La forme choisie pour réaliser le Cr dans ce cas sera zéro ou bien un pronom de 3^{ème} personne inaccentué (clitique en français). Ce Cr est alors considéré comme étant le Cr le plus probable de l'énoncé suivant. Dans (3a) :

- (3) a *Susan a offert un hamster à Betsy.*
 b *Elle lui a rappelé que les hamsters étaient assez sauvages.*

Le Ca contient les trois entités de discours « discrètes » suivantes : 'Susan' > 'Betsy' > 'le hamster que Susan a offert à Betsy'¹⁰, classés dans l'ordre hiérarchique indiqué. Les énoncés initiateurs du discours comme (3a) par définition ne comprennent pas de Cr, car ils créent un contexte mais ne l'élaborent pas (voir aussi Lambrecht, 1994 : 129). Donc (3a) n'en fournit pas. Dans l'énoncé suivant, (3b), le Cr est exprimé par le pronom sujet *elle* qui sélectionne le référent féminin le plus élevé à l'intérieur du Ca de (3a), c'est-à-dire 'Susan'. Cependant, il existe un autre pronom dans (3b), à savoir *lui* (datif) : ce fait ne contredit-il pas l'une des hypothèses majeures du Centrage, qu'il n'y a qu'un Cr par énoncé ?¹¹ Grosz et alii (1995) abordent ce problème potentiel pour la théorie en faisant observer que les pronoms ne codent pas *seulement* les entités topiques de discours locaux - autrement dit, le membre classé premier du Ca de l'énoncé précédent ; et, de plus, que les énoncés ultérieurs révèlent que le référent de ce pronom objet, et non sujet (dans ce cas-ci, 'Betsy') n'est pas le topique de ce segment de discours. Si par exemple, la séquence (3a,b) est suivie de (3c) :

- (3) c *Elle a demandé à Betsy si elle aimait ce cadeau,*

la possibilité d'employer de manière naturelle un nom propre à la place du pronom objet indirect pour reprendre cette entité montre que ce référent ne peut pas être le Cr de l'énoncé.

II Rétenion du Cr antérieur. Comme dans la Continuation du Cr antérieur, le Cr d'un troisième énoncé ((4c) ci-dessous) est toujours le membre du Ca de l'énoncé précédent classé premier (son Cp), mais cette entité ne correspond pas au Cp de ce troisième énoncé. Dans ce cas donc, le Cr du troisième énoncé est simplement "retenu" depuis celui du précédent, et on ne prédit pas qu'il restera le Cr de l'énoncé suivant ; plutôt, on anticipe un déplacement de

¹⁰ Comme le signalent Gordon et alii (1993 : 313, n. 2), il importe d'inclure en tant que membre du Ca l'événement, l'état ou le processus (autrement dit, 'l'éventualité') que cet énoncé désigne. Ici, ce serait 'le fait d'offrir un hamster à Betsy, de la part de Susan'. Cette entité de discours pourrait par la suite être reprise via le pronom *ce/cela/ça* dans un énoncé subséquent potentiel suivant immédiatement (3a), l'énoncé initial de ce texte :

- (3) b' *C'était un beau geste.*

Cf. aussi les résultats expérimentaux obtenus par Maes (1997 : 218-219), pour qui un référent abstrait est *ipso facto* plus difficile d'accès, en termes cognitifs, qu'un référent concret (animé ou inanimé). Le référent de *c'* dans (3b') sera donc classé en dernière position sur l'échelle des Ca de (3a).

¹¹ Gundel (1998) répond par l'affirmative à cette question : pour elle, il peut bien y avoir plus d'un Cr par énoncé. Mais voir également la note 12 à cet égard.

centre (i.e. de topique local), autrement dit, on prévoit que c'est le membre classé premier du Ca du troisième énoncé qui va remplir cette fonction. Exemples :

- (4) a *Susan a offert un hamster à Betsy.*
 b *Elle lui a rappelé que les hamsters étaient assez sauvages.*
 c *Betsy lui a dit qu'elle aimait beaucoup ce cadeau.*

Dans (4c), le nom propre *Betsy* apparaît en position sujet, ce qui est une manière de signaler que le Cr actuel est sur le point de changer, le référent 'Betsy' devenant à présent le membre classé premier du nouveau Ca. Dans la variante (5a-c) de ce fragment de discours présentée par Grosz et alii (1995) :

- (5) a *Susan a offert un hamster à Betsy.*
 b *Elle lui a rappelé que les hamsters étaient assez sauvages.*
 c *#Elle a dit à Susan qu'elle aimait beaucoup ce cadeau.*
 (Le dièse préfixé à (5c) vient de moi).

le fait de pronominaliser le référent 'Betsy', en plaçant ce pronom en position sujet, et en codant 'Susan' sous la forme d'un nom propre en position objet indirect, est très nettement incohérent. A remarquer que dans (5c), ce que j'appelle au chapitre 3 dans Cornish (1999) le 'segment indexical' (la prédication anaphorique) oriente la référence du pronom sujet vers 'Betsy' et non vers 'Susan'.

Comme l'indiquent Grosz et alii (1995), si les référents 'Susan' et 'Betsy' étaient équivalents en termes de topicalité dans ces variantes, des différences du genre de celles qui viennent d'être notées ne se produiraient pas. La conclusion qui s'impose est donc, que 'Susan' est le Cr dans les énoncés (b) de chacune des variantes présentées¹².

III Déplacement du Cr : 1. Déplacement "en douceur". Ici, le déplacement préfiguré dans la configuration "Rétention" a bien lieu. Ainsi, dans (6d) ci-dessous ::

- (6) a *Susan a offert un hamster à Betsy.*
 b *Elle lui a rappelé que les hamsters étaient assez sauvages.*
 c *Betsy lui a dit qu'elle aimait beaucoup ce cadeau.*
 d *Elle a dit que c'était tout à fait ce qu'elle voulait.* ((6d) est de moi)

le Cr de ce quatrième énoncé est différent de celui du deuxième ainsi que du troisième, bien que le Cr du quatrième soit identique au Cp de ce même énoncé.

