

Chapitre III - La question de la référence en indexation

Comme précédemment la dimension lexicale - et sans doute plus encore -, la question de la référence¹ en indexation n'est pas, en tant que telle, présente dans le discours classique. On retrouve sur ce point une impossibilité directement liée à l'adoption implicite du modèle objectiviste (*supra*). Dans ce cadre, l'indexation ne peut en effet penser son « extérieur », ou plutôt elle ne peut le penser que sous une seule et même forme : celle du concept.

La vision de la langue et du lexique qui sous-tend l'approche classique de l'indexation va en effet de pair avec une certaine vision du monde². Il importe donc, à nouveau, de dégager l'arrière-plan théorique implicite afin de faire émerger, par déplacement de modèles, ce que le discours classique ne montre pas : les processus de référenciation³ à l'œuvre en indexation.

En dépit de son manque de visibilité, du moins dans le discours normatif⁴, le rôle de la référence est pourtant déterminant en indexation. C'est Michel Le Guern et les membres de l'équipe SYDO qui ont, les premiers, attiré l'attention sur cet aspect de l'indexation. En effet, ils ont mis en valeur que, si l'indexation manipule des **mots**, c'est pour renseigner sur les **choses** :

Les systèmes documentaires, à de rares exceptions près, ont pour finalité de fournir à l'utilisateur des renseignements sur les choses, et non sur les mots.
Le Guern, 1991a, p. 22.

La problématique documentaire devient, dans cette perspective, celle de la mise en relation entre un objet du monde et le document (ou les documents) qui apportera

¹. La notion de référence peut être entendue, en première approximation, comme « la propriété d'un signe linguistique lui permettant de renvoyer à un objet du monde extra-linguistique, réel ou imaginaire ». Définition reprise du Dictionnaire de linguistique et des sciences du langage 1994, p. 404.

². Voir Rastier 1994, p. 29 : « Le maintien de la triade sémiotique garde la sémantique sous la dépendance d'une ontologie, seule capable de relier les mots et les choses, par la médiation de concepts [position mentaliste] ».

³. *Ibid.*, p. 19 : « Ce que nous appelons ici *référence* n'est pas un rapport de représentation à des choses ou à des états de choses, mais un rapport entre le texte et la part non linguistique de la pratique où il est produit et interprété. [...] La référence ainsi définie ne relève pas de la représentation mais de l'action, telle qu'elle est structurée par une pratique ».

⁴. « Le thésaurus est considéré par les spécialistes de la documentation comme une représentation symbolique ne prenant pas en compte le référent. Mais le descripteur renvoie à des documents qui traitent de la réalité, cela revient à dire que le descripteur renvoie à cette même réalité par référent interposé, c'est-à-dire par le document », Mustafa-Elhadi 1992, p. 468.

(apporteront) des informations sur cet objet. C'est dans ce cadre que le descripteur peut être envisagé comme une unité de discours⁵ : nous y reviendrons.

Nous nous inscrivons dans le cadre de l'analyse proposée par Michel Le Guern, en privilégiant un point de vue linguistique sur la référence⁶. La problématique s'exprime alors sous la forme suivante : « comment le langage parvient-il à parler du réel ? »⁷, question qui aussitôt se dédouble⁸ :

- (i) Comment se réalise l'acte référentiel ?
- (ii) Quelles sont les propriétés référentielles qui permettent aux expressions de référer ?

Cette approche de la référence ne permet certes pas de tout dire sur la question de la référence en indexation ; notamment la linguistique ne peut rien dire de la « réalité » des éléments qu'elle manipule, réalité mentale (notion de concept) ou réalité phénoménale (notion de chose). Elle permet en revanche de pouvoir concevoir l'indexation comme une opération qui crée, *via* les mots, ses propres « choses », ses propres objets : document et descripteur.

Pour montrer comment l'indexation peut se concevoir sous l'angle de la référenciation, nous procéderons de la façon suivante, en trois étapes :

- nous dégageons, dans un premier temps, la zone de tension entre modèle documentaire et modèle linguistique sur le point qui concerne l'appréhension de la référence : la vision du monde que sous-tendent les deux modèles est radicalement différente (I) ;
- nous présentons ensuite l'approche de la référence dans le cadre du modèle linguistique : la façon dont les questions sont posées ainsi que les distinctions qui y sont établies nous fournissent des repères pour identifier les phénomènes référentiels en indexation (II) ;
- nous proposons ensuite une approche de la référence en indexation qui tient compte des hypothèses linguistiques précédemment dégagées : l'indexation apparaît, dans ce cadre, comme une opération de référenciation, c'est-à-dire comme une opération qui construit ses propres objets comme ses propres effets référentiels (III).

⁵. « La finalité du descripteur exclut que l'on puisse l'envisager en faisant abstraction de la valeur référentielle de ses occurrences dans le corpus », Le Guern 1984, p. 164.

⁶. Cependant, la référence constituant typiquement un concept scientifique interdisciplinaire, nous serons amenée à aborder quelques aspects des problématiques logique et philosophique de la référence.

⁷. Kleiber 1981, p. 11.

⁸. *Ibid.*, p. 13 ; les deux questions (i) et (ii) sont intimement liées : « l'acte de référence ne saurait être accompli si les expressions n'avaient pas des caractères référentiels propres ».

I- Conflit entre modèles de la référence

Modèle documentaire et modèle linguistique illustrent chacun l'une des deux branches de l'alternative philosophique sur la référence.

Exprimé en termes philosophiques, le débat sur la référence fait émerger deux positions distinctes, celle des réalistes et celle des nominalistes, qui constituent la base du conflit entre les deux modèles documentaire et linguistique en présence. Après avoir rappelé les grandes lignes de ce débat philosophique (I.1), nous présenterons, d'une part, les principales marques du modèle réaliste en indexation (I.2) et, d'autre part, les limites les plus remarquables de ce présupposé réaliste pour la description des faits d'indexation (I.3).

I.1- Les termes du débat sur la référence

Traditionnellement, l'histoire de la philosophie situe la primeur du débat sur la référence au X^e siècle, en pleine « querelle des universaux » qui cristallise alors l'ensemble de la réflexion philosophique sur le langage, liant à la fois logique et théologie.

Le débat philosophique sur la référence s'organise⁹ autour de deux positions adverses, celle des Réalistes et celle des Nominalistes :

- pour les Réalistes, « les mots font directement référence à la réalité objective, soit permanente, soit impermanente, mais toujours réelle et réellement différenciée »¹⁰. C'est une position qui reprend la thèse d'Aristote¹¹, dite essentialiste, dans le sens où l'existence est appréhendée en termes d'essence ;
- pour les Nominalistes, la réalité n'est perceptible que par les mots : le concept n'est rien de plus qu'un nom, « qu'une construction mentale sans réalité extérieure »¹². C'est pourquoi nous ne pouvons pas connaître la réalité en soi, dans son essence, mais seulement dans « les représentations (les dénominations) par lesquelles nous percevons les phénomènes dont nous expérimentons les sensations »¹³. C'est une position dite

⁹. Il s'agit là d'une présentation où l'on radicalise les positions de chacun ; une approche plus complète peut se trouver, par exemple, dans De Libera [1993].

¹⁰. Zimmermann 1989, p. 402.

¹¹. Aristote n'illustre qu'un aspect du réalisme antique ; c'est surtout la lecture d'Aristote par les médiévaux que l'on retient ici.

¹². Zimmermann 1989, p. 402.

¹³. Le Moigne 1995, p. 45.

existentialiste, illustrée notamment par Guillaume d'Ockham, qui tient que s'il y a des propriétés communes à plusieurs éléments, on ne peut, pour autant, en inférer la réalité de ces éléments.

D'option philosophique, la position réaliste s'est peu à peu imposée comme épistémologie « officielle », pour reprendre les termes de Le Moigne¹⁴, notamment sous la forme du modèle « cartésiano-positiviste ».

C'est ainsi que la position réaliste s'est maintenue comme le modèle du sens commun, comme une *doxa* utile et utilisée dans la mesure où elle fournit un mode d'appréhension aisé et immédiat du rapport entre les mots et les choses¹⁵. Rien d'étonnant, là encore, à ce que le modèle documentaire épouse la forme de cette *doxa*, comme nous le verrons ci-après.

Face à ce modèle réaliste, dominant, le modèle nominaliste, moins représenté : ce sont surtout les linguistiques et la sémiologie qui, reprenant la position nominaliste, ont reconstruit une « opposition » qui prend alors le nom de constructivisme¹⁶. Cette épistémologie constructiviste se fonde sur les deux hypothèses suivantes :

Une hypothèse relative au statut de la réalité connaissable, qui pour être connue doit pouvoir être cognitivement construite ou reconstruite intentionnellement par un observateur-modélisateur ; et une hypothèse relative à la méthode d'élaboration ou de construction de cette connaissance qui ne fera plus appel à une « norme du vrai » (par déduction programmable) mais à une « norme de faisabilité » (par intuition re-programmable).

Le Moigne, 1995, p. 41.

Cependant, si ce sont les linguistes qui ont, en grande partie, contribué à raviver le débat sur la référence¹⁷, il n'y a pas, pour autant, consensus absolu entre les écoles : des différences se notent tout autant entre linguistes qu'entre linguistes et logiciens ou linguistes et philosophes¹⁸. Néanmoins, on peut formuler de façon générale la problématique de la référence en linguistique sous la forme suivante : s'il n'y a pas, dans la langue, de marque de référence, tout mot peut-il pour autant

14. Le Moigne [*Ibid.*, p. 4-11] parle plus volontiers d'épistémologie « institutionnelle ».

15. Voir Dubois et Mondada 1995, p. 274 : « La croyance en un monde extérieur est une propriété centrale de la "raison mondaine" ("mundane reason", Pollner 1987), qui donne une intelligibilité et une descriptivité à la réalité quotidienne, à ses représentations ordinaires, aux raisonnements de tous les jours ; qui permet en outre de traiter les contradictions ou les conflits entre des versions multiples et discordantes des "mêmes" réalités comme étant imputables à l'erreur ou à la folie ».

16. Le Moigne 1995, p. 44-46.

17. Par exemple, l'hypothèse dite Sapir-Worf (chaque langue est spécifique et configure le monde à sa façon).

18. Une présentation exhaustive du débat dont la référence a fait l'objet aussi bien en philosophie qu'en logique dépasserait largement le cadre de ce travail : on peut trouver une synthèse et des références bibliographiques dans Nef 1991, p. 85-109.

pointer indistinctement sur n'importe quel objet ? C'est principalement autour de cette question que les linguistiques abordent la question de la référence.

Sur la base de ces présupposés philosophiques, on peut faire apparaître ce qui oppose le modèle documentaire et le modèle linguistique : d'un côté, un monde construit dont il s'agit de rendre compte ; de l'autre, un monde à construire selon des procédés dont il faut rendre compte. L'ambition est radicalement différente.

I.2- Traces du modèle réaliste en indexation

On trouve, dans le discours classique, deux types de « trace » du modèle réaliste : l'un concerne la stabilité de la relation entre les mots et les choses (A), l'autre la pré-existence des documents en cause dans cette relation (B).

A- Stabilité de la relation entre les mots et les choses

Dans le modèle réaliste, le langage est définitivement coupé du monde, comme le monde du langage : d'un côté, existe un monde discrétisé en objets, incarnations d'essences (ou concepts), de l'autre existent des mots, expressions d'essences (ou concepts). La langue est là, qui se pose comme instrument de mise en relation entre mots et choses opérant *via* les concepts.

Le monde existant préalablement à toute appréhension, il est le même pour tous les sujets parlants : c'est pour cela que l'on peut transmettre ses éléments. Le seul problème vient des mots, qui ne sont pas toujours aptes à les transmettre correctement.

Le modèle réaliste, qui suppose à la fois une stabilité des « choses » et une totale extériorité du monde par rapport au langage, constitue le modèle de « référence » de l'indexation :

- l'indexation se donne en effet comme une procédure chargée de maintenir la stabilité des choses du monde au cours de leur transfert des auteurs vers les lecteurs ;
- elle se pose à ce titre comme une simple manipulation d'objets lui préexistant - objets du monde et objets textuels¹⁹ - qui ne sont pas, dans cette approche, distingués²⁰.

¹⁹. Ou plus généralement objets documentaires.

²⁰. Cf. Maniez 1993, p. 254 : l'indexation se justifie parce qu'il existe d'un côté des documents et de l'autre des besoins d'information. Dans ce cadre, l'indexation est ce qui doit indiquer « à un

Dans le cadre réaliste, l'enjeu de la référence en indexation se donne essentiellement sous la forme suivante : comment assurer une stabilité du référent d'un bout à l'autre de la chaîne documentaire ? Cette question, qui se pose uniquement dans un cadre réaliste, reçoit naturellement une réponse de type réaliste. À la problématique de la stabilité référentielle correspond en effet celle de la stabilité linguistique : comme le monde est stable, il suffit de désigner ses objets toujours par les mêmes mots. C'est là le fondement même des langages documentaires : pour être sûr de désigner toujours le même objet, on doit se mettre d'accord pour employer toujours le même mot²¹.

B- Pré-existence des objets documentaires

Le modèle réaliste ne permet pas de poser véritablement la question des types de référents, d'objets (mondains et/ou textuels), sur lesquels pointe l'indexation.

La question du type d'objet à indexer est en effet peu traitée dans le discours classique, comme si la problématique du document, comme celle de l'information, ne concernaient pas l'indexation elle-même.

Le document est en effet un « donné », l'« input » de l'indexation : il est en cela constitué en dehors d'elle²². De même l'information²³ est conçue comme étant extérieure au processus même de l'indexation ; on ne sait pas bien si elle constitue son « input » ou son « output », sans doute les deux.

On peut remarquer que le modèle réaliste imprime profondément sa marque en indexation :

- c'est lui qui détermine les outils de l'indexation, et notamment la forme des langages documentaires : hors de ce modèle, est-on

demandeur les documents relatifs au sujet qui l'intéresse c'est-à-dire à la portion de la réalité sur laquelle se focalise sa curiosité ». C'est nous qui soulignons la formulation de cette assimilation entre objet textuel et objet du monde.

²¹. Cette analyse du langage documentaire comme présupposé réaliste se retrouve chez d'autres auteurs, notamment chez Turner [1990, [p. 2]] : « L'idée que la construction d'un langage artificiel peut conduire à réduire l'ambiguïté du langage naturel est fondée sur une **théorie positiviste du langage**. Selon ce point de vue, la signification des mots vient de ce qu'ils nomment. [...] Le travail de création d'un langage artificiel consiste à ramener la diversité de signes à un symbolisme unique. Le langage documentaire est, en théorie, une liste de symboles ayant chacun un rapport intrinsèque avec la chose qu'il représente ».

²². La norme [Z 47-102 (1978), p. 231] appréhende en effet le document comme un objet déjà constitué : c'est l'« ensemble d'un support d'information, des données enregistrées sur ce support et de leur signification, servant à la consultation, l'étude, la preuve, etc. ».

²³. Dans les termes de la norme, l'information est entendue comme la « signification que l'homme attache aux données au moyen de conventions connues utilisées dans sa représentation » ; mais la notion de donnée y est parallèlement définie de façon circulaire comme la « représentation codée d'une information ». Vocabulaire de l'indexation 1987.

encore obligé de soutenir une corrélation entre stabilité référentielle et stabilité linguistique ?

- c'est lui qui autorise une mise à l'écart des objets de l'indexation du champ même de l'indexation : hors de ce modèle, il devient nécessaire de définir d'une part et de distinguer d'autre part les principaux objets de l'indexation, notamment le document et l'information.

On voit ce que suggère une remise en cause du modèle réaliste sous-jacent au discours classique : un changement radical dans l'approche de la notion d'indexation et un important travail de définition de ses objets. Malgré le caractère ambitieux de cette tâche, il nous semble important de l'amorcer, le modèle réaliste se révélant fortement inadéquat pour rendre compte des faits d'indexation.

I.3- Limites du modèle réaliste en indexation

Le présupposé réaliste de la pré-existence d'un monde stable et déjà discrétisé, tel qu'il est assimilé dans le modèle classique de l'indexation, laisse sans réponse deux ordres de problème auxquels se trouvent régulièrement confrontées les pratiques d'indexation :

- la variabilité des objets d'indexation : l'objet d'indexation, ou encore le document, est assimilé, dans le modèle réaliste, à un objet du monde comme les autres : il existe et est prêt à être indexé. Or, pour un monde donné, est-ce que ce sont toujours les mêmes objets qui sont retenus pour être indexés ?
- la variabilité des termes d'indexation : bien que le monde soit discrétisé en entités stables et que les langages documentaires donnent de quoi représenter de façon univoque chacune de ces entités, la pratique de l'indexation se heurte encore à la variabilité des termes d'indexation²⁴.

Ce sont là deux aspects de la pratique d'indexation qui contredisent les présupposés réalistes adoptés dans le discours classique et qui indiquent les limites de ce modèle pour une approche des faits d'indexation. Nous précisons ces deux aspects ci-dessous.

²⁴. Cf. norme Z 47-102 (1978), p. 6 : « Dans le cas idéal, pour un système donné, l'indexation d'un document devrait être identique quel que soit l'indexeur. Elle ne devrait pas non plus varier dans le temps pour un même indexeur si l'outil documentaire n'a pas été modifié ».

I.3.1- La variabilité des objets d'indexation : la question du document

Contrairement à la définition que propose la norme AFNOR²⁵ et suivant plutôt une proposition de Gardin²⁶, nous entendrons, dans un premier temps, par document tout objet retenu pour être indexé.

Cette définition posée, la question qui se pose est la suivante : qu'est-ce qu'un objet « indexable » ? Autrement dit, comme le souligne Odile Le Guern, avant la question du choix des termes d'indexation, l'indexeur se pose bien celle du « choix des objets représentés », de « ce qu'il va choisir de montrer (index) »²⁷. Dans la littérature consultée, on trouve peu de précisions sur ce point. Quand la question est évoquée, c'est pour dire qu'elle n'a justement pas été traitée²⁸.

Il nous semble que la question du choix des objets documentaires se pose à deux niveaux différents :

- (i) qu'est-ce qui, parmi l'ensemble de la production éditoriale, préside au choix d'une source documentaire* ?
- (ii) qu'est-ce qui, pour une source documentaire donnée, détermine le choix de la partie de la source effectivement indexée ?

La question (i), qui relève de ce qu'il est convenu d'appeler une « politique d'acquisition »²⁹, engage l'organisation de l'espace des documents dans son entier : elle sera traitée sous cet angle dans le chapitre IV.

La question (ii) renvoie à ce que l'on appelle généralement la sélection documentaire³⁰ : c'est elle qui constitue le document proprement dit ; c'est d'elle dont il sera prioritairement question dans ce paragraphe.

Les principes de sélection les plus couramment exprimés reposent sur l'idée qu'il importe de respecter la source éditoriale telle qu'elle se donne : l'indexeur doit rester « collé » à la réalité éditoriale. Or celle-ci est, heureusement, plus malléable qu'il n'y paraît.

Pour mesurer la latitude que les indexeurs prennent par rapport à la « réalité

25. Voir le glossaire à l'entrée « document ».

26. Gardin [1967] propose de considérer le document comme « tout objet au sens large (objet concret, image, texte, etc.) considéré comme unité d'analyse et/ou de référence, dans les travaux d'indexation ».

27. Odile Le Guern 1989, p. 427

28. Suzanne Bertrand-Gastaldi [1993, p. 161] rappelle la problématique, « restée jusqu'alors sans réponse », que formule Coates en 1979 : « How does an indexer determine what is or what is not indexable ? ».

29. Voir par exemple Calenge 1994.

30. Ou encore la « couverture documentaire » : ces notions renvoient aux « domaines » que traitent prioritairement une bibliothèque ou un centre de documentation. Cet aspect de la sélection documentaire reste le plus souvent à la discrétion des organismes documentaires : on ne trouve que peu de règles formalisées expliquant les choix effectués. Nous y revenons dans le chapitre IV.

éditoriale », nous avons réalisé une enquête auprès de dix organismes documentaires³¹ pour observer comment, à partir d'un même « monde », d'une même « réalité éditoriale », différents indexeurs opéraient leur découpage en documents.

A- Expérimentation

Dans l'expérience que nous avons menée, nous entendons :

- par « source », tout article du journal *Le Monde* : toute unité textuelle typographiquement circonscrite le plus souvent entre un titre et une signature ;
- par « document », le segment textuel affecté d'un (ou de plusieurs) descripteur(s).

En examinant le rapport entre sources sélectionnées et documents indexés³², on repère les trois cas de figure suivants :

- (i) six des dix organismes documentaires sondés ont sélectionné plus de textes qu'ils n'ont indexé de documents : autrement dit, la création du document consiste ici essentiellement en un **regroupement** de différentes sources ;
- (ii) deux des organismes documentaires ont indexé plus de documents qu'ils n'ont sélectionné de sources : dans ce cas, la même source subit un **éclatement** en objets documentaires différents ;
- (iii) seuls deux des dix organismes documentaires étudiés ont indexé autant de documents qu'ils avaient sélectionné de sources : autrement dit, le type de source sélectionnée constitue, pour ces organismes, un objet d'indexation pertinent.

Cette première observation montre que l'indexation ne manipule pas toujours les objets textuels (les sources) tels qu'ils lui pré-existent mais qu'elle crée ses propres objets.

Ce constat est renforcé par un examen plus minutieux qui montre que, à partir d'une même source (un même article *du Monde*), les différents organismes documentaires créent des documents différents.

La diversité des cas de figure est bien représentée par le traitement documentaire de deux articles du *Monde* portant, pour le premier, sur la chaîne de télévision Arte et, pour le second (un encadré) sur Jean-Marie Cavada (alors président de La

³¹. L'annexe 2 détaille les participants et les consignes de l'expérience réalisée.

³². Voir annexe 3.

Cinquième, promu président du GIE Arte-La Cinquième)³³ :

- soit les deux articles constituent chacun un document ;
- soit un seul des deux articles est retenu comme document ;
- soit les deux articles sont regroupés en un document ;
- soit les deux articles sont regroupés dans une même revue de presse (comportant d'autres « sources ») ; et c'est elle qui constitue un document (qui est affectée de descripteurs) ;
- soit l'un des deux articles est intégré dans une revue de presse qui, elle, forme un document particulier, indexé en tant que tel.

On observe également une variabilité de « mise en document » pour une même source au sein d'un même organisme documentaire ; ainsi le centre de documentation de la *Fondation nationale des sciences politiques* procède-t-il en deux temps :

- dans un premier temps, une source du *Monde* est constituée en document, c'est-à-dire, pour cet organisme documentaire, sélectionnée, affectée d'un indice de classification et intégrée dans un dossier documentaire ;
- dans un second temps, c'est le dossier documentaire lui-même, composé de coupures de presse de source hétérogène, qui constitue à son tour un « document » : il est, dans son intégralité, indexé dans un autre langage documentaire (Rameau en l'occurrence) et intégré comme **une** seule « référence » (au sens bibliothéconomique du terme) dans le catalogue de la bibliothèque.

Ces différentes observations, issues d'une expérience dont la portée exemplaire n'est, par ailleurs, pas prouvée³⁴, montrent cependant que le document, loin d'être un « donné », est un « construit », ou, mieux, un « état ». Il semble, en effet, que tout texte ou ensemble de textes peut fonctionner, à un moment donné, comme « un document »³⁵.

L'expérience menée nous fournit une parfaite illustration de ce dernier point (le

³³. *Le Monde* 1/12/1994, p. 10.

³⁴. Il faudrait qu'une expérimentation similaire soit effectuée à partir d'autres types de sources, issues par exemple de la littérature scientifique et technique. Retrouve-t-on, sur des domaines « étroits », cette même variabilité dans la « mise en documents » à partir d'une même « source » ?

³⁵. Voir, sur ce point, la notion de « document par intention » [Benoît (dir.) 1992] : « est document tout support d'information voulu comme tel ».

document conçu par « intention »). C'est à nouveau la pratique d'indexation du centre de documentation de la *Fondation nationale des sciences politiques* qui illustre ce cas de figure. A été retenu, en effet, comme unité d'indexation, un « non-article », une « lettre ouverte à monsieur Jacques Chirac » signée d'une association de « professeurs et accompagnateurs des conservatoires municipaux de Paris »³⁶, qui se rapproche d'une publicité (au sens premier du terme) et qui n'est assurément pas un article. Ce dernier exemple souligne ce qui ressortait déjà des observations précédentes : le document se présente comme un texte (ou ensemble de textes) en usage, ou plutôt mis en usage.

Les données de cette expérimentation permettent de montrer, au moins sur le cas du traitement documentaire de la presse généraliste, que, pour une « réalité éditoriale donnée », il y a toujours plusieurs mises en document possibles et qu'à ce titre, le document ne peut être considéré comme un donné : il est nécessairement construit et construit par « intention ».

B- Conclusions de l'expérimentation et formulation d'hypothèses

L'ensemble des précédentes observations issues de notre expérimentation amène à formuler un ensemble de questions et d'hypothèses relatif au document d'une part et à l'indexation d'autre part.

Questions et hypothèses relatives à l'approche du document en indexation

Nous ferons l'hypothèse que le document est en indexation un « construit », ou plutôt un « reconstruit » à partir d'une source, dont il propose une utilisation : nous nous placerons dans une approche constructiviste du monde.

Les questions à résoudre seront alors de divers ordres :

- en quoi consiste la construction du document ? Qu'est-ce que le document retient de la source et qu'est-ce qu'il ne retient pas ? La diversité de la « mise en document » des sources induit-elle que l'« information » proposée est différente d'un organisme documentaire à l'autre ? Autrement dit l'opération de mise en document intervient-elle dans l'interprétation de la source proposée à l'utilisateur ?
- parmi l'ensemble des sources existantes, qu'est-ce qui guide le

³⁶. *Le Monde* 1/12/1994, p. 14.

choix d'une source ? Sur quels critères une source peut-elle être utilisée comme un document ? Quelles sont les propriétés que doit avoir une source pour fonctionner comme document ?

- quel est le rapport entre le monde des choses et le monde des documents créés en indexation ? En quoi les documents, plus que les « sources », sont-ils à même de renseigner un utilisateur sur les « choses », pour reprendre les termes de Michel Le Guern ?

Questions et hypothèses relatives au processus d'indexation

Si l'indexation ne porte pas sur des objets existants mais qu'elle construit ses propres objets, il semble important de la redéfinir en intégrant cette dimension de la création du document : on dira à ce titre que le processus de l'indexation comprend une phase de mise en document. Cette phase, classiquement abordée sous l'angle de la sélection documentaire, est traditionnellement conçue comme distincte de celle de l'indexation.

Cependant, la proximité entre ces deux opérations a pu être relevée, par Le Loarer par exemple :

Le processus d'indexation peut être lié à (ou suivre) celui qui est relatif à la sélection des documents à indexer :

- parmi les ouvrages reçus ou acquis dans un centre de documentation ou une bibliothèque, quels sont ceux qui font l'objet d'une indexation ?
- dans un fascicule de périodique, indexe-t-on tous les articles ou seulement certains ?
- dans des actes de congrès traite-t-on toutes les communications ou seulement certaines ?

Le Loarer, 1994, p. 152.

Nous faisons ici hypothèse que l'opération d'indexation inclut une phase de sélection et de construction de ses objets.

Cette phase est réalisée *a priori* sur la base d'une présomption, qui relève moins d'un souci des besoins de l'utilisateur que d'une « vision du monde » que l'on estime pertinente à fournir à un utilisateur³⁷. S'il est vrai que l'indexation doit se penser dans les termes d'une « analyse prévisionnelle »³⁸, c'est moins **en aval** de l'indexation, au niveau des termes à choisir, qu'**en amont** de l'indexation au niveau des sources à retenir et des documents à construire.

Notre approche de l'indexation nous conduit à reformuler la notion de prédiction telle qu'elle se conçoit habituellement en indexation³⁹ : elle n'affecte plus les

37. C'est là où l'indexation rejoint la problématique de la catégorisation qui établit une discrétisation en « domaines » ; la détermination de ces domaines repose nécessairement sur une certaine perception d'un monde, sur une « idéologie », voir sur ce point Escarpit 1991, p. 156.

38. *Ibid.*, p. 167.

39. La notion de prédiction est surtout pensée dans le discours classique en termes de

« mots »⁴⁰ mais les « choses ». L'indexation apparaît en ce sens comme un acte de discrétisation du monde qu'elle propose de faire voir sous un certain angle (celui des documents qu'elle retient et qu'elle organise). Dans ce cadre, et en opposition au modèle réaliste, la variabilité des objets d'indexation n'est plus un problème, mais au contraire la marque distinctive de l'indexation : on attend de l'indexation qu'elle puisse donner un sens à la partie du « réel » qu'elle choisit de montrer⁴¹.

1.3.2- Variabilité des termes d'indexation : la question de la stabilité référentielle

La variabilité des termes d'indexation⁴², c'est-à-dire, par exemple, le constat que deux indexeurs d'un organisme documentaire donné choisissent, parmi les termes d'un même langage documentaire, deux descripteurs différents pour indexer un même document, ne pose un problème que dans le cadre réaliste. Or c'est précisément dans ce cadre qu'il ne peut être résolu.

En effet, la notion de variabilité de l'indexation repose sur une mesure de la disjonction entre stabilité référentielle (incarnée dans le document, vu comme donnée stable pourvue d'une signification⁴³) et stabilité linguistique (incarnée dans le descripteur, représentant lui aussi d'un sens unique⁴⁴). Que stabilité référentielle (stabilité des « choses ») et stabilité linguistique (stabilité des « mots ») se correspondent relève d'une conception réaliste de la signification⁴⁵ qui ne peut expliquer les cas de disjonction qu'en termes de faute ou d'erreur⁴⁶.

Hors d'un cadre réaliste, la problématique de la stabilité référentielle ne se pose pas,

« représentation » lexicale du « contenu ». Ainsi Fugmann [1993], qui a particulièrement développé cet aspect au point d'en faire un axiome de l'indexation, parle-t-il de la « prédictibilité de la représentation ».

⁴⁰. Toute la problématique des langages contrôlés relève de la croyance en la nécessité et en la possibilité de prévoir la forme linguistique la plus susceptible d'être utilisée à la recherche.

⁴¹. C'est la notion d'« intention de sens » que défend Batime pour la construction des systèmes d'information. Batime 1995, p. 19-25.

⁴². Si le problème de la variabilité en indexation est beaucoup plus large [voir Bertrand 1993, pour une revue de la question], la variabilité dans les termes d'indexation constitue néanmoins l'aspect le plus étudié (selon Bertrand 1993, p. 20).

⁴³. Voir le glossaire pour la définition normative du document.

⁴⁴. *Id.*

⁴⁵. Apothéloz et Reichler-Béguelin 1995, p. 202 : une conception réaliste de la signification est une conception « dans laquelle le signifié se réduirait à une relation rigide non manipulable par les sujets parlants, entre la langue et le monde ».

⁴⁶. Voir par exemple la remarque de Dubois 1995, p. 90 : les erreurs, décalages, difficultés ou inadéquations rencontrés par les professionnels sont imputés à la fois aux outils (rustiques mais perfectibles) et aux humains « à l'évidence condamnés aux imperfections ».

ou du moins ne se pose plus dans les mêmes termes⁴⁷. Elle apparaît non plus sous l'angle de la production (il ne s'agit plus de produire une unité linguistique « stable » pour exprimer toujours le « même » référent) mais sous l'angle de la réception ; d'un point de vue interprétatif peuvent se créer des effets de stabilité référentielle⁴⁸. Là encore, dans le cadre réaliste, où le monde est donné comme stable et discrétisé en objets de la même façon pour tous les sujets parlants, la différence des points de vue production/indexeur et réception/utilisateur ne peut apparaître. Or ce point de vue est déterminant quand il s'agit de penser la notion de stabilité référentielle en indexation : il permet notamment de distinguer fin et moyens.

Si l'indexation peut se donner pour fin d'assurer une relation référentielle stable, cet objectif se réalise nécessairement au niveau de la réception (de l'utilisateur) et non à celui de la production (de l'indexeur). En effet, l'indexeur, qui ne dispose que d'unités lexicales hors emploi, ne peut à proprement parler établir de relation référentielle, celle-ci ne se créant qu'en discours :

Les noms, en effet, ne trouvent pas leurs référents tout prêts dans la nature, mais ils doivent pour ainsi dire les construire à chaque fois dans la communication, en découpant les classes d'objets dans le tissu de l'expérience.
Formigari, 1992, p. 448.

La question de savoir si c'est toujours le même référent que le mot construit « dans la communication » ou bien si celui-ci varie en fonction de chaque discours, de chaque contexte pourra être abordée dans le cadre d'une approche linguistique de la référence. En effet, de ce point de vue, on peut interroger la morphologie du descripteur.

Les noms propres sont, comme nous le verrons, particulièrement à même d'établir une relation référentielle stable. Sur ce point, il n'est pas indifférent que les premières pratiques d'indexation se soient constituées autour du nom propre⁴⁹ ; mais il n'est pas indifférent, non plus, que les pratiques d'indexation aient plus tard utilisé aussi des noms communs, qui, eux, ne peuvent, en règle générale, établir de relation référentielle stable.

C'est donc du point de vue de la réception, de l'interprétation que se créent, notamment par le biais des noms propres des effets de stabilité référentielle :

⁴⁷. Dubois et Mondada 1995, p. 282 : « Notre argument consiste à dire que la "stabilité" résulte en fait d'un point de vue réaliste qui relie les catégories à des propriétés du monde - comme si l'objectivité du monde produisait la stabilité des catégories - au lieu de les relier à des discours socio-historiques et à des procédures ancrées culturellement ».

⁴⁸. Voir Dubois et Mondada [*Ibid.*, p. 273 notamment] qui parlent d'« effet stabilisateur des pratiques ».

⁴⁹. Voir Escarpit 1991, p. 153 : les catalogues d'Alexandrie offraient des entrées qui ont été d'abord les premiers mots d'un document « puis plus tard le nom de l'auteur quand l'imprimerie a tiré l'œuvre de l'anonymat ».

S'agissant de « personnages », ou tout autre objet-de-discours susceptible d'être redésigné sur la durée par le même nom propre ou par un désignateur peu contingent, on peut supposer que les opérations de référence sont au bénéfice d'un statut cognitif spécial. Du fait même des propriétés sémantiques des désignateurs dits rigides, un tel objet-de-discours **apparaîtra comme stable** en tant qu'objet mémoriel : son identité **pourra donner l'illusion** d'être indépendante des prédications dont il fait l'objet, quand bien même les connaissances qu'on a de lui évoluent au fil du discours. Apothéloz et Reichler-Béguelin, 1995, p. 266 (c'est nous qui soulignons).

La variabilité des termes d'indexation apparaît dans le discours classique comme un problème sans réelle solution : on finit par adopter une position fataliste où l'on invoque inmanquablement l'imperfection humaine. Or ce problème ne se pose que dans le cadre du modèle réaliste implicitement adopté.

I.4- Conclusions intermédiaires

Au début de ce paragraphe, nous cherchions à dégager la « zone de tension » entre modèle documentaire et modèle linguistique sur la question de la référence. En reprenant, de façon schématique, l'alternative classique en philosophie entre Réalistes et Nominalistes, nous avons proposé d'identifier la zone de conflit entre les deux modèles en termes de vision du monde : dans le modèle documentaire classique d'inspiration réaliste, le monde est donné ; dans le modèle linguistique d'inspiration nominaliste, le monde est construit.

Au-delà de simples positions de principe, quelle est l'incidence de l'adoption de l'un ou l'autre des deux modèles ?

Nous avons cherché à montrer que l'adoption implicite du modèle réaliste en indexation se révélait, pour qui cherche à analyser les faits d'indexation, à la fois limité et opacifiant : limité, dans le sens où ses présupposés se trouvent pris en défaut par la pratique sans que le modèle ne puisse proposer ni une explication ni une solution ; opacifiant, dans la mesure où fin et moyens de l'indexation sont, dans le modèle réaliste, peu distingués, ce qui conduit au « bricolage » de langages documentaires conçus pour réaliser une double stabilité, référentielle et linguistique, alors même qu'il s'agit là d'effets de nature différente. L'effet de stabilité linguistique du descripteur relève du lexique (elle vient de son autonomie lexicale) ; l'effet de stabilité référentielle relève du discours (elle vient de l'interprétation).