A strictement parler, (5c) ci-avant constitue ce que Walker et alii (1998 : 5-6) appellent un 'Déplacement en douceur' par rapport au focus d'attention qui caractérise l'énoncé immédiatement précédent, (5b). Cependant, il semblerait intuitivement que la transition entre les énoncés (5a-b), d'une part, et (5c) de l'autre, provoque une rupture dans la continuité topicale bien plus marquée que dans le cas "canonique" du Déplacement en douceur illustré par la transition entre (6c et d) ci-dessus, où le rapport entre les deux derniers énoncés semble au contraire parfaitement cohérent. La différence essentielle entre ces deux cas de figure, c'est que dans (6c), il y a eu 'Rétention' du Cr des énoncés précédents, ce qui

¹² L'existence, même à l'intérieur d'une seule proposition, de topiques multiples, est admise par Lambrecht (1994 : 147-50). Là où ceci se produit, cependant, ces topiques sont reliés les uns aux autres. La finalité des énoncés contenant des topiques multiples est précisément, selon Lambrecht, d'informer l'allocutaire de la nature de cette relation. L'un de ces topiques est primaire, les autres étant secondaires, tertiaires, et ainsi de suite.

prépare l'allocutaire au changement de topique à venir ; alors que dans (5c), ceci n'est pas le cas - le pronom sujet *elle* dans (5c) étant compris initialement comme reprenant le Cr de (5b), codé par un pronom sujet identique.

Le dernier des quatre types de transition entre énoncés au sein d'un segment, à savoir, le 'Déplacement brutal', est spécifiée en IV ci-dessous :

IV Déplacement du Cr : 2. 'Déplacement brutal'. Le Cr d'un énoncé donné n'est pas le membre classé premier de son Ca (autrement dit, le Cp de cet énoncé) ; et de plus, ce Cr n'est pas identique au Cr de l'énoncé immédiatement précédent. On « brûle les étapes » d'une façon plus nette encore que dans la configuration 'Déplacement en douceur', car le nouveau topique local n'est pas le Cp de l'énoncé qui l'évoque. Un exemple d'un déplacement brutal est fourni par di Eugenio (1996 : ex (2), p.2) :

- (7) a *Jean est un type sympa.*
- b *Il a rencontré Marie hier.*
- c.i *#Lucie était avec elle.*
- c.ii *Elle était avec Lucie* (ex. (2c.iii) de di Eugenio, 1996)
 (C'est moi qui attache le dièse à (7c.i) - FC)

Dans (7a), le référent 'Jean' est le seul membre du Ca, le SN *un type sympa* étant un nominal-prédicatif appliqué à 'Jean', et non une expression référentielle introduisant un autre individu dans le discours (comparez : *Jean a rencontré un type sympa*, où ce même SN pourra avoir cette valeur). Cependant, comme (7a) figure à l'initiale de ce segment de discours, il ne comporte pas de Cr. Dans (7b), c'est 'Jean' (en tête du Ca constitué par 'Jean' et 'Marie') qui en est le Cr : ce référent, réalisé dans l'énoncé immédiatement précédent, est réalisé dans celui-ci par le biais d'un pronom sujet, et l'autre référent par un SN. Ce second énoncé représente donc une Continuation par rapport à (7a).

Cependant, dans (7c.i), on trouve en position sujet un nom propre référant à une entité non encore introduite dans les énoncés précédents, comme si elle était le topique de cet énoncé : et le référent exprimé par un pronom ('Marie' - *elle*) se trouve dans une position inférieure sur le Ca établi par (7c.i). Donc, 'Lucie' dans (7c.i) n'était pas le Cr de l'énoncé précédent, et ce référent n'est pas non plus le Cr de (7c.i), tout en étant son Cp - car le fait que 'Marie', qui est évoquée dans l'énoncé immédiatement précédent, soit désignée dans (7c.i) par un pronom indique que c'est ce référent qui représente son Cr.

L'effet d'incohérence produit par cette réalisation serait atténué si (7c.i) pouvait être interprété comme instituant un nouveau segment de discours (voir également Schnedecker, 1997). La version alternative proposée par di Eugenio (1996) sous (7c.ii) constituerait un déplacement en douceur, car le Cr réalisé par la référence pronominale, à présent en position sujet, correspondrait au Cp de cet énoncé, mais pas au Cr de (7b) ; ce référent déjà mentionné dans (7b) est donc, dans une certaine mesure, une information "ancienne" dans ce discours. Dans (7c.ii), le référent nouveau, 'Lucie', est réalisé dans une position 'périphérique' en tant que SN complément d'une préposition, et de ce fait, dans une position inférieure à 'Marie' dans le Ca de cet énoncé.

Schématiquement, ces quatre configurations se présenteraient comme dans le Tableau 1 :

TABLEAU 1 (TABLE 1.1. *Centering transition states, Rule 2* : Walker et alii, 1998, p.6)

| | Cr (E_i) = Cr (E_{i-1}) OU Cr (E_{i-1}) = [?] | Cr (E_i) \neq Cr (E_{i-1}) |
|----------------------------------|--|--------------------------------------|
| Cr (E_i) = Cp (E_i) | Continuation | Déplacement en douceur |
| Cr (E_i) \neq Cp (E_i) | Retention | Déplacement brutal |

L'une des hypothèses fondamentales du Centrage, c'est que le niveau de cohérence manifesté par une séquence textuelle, et donc la facilité avec laquelle elle sera traitée par l'auditeur ou le lecteur, est fonction de la mesure dans laquelle cette séquence est fonction de sa conformité aux contraintes et aux règles suivantes (cf. Walker et alii (eds.), 1998 : 3-4) :

Contraintes

Pour chaque énoncé E_i dans un segment de discours D constitué des énoncés E_1, \dots, E_m :

1. Il y a exactement un centre rétroactif $Cr(E_i, D)$ ¹³.
2. Chaque membre de la liste des centres anticipateurs, $Ca(E_i, D)$, doit être réalisé dans E_i .
3. Le centre $Cr(E_i, D)$ est le membre classé premier du $Ca(E_{i-1}, D)$ qui est réalisé dans E_i ¹⁴.

Le Centrage comprend également deux "règles", énoncées comme suit :

Règle 1 Pour chaque E_i dans un segment de discours D constitué par les énoncés E_1, \dots, E_m : si un membre quelconque du $Ca(E_{i-1}, D)$ est exprimé par le biais d'un pronom dans E_i , alors le $Cr(E_i, D)$ l'est également.

Règle 2 En ce qui concerne les relations entre énoncés, des séquences continuant le Cr de l'énoncé précédent (voir I 'Continuation du Cr ' ci-dessus) sont préférées par rapport à des séquences qui ne font que le retenir (voir II 'Rétention du Cr ' ci-dessus), et des séquences qui retiennent le Cr précédent sont préférées par rapport à des séquences qui le déplacent (voir III 'Déplacement du Cr ' ci-dessus). La transition de type III 'Déplacement en douceur' est préférée par rapport à la transition de type IV 'Déplacement brutal'¹⁵.