Pour concevoir les différents aspects de la référence en indexation - ceux qui

concernent le document d'une part et ceux qui concernent l'effet de stabilité référentielle d'autre part -, nous devons au préalable disposer d'un modèle qui nous permette de traiter ces questions de façon explicite. C'est à ce titre que nous sollicitons une approche linguistique de la référence (paragraphe II, suivant), qui nous permettra de proposer une approche documentaire explicite de la référence (paragraphe III de ce chapitre).

II- L'approche de la référence dans le modèle linguistique

Le modèle linguistique⁵⁰ formule la problématique de la référence sous un angle qui, en distinguant les objets et en reformulant les questions, nous permet d'approcher de façon plus précise les phénomènes référentiels en indexation.

Nous présentons succinctement la formulation de la problématique de la référence en linguistique, puis nous nous focalisons ensuite sur la question centrale de la référence en linguistique : celle du rapport entre le sens et la référence.

II.1- Formulation de la problématique de la référence en linguistique

Comment nous l'avons mentionné dans le premier chapitre et comme il est apparu dans le chapitre II⁵¹, la méthodologie d'analyse en linguistique ne permet pas de traiter directement de faits massifs comme celui de la référence.

La question de la référence est en linguistique généralement abordée sous l'angle de phénomènes de langue précis et jugés exemplaires : c'est le plus souvent par l'étude des anaphores que sont posés les problèmes de référence. Nous esquissons rapidement ce cadre privilégié de l'étude de la référence en linguistique. Nous présentons ensuite les principales options méthodologiques de l'analyse linguistique de la référence : il s'agit, d'une part, de distinguer les objets traités (types de référents et acte de référence) et, d'autre part, de sérier les aspects de la problématique de la référence que la linguistique peut traiter (qu'est-ce que la linguistique peut dire de la référence en restant dans son domaine de compétence ?).

II.1.1- Cadre privilégié pour l'étude de la référence en linguistique

Qu'il s'agisse de Milner [1976], de Corblin [1995] ou encore d'Apothéloz et Reichler-Béguelin [1995], c'est en général le problème de l'anaphore qui est retenu

⁵⁰. L'emploi du singulier ne signifie pas qu'il y a consensus, en linguistique, sur l'approche de la référence ; sur ce point Kleiber 1981 donne, nous semble-t-il, une bonne vue d'ensemble des différentes positions linguistiques sur la référence. Reste que pour l'opposition que nous traitons dans ce chapitre, nous pouvons aborder l'approche linguistique de la référence dans son ensemble puisque, quelles qu'elles soient, les positions linguistiques contemporaines se situent toutes dans un cadre non réaliste, même si elles ne le font pas exactement dans les mêmes termes, voir notes 53 et 54. Sur ce point, Apothéloz et Reichler-Béguelin 1995, p. 240 : « Le réalisme n'intéresserait en soi que la critique philosophique si, au plan linguistique, il ne conduisait en droite ligne à des conceptions aujourd'hui intenable en matière de sémantique lexicale. [...] Cette position va forcément de pair avec une conception nomenclaturiste du lexique dont Saussure a fait la critique que l'on sait ».

⁵¹. Dans le chapitre II (§ II.2.1), nous avons vu que, dans le cadre retenu, l'analyse de la signification lexicale se faisait sur des pans délimités du matériel lexical, non pas mot par mot, mais par type de mots ; à ce titre, nous avons pris, comme exemple d'analyse de la signification lexicale, le domaine que travaille D. Corbin : celui des mots construits.

pour aborder la question de la référence⁵². Parmi les différentes problématiques qui relèvent d'une analyse des anaphores, c'est celle des anaphores « évolutives » qui semble le mieux poser les problèmes de la référence.

Pour simplement situer le cadre des questions que pointent les anaphores évolutives, prenons l'exemple le plus « fameux » de l'emploi des anaphores évolutives, celui de la recette du poulet :

Tuez un poulet actif et bien gras. Préparez-le pour le four. Coupez-le en quatre morceaux et faites le rôtir avec du thym pendant une heure.

La question qui est posée à partir de ce texte est la suivante : à quoi renvoient les quatre anaphores « le » ? D'un point de vue strictement grammatical, les quatre anaphores « le » devraient renvoyer à la première mention du groupe nominal « un poulet actif et bien gras » ; du point de vue intuitif de la compréhension, les quatre anaphores « le » renvoient plutôt à des états successifs de l'objet « poulet ». La question est de savoir ce qui « évolue » au cours de ce texte : le référent-objet « poulet » ou bien le référent-mot « poulet » ?

Encore faut-il poser une distinction entre ces deux types de référent et, à partir de là, se demander ce qui ressort du travail du linguiste : l'examen de l'évolution des choses ou bien l'examen de l'évolution de la saisie des choses par les sujets parlants.

C'est en ce sens que l'étude précise des phénomènes particuliers d'anaphore peut permettre de tenir des propositions plus générales sur la référence d'un point de vue linguistique, comme le rappellent clairement Apothéloz et Reichler-Béguelin :

La problématique des référents évolutifs ne trouve à nos yeux d'intérêt qu'à condition d'être replacée au sein de celle, plus générale, de l'évolution de la référence et de la catégorisation ; il s'agit alors d'envisager la globalité des paramètres qui conditionnent la gestion de l'acte référentiel par un sujet plongé dans une situation de communication concrète.

Apothéloz et Reichler-Béguelin, 1995, p. 266.

Il va de soi que les distinctions à faire (les différents types de référent) et que les questions à sérier (relatives à l'évolution de la référence) ne peuvent s'exprimer que dans un cadre philosophique non réaliste : est-ce pour autant un cadre d'analyse résolument nominaliste ? Les propos de Milner sont sur ce point plus que nuancés⁵³ ; ceux de Apothéloz et Reichler-Béguelin ne sont pas dénués

⁵². Par exemple, Apothéloz et Reichler-Béguelin 1995, p. 243-244 : « Les virtualités innovantes de l'acte référentiel transparaissent bien dans l'emploi des anaphores lexicales, domaine où la latitude de choix dans les moyens linguistiques utilisés est accrue par le fait même que l'objet désigné est déjà identifié et en général dénommé dans le modèle du monde construit par le discours ».

⁵³. La question est abordée entre autres dans Milner 1989, principalement aux pages 67-68, 138-139, 168-169.

d'ambiguïté⁵⁴.

Ne pouvant entrer dans la délicatesse de ces débats, nous considérerons que l'étude linguistique de la référence s'énonce dans un cadre philosophique non réaliste (au sens « médiéval » du terme, tel qu'abordé au § I.1 de ce chapitre).

II.1.2- Distinguer les objets : les dimensions du référent

Que faut-il entendre par « référent », par « objet extra-linguistique » ?

En première approximation, nous avons proposé de comprendre la référence comme « la propriété d'un signe linguistique lui permettant de renvoyer à un objet du monde extra-linguistique, réel ou imaginaire »⁵⁵. Cette définition présente l'inconvénient de pousser à identifier l'opposition domaine linguistique / domaine extra-linguistique à l'opposition mot / chose ; Berrendonner appelle à plus de nuance :

On peut considérer le monde extra-linguistique non plus comme un référent absolu mais comme le lieu de la manifestation du sensible, susceptible de devenir la manifestation du sens humain, c'est-à-dire de la signification pour l'homme, et traiter en somme le référent comme un ensemble de systèmes sémiotiques plus ou moins implicites.

Berrendonner, 1978, p. 21.

En effet, il est tout à fait possible de postuler la référence, c'est-à-dire postuler que la langue parle DE (quelque chose), sans doute même doit-on le faire⁵⁶, sans pour autant postuler l'existence, ou encore la réalité ontologique, de ce dont on parle.

Il faut pour cela établir les différentes dimensions du référent, selon le point de vue de l'analyse considéré. Notamment, il importe de distinguer référents mondains d'une part et référents discursifs d'autre part⁵⁷ :

- la notion de référent mondain correspond aux choses extra-linguistiques de la réalité mondaine⁵⁸ ;

54. « Cette option théorique [non réaliste] ne signifie pas bien entendu que, pour nous, les échanges langagiers se dérouleraient uniquement au plan d'une sémiologie déconnectée de la réalité, où la notion de référent se trouverait purement et simplement évacuée ou - ce qui revient au même - identifiée au signifié linguistique. D'une part, nous pensons que l'identité des objets-de-discours intègre forcément certains paramètres référentiels (au sens extensionnel du terme) ; d'autre part, il est bien entendu que l'interprétation des expressions référentielles sollicite constamment notre connaissance et notre expérience des propriétés du monde " réel " ». Apothéloz et Reichler-Béguelin 1995, p. 240.

55. Définition reprise du Dictionnaire de linguistique et des sciences du langage 1994, p. 404.

56. Voir Bonhomme 1987, p. 31 : sans la référence, « le langage se dissout dans le non-sens car, comme le dit excellemment Benveniste, parler revient toujours à parler DE. C'est précisément dans ce DE que se déploie la référence qui procure au langage des objets (au sens logique) sur lesquels il fonctionne et qui établit par la suite de nouveaux objets de discours ».

57. Apothéloz et Reichler-Béguelin 1995, p. 239.

58. *Id.*

- la notion de référent discursif capte la notion de « représentation » que les sujets parlants se font des choses mondaines *via* l'activité langagière. Il est d'usage, notamment pour marquer la distinction avec la notion classique de référent, de préférer alors le terme « objet-de-discours » : cette notion empruntée au logicien Grize peut être entendue comme une « représentation alimentée par l'activité langagière »⁵⁹.

Reste qu'existe nécessairement un lien entre un référent mondain et les représentations dont il est l'objet. Cependant ce lien n'est pas évident à capter, surtout au niveau d'où se place le linguiste. En effet, dès qu'ils entrent dans un discours, les référents (ou encore les *realia*) deviennent nécessairement des objets de discours :

C'est qu'une fois promu au statut d'objets-de-discours, ou assimilés à une quelconque pratique sociale, l'identité [des] *realia* devient le produit d'une interaction entre le sujet humain et son environnement. On ne peut plus dès lors se contenter de parler d'eux uniquement comme des *référents* au sens mondain du terme, dans la mesure où ces objets ont acquis le statut de construits culturels, et où par conséquent leur « essence » comporte forcément un paramètre anthropologique. Apothéloz et Reichler-Béguelin, 1995, p. 239.

Ainsi l'objet d'analyse du linguiste ne peut être que l'objet de discours, sans quoi il cesse de faire de la linguistique et « court le risque de s'égarer dans une recherche sans fin sur ce que sont ou pas les essences des *realia* ». Cette recherche relèverait typiquement du champ philosophique :

Quant à savoir quels attributs sont constitutifs de l'identité profonde des réalités désignées, lesquels peuvent être modifiés ou supprimés sans que cette identité soit atteinte voire détruite, lesquels relèvent d'une identité « qualitative », « individuelle » ou « sortale ». [...] Il s'agit là de problèmes philosophiques. Nous nous estimons quant à nous incompetents pour spéculer sur l'"essence" des *realia* susceptibles d'entrer à titre d'objets dans les pratiques langagières. Apothéloz et Reichler-Béguelin, 1995, p. 239.

Dans ce cadre d'analyse de la référence ainsi comprise, les objets-de-discours, dont on fait l'hypothèse qu'ils entretiennent un lien avec les référents mondains, ne pré-existent pas au discours : créés uniquement par le discours, ils n'ont à ce titre aucune stabilité référentielle (de désignation) ; l'instabilité est proprement constitutive de ce type d'objet :

Les dits objets-de-discours ne préexistent pas « naturellement » à l'activité cognitive et interactive des sujets parlants, mais doivent être conçus comme produits - fondamentalement culturels - de cette activité. Apothéloz et Reichler-Béguelin 1995, p. 229.

C'est ainsi que ce modèle d'analyse montre que, ce qui évolue, dans les cas

⁵⁹. *Id.*

d'anaphores évolutives, ce n'est pas le référent mondain, c'est l'objet-de-discours, mais c'est là sa propriété même ; autrement dit, la question des anaphores évolutives doit être reformulée :

Force est de constater que le problème des référents évolutifs n'en est pas un : tout objet-de-discours est, par définition, évolutif, car chaque prédication le concernant modifie son statut informationnel en mémoire discursive.
Apothéloz et Reichler-Béguelin, 1995, p. 240.

II.1.3- Sérier les questions : les problématiques linguistiques de la référence

Compte tenu, d'une part, de la distinction établie entre référent mondain et référent discursif et, d'autre part, de la propriété dégagée d'instabilité des objets de discours, comment peut se formuler la problématique de la référence en linguistique ? On dégage ci-dessous quelques-unes des questions que la linguistique se propose de traiter sur la question de la référence.

(i) Concernant l'étude du référent discursif, il ne peut s'agir, d'un point de vue linguistique, que d'une étude de l'évolution de la référence, c'est-à-dire des stratégies que mettent en œuvre les sujets parlants pour faire évoluer les objets de leur discours⁶⁰. Ce type d'étude relève de ce que Rastier a proposé d'appeler la référenciation, entendue comme « le rapport entre le texte et la part non linguistique de la pratique où il est produit et interprété »⁶¹. Ce type d'approche s'oppose aux approches classiques de la référence, comme le signalent Dubois et Mondada :

Au lieu de partir du présupposé d'une segmentation a priori du discours en noms et du monde en entités objectives, et ensuite, de questionner la relation de correspondance entre l'une et l'autre - il nous semble plus productif de questionner les processus de discrétisation eux-mêmes.
Dubois et Mondada, 1995, p. 275.

L'étude du **processus de discrétisation** consiste à analyser la façon dont les sujets parlants construisent leurs objets (objets de discours sur un plan linguistique, catégories sur un plan cognitif⁶²).

(ii) De façon corollaire, une analyse en termes de référenciation engage à une analyse des **processus de stabilisation** obtenus au cours de l'activité des sujets parlants. L'instabilité reste la donnée de départ à partir de laquelle sont étudiés les

⁶⁰. *Ibid.*, p. 264-265.

⁶¹. Rastier 1994, p. 19.

⁶². Dubois et Mondada 1995, p. 276.

effets de stabilité référentielle :

Nous aimerions en outre souligner qu'au lieu de présupposer une stabilité a priori des entités dans le monde et dans la langue, il est possible de reconsidérer la question en partant de l'instabilité constitutive des catégories à la fois cognitives et linguistiques ainsi que leurs processus de stabilisation.
Dubois et Mondada, 1995, p. 275.

L'étude des effets de stabilisation se place, non plus du côté du sujet parlant qui produit et contrôle son discours, mais du côté de l'interlocuteur ou du lecteur qui reçoit, interprète le discours⁶³.

L'approche linguistique de la référence, qui situe généralement son point de départ dans l'étude des anaphores, propose une méthode d'analyse qui distingue les dimensions du référent, et notamment les dimensions mondaine et discursive, et qui circonscrit les marges du problème : du point de vue de la production, il s'agit d'étudier la construction de la référence, les modalités de discrimination du monde ; du point de vue de la réception, il s'agit d'étudier les effets de stabilisation obtenus par l'activité des sujets parlants au sein d'une pratique.

L'approche linguistique de la référence accorde donc une place centrale à l'acte de référence, acte à partir duquel sont étudiées les propriétés référentielles proprement dites des expressions linguistiques⁶⁴. Sur ce dernier point, la question centrale pour le linguiste consiste à déterminer sur quelle base se construit la référence, l'objet de discours. La problématique est alors la suivante : si la référence se construit sur une base lexicale, comment rendre compatible la stabilité de la signification lexicale avec l'instabilité des objets de discours qu'elle permet de construire ?

II.2- La question du rapport entre sens et référence

La question du rapport entre sens et référence concerne non pas l'acte de référence lui-même mais les expressions linguistiques susceptibles de contribuer à la réalisation de cet acte. Cette question se situe donc en amont de l'acte référentiel et par conséquent en amont de la constitution des objets de discours proprement dits.

Après avoir présenté les enjeux que sous-tend la question du rapport entre sens et référence, on propose, en reprenant les arguments de Milner, de considérer le

⁶³. Apothéloz et Reichler-Béguelin 1995, p. 266-267 : « Mais il faut bien voir que l'effet de coréférence [l'un des aspects de la stabilité référentielle] résulte alors davantage des investissements interprétatifs du décodeur, que des transformations subies ou non subies, concrètement ou sémiotiquement, par le référent discursif ».

⁶⁴. Comme le souligne Kleiber [1981, p. 13 par exemple], il importe de distinguer, dans la construction de la référence, l'acte de référence lui-même et les moyens par lesquels il s'effectue.

groupe nominal comme l'atome référentiel minimal. Pour finir, on donne deux exemples d'analyse linguistique qui montrent comment la référence se construit sur la base de la signification lexicale.

II.2.1- Enjeux du rapport entre sens et référence

L'enjeu du rapport entre sens et référence s'inscrit dans le cadre des problématiques suivantes :

- (i) quelles sont les expressions linguistiques dotées de pouvoir référentiel ? De façon générale, les linguistiques s'accordent pour voir dans le groupe nominal l'atome référentiel minimal ;
- (ii) si la construction de la référence s'effectue en discours, est-elle entièrement dépendante du discours, c'est-à-dire complètement déconnectée de la langue ? Sur ce point, il n'y a pas de consensus sur ce qui déclenche la construction référentielle.

On peut en effet dégager deux principales positions :

- (i) pour les uns, ce sont les pratiques sociales qui sont à l'origine de la création de la référence. C'est une position défendue par exemple par Rastier⁶⁵. De façon moins nette, Reichler-Béguelin et Apothéloz soutiennent eux aussi cette hypothèse⁶⁶ ;
- (ii) pour les autres, comme Milner, dans la mesure où n'importe quel mot ne permet pas de construire n'importe quel objet de discours⁶⁷, on est obligé de postuler que « quelque chose » guide la construction du référent ; on fait l'hypothèse que ce « quelque chose » est la signification lexicale. C'est dans ce cadre que se pose la question du rapport entre sens et référence et l'essentiel de ses problématiques. En effet, compte tenu d'une part de la stabilité de la signification lexicale et d'autre part de l'instabilité des objets de discours, on doit supposer que le sens des unités lexicales est sous-déterminé par rapport à la référence, autrement dit que la signification d'une unité lexicale n'épuise pas ses possibilités interprétatives et référentielles⁶⁸.

La question du rapport entre sens et référence devenant celle d'un rapport de sous-détermination dessine du même coup un espace de

⁶⁵. Rastier [1994].

⁶⁶. Apothéloz et Reichler-Béguelin 1995, p. 266.

⁶⁷. Ou alors c'est la langue elle-même qui est niée : « De même qu'un Œdipe libre d'épouser sa mère, une langue où tout pourrait se dire est une contradiction dans les termes », Milner 1978, p. 10.

⁶⁸. Marandin 1992a, et précédemment chapitre II § II.2.1.

« jeu » interprétatif, espace dans lequel peuvent intervenir les pratiques sociales.

La construction de la référence établie sur la base de la signification lexicale nous paraît en cela constituer un modèle pertinent permettant de penser à la fois le rôle discriminant des expressions linguistiques dans le découpage référentiel (n'importe quelle expression linguistique ne construit pas n'importe quel référent) et le rôle discriminant des pratiques sociales dans la fixation de la référence proprement dite, les pratiques sociales intervenant mais à un autre niveau et de façon seconde.

II.2.2- L'atome référentiel minimal : le groupe nominal

De façon générale, les linguistes s'accordent à reconnaître que seuls les groupes nominaux⁶⁹ (terme abrégé par GN désormais) réfèrent. Chaque école linguistique propose une approche particulière du GN : nous en aborderons deux dans le chapitre V, lorsqu'il apparaîtra nécessaire d'examiner les propriétés que l'on peut attribuer au GN.

À ce stade de l'exposé, on ne définira pas le GN à proprement parler. On en décrira simplement l'apparence⁷⁰ en disant qu'un GN, c'est une séquence dont la forme minimale est : SPEC (N), où SPEC correspond à un spécifieur (qui peut être de type article, démonstratif ou quantifieur) et où N est une unité lexicale de catégorie nominale qui, en position de « tête » dans le groupe, lui donne son nom : groupe nominal. L'unité lexicale de catégorie N peut se voir complétée par des groupes adjectivaux, des groupes prépositionnels et des propositions (relatives et complétives). Seront donc identifiées comme des GN les séquences suivantes : « cet homme, le grand homme, un homme vert, l'homme qui rit, trois hommes aux pistolets d'or, un grand homme vert qui rit, etc. ».

Le GN est un individu linguistique qui n'existe qu'en discours (il est construit par la syntaxe) et s'oppose en cela à l'unité lexicale de catégorie N (comme « homme ») qui, elle, appartient au lexique. La question du rapport sens et référence se pose donc à travers les relations entre une unité lexicale de catégorie N et le GN dont elle constitue la « tête » : la première de ces unités est dotée d'une signification lexicale (nécessairement stable), le second constitue un objet de discours (nécessairement instable).

⁶⁹. Ou plutôt, s'il existe, le statut référentiel d'autres séquences, comme les séquences verbales, n'est pas de même nature que celui des séquences nominales, Milner 1976, p. 63.

⁷⁰. En reprenant la formalisation traditionnelle en linguistique issue de la théorie X-barre, établie par Chomsky. On ne présente pas ici les discussions dont a fait l'objet ce formalisme.

Pour capter le rapport qu'entretiennent N et GN du point de vue de la référence, nous retiendrons, malgré les critiques dont elle peut faire l'objet⁷¹, la distinction qu'établit Milner entre « référence actuelle » et « référence virtuelle » dans la mesure où ce modèle met l'accent sur ce qui rapproche et maintient distincts, du point de vue de la référence, signification lexicale et référent discursif.

Dans le cadre du modèle de Milner, une unité lexicale hors emploi de catégorie N, par exemple « homme », n'a qu'une « référence virtuelle », qui est sa signification. Un GN, une unité de discours, a, lui, une « référence actuelle », propre à construire des référents discursifs (par exemple, « un homme passe », « tout homme est mortel », etc.). La référence virtuelle, le sens d'une unité lexicale, « pèse » sur la construction du référent discursif⁷² : c'est sur ce point que s'articulent langue et discours. En effet, ce n'est pas parce que la référence se construit en discours qu'elle est nécessaire déclenchée par le discours : on devrait sinon pouvoir construire n'importe quel objet de discours à partir de n'importe quel mot.

Observant au contraire que « n'importe quelle séquence nominale n'est pas associée à n'importe quel segment » de réalité, Milner⁷³ relie la question du rapport sens/référence à celle de la propriété de discrimination des unités lexicales⁷⁴ : les unités lexicales se distingueraient entre elles notamment en fonction du type de « segment de réalité » qu'elles pourraient désigner. Toutes les unités lexicales ne sont donc pas équivalentes du point de vue du référent qu'elles permettent de construire. Le mécanisme est décrit ainsi par Milner :

Une unité lexicale étant choisie, certains segments sont d'emblée éliminés en tant que références possibles ; en ce sens, à chaque unité lexicale individuelle, est attaché un ensemble de conditions que doit satisfaire un segment de réalité pour pouvoir être la référence où interviendrait cruciallement l'unité lexicale en cause. Cet ensemble de conditions décrit donc un *type* (ou si l'on veut une *classe*) de référence possible ; il est distinct des segments de réalité, mais pèse sur eux. Pour exprimer cette situation, on pourrait recourir aux termes suivants : le segment de réalité associé à une séquence est sa *référence actuelle* ; l'ensemble de conditions caractérisant une unité lexicale est sa *référence virtuelle*.

Milner, 1976, p. 64.

Notons qu'en dépit de formulations (telles que « segment de réalité ») qui ont pu faire dire que Milner, malgré ses ambitions, n'échappait pas au « problème ontologique des universaux », cette proposition en termes de références actuelle et

⁷¹. Par exemple, Kleiber 1981, Tyvaert 1995.

⁷². Dans le cadre d'un autre modèle, Berrendonner a pu soutenir de semblables propositions : « Grâce aux référents [entendus ici comme « syntagmes nominaux définis »], la langue contient la mention de certaines de ses conditions d'emploi », Berrendonner 1978, p. 47.

⁷³. Milner 1976, p. 63.

⁷⁴. Milner parlera plus tard [1989] de « facteurs individuation lexicale » et proposera un modèle de l'individuation lexicale à trois composantes, Milner 1989, p. 324 et suiv.

virtuelle reste dans un cadre non réaliste. En effet, la notion de « classe » ou de « type » est chez Milner de nature artefactuelle ; une classe n'est pas constituée à partir des éléments du « réel », seul le discours peut l'actualiser, et ce n'est plus alors le même type d'unité qui est mis en cause :

Une unité lexicale ne peut avoir de référence actuelle que si elle est employée [...]. Mais, d'autre part, si l'on considère les emplois eux-mêmes, ce ne sont pas aux unités lexicales comme telles que sont associés des segments de réalité, mais bien aux **groupes nominaux** pris dans leur ensemble.
Milner, 1976, p. 64 (c'est nous qui soulignons).

L'« atome référentiel, c'est donc le groupe nominal » et non le nom, le N, qui en constitue la tête.

Milner propose d'observer la différence entre référence actuelle et référence virtuelle à partir de différents types de reprises anaphoriques. Quand l'anaphore reprend un N (une référence virtuelle), l'anaphorique ne construit pas les mêmes objets de discours que ceux constitués par la source de l'anaphore ((1) ci-dessous) ; quand l'anaphorique reprend un GN (une référence actuelle), c'est le même référent discursif qui est visé (2) :

(1) J'ai vu dix lions et toi tu en as vu quinze
« en » reprend ici un N « lions », c'est-à-dire une référence virtuelle : il ne s'agit pas des mêmes référents visés (il ne s'agit pas des mêmes lions) ;
(2) J'ai capturé dix des lions et toi tu en a capturé quinze
« en » reprend ici un GN « les lions », c'est-à-dire une référence actuelle : il s'agit des mêmes référents (c'est le même ensemble de lions qui est visé).
Milner, 1976, p. 64.

C'est cet effort d'articulation et de distinction entre construction du sens et construction de la référence qui caractérise, nous semble-t-il, l'apport de Milner. En effet, comme le signale d'ailleurs Milner lui-même⁷⁵, la dualité qu'il propose s'apparente à celle précédemment proposée par Frege. Mais si Frege distingue d'une part la « Bedeutung » (dénotation) et le « Sinn » (sens), c'est pour les opposer⁷⁶, alors que Milner s'attache à montrer comment les deux se déterminent l'un par rapport à l'autre ; c'est pourquoi, précise-t-il, « il semble préférable d'utiliser une terminologie qui ne dissimule en rien l'articulation »⁷⁷.

L'autre apport de Milner consiste à spécifier la « plasticité »⁷⁸ particulière des GN en discours, qui leur permet de désigner différents types d'objet :

⁷⁵. Milner 1989, p. 341.

⁷⁶. C'est l'exemple bien connu des deux expressions « étoile du matin » et « étoile du soir » qui, tout en ayant un sens différent, désignent le même référent. Frege 1971 [1892], p. 102-126.

⁷⁷. Milner 1978, p. 26.

⁷⁸. Nous revenons dans le chapitre V sur cette notion de plasticité.

Bien qu'un nom ordinaire puisse désigner des individus totalement distincts suivant les énoncés, il reste toujours possible de définir de manière générale la classe des êtres dont ce nom est la désignation et inversement d'exclure *a priori* des êtres qui ne pourront jamais être désignés par lui.
Milner 1978, p. 198-199.

Milner pose ce faisant la possibilité de définir la référence d'un nom hors contexte : c'est ce qu'il nomme son « autonomie référentielle », qui correspond approximativement à la notion logique de classe⁷⁹. Son propos vise ici à distinguer la relation référentielle établie par le nom et celle établie par le pronom, soit les interprétations de (4) et de (5) ci-dessous :

- (4) le livre est beau
- (5) il est beau

Dans (4), l'interprétation est possible « même si l'on ignore de quel livre il s'agit. [...] Inversement, l'énoncé (5) n'est, dans les mêmes conditions, absolument pas interprétable »⁸⁰. La différence tient à ce qu'en (4), la référence virtuelle de *livre* est déterminée hors emploi : à ce titre *livre* est dit référentiellement autonome⁸¹. L'unité *livre* correspond à une classe qui circonscrit le type d'objet singulier auquel il est possible de référer en discours. Par opposition, « sera donc non autonome une unité dont la référence virtuelle ne peut être définie sans mentionner l'unité elle-même en tant qu'elle est énoncée dans un énoncé singulier »⁸². Le cas typique des unités non autonomes est représenté par les pronoms. Ainsi la référence virtuelle du pronom *je* ne peut être autonomisée que par rapport à son emploi dans un énoncé, elle est suspendue à l'énonciation ; il y a, dans ce cas, circularité de la référence virtuelle : la propriété définissante du référent est elle-même suspendue à l'énonciation⁸³.

Il importe de noter que l'autonomie référentielle dont parle ici Milner porte sur **le rapport entre l'unité lexicale et l'énoncé où elle est insérée**, et non sur le rapport entre l'unité et le référent qu'elle désigne : on ne sort donc pas du cadre non réaliste et cette hypothèse ne contredit pas celle « de la construction des objets de discours »⁸⁴.

79. Voir aussi sur ce point le chapitre V.

80. Milner 1978, p. 199.

81. Une unité est dite référentiellement autonome quand « les conditions de possibilité de désignation sont indépendantes de l'énoncé particulier où l'unité est employée : l'unité ne tire sa capacité référentielle que d'elle-même », *Ibid.*, p. 333.

82. Milner 1976, p. 65.

83. Milner 1978, p. 333-334.

84. Kleiber reconnaît sur ce point que la solution de Milner reste satisfaisante dans la mesure où elle « montre qu'un item lexical, tout en n'entretenant pas de relation directe avec des êtres ou objets précis de la réalité, est malgré tout en "prise" avec la référence par le biais de conditions d'application référentielle. En second lieu, elle ne nécessite nul engagement ontologique, puisqu'en parlant de "segments de réalité", elle ne place pas en première ligne la question de l'existence des

Les propositions de Milner articulant sens et référence par le biais des notions de référence virtuelle et référence actuelle présentent l'avantage, nous semble-t-il, de pouvoir circonscrire le type d'unité linguistique minimale permettant de référer⁸⁵ : c'est une molécule syntaxique de type GN, construite à partir d'un atome syntaxique de type N.

Une fois ces distinctions établies entre unité lexicale et groupe nominal, comment peut-on penser, de façon plus précise, l'inscription de la référence dans une unité lexicale ? En quoi peut consister la référence virtuelle d'une unité lexicale hors emploi ?

Comme précédemment dans le chapitre II, nous emprunterons à D. Corbin⁸⁶ les éléments de sa démonstration concernant le cas particulier des mots construits⁸⁷ ; ceux-ci présentent en effet l'avantage de faire « voir » le rapport entre sens et référence :

Les sens des mots construits sont des exemples privilégiés pour observer les relations entre le sens linguistique et les catégories référentielles : dans la mesure où ces mots sont construits par des opérations linguistiques, leur sens peut être calculé de façon proprement linguistique, indépendamment des catégories que ces mots dénomment.

Corbin et Temple, 1994, p. 6.

II.2.3- Construction de la référence sur la base de la signification lexicale

Rappelons que la problématique du rapport sens/référence repose sur deux principales questions :

- comment expliquer que, pour une forme lexicale donnée, il y ait stabilité de la signification lexicale d'une part et instabilité des référents discursifs d'autre part ?
- comment penser un modèle de construction de la référence sur la base de la signification lexicale qui inclut l'activité des sujets

référents », Kleiber 1981, p. 20.

⁸⁵. Dans un autre cadre, qui se fonde sur des arguments de nature logique et de nature linguistique, Michel Le Guern propose une description du groupe nominal qui permet de le faire voir également comme l'unité référentielle minimale ; nous revenons sur cette description dans le chapitre V.

⁸⁶. Bien que Corbin tienne à marquer ses distances avec le modèle proposé par Milner, voir par exemple Corbin et Temple 1994, p. 91, n. 4.

⁸⁷. « La spécificité d'un mot construit par rapport à un mot non construit est que l'interprétation sémantique du premier est compositionnelle par rapport à sa structure interne », Corbin 1990, p. 176. Ainsi si « maisonnette » est un mot construit, « omelette » n'en est pas un (« omel- » ne constitue pas une base suffixable en français).

parlants (les pratiques discursives, les stratégies de désignation, etc.) ?

Le modèle de D. Corbin propose un cadre de réponse à ces deux questions. En effet, la base de son programme de recherche⁸⁸, mené dans le domaine de la morphologie dérivationnelle, consiste à rendre compte à la fois des unités attestées (constituant le « lexique conventionnel », illustré par le dictionnaire) et des unités possibles (constituant le « lexique dérivationnel », produisant des unités dont l'usage ne s'est pas emparé, ou pas encore emparé).

Sa position suppose que c'est la langue qui régit, ou programme, les désignations référentielles, dont certaines peuvent être jugées, pour des raisons sociales ou culturelles, « inacceptables » et donc n'être jamais employées (du moins en synchronie).

Ainsi, dans son modèle :

- la signification lexicale d'une unité peut permettre la construction de référents de nature hétérogène, l'hétérogénéité étant elle-même « programmée » ; autrement dit, l'hypothèse d'un conditionnement linguistique n'oblige en rien à penser un rapport univoque et constant entre sens et référent : c'est l'exemple donné ci-dessous de l'analyse du mot « chinois » ;
- le rôle des pratiques sociales consiste moins, dans ce cadre, à créer la référence qu'à la révéler : cette hypothèse permet donc d'intégrer l'influence des sujets parlants dans la fixation de la référence. Le lien sens/référence est maintenu dans son « indétermination », laissant une latitude de décodage ; c'est l'exemple donné ci-dessous de l'analyse du mot « fenouillette ».

Ce modèle qui défend le rôle du sens lexical dans la construction de la référence n'enlève donc rien à la contingence observable et observée des objets de discours ; il lui donne au contraire un cadre qui la révèle.

On reprend ci-dessous, de façon schématique, les différents niveaux que pose Corbin pour penser l'articulation entre sens et référence (A). On présente ensuite deux exemples d'analyse qui illustrent chacun un point particulier : le premier, l'analyse de « chinois », montre comment le même sens linguistique peut construire des référents distincts (B) ; le second, l'analyse de « fenouillette », met en valeur le rôle des pratiques sociales dans la construction des référents (C).

⁸⁸. Corbin 1987.

A- Distinction des niveaux dans le modèle de Corbin

Notons d'emblée que, si, comme Milner, Corbin considère qu'il y a distorsion entre sens et référence, elle ne postule pas, comme lui, un rapport direct entre les deux. Elle pose en effet un niveau intermédiaire entre ce qu'elle nomme les « catégories sémantiques » (définies en intension) et les « catégories référentielles »⁸⁹ (définies en extension) : c'est le niveau des « catégories pré-référentielles », définies en intension et en extension, qui « représentent le résultat des découpages conceptuels que les propriétés sémantiques permettent d'opérer et de dénommer ». C'est seulement lorsqu'elles sont « adaptées en fonction de notre appréhension pragmatique et culturelle des choses » qu'elles deviennent des catégories référentielles. Autrement dit, le modèle est prêt à accueillir des unités bien construites du point de vue de la langue mais bloquées au niveau pré-référentiel pour des raisons non linguistiques mais culturelles ou sociales.

C'est dans le cadre de ces distinctions que Corbin démontre, de façon parallèle, d'une part l'autonomie du sens lexical par rapport à la référence et d'autre part le conditionnement sémantique de la référence :

Défendre l'idée que les catégories sémantiques ne sont pas nécessairement isomorphes aux catégories référentielles et que ne leur correspondent pas nécessairement des catégories référentielles présuppose qu'il existe des catégories sémantiques associées aux mots, donc un sens lexical, [qui] détermine et conditionne, conjointement à d'autres facteurs, leur interprétation contextuelle.

Corbin et Temple, 1994, p. 7.