Ces règles suggèrent que la continuité topicale est la situation de discours par défaut (non-marquée), la discontinuité topicale étant marquée (voir Conte, 1994, Cornish, 1999). En effet, comme le remarquent Grosz et alii (1995 : 215), des déplacements de topique fréquents aboutissent à l'impression que la cohérence discursive locale est moins grande.

3. LA MISE À L'ÉPREUVE EXPÉRIMENTALE DES HYPOTHÈSES DU CENTRAGE

De nombreuses expériences ont été conduites pour tester les prédictions de la théorie du Centrage. La plupart d'entre elles ont pour base de courts textes narratifs construits qui sont présentés phrase par phrase sur un écran de micro-ordinateur. Les temps de lecture des phrases cibles¹⁶ sont enregistrés et mesurés, et parfois également on demande au sujet de

¹³ En fait, cette contrainte ne porte que sur les énoncés non-initiaux, ou élaborateurs de contexte, et non sur l'énoncé initial, créateur de contexte, où le topique local ou global n'est pas encore installé.

¹⁴ En réalité, cette définition du Cr ne convient que pour la relation de transition entre énoncés la plus favorisée, à savoir, celle de Continuation (voir *infra*).

¹⁵ Plusieurs des contributions à Walker et alii (eds.) (1998) soulignent que ce dernier type de transition est rarissime dans les discours naturels ; cependant, voir l'exemple français attesté présenté sous (11) à la section 4 pour une illustration.

¹⁶ Les phrases dans lesquelles telle ou telle des entités de discours évoquées au cours de ces textes est reprise par le biais de l'emploi soit d'un pronom, soit d'un nom propre répété.

vérifier des affirmations portant sur le texte afin d'assurer qu'il lit bien les textes en s'efforçant de les comprendre. Les temps de lecture de ces affirmations sont également enregistrés et mesurés, et la proportion des réponses exactes à la question "vraie?" / "fausse?" est calculée. L'idée de base, c'est que toute intégration du contenu d'un énoncé dans le contexte créé par l'interprétation de l'énoncé précédent et qui implique un centre qui n'est ni le Cp ni le Cr de ce dernier énoncé, exigera des efforts cognitifs supplémentaires lorsque ce centre est désigné à l'aide d'un pronom inaccentué ou d'un zéro. De même, si l'intégration se fait avec un centre qui est soit le Cr, soit le Cp de l'énoncé précédent (voire même les deux à la fois), une complication sera introduite si ce centre est réalisé par un marqueur *autre* qu'un pronom inaccentué. Dans ces deux cas, on prédit que cet effort cognitif supplémentaire se traduira par un temps de réaction sensiblement plus lent et que, donc, il sera mesurable par ce biais.

Je ne vais pas m'attarder sur ces expériences (pour une présentation critique assez complète, voir Cornish, 1999 : 172-82), mais pour l'essentiel, elles confirment les prédictions du Centrage, tout en apportant des précisions permettant de les étendre et de les affiner. Comme prédit, l'emploi en position sujet dans un énoncé non-initial au sein d'un segment, d'un nom propre répété codant un référent dont le niveau de focus est élevé (c'est le Cr de cet énoncé) donne lieu à des temps de lecture nettement plus grands par rapport à ceux enregistrés lors de l'emploi d'un pronom dans la même position. C'est la fameuse "pénalité du nom répété" (Gordon et alii, 1993 : 322). (8a-d) (Gordon et alii, 1993 : 317-18, exp. 1) en fournissent une illustration :

- (8) a *Bruno était la terreur du voisinage.*
 b *Il/Bruno pourchassa Tommy jusqu'à la maison depuis l'école, un jour.*
 c *Il/Bruno le regarda/regarda Tommy se cacher derrière un arbre
 et commencer à pleurer.*
 d *Il/Bruno cria après lui/Tommy si fort que tous les voisins s'attroupèrent.*

Les résultats de cette expérience montrent qu'il y a un accroissement sensible du temps de traitement avec le nom répété même là où aucun autre référent n'était réalisé sous la forme d'un pronom dans ces énoncés non-initials. Comme prédit, les temps de lecture des énoncés non-initials contenant un nom répété en position sujet sont significativement plus longs que ceux des deux autres conditions ('pronom-nom' et 'pronom-pronom'). Les résultats (respectivement, 8460ms, 7624ms, et 7581ms) indiquent que la réalisation du second référent humain, 'Tommy', par un nom propre ou par un pronom ne donne pas lieu à des écarts aussi marqués. Donc c'est bien le pronom lui-même qui contribue à la cohérence locale de la prédication anaphorique en liaison avec son contexte discursif.

Une façon d'expliquer les temps de lecture plus élevés dans la condition 'nom-nom' serait de poser l'existence, chez l'allocutaire, de l'une ou l'autre des deux démarches suivantes : étant donné que le pronom attendu codant le Cr n'a pas été employé, le sujet mobiliserait des ressources cognitives supplémentaires, soit afin d'instancier un nouveau référent ayant le même nom que celui du référent topical ; soit en interprétant cette répétition comme un signal pour ouvrir un nouveau sous-topique, une unité de discours nouvelle où le référent le plus saillant de l'unité précédente exige d'être réintroduit en même temps que de nouvelles coordonnées de temps, d'espace et de perspective. Cependant, une telle inférence n'est pas justifiée par la suite du texte, car les trois événements décrits par (8b-d) se laissent facilement subsumer comme des instances particulières sous la généralisation établie par l'énonciation de (8a) – et donc, interpréter comme faisant partie d'une seule et même unité de discours.

Lorsque le nom répété continuant le Cr de l'énoncé précédent se trouvait dans un contexte discursif du type *déplacement* et non *continuation* du topique local, la pénalité de nom répété associée au sujet de la phrase critique était fort mince (42 msec.). On ne pourra donc pas le qualifier de "pénalité". Cette situation est illustrée par (9) (extrait de Gordon et alii, 1993 : 333-4, exp. 4) :

(9) Paire de phrases initiale

1. *Georges sauta de derrière un arbre et fit peur à Debbie.*
2. *Il fut surpris par sa réaction hystérique.*

Condition 'Continuation'

3. *Il/George ne pense jamais aux réactions possibles des autres.*
4. *Les farces ne sont pas drôles pour tout le monde.*

Condition 'Déplacement'

- 3'. *Elle/Debbie hurle fortement et s'enfuit.*
- 4'. *Les farces ne sont pas drôles pour tout le monde.*

Paire initiale alternative

- 1'. *Debbie fut la victime d'une autre farce de la part de Georges hier.*
- 2'. *Elle fut terrifiée par son attaque-surprise de derrière un arbre.*

(9.3') et (9.4') représentent la "Condition Déplacement" de cette expérience, où la suite du discours porte non plus sur George, mais sur Debbie. Le petit écart dans les temps de lecture entre le pronom et le nom répété dans cette condition montre bien l'orientation naturelle du marqueur d'accessibilité basse qu'est le nom propre vers un référent non focalisé dans le contexte discursif. Mais il montre également que les pronoms inaccentués en position sujet ne sont pas restreints à reprendre le Cr de l'énoncé précédent : *elle* dans (9.3') paraît tout à fait naturel (à la fois dans la version anglaise et française) et ne semble pas donner lieu à ce qu'on pourrait appeler "une pénalité de pronom", telle que celle qui pourrait être provoquée par l'emploi du même pronom sujet dans (5c) ci-dessus. La paire de phrases initiale alternative 1' et 2' est conçue de telle sorte que la même phrase critique (3') puisse apparaître à la fois dans les conditions 'Continuation' (lorsqu'elle fait suite à 2') et 'Déplacement' (lorsqu'elle fait suite à 2).

La différence essentielle entre les deux contextes discursifs à l'œuvre dans (5) (voir sous la sous-section 2.4.II ci-dessus) et dans (9) est que dans (9), nous sommes en présence de référents humains de genre différent, alors que dans (5), ces référents sont de même genre. Étant donné que le Cr est maintenu dans (5b) ('Susan'), on s'attend à ce que l'emploi d'un pronom convenant à ce référent en position sujet de la phrase immédiatement suivante soit interprété comme reprenant ce référent. Ce n'est qu'au moment où le contenu prédicatif de la proposition anaphorique dans son ensemble aura été saisi que cette décision initiale sera modifiée, modification qui se traduira par un coût en terme de ressources cognitives. Mais ce genre de changement de décision en mi-parcours n'a pas lieu d'être dans le cas de (9.3') par rapport à (9.2), étant donné que le genre marqué sur le pronom *elle* ne lui permet de reprendre qu'un seul des deux référents humains en lice à ce point du discours. Ceci expliquerait donc les différences entre les deux cas de figure. (Voir aussi l'interprétation qu'avancent Hudson D'Zmura & Tanenhaus, 1998 des résultats fournis par leur Expérience 2, qui a trait justement à la valeur différentielle des traits de genre portés par un pronom)¹⁷.

¹⁷ Dans Fossard (1999), cependant, qui utilisait des matériaux français où les référents concernés n'étaient pas ambigus du point de vue du genre, les résultats indiquaient une nette différence en temps de lecture, selon que le pronom référait à l'entité en focus, ou non.

Fossard (1999), dans le compte-rendu d'une expérience très semblable à celles qu'on vient de voir, fait état de résultats très voisins, mais en utilisant des matériaux français. Pour M. Fossard, il y a très nettement un effet de "Pénalité du nom répété" lorsque le Np en question redétermine le Cr au sein du même segment de discours, et un effet de "pénalité du pronom" lorsqu'un pronom clitique reprend un référent plus bas que le Cp sur l'échelle. Cependant, une comparaison des temps de lecture des deux noms propres répétés (il y a toujours deux individus humains, de sexe opposé, dans ses passages) en fonction, soit du Cr, soit d'un centre placé plus bas que le Cp sur le Ca, montre que la différence entre ces deux emplois en temps de lecture moyen pondéré par le nombre de caractères n'est pas significative (40,6ms vs. 41,9ms). Elle en déduit que, à la différence du pronom (clitique), le nom propre n'est pas spécialisé, marqué pour la référence focalisée (soit en direction du référent focalisé, soit en dehors de lui). Voici un exemple d'une des séquences qu'elle a utilisées :

- (10) 1. *Les patrons de l'usine faisaient une grande réunion.*
 2a *Alice_i exigeait le rapport de Patrice pour (PRO_i) donner un avis.*
 2b *Patrice_i exigeait le rapport d'Alice pour (PRO_i) donner un avis.*
 3a *Pendant une heure, elle parla sans arrêt.*
 3b *Pendant une heure, Alice parla sans arrêt.*

Affirmation : *La femme est bavarde. (vrai/faux)*

De même, di Eugenio (1998) et Iida (1998) font valoir que, dans le cas de l'italien et du japonais, respectivement, l'emploi de pronoms zéro¹⁸ en position sujet est cohérent alors que leur référent n'est pas le Cr de l'énoncé immédiatement précédent, à condition que certains aspects de la prédication anaphorique permettent de désambiguïser la référence de ces marqueurs : l'accord marqué sur le participe passé du verbe hôte, ainsi que le placement de clitiques objets (sur le verbe fini ou sur le verbe non-fini de son complément) dans le cas de l'italien ; et la sémantique lexicale du verbe hôte, le choix de valeurs pour le temps et l'aspect grammatical, ainsi que certains phénomènes d'accord, dans celui du japonais. Cependant, dans di Eugenio (1996 : 4), l'auteur indique que dans son corpus écrit de 12,000 mots, 70% des sujets zéro se trouvaient dans la configuration 'Continuation'; et elle concède (p. 5) également que, là où la forme sujet zéro est employée pour les configurations 'Déplacement' ou 'Etablissement d'un centre', le discours qui en résulte est parfois "légèrement moins cohérent".

4. LA THEORIE DU CENTRAGE : UN BILAN CRITIQUE

Le Centrage met en évidence les interconnexions étroites qu'il y a entre la "centration" (ou mise en focus des référents évoqués dans un segment de discours) et la structure discursive de ce segment. L'allocutaire d'un énoncé initial dans un segment, connaissant (tacitement) les principes du Centrage et les principes sous-tendant la communication de façon plus générale, va "projeter" certaines structures hiérarchiques de discours, anticipant ainsi le cadre intégrateur qui permettra d'ancrer le contenu des énoncés ultérieurs. La notion d'un "énoncé topical", initiateur de segment, évoquant un ensemble de centres anticipateurs partiellement ordonné, fait évidemment partie de ces procédés de traitement "descendants" ('top down').

¹⁸ Formes qui sont équivalentes dans ces deux langues, du point de vue de leur fonctionnalité discursive, aux pronoms clitiques (inaccentués) du français.

L'usager procède en même temps à un traitement "rétroactif" ou de contrôle, au fur et à mesure que les énoncés suivants arrivent dans la mémoire de travail. Ce type de traitement vise à confirmer ou à infirmer les prédictions du traitement projectif initial (top down), en essayant d'y intégrer le contenu prédicationnel d'un énoncé anaphorique - de façon à l'étendre (de préférence, selon la règle 2 du Centrage concernant les transitions entre énoncés). La forme et le contenu de cet énoncé anaphorique (y compris les instructions fournies par les connecteurs de discours ainsi que par la forme de l'expression anaphorique choisie) vont lui donner l'instruction de relier son contenu propositionnel avec le Cr de l'énoncé précédent (Continuation) ou d'instituer un déplacement en fonction d'un centre classé plus bas sur le Ca de l'énoncé précédent ou des énoncés précédents.