B- L'analyse du mot « chinois » : sens unique et multiplicité référentielle

Le nom « chinois » est un mot construit par conversion à partir d'un adjectif⁹⁰ : « ce procédé de construction de mots illustre l'hypothèse qu'à des catégories sémantiques homogènes peuvent correspondre des catégories référentielles conçues comme hétérogènes »⁹¹. On constate en effet que le mot « chinois » renvoie à des catégories référentielles différentes : personne ; langue ; petite orange amère ; passoire.

Corbin et Temple proposent de voir dans le cas des noms construits par conversion à partir d'adjectifs des cas de polyréférence, à rapporter non à des sens différents mais à des sens uniques, à des principes organisateurs uniques subsumant les

⁸⁹. Corbin et Temple sont plus précises : les catégories référentielles « forment un sous-ensemble des catégories conceptuelles dénommables par les unités lexicales, c'est-à-dire celles qui sont nommables de façon fixe et codée par une expression linguistique », Corbin et Temple 1994, p. 6.

⁹⁰. Corbin et Temple 1994, p. 12-13, qui analysent, dans les mêmes termes, les mots « simple » et « bleu ».

⁹¹. *Ibid.*, p. 13.

différentes catégories référentielles⁹².

Le point de la démonstration constitue un article en soi : nous ne rentrerons pas dans le détail de l'analyse qui nécessiterait en outre d'exposer l'appareillage descriptif qui constitue le modèle d'analyse de Corbin [1987]. De façon schématique, on peut dire que le sens unique des noms construits par conversion d'adjectifs repose sur le sens de la base lexicale à l'origine de la conversion : la propriété sémantique retenue de l'adjectif au terme de la conversion constitue un « type », une propriété saillante, qui doit servir de propriété dénominative aux entités susceptibles d'être dénommées par le mot construit par conversion de l'adjectif.

C'est sur la base d'une propriété saillante de l'adjectif que le mot construit peut renvoyer à des catégories référentielles hétérogènes.

C- L'analyse du mot « fenouillette » : intervention des pratiques sociales dans la fixation de la référence

Cet exemple illustre le cas où le sens d'un mot construit (les catégories sémantiques qu'il construit) ne peut pas servir de support à la dénomination de catégories référentielles, ceci pour des raisons sociales et culturelles ; il faut, pour qu'elle puisse être employée, qu'une dénomination respecte nos modes d'organisation conceptuels.

La démonstration de ce point étant assez complexe, nous n'en reprenons que les grands traits. Corbin et Temple étudient le cas de « fenouillette »⁹³, dont la définition usuelle est : « petite pomme grise dont le parfum rappelle celui du fenouil ». Il s'agit d'expliquer le rapport entre la base du mot construit (fenouil) et le référent désigné (pomme), rapport visiblement incongru puisqu'« il ne semble pas possible en français de dénommer un végétal comestible par le nom d'un autre végétal comestible », à moins de construire une expansion, comme on le voit dans les unités de type « poire *d'avocat* » (fruit de l'avocatier) ou « poire *de terre* » (nom régional du topinambour). Cette impossibilité est à l'origine du report de la signification de « fenouillette » sur l'« odeur de fenouil » : c'est une propriété stéréotypique de « fenouil » qui est alors sélectionnée. Or la catégorie référentielle des « odeurs » n'existe pas en français : « on ne trouve pas dans le lexique français de dénominations de catégories rassemblant différents objets ayant le même parfum ». Dès lors, pour expliquer l'acception usuelle de « fenouillette », il faut

⁹². *Id.* Il y a, dans ces cas, une catégorie préréférentielle unique et des catégories référentielles différentes : « Les catégories préréférentielles correspondant aux catégories sémantiques définies par les sens des mots construits ne sont pas isomorphes aux catégories référentielles dénommées par les mots ».

⁹³. *Ibid.*, p. 17-24.

supposer qu'à une catégorie sémantique (ici « odeur ») ne correspond pas nécessairement une catégorie référentielle du même type :

Il existe des catégories sémantiques dont certaines propriétés ne peuvent servir de support à la dénomination de concepts. [...] La base *fenouil* sur laquelle s'applique *-ette* pour construire *fenouillette* n'est pas un mot, c'est une forme donnant corps à l'une des propriétés sémantiques du sens métaphorique de fenouil - celle qui renvoie au parfum -, propriété qui ne peut pas servir à organiser une catégorie référentielle. Corbin et Temple, 1994, p. 24.

L'une des conclusions que permet de formuler ce traitement de « fenouillette » est que « bien que les mots aient un sens et que ce sens permette de référer, les mots ne sont pas directement des dénominations »⁹⁴ : ils ne le sont que dans la mesure où ils respectent nos modes d'organisation du monde.

Les exemples d'analyse de « chinois » et de « fenouillette » proposés par Corbin montrent que la construction de la référence est toujours doublement configurée : et par le sens lexical et par les contingences discursives. Compte tenu de cette dualité, le sens est nécessairement sous-déterminé par rapport à la référence⁹⁵. C'est cette sous-détermination qui constitue l'espace de « jeu », interprétatif et désignatif, qui revient au sujet parlant. La notion de sous-détermination montre que l'hypothèse de l'instabilité des objets de discours peut être compatible avec celle de la contrainte sémantique sur la construction de la référence.

Chacun dans leur approche, Milner et Corbin s'attachent à faire apparaître à la fois une disjonction nette entre sens et référence et une contrainte sémantique sur la construction de la référence. En ce sens, le rapport entre sens et référence se donne comme un rapport de sous-détermination : la langue et la référence n'étant pas coïncidentes, la construction de la référence n'est qu'en partie réalisée par la langue.

II.3- Conclusions et mise en perspective

Dans le paragraphe I de ce chapitre, nous avons tenté de dégager, au-delà des présupposés réalistes du discours classique, les zones de manifestation de la référence en indexation : à ce titre, le document, en tant qu'élément constitutif de l'« extérieur » de l'indexation, nous est apparu comme un objet construit, distinct de la « réalité éditoriale ». Parallèlement, nous nous étions interrogée sur la stabilité référentielle attachée, dans le discours classique, aux unités d'indexation :

⁹⁴ *Ibid.*, p. 25.

⁹⁵ *Ibid.*, p. 17.

hors du modèle réaliste, ce type de stabilité référentielle nous a paru relever davantage d'un principe de réception que d'un principe de production.

Afin de pouvoir clarifier ces aspects de la référence en indexation, nous avons cherché à nous dégager du modèle réaliste. Nous avons alors sollicité le modèle de la référence sur lequel s'appuie l'approche linguistique. Dans ce cadre, est apparue une série de distinctions :

- (i) entre éléments de la référence : l'acte de référence doit être distingué des propriétés référentielles des expressions linguistiques ;
- (ii) entre types de référent : l'approche linguistique établit une distinction entre référent mondain et référent discursif ;
- (iii) entre expressions linguistiques : seul le GN est susceptible de construire un objet de discours, mais cette construction passe par l'exploitation de la signification de l'unité lexicale de catégorie N qui en constitue la tête ;
- (iv) entre sens et référence : tout en restant distincts, sens et référence établissent entre eux un rapport de sous-détermination ;
- (v) entre langue et usage : la référence, potentiellement déterminée en langue, se fixe par l'usage, par l'intervention des sujets parlants.

L'établissement de ces distinctions, propres à une approche non réaliste de la référence, permet, « rétroactivement », de préciser ce qui, dans le discours classique sur l'indexation, relève du modèle réaliste et ce qui, par conséquent, ne pourra apparaître dans le cadre du nouveau modèle de la référence en indexation proposé ci-après.

Ainsi, la distinction (iv) contrecarre l'hypothèse du discours classique concernant la correspondance entre stabilité référentielle et stabilité linguistique en indexation : ce n'est pas parce que l'on emploie le même mot que l'on désigne la même chose, même si ce mot est un « descripteur ». L'approche linguistique de la référence montre en effet qu'un accord de désignation ne suppose pas un « accord » de signification⁹⁶ : le sens d'une unité lexicale s'établit bien en deçà de l'intervention des sujets parlants et reste sous-déterminé.

Si tous les mots ne découpent pas la réalité de la même façon, tous les mots ne se valent pas. À ce titre, la notion du descripteur comme « terme préférentiel » (comme terme retenu de préférence à d'autres tenus pour équivalents) devient problématique. Nous sommes amenée à nous interroger sur l'aspect de la sélection

⁹⁶. Tyvaert 1994, p. 48.

des termes en indexation : si un objet de discours n'est construit qu'au travers de saisies multiples opérées grâce aux sens différents des unités lexicales, l'indexation doit-elle restreindre le nombre des accès lexicaux aux documents, c'est-à-dire restreindre les possibilités de saisie multiple ?

La distinction (iii) souligne la différence entre N et GN du point de vue de leurs propriétés référentielles : un N, une unité lexicale hors emploi, ne peut que désigner une classe de référents possibles ; un GN, l'unité lexicale en discours, désigne, lui, un élément singulier de cette classe. C'est lui qui peut établir une relation référentielle stable ; or les descripteurs se donnent généralement sous la forme linguistique de N.

Le point (v) fait apparaître la dimension du discours en indexation : elle est complètement absente des approches normatives, mais se dégage nettement sous l'angle de l'approche linguistique de la référence. En effet, comme le rappellent Apothéloz et Reichler-Béguelin, « le problème du choix des dénominations ne doit pas être pensé dans le rapport entre la langue et le monde mais à l'intérieur du discours »⁹⁷. La reformulation de la référence en indexation impose donc de circonscrire cette dimension discursive en indexation.

Si les points (iii), (iv) et (v), dégagés d'une approche linguistique de la référence, nous engagent à introduire la dimension du discours en indexation et à redéfinir la morphologie du descripteur⁹⁸, les points (i) et (ii) nous permettent d'ores et déjà d'approcher ce que peut être la construction de la référence en indexation.

⁹⁷. Apothéloz et Reichler-Béguelin 1995, p. 266.

⁹⁸. Voir les chapitres IV et V.

III- La construction de la référence en indexation

Dans ce paragraphe, nous tenterons de décrire le processus de l'indexation en l'inscrivant dans le cadre du modèle d'approche linguistique de la référence (approche non réaliste) tel que nous l'avons précédemment décrit.

L'adoption de ce cadre nous conduit à distinguer :

- différents aspects de la référence : l'acte référentiel et les propriétés référentielles des expressions linguistiques utilisées en indexation ;
- différentes dimensions du référent, notamment les dimensions mondaine et discursive.

Pour penser ces différents aspects de la référence, il importe de préciser d'abord l'acte de référence effectué en indexation, duquel dépendent les propriétés référentielles des expressions linguistiques utilisées en indexation.

L'acte de référence est, en indexation, double :

- du côté de l'indexeur, il s'établit de la réalité éditoriale à la réalité documentaire. Le monde de référence des indexeurs n'est pas, pour reprendre les termes de Ricœur⁹⁹, le « monde ambiant » mais le « quasi-monde des textes ». L'acte de référence consiste alors en un **acte de discrétisation** sur ce monde des textes, qui permet de construire un univers proprement documentaire : l'univers des documents. Cet acte de discrétisation, s'il peut se traduire en « mots », par l'assignation de descripteurs, ne suppose pas en soi une nomination à proprement parler : la dénomination peut n'intervenir qu'après, si l'on veut indiquer expressément la base sur laquelle des documents ont été rapprochés. À ce niveau, l'indexeur cherche à donner accès à des **classes de référents discursifs** et non à des objets de discours particuliers à ces classes. Il utilisera pour ce faire, des unités lexicales hors emploi, pourvue d'une référence virtuelle ;
- du côté des utilisateurs, l'acte de référence s'établit de la réalité documentaire à la réalité éditoriale, sur la base de leurs représentations de la réalité mondaine : ce sont elles qui guident la recherche de l'utilisateur à travers l'univers documentaire. L'acte de référence consiste ici en un **acte de stabilisation du monde de référence** : il s'agit de construire, à travers l'exploration de l'espace documentaire,

⁹⁹. Ricœur 1986, p. 141.

des objets de discours perçus comme stables (*i.e.* « parlant de la même chose »). Les référents discursifs n'intéressent les utilisateurs que par les représentations des référents mondains auxquels ils renvoient : l'utilisateur cherche à capter des représentations discursives du monde. Pour cela, il opère différentes saisies sur le monde des textes lui permettant de construire de proche en proche ses objets de discours : l'utilisateur traverse les différentes classes de documents constituées par le biais d'un même individu linguistique, **un groupe nominal**, qui peut alors apparaître comme un thème de discours.

La proposition de ce modèle de la référence documentaire nous écarte de façon radicale des approches classiques de l'indexation et du descripteur :

- l'indexation ne peut plus être tenue pour une opération symétrique : les objets construits par les indexeurs (des classes de discours) ne correspondent pas à ceux construits par les utilisateurs (des objets de discours) ;
- le descripteur ne peut donc plus être conçu de façon homogène. La référence du descripteur apparaît en effet sous une double forme : le descripteur des indexeurs pointe sur une source/des sources *via* un document. Le descripteur des utilisateurs pointe sur un objet mondain *via* des objets de discours. Dans les deux cas, la référence d'un descripteur est à chercher, non dans le « mot » lui-même, mais dans les discours, les textes, auxquels il donne accès.

L'approche documentaire de la référence conduit à postuler l'existence d'un lien entre référent mondain et référent discursif, lien qui repose sur une croyance, celle de partager un univers réel commun¹⁰⁰. À ce titre, l'indexation peut être considérée comme l'opération qui crée cet « univers réel commun », ou du moins perçu comme tel. En effet, l'indexation, appréhendée sous l'angle de la référenciation, doit permettre d'établir une correspondance entre les différents espaces de représentations, celui des auteurs, celui des indexeurs et celui des utilisateurs : les sources à partir desquelles les indexeurs construisent des documents constituent des représentations, des saisies de la réalité mondaine ; ces saisies sont ressaisies par les indexeurs qui construisent des documents ; ces documents sont à nouveau

¹⁰⁰. Michel Le Guern (communication personnelle) : « Que la référence porte en fait sur des objets de discours, c'est vrai ; mais l'indexation n'existerait pas s'il n'y avait pas dans l'esprit des indexeurs et des utilisateurs la croyance en un univers commun réel, les objets de cet univers, que ce soient des objets concrets ou des abstractions, étant identifiés de manière plus ou moins floue avec les objets de discours ».

ressais par les utilisateurs qui construisent leurs objets de discours, autres représentations de la réalité mondaine.

L'indexation ne se justifie que dans la mesure où elle peut assurer une certaine continuité entre ces différentes formes de représentations. Nous essayerons de dégager, dans le chapitre IV, des stratégies d'exposition propres à l'indexation permettant de créer un « espace commun » de représentations.

Auparavant, nous nous proposons ci-dessous de préciser comment l'indexation peut se comprendre comme un acte de référencement, c'est-à-dire comme un acte établissant un **double processus de discrétisation et de stabilisation**. C'est sous ces deux rapports que l'on étudiera :

- (i) d'une part, la construction des référents de l'indexation, c'est-à-dire le passage du texte (ou source) au document : il s'agit là d'une étude du processus de discrétisation qui relève de l'acte de référence réalisé par les indexeurs ;
- (ii) d'autre part, la construction de l'effet de stabilité référentielle, c'est-à-dire l'utilisation documentaire du « désignateur rigide ». Il s'agit là d'une étude de la stabilisation référentielle (première approche¹⁰¹), qui porte non plus sur l'acte de référence mais sur les propriétés référentielles des expressions linguistiques utilisées en indexation.

III.1- Construction du document en indexation

Nous souhaitons ici étayer l'hypothèse que, en tant que processus de construction de la référence, l'indexation construit son référent (le document), et intègre, ce faisant, la procédure de sélection documentaire au cœur même de son mécanisme de fonctionnement.

Cette hypothèse a été formulée sur la base des résultats de l'expérimentation de « mise en document » que nous avons menée¹⁰² : objet textuel et objet documentaire ne sont pas apparus, dans tous les cas, isomorphes. Ce décalage peut être appréhendé de deux façons :

- dans les termes de l'approche « instrumentale » : la sélection documentaire n'est que le résultat de contraintes liées aux besoins des utilisateurs ;

¹⁰¹. L'étude sera approfondie dans le chapitre V.

¹⁰². Voir, dans ce chapitre, § I.3.1.

- dans les termes de l'approche « procédurale » : la sélection documentaire se comprend comme un principe de fonctionnement documentaire (du document). La transformation d'un objet en document constitue le propre de l'indexation, entendue comme mise à disposition du savoir (ou plutôt mise à « construction » du savoir). C'est là une piste de recherche proposée par Escarpit¹⁰³ et que nous retenons dans cette étude.

Pour caractériser la construction du document à partir de sources, nous invoquerons des mécanismes de transformation permettant, d'une part, de distinguer ce qui change dans le passage de la source au document (le contexte*, selon notre hypothèse) et, d'autre part, d'indiquer le rôle du document par rapport à la source (rôle d'interprétant, selon nous¹⁰⁴).

Pour étudier ces deux aspects, nous adopterons une approche par point de vue :

- nous étudierons d'abord la construction du document du point de vue de la source (III.1.1) ; pour montrer que c'est le contexte de la source qui est modifié, nous utiliserons des éléments descriptifs issus de la théorie de l'énonciation¹⁰⁵ ;
- nous étudierons ensuite la construction du document vue du côté du document (III.1.2) : pour caractériser le rôle que joue le document par rapport à la source, nous utiliserons, là encore, une théorie, celle de Peirce, et plus particulièrement la notion d'interprétant qu'il propose.

Sans le recours à ces emprunts théoriques, les processus à l'œuvre dans la construction du document en indexation restent peu visibles.

Enfin, en III.1.3, nous proposerons une schématisation du passage de la source au document.

¹⁰³. « Ce qui est certain, c'est que le pas décisif a été franchi lorsque l'homme a institué le *document*, cumulation de traces fixes et permanentes [...] où les réponses données, en feedback, à travers le temps, aux expériences antérieures restent disponibles pour une lecture, c'est-à-dire pour l'exploration libre de toute contrainte événementielle ou chronologique, en fonction du projet et de la stratégie destinée à le réaliser. En d'autres termes : il y a constitution d'un savoir », Escarpit 1991, p. 62-63.

¹⁰⁴. Nous remercions Michel Le Guern pour nous avoir guidée sur cette voie.

¹⁰⁵. Suivant en cela aussi une proposition de Michel Le Guern : qu'il en soit ici remercié. L'énonciation sera ici envisagée comme l'étude des « allusions qu'un énoncé fait à l'énonciation, allusions qui font partie du sens même des énoncés », Ducrot et Schaeffer 1995, p. 603.

III.1.1- La construction du document vue du côté de la source

A- Conjecture : la source vue comme une énonciation

En première approche, on peut définir l'énonciation¹⁰⁶ comme « l'événement historique constitué par le fait qu'un énoncé a été produit »¹⁰⁷ ; en cela, l'énonciation suppose un énonciateur et une situation d'énonciation. L'énoncé, en tant que produit de l'énonciation, porte, en lui, sous forme de « traces », les marques de la « subjectivité » de l'énonciation¹⁰⁸ : c'est à ce titre qu'un énoncé peut être interprétable, même coupé de ses instances de production initiales, mais aussi réinterprétable en fonction des nouvelles situations de discours dans lesquelles il apparaît.

Ce cadre rapidement posé, on propose de voir la source comme une « énonciation »* et le document comme un « énoncé »* dont, par ailleurs, certains traits énonciatifs, tels que l'énonciateur (l'auteur d'un article par exemple) et le contexte de production (le nom du journal d'où est issu un article par exemple), peuvent être notifiés ailleurs dans le système d'information¹⁰⁹.

La « mise en document » aurait donc pour effet d'extraire, en partie, l'énoncé de sa situation d'énonciation, pour le réintroduire dans un autre espace d'interprétation (une collection documentaire). La mise en document opère donc un changement de contexte, qui modifie les caractéristiques de la source : la source devenant autonome par rapport à son utilisation initiale est ainsi susceptible de connaître d'autres usages que ceux pour lesquels elle avait été conçue¹¹⁰.

Escarpit procède à une analyse de la constitution du document dans des termes semblables¹¹¹, mise à part que, dans son cadre, la notion de source est approchée en termes d'« événement », qu'il s'agit de stabiliser par une transformation en document¹¹².

¹⁰⁶. Ici entendue sous l'angle de l'acte d'énonciation, et non plus sous l'angle de la théorie de l'énonciation, comme précédemment.

¹⁰⁷. Ducrot et Schaeffer 1995, p. 603.

¹⁰⁸. Kerbrat-Orecchioni 1980.

¹⁰⁹. Nous reprenons ce point ci-après dans le chapitre IV.

¹¹⁰. Il importe sur ce point de relever qu'il n'existe aucun texte, aucune photographie, aucun disque, etc., spécifiquement dédiés à l'indexation ou à la recherche documentaires : c'est un **usage documentaire** des sources qui est créé. Voir, sur ce point, Varet 1995 : « Le livre devient un document lorsqu'il est invoqué comme texte à l'appui ».

¹¹¹. Escarpit 1991, chap. 8, p. 121-147.

¹¹². D'autres auteurs insistent sur cet aspect du document comme « stable » ; ainsi Turner [1994], par exemple : la caractéristique principale du document est, selon lui, « d'avoir une certaine immuabilité lui permettant de traverser l'espace et le temps ». Dans le cadre de notre approche, cette caractéristique du document peut s'analyser en termes d'**effet stabilisateur** réalisé par la pratique d'indexation. Voir aussi sur ce point Escarpit [1991, p. 125] qui montre qu'un document n'est pas, à proprement parler, stable : « Un objet stable reste emporté par le temps, la stabilité n'étant que le non-changement des relations qui définissent sa configuration caractéristique (son *pattern*) aux yeux d'un observateur emporté par un temps qui est, par convention pratique,

Comme le fait remarquer Escarpit, introduire la dimension de l'énonciation dans l'approche du document permet de mettre l'accent sur la source de l'information¹¹³, et non plus, comme dans les modèles mécanistes, uniquement sur le canal. Sur ce point, Escarpit propose la notion de « semi-document » pour capter les cas où la sélection documentaire n'opère que sur le canal :

En fait, la machine M3¹¹⁴ n'est qu'un simple codeur-décodeur du temps. Elle n'agit ni au niveau de la source ni au niveau du destinataire. Elle agit au niveau du canal. Elle se contente de coder en synchronie documentaire la diachronie événementielle. Ce qu'elle restitue est simplement une image répétitive (sonore ou visuelle) de l'événement, soumise à la loi de dégradation de l'information. [...] Ce n'est donc qu'un semi-document.
Escarpit, 1991, p. 126.

Avec la mise en valeur de cette notion de « semi-document » apparaît en propre la spécificité du document : le document est ce qui permet de lire une source non plus comme « la réactivation de l'événement »¹¹⁵ (niveau du canal, codage/décodage) mais comme la « production d'une information nouvelle »¹¹⁶. La source est **disponible** pour de nouvelles lectures et de nouveaux usages¹¹⁷.

Si l'on définit une source documentaire comme une énonciation susceptible d'être décontextualisée (transformée en énoncé), tous les objets du « quasi monde des textes » constituent-ils, au même titre, des sources documentaires ? De quels types de propriété doit être dotée une source documentaire ?

B- Propriétés de la source

Pour qu'elle puisse être transformée en document, une source doit, nous semble-t-il, être dotée de deux propriétés :

- (i) une propriété d'autonomie par rapport à son contexte de production ;
- (ii) une propriété d'usage : une source doit pouvoir être détournée de son usage initial.

supposé le même ».

¹¹³. Voir le schéma de la communication présenté dans le chapitre I, § I.2.

¹¹⁴. La machine M3 correspond, dans la typologie d'Escarpit, à une « machine à mémoire » de type magnétophone ; dans le contexte, elle s'oppose à une machine de type M6, dite « machine à langage ». Escarpit 1991, p. 109.

¹¹⁵. *Ibid.*, p. 126.

¹¹⁶. *Id.*

¹¹⁷. *Id.* : « La *disponibilité* pour un balayage volontaire et non forcément synchrone, non plus au niveau du canal (codage-décodage) mais au niveau de la source et du destinataire qui en ont chacun de son côté l'initiative, est donc ce qui caractérise la document ».

(i) Propriété d'autonomie de la source

Il est constant de remarquer que « l'information documentaire ne représente qu'une partie de l'information en circulation, la partie qui peut être détachée de ses contraintes originelles »¹¹⁸.

L'information en circulation se distingue, selon nous, de l'information documentaire sur deux points : elle n'est autonome ni d'un point de vue physique (pas de support) ni d'un point de vue « logique » (pas de contexte de production explicite : on pourrait dire qu'en cela l'information en circulation reste « suspendue » à l'acte d'énonciation et qu'elle n'existe pas en dehors d'elle).

Escarpit insiste surtout, dans l'approche du document qu'il propose, sur l'autonomie physique propre au document et qui lui vient du fait qu'il constitue un écrit¹¹⁹. La dimension de l'écrit apparaît dans son modèle comme proprement définitoire du document : on comprend alors pourquoi le document doit toujours, en indexation, comporter un texte¹²⁰.

Mais l'autonomie physique ne nous semble pas être le seul aspect de l'autonomie que la source doit présenter. Les interrogations récentes des professionnels sur le caractère documentaire ou pas (« indexable » ou pas) des « flux d'informations » transmis, par écrit, sur le réseau Internet¹²¹ mettent l'accent sur l'importance, dans la construction documentaire, de pouvoir circonscrire un ensemble de paramètres de production de l'information (lieu, temps, acteur, notamment). Ces flux d'information posent la question de ce que nous avons appelé l'« autonomie logique » d'une source : il paraît en effet que, pour être décontextualisée, une source doit bel et bien avoir été constituée dans un contexte à même de laisser une « trace » dans la source elle-même.

Sur ce point, l'analyse documentaire des images est éclairante¹²² : autant une image fixe peut, moyennant l'attribution d'une légende par exemple, être constituée en document autonome. Autant une image mobile, extraite d'une séquence filmique, pourra être difficilement constituée en document : le texte d'accompagnement devrait alors fonctionner comme une description du contexte des autres images de la séquence ; ce n'est pas là son rôle en indexation. En effet, l'autonomie « logique » de la source, qui signifie que le document construit garde une trace du contexte de production initial, est ce qui permet de **contraindre** les

¹¹⁸. Salaün 1991, p. 139 et suiv.

¹¹⁹. La « fonction documentaire » correspond, pour Escarpit, à la stabilisation d'un message « sur un support qui le rend indépendant du temps et synchroniquement disponible », Escarpit 1991, p. 124.

¹²⁰. Voir notre interrogation à la fin du chapitre I.

¹²¹. Par exemple les échanges écrits dans une liste de discussion.

¹²². On reprend ici l'analyse que propose Odile Le Guern [1989, p. 428].

détournements d'usage que le document doit nécessairement autoriser pour permettre la production de nouvelles « informations ».

(ii) Propriété d'usage de la source

Une source doit pouvoir être utilisée moins pour l'usage explicite pour lequel elle a été conçue que pour d'autres usages¹²³, dont tous ne peuvent être connus ou répertoriés : c'est ici l'espace d'inventivité de l'utilisateur qui ne concerne pas nécessairement l'indexeur. La mise en document apparaît alors comme ce qui doit circonscrire l'espace de détournement possible d'une source.

La propriété d'usage multiple d'une source peut être appréhendée en termes de rentabilité ; ainsi du document iconographique :

Si l'on pose le problème en termes de rentabilité, l'image rentable, c'est-à-dire celle qui serait susceptible d'être intégrée à un très grand nombre de discours, serait celle qui pourrait être facilement décontextualisée, qui pourrait apparaître comme autonome et indépendante par rapport au contexte que constitue l'organisation séquentielle qui l'intègre.

O. Le Guern, 1989, p. 428.

Toutes les sources ne sont pas au même titre rentables, c'est-à-dire décontextualisables et recontextualisables dans le cadre d'usages différents ; ainsi, pour reprendre la boutade d'U. Eco, seul Proust est capable de mener une lecture « psychédélique » de l'indicateur des chemins de fer¹²⁴.

En effet, il y a sans doute des types de source qui sont moins rentables que d'autres : là encore, toute une étude précise reste à mener qui indiquerait, notamment par l'établissement d'une typologie par « genres » de production, un « degré » de détournement possible.

Pour préciser ce sur quoi porte, en indexation, la mise en document, nous avons sollicité quelques-uns des éléments d'une théorie de l'énonciation : à travers ce modèle, on a proposé de considérer la mise en document comme une opération portant sur le contexte d'énonciation, ou encore le contexte de production. De là peuvent se formuler les propriétés discriminantes des sources documentaires. Comment se passe la construction du document, vue du côté du document lui-même ?

¹²³. Par exemple, un brevet, dont la finalité première est de protéger un inventeur, peut être, documentairement utilisé, dans une base de brevets, pour permettre d'informer sur les découvertes récentes.

¹²⁴. Eco 1985 [1979], p. 74 : « Proust pouvait lire l'horaire des chemins de fer et retrouver dans les noms des localités du Valois les échos doux et labyrinthiques du voyage de Nerval à la recherche de Sylvie. Mais il ne s'agissait pas d'interprétation de l'horaire, c'était l'une de ses utilisations légitimes, presque psychédélique ».

III.1.2- La construction du document vue du côté du document

Comme précédemment, nous partirons d'une conjecture (ici, le document vu sous l'angle de l'interprétant) pour arriver à déterminer, ne serait-ce que partiellement, des propriétés, ici celles du document. Parallèlement, nous tenterons de préciser le rapport qu'établit le document, en tant qu'interprétant, avec la source : est-ce un rapport d'interprétation ? est-ce un rapport d'utilisation ?

A- Conjecture : le document comme interprétant de la source

Pour capter le rôle du document par rapport à la source, il paraît fructueux de recourir au modèle proposé par Peirce. En reprenant ses termes, on dira que le document est l'interprétant de la source, c'est-à-dire non l'interprète mais « le moyen que celui-ci utilise pour effectuer son interprétation »¹²⁵. Il constitue en cela les conditions d'interprétation d'une source : on se souvient en effet que l'interprétant est « la règle qui permet au *representamen* de renvoyer à un objet »¹²⁶.

On constate en effet souvent ce paradoxe qui consiste à indexer un document non pour retrouver le document lui-même mais les informations qu'il contient¹²⁷. Les informations, tout comme le document, n'existant pas en soi, hors d'une lecture ou d'un usage¹²⁸, l'indexation ne peut directement pointer sur elles ; en revanche l'indexation peut donner à interpréter les « données » d'une source en « informations » *via* un document.

Ainsi le référent de l'indexation est-il bien le document qui, en tant qu'interprétant, désigne¹²⁹ sa source : par le document, l'indexation crée un espace intermédiaire de représentations communes où les utilisateurs peuvent interpréter les sources et construire leurs objets de discours, nous y reviendrons dans le chapitre IV.

Si, comme dans le cas du semi-document, l'indexation portait directement sur la source, il n'y aurait pas de décontextualisation possible et donc la construction des objets de discours ne serait que la reconstruction des objets du discours du texte. Dans le cas de l'indexation du document, la situation est différente : l'information ne se conçoit plus comme « donnée puis comme ressource stockable puis

125. Peirce *in* Everaert-Desmedt 1990, p. 40.

126. Voir précédemment, chapitre II (§ III.2), la présentation du modèle sémiotique de Peirce.

127. Cf. Dachelet 1990 : « On a renoncé à l'espoir de fournir à l'utilisateur LA réponse à LA question posée. Une réponse, c'est aujourd'hui un document ou un ensemble de documents que l'utilisateur estimera pertinent ».

128. *Id.* : « Une base de données ne fait que stocker des données, données dont c'est à l'utilisateur lui-même de décider lesquelles constituent les informations qu'il cherche ».

129. Il nous semble en effet que l'interprétant est, dans le cas du document, de type indiciaire.

consommable, [mais] comme processus »¹³⁰. C'est pour cette raison que si les descripteurs sont des accès aux documents, ils sont des accès à l'information et non des accès d'information, c'est-à-dire informatifs par eux-mêmes. Interprétant et non interprète, l'opération de mise en document organise donc un accès sous-déterminé (d'un point de vue interprétatif) aux sources.

Sur ce point, il apparaît que les différents types de « mise en document » des sources¹³¹ ne modifient pas nécessairement l'information proposée à l'utilisateur : plusieurs conditions d'interprétation différentes peuvent en effet permettre le même type de lecture.

Cette approche du document comme interprétant de la source permet de souligner que les objets documentaires ne sont pas, en indexation, des objets « interprétés » : il semble plutôt que les sources sont présentées aux utilisateurs (les interprètes) comme interprétables dans un cadre, dans un contexte que le document précise. Ce peut être, par exemple, pour un document constitué de plusieurs sources, les sources elles-mêmes qui constituent les unes pour les autres leur propre contexte d'interprétation ; ce peut être aussi les descripteurs, qui peuvent permettre d'interpréter une source par référence aux autres sources (textes) dont le mot porte la trace.

Vu sous l'angle de l'interprétant, le document est appelé à être redéfini aussi bien dans sa « nature » que dans sa « fonction ». Du point de vue de sa « nature », le document se laisse voir comme un « énoncé », inséré dans de nouveaux contextes d'énonciation. Pour ce qui est de son rôle, le document en tant qu'interprétant d'une source, s'il en donne les conditions d'interprétation, peut aussi en fixer les règles d'utilisation : en effet, si le propre d'une source est d'être détournée, encore faut-il que ce détournement puisse être contrôlé. Mais comment ?

Préciser le rôle que tient le document en tant qu'interprétant de la source conduit à interroger ce que l'indexation cherche à transmettre *via* un document : des conditions d'interprétation ou bien des modes d'utilisation des sources ?

¹³⁰. Batime 1995. Cf. aussi Capurro 1992 : « Information is not the end product of a representation process, or something being transported from one mind to the other, or, finally, something separated from a capsule-like-subjectivity, but an existential dimension of our being-in-the-world-with-the-others ».

¹³¹. Voir les résultats de notre enquête rapportés dans le § I.3.1 de ce chapitre.

B- Indexation : mode d'interprétation ou mode d'utilisation ?

Pour discuter la proposition suivante de Valéry : « il n'y a pas de vrai sens d'un texte », Eco¹³² propose d'établir la distinction suivante entre les deux notions d'« interprétation » et d'« utilisation » :

- l'interprétation d'un texte consiste à « respecter le monde possible » décrit dans le texte ;
 - l'utilisation d'un texte se marque par l'introduction dans le discours d'« informations extratextuelles ».
- Eco 1992 [1990], p. 39.

Ce qui distingue essentiellement interprétation et utilisation¹³³, c'est le type de contexte, interne ou externe au texte, dans le cadre duquel on se propose de lire, ou dans le cadre duquel on donne à lire, un texte : en effet,

Un texte n'est pas autre chose que la stratégie qui constitue l'univers de ses interprétations - sinon légitimes - du moins légitimables. Toute autre décision d'utiliser librement un texte correspond à une décision d'**élargir l'univers de discours**. La dynamique de la sémosis illimitée ne l'interdit pas, au contraire, elle l'encourage. Mais il faut savoir ce que l'on veut : faire subir un entraînement à la sémosis ou interpréter un texte.

Eco 1985 [1979], p. 74. (C'est nous qui soulignons).

L'approche du document en termes d'interprétant favorise, semble-t-il, **un mode utilitaire** des sources ; ceci semble cohérent avec notre autre conjecture, où la source se définit en fonction de ses possibilités de détournement. Cependant, et notamment parce que l'indexation doit maintenir un lien entre les différents espaces de représentations (des auteurs, des indexeurs, des utilisateurs), il faut pouvoir stopper « l'entraînement de la sémosis » ; autrement dit, il faut pouvoir circonscrire l'univers de discours¹³⁴, c'est-à-dire l'espace des utilisations qu'il est nécessaire de proposer aux utilisateurs : c'est en ce sens que nous proposerons la notion de discours documentaire* (*infra*, chapitre IV).

Mais avant, nous devons préciser le rapport qu'entretiennent utilisation et interprétation (B1). De ce point de vue s'éclaire la problématique du traitement documentaire de l'image (B2).

B1- Quelles sont les relations qu'entretiennent utilisation et interprétation d'un texte ? Sont-elles exclusives l'une de l'autre ?