Les attentes souvent fortes concernant la structure du discours à venir ainsi que le développement du topique local que les énoncés donnés peuvent créer, font qu'au niveau de l'emploi réel, la résolution des anaphoriques se fera pour la plupart relativement rapidement, inconsciemment et sans effort particulier. Ainsi, le fait même que des anaphoriques de type, voire de forme, identiques, peuvent co-apparaître dans un énoncé donné n'a pas nécessairement des conséquences négatives quant à l'intercompréhension¹⁹.

En général, les prédictions du Centrage s'avèrent correctes lorsqu'on examine de courts textes (généralement écrits, normés) fabriqués par les chercheurs "centragistes" ou par les psycholinguistes. Mais il existe des situations qui ne sont pas prédites par ces hypothèses (nous allons en voir deux dans un instant, soulevées, entre autres, par Kehler, 1997 et par moi-même). Il faut prendre conscience des limitations de la portée du Centrage, de même que de sa puissance prédictive. Celui-ci se préoccupe (ou s'est préoccupé jusqu'à présent) exclusivement de segments de discours locaux, de courtes séquences d'énoncés introduites par un énoncé topical (ou "thétique"). Les référents de discours (ou centres potentiels) sont en effet limités à ceux qui sont explicitement introduits par des SN et qui désignent des entités de premier ordre (personnes ou objets matériels). Aucune prévision n'est faite pour des référents introduits dans le discours via une inférence, via un signal non langagier, ou par la perception d'un objet, d'une situation ou d'un événement. D'autre part, les connections entre énoncés au sein d'un segment de discours sont établies linéairement, chaque énoncé se reliant à celui qui le précède immédiatement (alors que, parfois, un énoncé peut s'orienter vers un énoncé plus antérieur dans le segment, voire même à un énoncé dans un segment différent)²⁰. De plus, la théorie ne dit pas grand-chose de la variété de relations rhétoriques en fonction desquelles les énoncés membres d'un segment peuvent être intégrés.

À présent, appliquons la théorie à un fragment de texte attesté relativement complexe, qui contient un marqueur indexical (*elle*, sujet de la dernière phrase), qui pose à certains lecteurs de légères difficultés de résolution.

- (11) "...¹*il est terrible cet appel au jugement de Dieu sur le peuple élu, présenté comme infidèle !* ²*On croit réentendre le cri - "Que son sang retombe sur nous et sur nos enfants !" - poussé, vingt ans plus tôt, par la plèbe de Jérusalem au pied d'une croix romaine.* ³*Pilate lui avait donné à choisir entre le délinquant Barrabas et ce Jésus le Nazaréen qui voulait libérer les juifs de l'occupant*

¹⁹ La première d'une série de cinq expériences en laboratoire que je suis en train de préparer avec l'aide d'une équipe de psycholinguistes britanniques dirigée par Alan Garnham, du Laboratory of Experimental Psychology de l'Université de Sussex en Grande Bretagne, met à l'épreuve cette hypothèse (voir aussi les deux premières expériences rapportées par Hudson-D'Zmura et Tanenhaus, 1998, que je discute dans la sous-section 6.4.1 de Cornish, 1999).

²⁰ Voir M. Walker ici même pour plusieurs exemples attestés de ce cas de figure.

romain. ⁴Elle l'avait laissé mourir piteusement sur une potence de bois." (*Le Monde*, 23.7.99, p. 10)

Le texte est divisé en propositions au moyen d'une numérotation pour chacune d'entre elles. A partir de la proposition 2, la structure de chaque proposition consécutive serait la suivante, selon le Centrage :

(11) a *Analyse de (11) en termes du Centrage*

2. Ca : 'le cri poussé vingt ans plus tôt par la plèbe de Jérusalem au pied d'une croix romaine' > 'la plèbe de Jérusalem [à l'époque]' > 'Jérusalem' > 'le pied d'une croix romaine' > 'une croix romaine' ; [Cr = 'le cri...romaine'] ; *Type de Transition* : Continuation.

3. Ca : 'Pilate' > 'la plèbe de Jérusalem...' > 'le délinquant Barrabas' > 'Jésus le Nazaréen' > 'les juifs' > 'l'occupant romain' ; [Cr = 'la plèbe de Jérusalem'] ; *Type de Transition* : Déplacement brutal.

4. Ca : 'la plèbe de Jérusalem' > 'Jésus le Nazaréen' > 'une potence de bois' ; [Cr = 'la plèbe de Jérusalem'] ; *Type de Transition* : Continuation.

Le SN défini complexe *le cri poussé, ...romaine* dans la proposition 2, bien que n'étant pas un pronom ou une forme zéro²¹, est considéré comme codant le Cr de cette proposition en vertu de la relation de coréférence qu'il entretient avec le SN démonstratif topical *cet appel au jugement de Dieu sur le peuple élu, présenté comme infidèle* de la proposition 1. C'est donc en fonction de cette relation, ainsi que du fait que ce référent est le Cp de la proposition 2, que la relation de Continuation est postulée entre ces deux propositions. Dans la proposition 3, nous avons une configuration qui correspond à la relation de Transition 'Déplacement brutal' par rapport à la proposition 2 : en effet, le Cr est le référent 'la plèbe de Jérusalem', car ce référent est codé au moyen d'un pronom (*lui*, datif), et aucun autre membre du Ca de 3 ne l'est également. De plus, ce Cr n'est pas le Cp de cette proposition (ce statut dévoluant à 'Pilate'), et ne correspond pas non plus au Cr de E_{i-1} , en l'occurrence 'le cri...romain'. Dans la proposition 4, la configuration pertinente correspond exactement au type de transition 'Continuation', 'la plèbe de Jérusalem' étant à la fois le Cr de cette proposition (codé par un pronom), son Cp (la position sujet lui conférant le placement en tête du Ca de cet énoncé), et reprenant le Cr de la proposition précédente (p. 3). Cependant, le caractère légèrement marqué (pour certains lecteurs de ce passage, dont moi-même, du moins) de l'emploi du pronom *elle* dans la proposition 4 pour reprendre 'la plèbe de Jérusalem', pourrait être expliqué par le facteur suivant : le Cr de E_{i-1} (la proposition 3) que ce pronom reprend a lui-même été établi via un Déplacement brutal par rapport à la proposition 2. Il serait de ce fait moins bien 'ancré' comme topique local que s'il avait fait l'objet d'une Continuation par rapport à cette proposition. Cependant, une telle relation 'non-adjacente' entre énoncés au sein d'un segment de discours n'est pas prévue par le Centrage, les Cr n'étant créés ou 'hérités' (sous le type de transition 'Continuation') que depuis l'énoncé immédiatement précédent.