¹³² Eco 1985 [1979] et Eco 1992 [1990].

¹³³ On peut tout à fait considérer, à la suite de Bourdieu [1982] par exemple, que seul l'usage fait le sens et donc que l'interprétation d'un texte n'est rien d'autre que l'utilisation qui en est faite. Cette position revient, nous semble-t-il, à nier le rôle de la langue dans les faits d'interprétation. C'est pourquoi nous ne la retiendrons pas dans cette recherche.

¹³⁴ Eco 1985 [1979], p. 45 : la limite logique à l'entraînement sémiotique est l'univers de discours, « un univers de faits limité ».

Sur ce point, Umberto Eco note que, si « leur libre utilisation n'a rien à voir avec leur interprétation [...] toute lecture est toujours un mélange des deux », dans la mesure où est toujours présupposée « une référence au texte-source, du moins en tant que prétexte »¹³⁵.

C'est alors le type de lecture qui est fait de la source qui détermine le rôle que joue le document en tant qu'interprétant : mode d'interprétation ou mode d'utilisation d'une source ?

C'est ici à la distinction établie par Escarpit¹³⁶ entre lecture objective et lecture projective que nous nous référons :

- dans la lecture objective, on cherche à « épuiser l'entropie du texte, c'est-à-dire à énoncer toute l'information qu'il contient, à le rendre entièrement connu, de sorte que tout nouveau balayage ne produira que des événements prévisibles »¹³⁷. En reprenant la problématique d'Eco, on peut dire que ce type de lecture fige la source dans l'une de ses utilisations ;
- dans la lecture projective, il s'agit d'effectuer une « mise en mémoire »¹³⁸ d'une source dans un document de façon à ce qu'elle reste « disponible pour une lecture c'est-à-dire pour une exploration libre de toute contrainte événementielle ou chronologique »¹³⁹. Cette lecture consiste à réaliser une double opération : « sélectionner » et « associer »¹⁴⁰. Par la sélection, cette lecture introduit le facteur « oubli », que nous traduisons comme étant l'oubli des contextes de production d'une source. Par l'association, cette lecture introduit des « possibilités d'évocation analogique »¹⁴¹. Dans les termes d'Eco, on peut dire que cette lecture fournit la possibilité d'utilisation diverse mais réglée d'un texte (contraintes liées à celles de l'analogie) .

Dans les deux types de lecture, il y a, avant toute décision d'utilisation, nécessairement une interprétation de la source ; cette interprétation pèse sur les utilisations possibles. En ce sens, l'utilisation est toujours seconde et toujours

135. Eco 1992 [1990], p. 46-47.

136. Escarpit 1991, p. 129.

137. *Id.*

138. *Ibid.*, p. 150.

139. *Ibid.*, p. 63.

140. *Ibid.*, p. 152.

141. Par l'analogie se crée la « production d'un nouvel événement sous l'effet d'un nouveau stimulus qui, d'une part, suscite une réponse libre, d'autre part, peut réactiver par analogie et non plus par identité la trace d'une expérience passée », Escarpit 1991, p. 151.

déterminée par l'interprétation, qui, elle, est toujours première, sauf dans le cas particulier des objets non textuels où l'utilisation ne nécessite pas une « interprétation » proprement dite. C'est, nous semble-t-il, pour cette raison que les objets non textuels ne deviennent documents, candidats à l'indexation, qu'une fois affectés d'un texte qui, lui, peut faire l'objet d'une interprétation. Nous abordons succinctement cette problématique ci-dessous.

B2- La problématique du rapport utilisation/interprétation se trouve au cœur du traitement documentaire de l'image¹⁴². En effet, l'utilisation de l'image peut complètement supplanter l'interprétation¹⁴³ : se posent, alors, de façon cruciale, le problème de l'absence de contraintes interprétatives et le danger du détournement incontrôlé des images. On connaît l'exemple fameux de l'utilisation intempestive d'une des photographies de Robert Doisneau¹⁴⁴ : la scène photographiée est celle d'une discussion, dans un café de la rue de Buci à Paris, entre une jeune étudiante et son professeur d'architecture. Utilisée une première fois comme support d'illustration pour une campagne anti-alcoolique, la même image apparaît, peu de temps après, dans un magazine « à scandale » avec pour légende : « Prostitution aux Champs-Élysées ». Reprenant ce même exemple, Odile Le Guern souligne que seule la formulation linguistique (légendes et/ou formules d'indexation) peut neutraliser l'interprétation et contraindre ainsi les utilisations d'une image¹⁴⁵. Autrement dit, l'image seule, décontextualisée, ne supporte aucune contrainte d'utilisation ; pourvue d'un texte, l'image est contextualisée et c'est cette contextualisation (par le texte) qui établit des contraintes d'utilisation.

La problématique de l'indexation des images mise en lumière par Odile Le Guern montre que, lorsqu'il n'y a pas de texte, interprétation et utilisation ne peuvent être distinguées : il n'y a pas alors d'indexation possible ; ce qui dessine en creux le processus de l'indexation : il repose sur la distinction des deux modes d'interprétation et d'utilisation. Pour proposer des utilisations des sources (soit unique dans le cas de la lecture objective, soit multiple dans celui de la lecture projective), l'indexation doit exploiter préalablement la nature textuelle, ou plutôt discursive, des documents (niveau de l'interprétation).

¹⁴². Pris ici comme exemple d'objet non textuel soumis à l'indexation. Il faudrait analyser, dans le détail, le comportement des autres objets non textuels du point de vue du rapport utilisation/interprétation.

¹⁴³. Conférence Hudrisier ENSSIB (Villurbanne), mars 1996.

¹⁴⁴. Voir Hudrisier 1984 et O. Le Guern 1989.

¹⁴⁵. O. Le Guern 1989, p. 427.

L'emprunt, au modèle de Peirce, de la notion d'interprétant nous permet de dégager une problématique, peu abordée dans le discours classique : l'indexation vise-t-elle, *via* son référent, à fournir des modes d'interprétation ou des modes d'utilisation des sources ?

Nous avons fait l'hypothèse que, si l'indexation devait plutôt fournir des utilisations des sources, elle devait se donner le moyen de contraindre ces utilisations : en ce sens, nous avons fait l'hypothèse que l'indexation créait un niveau de discours particulier - le discours documentaire - dans lequel les utilisations des documents pouvaient être contrôlées. Dans ce cadre, quelles sont les propriétés dont doit être pourvu le document ?

C- Propriétés du document

Nous ne faisons ici qu'amorcer une réflexion sur les propriétés dont doit être pourvu le document pour « fonctionner comme » l'interprétant d'une source. Nous en proposerons des prolongements dans le chapitre IV.

(i) Propriété de recontextualisation

Si le document est construit sur la base de la décontextualisation d'une source, il doit être recontextualisé, sous peine de rester au stade du semi-document : il s'agit là d'un « effet de compensation » dans les termes d'Escarpit¹⁴⁶. La recontextualisation du document doit se faire de façon à permettre de créer des « évocations analogiques » : un document n'est tel que s'il permet d'établir des liens avec d'autres documents sous une forme qui ne soit pas celle de la répétition à l'identique. C'est pourquoi tous les organismes documentaires ne construisent pas nécessairement, à partir d'une même « réalité éditoriale », les mêmes documents¹⁴⁷ : tout dépend des autres documents déjà présents. Ces documents constituent un « contexte » les uns pour les autres ; sur ce point, la recontextualisation d'une source qui lui permet de fonctionner comme un document repose sur une mesure de « compatibilité »¹⁴⁸ entre la source et les documents déjà sélectionnés.

¹⁴⁶. Escarpit 1991, p. 125 : « On notera que le temps, dont l'effet est compensé lors de la constitution du document, doit être réintroduit sous forme de *mouvement* pour que l'information soit restituée au destinataire ».

¹⁴⁷. Voir précédemment, dans ce chapitre § I.3.1., le résultat de notre expérimentation.

¹⁴⁸. Ce point est développé dans le chapitre IV.

(ii) Propriété de stabilité

En mettant en regard document et événement, Escarpit souligne l'effet de stabilisation que peut donner un document. Si le document a pour propriété celle de cette stabilité, il ne l'a que dans le cadre de l'espace documentaire¹⁴⁹, qui stabilisant la sémiosis, stabilise du même coup un état du monde, qui peut alors se donner comme « réalité commune » aux indexeurs et aux utilisateurs. C'est parce que la propriété de stabilité n'est valable que dans le cadre de l'espace, du discours documentaire, qu'elle nous semble être une propriété distinctive du document.

Nous avons, dans un premier temps (§ III.1.1 et § III.1.2 ci-dessus), tenté de distinguer les différents processus en jeu dans la construction du document en indexation, en adoptant deux points de vue différents :

- le point de vue de la source, abordé sous l'angle de la théorie de l'énonciation, laisse apparaître la notion de contexte (contexte dans lequel une source est produite) comme ce qui distingue une source d'un document ;
- le point de vue du document, appréhendé dans le cadre du modèle de Peirce, a permis, lui, de préciser ce qui rapproche source et document : ils se trouvent dans un double rapport d'interprétation et d'utilisation.

Nous proposons désormais de synthétiser, sous la forme d'un schéma, le processus de la création d'un document à partir d'une source. Par cette schématisation, nous montrerons que la construction du document telle que nous l'entendons fait partie intégrante du processus de l'indexation.

III.1.3- La construction du document : une opération à double détente

Nous commençons par exposer notre schématisation de la construction du document en indexation (A), que nous commentons ensuite sous trois aspects :

¹⁴⁹. Espace artefactuel de non-changement dans les termes d'Escarpit [1991, p. 125] : « Un objet stable reste emporté par le temps, la stabilité n'étant que le non-changement des relations qui définissent sa configuration caractéristique (son *pattern*) aux yeux d'un observateur emporté par un temps qui est, par convention pratique, supposé le même ».

(B) commentaire sur la notion de contexte : la schématisation proposée repose sur la notion de contexte, notion floue s'il en est, que nous essayerons néanmoins de préciser pour ce qui concerne son rôle en indexation ;

(C) commentaire sur la notion de processus : cette schématisation suppose une compréhension de la notion de processus en indexation au sens « fort » ; il s'opère, en indexation, notamment par la création du document, une transformation d'objets ;

(D) commentaire sur la visibilité de ce processus : cette schématisation rend visible un aspect de l'indexation (la mise en document) qui ne bénéficie généralement que d'une très faible visibilité : nous aborderons pour finir cet aspect.

A- Une opération à double détente

Si la mise en document opère une « recontextualisation » dans un cadre particulier à préciser, une opération, symétrique et antérieure, de « décontextualisation », doit être postulée. La phase de décontextualisation, qui correspond à la notion classique de « sélection documentaire », est donc indissociablement liée aux conditions d'existence du document. Dès lors, notre thèse consiste à appréhender l'indexation comme un processus à double détente, composé de deux opérations symétriques et inverses de contextualisation¹⁵⁰.

Pour fixer les idées, on pose, en première approximation, la schématisation de notre thèse sous la forme suivante :

¹⁵⁰. On est assez proche en cela de la position défendue par Escarpit qui parle, lui, de déconstruction et de reconstruction du texte, Escarpit 1991, p. 166 : l'analyse documentaire « doit procéder à une réécriture du texte qui n'est pas simplement linéaire ou quantitative, mais qui comporte une déconstruction, puis une reconstruction systématique du texte ».

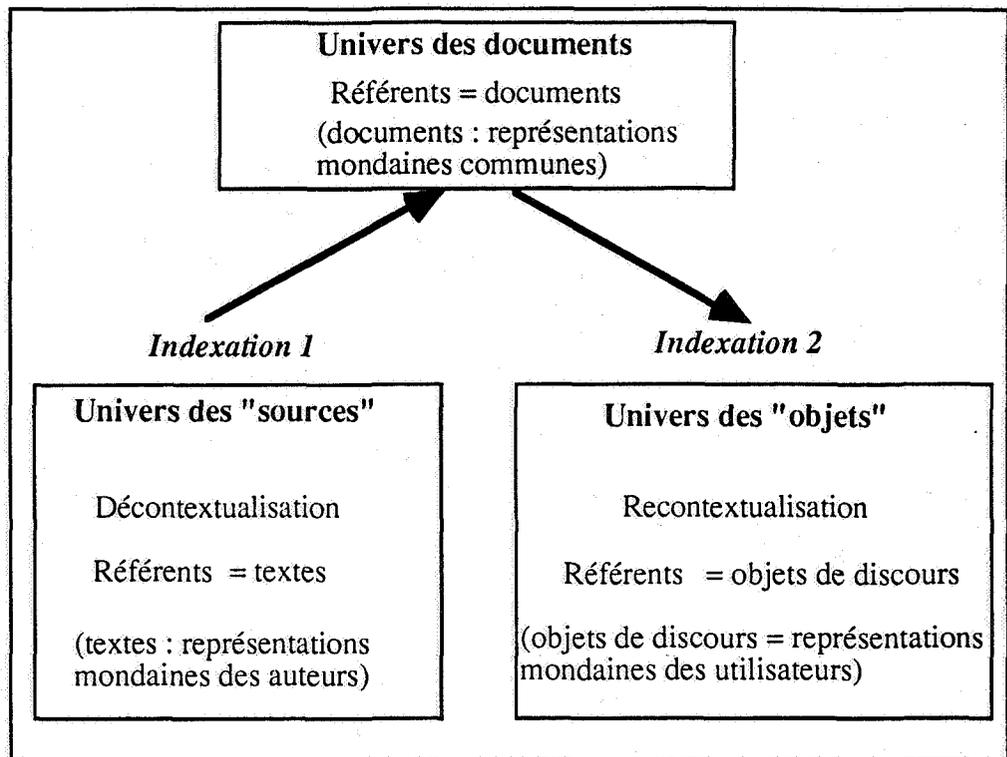


Figure 3- Mécanisme « à double détente » de l'indexation

Dans ce schéma, l'indexation crée son référent (le document) en procédant à une double opération de contextualisation :

- la première est une décontextualisation : la source est détournée de son utilisation initiale ;
- la seconde est une recontextualisation : la source prend de nouvelles potentialités interprétatives, intégrée dans un ensemble que l'on nommera « discours documentaire ». Dans ce cadre se dispose l'ensemble des usages possibles d'un document (des détournements autorisés d'une source).

L'enjeu de l'indexation est donc de réaliser un changement de contexte : cette notion, qui met en valeur le processus (au sens fort) de l'indexation, mérite d'être précisée.

B- Notion de contexte en indexation

La notion de contexte est, en linguistique, mal définie. Elle est, d'après Kerbrat-Orecchioni¹⁵¹, à la fois « problématique »¹⁵² et « indispensable »¹⁵³. Rastier, qui se livre à une semblable remarque¹⁵⁴, insiste, lui, sur la nécessité d'approcher la notion de contexte en termes de palier : syntagme, période, section, texte constituent chacun un « palier de contextualité ».

Rastier traite ce faisant la notion de contexte linguistique¹⁵⁵ (dit aussi co-texte) que l'on oppose généralement à la notion de contexte extra-linguistique¹⁵⁶ (dit aussi contexte situationnel)¹⁵⁷. Kerbrat-Orecchioni s'attache, elle, plus à ce deuxième type de contexte.

Ce que ces deux auteurs mettent en lumière, c'est que le contexte n'est pas un donné, mais un construit. À ce titre, on peut dire qu'un contexte, c'est « l'ensemble des représentations que les interlocuteurs ont du contexte »¹⁵⁸. En ce sens, un contexte est toujours « choisi », pour reprendre les termes de Rastier¹⁵⁹, et ce choix constitue un « acte décisif » :

Alors que le texte appartient au « donné » empirique, tel qu'on choisit de le décrire, le contexte est *choisi*, donc soumis comme tel à des conditions herméneutiques. Dans l'interprétation des mots, comme dans l'interprétation des textes, le choix du contexte est un acte décisif, qui doit être déterminé par une stratégie. (C'est nous qui soulignons).

Rastier 1994, p. 65.

Au risque de remplacer une notion floue par une autre notion floue, nous essayerons d'approcher la notion de contexte en indexation sous l'angle de la notion de choix stratégique.

Les questions qui se posent à une approche du contexte sont alors formulées dans d'autres termes :

151. Kerbrat-Orecchioni 1996, p. 39-60.

152. Latraverse cité in Kerbrat-Orecchioni 1996, p. 39 : « La notion de contexte est d'une telle souplesse et d'un accueil si généreux qu'il est difficile de considérer qu'elle a des frontières suffisamment établies pour jouer un rôle théorique non équivoque ».

153. *Id.* : « Il n'est guère de théorie ou d'approche sémantique des langues naturelles qui en fasse l'économie ».

154. Rastier 1994, p. 64 : « On ne dispose pas en linguistique de théorie générale du contexte. Si la notion de contexte est souvent évoquée, elle est rarement définie ».

155. Ce type de contexte peut être entendu comme l'« ensemble du texte qui entoure un élément de la langue (mot, phrase, fragment d'énoncé) et dont dépend son sens, sa valeur », Kerbrat-Orecchioni 1996, p. 40.

156. Ce type de contexte peut être entendu comme l'« ensemble des circonstances dans lesquelles s'insère un fait ». *Id.*

157. Le problème de cette distinction, même si elle est importante à maintenir sur un plan méthodologique, est, comme le remarque Kerbrat-Orecchioni, que les deux types de contexte sont en fait dans un rapport de « vases communicants », voir Kerbrat-Orecchioni 1996, p. 40.

158. *Ibid.*, p. 41.

159. De même pour Kerbrat-Orecchioni, la dimension du contexte à prendre en compte « dépend évidemment des cas ». *Ibid.*, p. 40.

(i) La question de la dimension du contexte à prendre en compte

Le contexte d'une source peut être approché par la notion de « contexte situationnel », contexte extra-linguistique des « circonstances ». Le contexte du document relève, lui, d'abord du contexte linguistique, ensuite du contexte situationnel.