Une autre critique des principes fondamentaux du Centrage mérite d'être évoquée à ce stade de la présentation. Elle est émise par Kehler (1997), et fait également intervenir la pertinence des relations entre énoncés non-adjacents dans un segment. Elle porte sur la difficulté qu'aurait la théorie à déterminer l'interprétation de pronoms en utilisant les règles 1 et 2 présentées dans la sous-section 2.4, et surtout le fait que, pour être véritablement

²¹ Voir Cornish (1998, section 3) pour une discussion des facteurs (discursifs ainsi que syntaxiques) qui favorisent le codage d'un topique local ou global d'une manière autre que pronominale.

opératoire, la théorie se voit obligée (bien qu'elle s'en défende) de tenir compte de *l'ensemble* de l'énoncé contenant un ou des marqueurs indexicaux. Il faut se rappeler à ce titre que le Centrage se fait fort de se servir de la structure du discours pour faire des prédictions précises sur l'entité qui sera le Cr de l'énoncé subséquent au sein d'un segment de discours, et cela en fonction des différents types de transition qui permettront de passer d'un énoncé au suivant. C'est donc une approche qui privilégie la simulation d'un type de traitement "descendant" ('top-down') en fonction de l'état du modèle du discours au moment où le marqueur référentiel est rencontré, et non "ascendant" ('bottom-up').

Un court segment de texte inventé qui illustre bien ces difficultés est (12) ((7) chez Kehler, 1997)²² :

- (12) a *The three candidates had a debate today.*
 b *Bob Dole began by bashing Bill Clinton.*
 c *He criticised him on his opposition to tobacco.*
 d *Then Ross Perot reminded him that most Americans are also anti-tobacco.*
 d' *Then Ross Perot slammed him on his tax policies.*
 a. Les trois candidats ont eu un débat aujourd'hui.
 b. Bob Dole a commencé par attaquer Bill Clinton.
 c. Il l'a critiqué à propos de son opposition au tabac.
 d. Alors Ross Perot *lui* a rappelé que la majorité des Américains sont également hostiles au tabac.
 d.' Puis Ross Perot *l'*a critiqué sévèrement à propos de sa politique fiscale.'

Comme Kehler, je trouve que le référent préféré du pronom *him* dans (12d) est 'Bob Dole', et non 'Bill Clinton', alors que celui du même pronom dans (12d') est 'Bill Clinton', et non 'Bob Dole'. La situation serait identique dans la version française (voir ci-dessus). Cependant, chaque version de ce passage ((12a-d) et (12a-d')) partage en commun le contexte (12a-c), autrement dit, le Cp (E_{10c}) et le Cr (E_{10c}) sont les mêmes pour chaque suite. De même, chaque suite contient un même nouveau sujet (*Ross Perot*, dont le référent sera le nouveau Cp) et un même pronom objet (dont le référent sera le nouveau Cr). Ainsi, puisque les référents Cr et Cp de l'énoncé précédent ((12c)) sont les mêmes, une approche telle que le Centrage ne saurait distinguer entre ces cas de figure : elle prédirait que le référent préféré du pronom *him* sera 'Bob Dole' dans chaque cas. Kehler suggère que ce sont les relations rhétoriques (voir Asher, 1993) de **narration** et de **parallélisme** qui, respectivement, permettent de donner lieu à ces interprétations divergeantes.

En effet, puisque (12b, c et d) seront perçus comme décrivant des actes (de parole) successifs dans un débat, le troisième acte (le rappel de Ross Perot) sera compris comme s'enchaînant sur l'avant-dernier, à savoir la critique vis-à-vis de Clinton émise par Bob Dole dans (12b-c). Et puisque, alternativement, (12b,c), d'une part, et (12d'), de l'autre, peuvent également être compris comme des critiques parallèles émises par des opposants à Clinton, la variante (12d') peut de ce fait être interprétée comme s'enchaînant, non sur (12c), l'avant-dernier énoncé dans une suite narrative, mais sur (12b+c), comme décrivant l'attaque de Ross Perot envers Clinton, et comme le pendant équivalent à l'attaque précédente de Bob Dole²³.

²² Il s'agit d'un bref résumé d'un débat télévisé entre les trois principaux candidats à la dernière élection présidentielle américaine (1996). Bob Dole était le candidat du parti Républicain (conservateur), et Ross Perot était un candidat indépendant, également conservateur.

²³ J'ai explicité ici la nature et la source de ces deux types d'interprétations, chose que Kehler ne fait pas.

Mais il est clair que l'un des déterminants de ces deux interprétations discursives différentes est le contenu prédicatif des prédictions anaphoriques respectives (voir Cornish, 1999 : ch. 3 pour une étude détaillée de celles-ci, ainsi que leur fonction dans la résolution des anaphores) : en effet, la complétive enchâssée dans (12d) contient un connecteur (*also* 'également') spécifiant une relation parallèle vis-à-vis de son contexte gauche ; il faut bien voir également qu'elle évoque l'idée d'une prise de position envers le tabac (...*most Americans are anti-tobacco*), ce qui fait que l'on s'attend à ce que l'énoncé dans son ensemble enchaîne sur l'énoncé immédiatement précédent, dans lequel il est question de l'attitude de Clinton envers le tabac. Dans (12d') par contre, la prédication anaphorique concerne le fait que "Ross Perot critique fortement x sur la politique fiscale de x". Il ne s'agit donc plus de l'attitude d'un candidat à l'élection présidentielle vis-à-vis du tabac, mais d'un tout nouveau thème électoral.

Kehler suggère (1997 : n. 8, p. 474) que le fait de faire appel à des facteurs de plausibilité sémantique (ou pragmatique) ne serait pas pertinent, "puisque'il est au moins aussi plausible que Perot critique sévèrement (*slam*) Dole pour sa politique fiscale qu'il ne le ferait vis-à-vis de Clinton." Cette remarque néglige toutefois un aspect important de la structure du discours de ce fragment, et une dimension de notre connaissance du monde concernant la conjoncture politique américaine qui est à l'arrière-plan de ce fragment : à savoir que les Républicains ainsi que Ross Perot étaient violemment opposés à la politique fiscale démocrate, trop "étatique" à leurs yeux. Sachant donc cela, le lecteur de (12d') serait sans doute peu enclin à penser que Ross Perot critiquerait violemment la politique fiscale prônée par le parti républicain, représenté par Dole, mais serait beaucoup plus porté à interpréter une telle critique comme visant le parti démocrate, et donc Bill Clinton.

L'autre facteur, tenant à la structure de discours de ce fragment, vient d'être évoqué : c'est le fait que le contenu prédicatif de la prédication anaphorique de (12d') (indiquant qu'il s'agit à ce point du débat de la "politique fiscale des candidats en lice"), conjointement au connecteur argumentatif ou narratif *Then*²⁴, aiguillent le lecteur vers un enchaînement avec le sous-segment qui correspond à l'interprétation de (12b+c), comme nous l'avons vu ci-dessus. Le contenu sémantique des deux suites (12d)/(12d') a donc un rôle éminemment pertinent à jouer dans la détermination de la résolution des pronoms dans chacune des suites alternatives du segment de discours (12a-c). L'analyse de cet exemple permet également de prendre conscience de l'importance de la structure du discours dans lequel un énoncé anaphorique survient, dans la détermination de la référence de son ou de ses marqueur(s) indexical ou indexicaux.

5. CONCLUSION

Pour être vraiment opératoire, le Centrage a donc besoin de prendre en compte la structure du discours à l'intérieur comme à l'extérieur des segments de discours, pour le traitement desquels elle était initialement conçue. Cette structure est tributaire non seulement de la distribution particulière des expressions référentielles qu'elle comporte, mais aussi de l'organisation des intentions du locuteur telle qu'elle se manifeste par le jeu des connecteurs, et des orientations ponctuelles qui peuvent être signalées par le contenu prédicatif des

²⁴ Observez la différence de traduction des deux *Then* de (12d) et de (12d') : *Alors* dans le premier cas, et *Puis* dans le second. Comme l'argumente Reyle (1998), le connecteur *alors* suppose une dépendance sémantico-pragmatique vis-à-vis de son contexte gauche, alors que *puis* établit une séparation entre le contenu des prédictions qu'il permet de relier. Comme me le signale Marc Plénat (comm. pers.), cette distinction conforte ici l'analyse de Kehler, transposée aux faits du français.

énoncés successifs dans un segment de discours²⁵. Les principes de segmentation du discours sont utilisés pour dégager les unités discursives minimales (les segments de discours), mais l'on se préoccupe trop peu au sein du Centrage de la structuration discursive interne de ces segments. Les deux facteurs qui viennent d'être mentionnés (le rôle des connecteurs et du contenu prédicatif des énoncés constituants) rendent accessibles l'existence de relations rhétoriques d'un type particulier (telles que celles de narration et de parallélisme que nous venons de voir à l'œuvre dans l'interprétation de (12)), relations dont l'importance semble être minimisée par le Centrage. Une prise en compte de la structure « centrationnelle » d'énoncés en amont du dernier énoncé (E_{i-1}) pour l'établissement du statut exact de la relation d'un énoncé vis-à-vis de son contexte discursif gauche est plus que nécessaire (voir les exemples (11) et (12) ainsi que leurs analyses). La structure discursive d'un segment de discours n'est pas constituée par les seules relations qui peuvent exister entre énoncés adjacents au sein de ces unités.

Une autre lacune qui doit rapidement être résolue par le Centrage est la suivante : principalement (et ce pour mieux prévenir le reproche de la circularité) celle qui concerne l'établissement des Ca. Comme nous l'avons brièvement vu dans la sous-section 2.3, d'autres critères que celui des seules fonctions grammaticales (voir l'échelle présentée en (2)) devraient être proposés – bien que, dans la pratique, nous ayons vu qu'elles permettaient de prédire des échelles de topicalité intuitivement satisfaisantes dans le cas du fragment de texte attesté (11) analysé plus haut. Une étude plus fine des phénomènes de topicalité et de focalité au sein de la prédication dans le cadre de l'approche que propose Dik (1997, vol. 1, ch. 13), par exemple, permettrait certainement d'y parvenir.

Un autre problème auquel on doit prêter attention concerne la question de la détermination de la nature de la transition 'Déplacement en douceur'. Nous avons vu en comparant les suites (5b-c), qui n'était pas cohérente, et (6b-c-d), qui l'était, que la seule configuration du Cr de E_{i-1} et du Cp et du Cr de E_i n'était pas à même de prédire une telle différence. On devrait donc reconnaître le rôle théoriquement crucial de la présence ou non de la transition 'Rétention', qui peut précéder ou non un 'Déplacement en douceur' ; car là où elle le précède (comme dans (6c) par rapport à (6d)), le Déplacement du topique local est effectivement « doux » (intuitivement parlant), alors que là où elle est absente (comme dans (5c) par rapport à (5b)), il ne l'est guère. Il s'agit là de l'existence (ou de l'absence) d'une relation interpersonnelle particulière entre interlocuteurs, en l'occurrence de la « prévenance » de l'énonciateur vis-à-vis de son allocutaire.

Enfin, comme nous l'avons remarqué dans la sous-section 2.2, il est également nécessaire que le Centrage puisse tenir compte des centres non-réalisés explicitement, qu'il s'agisse de l'introduction ou de la récupération des entités concernées au sein du discours²⁶. La note 8 ci-dessus suggère qu'on utilise l'ordre d'évocation dans le discours pour classer le premier type d'entités ; mais on peut également utiliser à cet effet le classement des référents concernés en fonction de l'ordre d'entité dont il est question (ordre 1 > ordre 2 > ordre 3 – voir la note 10).

La prise en compte de tous ces facteurs permettrait au Centrage de constituer une véritable théorie prédictive de la structure du discours, et de la cohérence relative des segments de discours, en fonction de l'emploi à tel ou tel endroit stratégique dans l'enchaînement textuel de marqueurs indexicaux de types particuliers.

RÉFÉRENCES BIBLIOGRAPHIQUES

²⁵ Cf. la "structure intentionnelle" chez Grosz & Sidner (1986).

²⁶ Voir à cet égard la tentative de Cote (1998) pour classer les centres réalisés « indirectement ».

- ARIEL, M. (1990), *Accessing Noun-Phrase Antecedents*, Londres : Routledge.
- ARIEL, M. (1996), "Referring expressions and the +/- coreference distinction", In T. Fretheim & J.K. Gundel (eds.), pp. 13-25.
- ASHER, N. (1993), *Reference to Abstract Objects in Discourse*, Dordrecht : Kluwer Academic.
- BOSCH, P. & VAN DER SANDT, R. (eds.) (1999), *Focus. Linguistic, Cognitive, and Computational Perspectives*, Cambridge : Cambridge University Press.
- BRENNAN, S.E. (1995), "Centering attention in discourse", *Language and Cognitive Processes* 10(2), 137-167.
- CONTE, M-E. (1994), "Discontinuity in texts", In S. Cmejrková & F. Stícha (eds.), *The Syntax of Sentence and Text*, Amsterdam : John Benjamins, pp. 195-204.
- CORNISH, F. (1995), "Référence anaphorique, référence déictique, et contexte prédicatif et énonciatif", *Sémiotiques* 8,31-55.
- CORNISH, F. (1999), *Anaphora, Discourse, and Understanding. Evidence from English and French*, Oxford : Clarendon Press.
- COTE, S. (1998), "Ranking forward-looking centers", in Walker, M.A. *et alii* (eds.), 55-69.
- DAVID, S. (1990), "Quelques remarques à propos du traitement automatique des relations anaphoriques", in Kleiber, G., & Tyvaert, J-E. (eds.), *L'anaphore et ses domaines, Recherches Linguistiques XIV*, 111-141.
- DEANE, P.D. (1992), *Grammar in Mind and Brain. Explorations in Cognitive Syntax*, Berlin/New York : Mouton-de Gruyter.
- DI EUGENIO, B. (1996), "The discourse functions of Italian subjects : a Centering approach", In Proceedings COLING96, Copenhagen, pp. 352-357.
- DI EUGENIO, B. (1998), "Centering in Italian", in Walker, M.A. *et alii* (eds.), 115-137.
- DIK, S.C. (1997), *The Theory of Functional Grammar, Part I : the structure of the clause* (2ème édition), Berlin : Mouton de Gruyter.
- FOSSARD, M. (1999), "Traitement anaphorique et structure du discours : étude psycholinguistique des effets du 'focus de discours' sur la spécificité de deux marqueurs référentiels : le pronom anaphorique *il* et le nom propre répété", *InCognito* 15, 33-40.
- FOX, B.F. (1987), *Discourse Structure and Anaphora : written and conversational English*, Cambridge : Cambridge University Press.
- FRETHEIM, T. & GUNDEL, J.K. (eds.) (1996), *Reference and Referent Accessibility*, Amsterdam : John Benjamins.
- GARROD, S., FREUDENTHAL, D., & BOYLE, E. (1994), "The role of different types of anaphor in the on-line resolution of sentences in discourse", *Journal of Memory and Language* 33:39-68.
- GERNSBACHER, M.A., & JESCHENIAK, J.D. (1995), "Cataphoric devices in spoken discourse", *Cognitive Psychology* 29, 24-58.
- GORDON, P.C., & CHAN, D. (1995), "Pronouns, passives, and discourse coherence", *Journal of Memory and Language* 34, 216-331.
- GORDON, P.C., GROSZ, B.J., & GILLIOM, L.A. (1993), "Pronouns, names, and the centering of attention in discourse", *Cognitive Science* 17, 311-347.
- GROSZ, B.J., & SIDNER, C.L. (1986), "Attention, intentions and the structure of discourse", *Computational Linguistics* 12, 175-204.
- GROSZ, B.J., WEINSTEIN, S., & JOSHI, A.K. (1995), "Centering : a framework for modeling the local coherence of discourse", *Computational Linguistics* 21(2), 203-225.
- GUNDEL, J.K. (1996), "Relevance Theory meets the Givenness Hierarchy : an account of inferrables", In T. Fretheim & J.K. Gundel (eds.), pp. 141-153.

- GUNDEL, J.K. (1998), "Centering Theory and the Givenness Hierarchy : towards a synthesis", in Walker, M.A. *et alii* (eds.), 183-198.
- GUNDEL, J.K. (1999), "On different kinds of Focus", ch. 15 in Bosch & Van der Sandt (eds.), pp. 293-305.
- GUNDEL, J.K., HEDBERG, N., & ZACHARSKI, R. (1993), "Cognitive status and the form of referring expressions in discourse", *Language* 69(2), 274-307.
- HOFFMAN, B. (1998), "Word order, information structure, and centering in Turkish", in Walker, M.A. *et alii* (eds.), 253-271.
- HUDSON-D'ZMURA, S., & TANENHAUS, M.K. (1998), "Assigning antecedents to ambiguous pronouns :the role of the center of attention as the default assignment", in Walker, M.A. *et alii* (eds.), pp. 199-226.
- IIDA, M. (1998), "Discourse coherence and shifting centers in Japanese texts", in Walker, M.A. *et alii* (eds.), pp. 161-180.
- KAMEYAMA, M. (1999), "Stressed and unstressed pronouns:complementary preferences", ch. 16 in Bosch & Van der Sandt (eds.), pp. 306-321.
- KEHLER, A. (1997), "Current theories of Centering for pronoun interpretation", *Computational Linguistics* 23 (3), 467-475.
- LAMBRECHT, K. (1994), *Information Structure and Sentence Form. Topic, focus, and the mental representation of discourse referents*, Cambridge : Cambridge University Press.
- MAES, A.A. (1997), "Referent ontology and focus of attention in discourse", *Journal of Semantics* 14, 207-235.
- SANFORD, A.J., & GARROD, S.C. (1981), *Understanding Written Language : Explorations in comprehension beyond the sentence*, Chichester : John Wiley.
- SCHNEDECKER, C. (1997), *Nom propre et chaînes de référence*, Paris : Klincksieck.
- SIDNER, C.L. (1983), "Focussing in the comprehension of definite anaphora", in Brady, M. & Berwick, R. (eds.), *Computational Models of Discourse*, Boston : MIT Press, pp. 267-329.
- TURAN, U.D. (1998), "Ranking forward-looking centers in Turkish : universal and language-specific properties", in Walker, M.A. *et alii* (eds.), pp. 139-160.
- WALKER, M.A. (1998), "Centering, anaphora resolution, and discourse structure", in Walker, M.A. *et alii* (eds.), pp. 401-435.
- WALKER, M.A. (2000), « Vers un modèle de l'interaction du Centrage avec la structure globale du discours », *Verbum* 22 (1), 31-58 (ce volume).
- WALKER, M.A., & PRINCE, E.F. (1996), "A bilateral approach to givenness : a hearer-status algorithm and a centering algorithm", in Fretheim, T., & Gundel, J.K. (eds.), pp. 291-306.
- WALKER, M.A., JOSHI, A.K., & PRINCE, E.F. (eds.) (1998), *Centering Theory in Discourse*, Oxford : Clarendon Press.