Le contexte du texte d'un document est en effet constitué des descripteurs affectés à ce document : on peut parler à ce titre de modification du contexte linguistique du document. Ces descripteurs fonctionnent comme des relais lexicaux entre plusieurs documents ; ils contribuent alors à créer un nouveau contexte d'interprétation, un nouveau contexte situationnel : le contexte d'un document est ainsi constitué des autres documents qui partagent les mêmes descripteurs. C'est sur ce point qu'apparaît nous semble-t-il la notion de stratégie documentaire : c'est elle qui détermine le choix du regroupement des documents comme celui des textes dans un même document ; ce choix s'exprime sous une forme lexicale, par les descripteurs, mais n'est que le résultat d'une stratégie qui ne se joue pas au seul niveau des mots. En ce sens, on dira que l'indexation opère le passage d'un contexte situationnel à un autre, par modification du contexte linguistique des documents (les descripteurs), et qu'elle obéit ce faisant à une stratégie que nous aurons à préciser.

(ii) La question du rapport entre un texte et un contexte

Kerbrat-Orecchioni met en valeur la relation dialectique qui caractérise le rapport entre texte et contexte : « le discours façonne le contexte autant que le contexte façonne le discours »¹⁶⁰, et ce, parce que texte et contexte peuvent être de même nature, précise Kerbrat-Orecchioni : « une unité donnée n'est pas en soi élément du texte ou du contexte »¹⁶¹.

Là encore, pour ce qui est des faits d'indexation, il nous semble important de relever que, si une collection documentaire peut se comprendre en termes de situation d'interprétation (de contexte situationnel) pour des utilisateurs, c'est parce que la stratégie documentaire permet aux documents qui s'y trouvent de se transformer et/ou de s'éclairer les uns les autres de façon « légitime », pour reprendre les termes d'Eco. Cet éclairage qui permet à un texte de se donner comme le contexte d'un autre texte repose, selon Kerbrat-Orecchioni, sur la notion d'« événement focal », mise en valeur par Duranti & Goodwin¹⁶². Cette notion d'« événement focal » qui permet, par le biais d'un texte T₁, de focaliser un aspect

¹⁶⁰. Kerbrat-Orecchioni 1996, p. 49.

¹⁶¹. *Ibid.*, p. 43.

¹⁶². « The context is thus a frame that surrounds the event being examined and provides resources for its appropriate interpretation », repris de Kerbrat-Orecchioni 1996, p. 43-44.

d'un texte T₂ (T₁ joue alors un rôle de contexte pour T₂) nous semble être à l'œuvre en indexation : nous y reviendrons dans le chapitre V.

(iii) La question de la production et de l'interprétation du contexte

Cette question nous semble pouvoir être également envisagée sous l'angle du choix stratégique.

Du côté des indexeurs, il y a production d'un nouveau contexte situationnel (univers des documents dans le schéma), production déterminée par une stratégie (que nous essayerons de définir) ; du côté des utilisateurs, il y a interprétation du contexte situationnel proposé par les documents, interprétation elle aussi déterminée par l'adoption d'une stratégie (ici une stratégie de recherche, qui ne relève pas de notre étude). Il importe peu que stratégie de production et stratégie d'interprétation des contextes soient ou pas de même nature ; il importe surtout, nous semble-t-il, de constituer les actes des différents acteurs de l'indexation sous forme de stratégie. En effet, comme le note Eco, « la coopération textuelle est un phénomène qui se réalise, nous le répétons, entre deux stratégies discursives et non pas entre deux sujets individuels »¹⁶³. Ce qu'Eco appelle ici « coopération textuelle », c'est l'acte de lecture. La stratégie de l'auteur est en indexation (dans l'univers des documents) mise au second plan, le premier plan étant occupé par la stratégie des indexeurs : c'est elle qu'il importe, selon nous, de dégager. L'indexeur se trouve pris alors dans le mouvement de la coopération textuelle, de façon plus décisive que ne le laisse supposer l'image de l'indexeur-médiateur. S'il est en effet souvent question, dans le discours classique, de la stratégie (ou des stratégies) des utilisateurs, il est peu question de la stratégie (ou des stratégies) des indexeurs.

Nous avons proposé d'approcher le contexte en indexation comme élément d'une stratégie plus globale, menée par les indexeurs pour donner à lire et à utiliser des sources dans des sens et des usages qui puissent être « nouveaux ». C'est cette stratégie qui permet de considérer le processus de l'indexation dans le sens fort de « transformation ».

C- Notion de processus

Aussi triviale qu'elle puisse paraître, l'intégration d'une source dans une collection documentaire ne s'effectue ni de façon évidente ni de façon neutre : par l'utilisation des deux modèles de l'énonciation et de l'interprétant, nous avons cherché à

¹⁶³. Eco 1985 [1979], p. 78.

montrer que cette intégration supposait une réelle transformation des objets. En effet, la modification du contexte situationnel qui signe le passage de la source au document introduit les éléments d'une lecture qui seuls permettent à une source d'être un document, c'est-à-dire de pouvoir être interprétée et utilisée à d'autres fins que celle qui lui avait été initialement attribuée. C'est en ce sens que l'indexation constitue un processus : nous nous attacherons désormais dans cette recherche à le préciser.

La difficulté tient à ce que ce processus n'est pas rendu visible : le remplacement d'un contexte situationnel par un autre ne se voit pas ; la phase de décontextualisation de la source reste peu visible, puisqu'au bout du compte, il y a toujours un contexte situationnel.

D- Défaut de visibilité

Le marquage le plus net de ce que l'on nomme ici la recontextualisation est l'« indexation », comprise dans le sens étroit d'attribution, à un document, de descripteurs : l'indexation ainsi entendue **matérialise**, en effet, sur un plan linguistique, les opérations effectuées en amont sur les sources, autrement dit la sélection documentaire.

Si l'on tente de se dégager des « effets » pour capter les « fondements », on peut soutenir l'hypothèse que la sélection documentaire (ou, plus exactement, la sélection des sources) relève du processus de l'indexation, même si la matérialisation y est reportée. On rejoint, ce faisant, les conclusions de Dubois et Mondada, émises suite à une expérience de tri de photographies :

Pour résumer ce point, on dira que, même à un niveau non explicitement verbal (le tri ne demandait aucune lexicalisation), la discrétisation du monde en catégories n'est absolument pas donnée *a priori*, mais varie selon les activités cognitives des sujets qui les opèrent.

Dubois et Mondada, 1995, p. 288.

De même que dans le compte rendu de cette expérience, on peut dire que, dans l'indexation, la discrétisation du monde en objets (la création des documents), qui n'est le plus souvent pas verbalisée, constitue une opération qui influe directement sur la décision *in fine* du choix de tel ou tel descripteur.

Les deux phases du processus de l'indexation n'offrent donc pas la même visibilité. En effet, seule la seconde phase (dite de recontextualisation) se réalise sous une forme linguistique, mais cette matérialisation ne correspond en fait qu'à l'établissement préalable et non verbalisé de liens posés entre les documents. Le fait que, dans le discours classique, la sélection documentaire ne soit pas si évidemment comprise dans le mécanisme de l'indexation tient sans doute à ce que la phase de

recontextualisation, en tant que phase de désignation, soit la seule à être dotée d'une verbalisation. À ne considérer que la face visible de l'indexation, on court cependant le risque de mésinterpréter le rôle des descripteurs en indexation : matérialisations d'une analyse documentaire réalisée en amont sur la base de principes de regroupement textuel, ils ne peuvent, à proprement parler, rendre compte du contenu d'une source ; ils peuvent, en revanche, donner à lire, à interpréter et à utiliser cette source.

Au début du paragraphe III, notre objectif était de préciser comment l'indexation pouvait se comprendre comme un acte de référencement, c'est-à-dire comme un acte établissant un double processus de discrétisation et de stabilisation.

Nous avons essayé de dégager, en III.1, le processus de discrétisation que l'indexation opère pour construire son « monde », celui des documents : nous avons établi que ce processus de discrétisation était réalisé en indexation sur la base d'une transformation de contextes, transformation qui donne la possibilité d'utilisations nouvelles.

Nous nous intéressons, dans le paragraphe suivant, au second aspect de l'acte de référencement : le processus de stabilisation. Nous montrerons alors comment l'indexation peut construire des effets de stabilité référentielle.

III.2- Construction de l'effet de stabilité référentielle en indexation

Pour appréhender l'indexation sous l'angle d'un processus de stabilisation, nous ne nous intéresserons pas, comme précédemment, à l'acte de référence lui-même mais aux propriétés référentielles des expressions linguistiques sollicitées dans la réalisation de cet acte.

Nous avons vu précédemment¹⁶⁴ que, s'il y avait stabilité référentielle par le biais d'une stabilité linguistique (cas où un mot renvoie à un objet et un seul), il ne s'agit là que d'un effet, produit de façon typique par l'interprétation d'un type d'unité linguistique particulier : le nom propre¹⁶⁵. C'est sur ce type d'unité que nous centrons nos propos dans cette partie de l'étude et ce, sous deux angles :

¹⁶⁴. Voir § I.3.2.

¹⁶⁵. Tout comme Gary-Pricur [1994, p. 7], on entendra ici, par noms propres, les « noms propres prototypiques », de type nom de personnes, nom de lieux, nom de marques, etc., laissant de côté les cas problématiques (comme les noms de peuples, écrits avec une majuscule). Par ailleurs, nous nous attacherons aux seuls cas où le nom propre est employé sans déterminant, ignorant les emplois du type : « Je croyais que Ramiz Alia pouvait devenir le Gorbatchev albanais » (repris de Gary-Pricur 1994, p. 38).

- nous chercherons d'une part à expliquer ce qui permet à un nom propre de créer un effet de stabilité référentielle : pour cela, nous empruntons au logicien Kripke¹⁶⁶ la notion de « rigidité » telle qu'il l'a établie¹⁶⁷ ;
- nous montrerons d'autre part que la pratique d'indexation, qui recourt largement à l'utilisation de noms propres, n'est pas sur ce point une pratique isolée. D'autres pratiques révèlent un recours massif à ce type d'unité : il y a là, nous semble-t-il, un élément susceptible de contribuer aux recherches portant sur l'analyse des effets stabilisateurs des pratiques¹⁶⁸.

On remarquera par ailleurs que l'effet de stabilisation que peut obtenir l'indexation, *via* l'utilisation de propriétés linguistiques particulières, ne rend que plus délicate la mise en faillite du modèle réaliste en indexation ; en effet, l'obtention de cet effet stabilisateur permet au discours classique de pouvoir, ne serait-ce que partiellement, réaliser son pari : faire correspondre aux mêmes mots toujours les mêmes choses. Reste que le discours réaliste ne peut expliquer pourquoi, dans tel cas, son pari est tenu et pourquoi, dans tel autre cas (quand les descripteurs sont des noms communs), il l'est moins : c'est précisément ce que nous permet de faire le modèle non réaliste de la référence que nous avons adopté.

II.2.1- Problématique de la « rigidité »

De nature logique et philosophique, et non linguistique¹⁶⁹, la réflexion de Kripke s'inscrit dans le cadre d'un vaste débat, dont nous ne rappelons ici que les principaux enjeux, en reprenant : la question initiale de ce débat et la solution proposée par Kripke (notion de rigidité) en A ; les corollaires au principe de la rigidité : la notion de « mondes possibles » (B), celle de désignateur rigide (C) et la dimension discursive du nom propre (D).

¹⁶⁶. Kripke 1982 [1972].

¹⁶⁷. Nous n'abordons pas ici les approches strictement linguistiques du nom propre, bien qu'elles se réfèrent toutes, plus ou moins explicitement, aux positions de Kripke. Il y a un débat entre ceux pour qui le nom propre peut faire l'objet d'une description sémantique (analyse en termes de présupposition) et ceux pour qui elle ne le peut pas (le nom propre n'a pas de signification lexicale) : pour une revue de la question, on peut se reporter par exemple à Kleiber [1981] et Gary-Prieur [1994].

¹⁶⁸. Dubois et Mondada [1995, p. 292-297] proposent d'engager une réflexion sur les moyens, cognitifs et linguistiques, par lesquels un effet de stabilisation du monde peut être obtenu.

¹⁶⁹. Il n'empêche que la notion de « désignateur rigide » a été largement reprise en linguistique, voir Gary-Prieur 1994, p. 14-25.

A- Question initiale : les propriétés référentielles du nom propre

Kripke s'oppose à Searle qui, s'inspirant de Wittgenstein, pose que les noms propres, comme les descriptions définies¹⁷⁰, ont un « sens », qui connote, pour les descriptions définies, une propriété bien déterminée et, pour les noms propres, un faisceau mouvant de propriétés : cette seule différence, de degré, entre descriptions définies et noms propres est également retenue par Frege et Russell qui traitent alors les noms propres comme des « descriptions déguisées ». Dans ce cadre, le référent d'un nom propre « est l'objet qui possède un nombre suffisant de propriétés qu'on associe à ce nom dans la communauté linguistique »¹⁷¹.

Kripke reprend la position de Mill pour soutenir une thèse inverse : il pose, comme lui, que les noms propres se distinguent des descriptions définies, en ce qu'ils désignent un objet **indépendamment des propriétés que cet objet possède**. Mais, alors que Mill utilise la notion de connotation pour définir, par la négative, le nom propre, Kripke conçoit le principe de « rigidité » pour définir de façon positive le nom propre et rendre compte de l'effet de stabilité référentielle propre à certaines descriptions définies.

C'est dans ce cadre que Kripke propose la notion de « désignateur rigide » qu'il oppose à celle de « désignateur accidentel ».

Par « désignateur », Kripke entend « un terme commun couvrant à la fois les noms et les descriptions »¹⁷², étant entendu que le nom est, dans ses termes, « un nom propre, c'est-à-dire le nom d'une personne, d'une ville, d'un pays, etc. »¹⁷³.

Dans ce cadre, Kripke distingue donc :

- les désignateurs rigides : « expression dont la dénotation* ne varie pas, quel que soit le monde considéré »¹⁷⁴ ;
- les désignateurs accidentels : « expression dont la dénotation varie selon le monde de référence, qui ne désigne pas le même objet dans tous les mondes possibles »¹⁷⁵.

170. Depuis Russell, les logiciens entendent par « descriptions définies » « les expressions comportant un nominal (nom, nom+adjectif, nom+relative, nom+complément) accompagné d'un article défini » ; des expressions comme « le livre » ou « le livre que j'ai prêté » sont analysées en logique comme des descriptions définies ; repris de Ducrot et Schaeffer 1995, p. 306.

171. Récanati 1983, p. 107.

172. Kripke 1982 [1972], p. 13.

173. *Id.*

174. Récanati 1983, p. 109.

175. *Id.*

B- Notion de « monde possible »

Le principe de rigidité est étroitement associé à la notion de « mondes possibles »¹⁷⁶ qui relève, elle, de la logique modale¹⁷⁷.

Soit les deux exemples suivants¹⁷⁸ contenant chacun l'opérateur modal « aurait pu » :

(1) Le président de la République aurait pu être un homme de gauche.

(2) Chirac aurait pu être un homme de gauche.

En (1), l'énoncé est ambigu : il réfère soit au président actuel (référence au monde réel¹⁷⁹) que l'on projette dans un « monde possible » où il aurait été de gauche, soit au président d'un monde possible (référence à un monde possible) qui, lui, serait de gauche¹⁸⁰. Selon le monde de référence considéré (monde réel ou monde possible), le référent varie en (1) : c'est pourquoi la description définie « le président de la République » est dite « désignateur accidentel ». En (2), il n'y a pas d'ambiguïté : que l'on parle du monde réel ou d'un monde possible, le référent est toujours Chirac¹⁸¹ ; en cela, le nom propre est un « désignateur rigide ».

Comme le signale Kripke, la rigidité du nom propre correspond à un primat du monde réel : c'est toujours le référent du monde réel qui est désigné même si est imaginée une situation contrefactuelle où le référent du nom propre ne présente aucune de ses propriétés.

Soit, par exemple¹⁸², l'énoncé (3) « Imaginons que Hitler ait été doux comme un agneau » ; c'est précisément parce que Hitler garde ses propriétés du monde réel, à savoir être responsable d'un génocide, que l'énoncé est interprétable¹⁸³.

176. La question de savoir si la notion de « monde possible » suggère nécessairement un engagement ontologique est ouverte ; Kripke n'en dit rien de précis. Certains linguistes commentateurs de Kripke proposent d'analyser la notion de « monde possible » en termes de discours : ainsi, selon Gary-Prieur, les mondes possibles « n'ont pas d'existence extérieure au discours qui les met en place », Gary-Prieur 1994, p. 21. C'est cette lecture que nous privilégions dans cette recherche, voir aussi le chapitre IV.

177. Il est hors du cadre de cette étude de préciser ce qu'on peut entendre par « logique modale » ; nous nous efforcerons simplement, dans le paragraphe qui suit, d'exposer les éléments pertinents à la compréhension de notre démarche.

178. Les exemples présentés ci-dessous ainsi que leurs commentaires sont repris de Récanati 1983, p. 108-109 ; les exemples sont, pour des raisons de commodité de présentation, mis au goût de l'actualité politique actuelle.

179. Ici encore, on peut considérer que la notion de « monde réel » est chez Kripke de nature discursive, voir Gary-Prieur 1994, p. 21 : « Même si ce n'est pas dit explicitement sous cette forme dans le texte de Kripke, je crois que c'est, en dernière analyse, l'acte d'énonciation qui fonde le " monde réel " auquel s'opposent les mondes possibles ».

180. Plus précisément, la première lecture porte sur l'objet lui-même (« le président de la république »), c'est une lecture *de re* ; la seconde lecture porte sur l'ensemble de l'énoncé, elle est dite *de dicto*. Nef 1991, p. 102.

181. Le nom propre neutralise l'ambiguïté des lectures *de re* et *de dicto*.

182. Repris de Récanati 1983, p. 109.

183. *Id.* On voit la différence avec l'exemple suivant : « Imaginons que le chef de l'Allemagne

Par conséquent,

Dire que la relation de désignation qui associe le nom propre à l'objet est « rigide », c'est dire qu'elle n'est pas affectée par le changement du monde de référence, et dire cela, c'est dire qu'elle n'est pas fonction des propriétés de l'objet, dans la mesure où ce sont ces propriétés qui varient d'un monde à l'autre.

Récanati, 1983, p. 110.

Cette permanence de la référence maintenue par le nom propre à travers tous les univers de référence imaginables n'est pas sans nous rappeler l'ambition de l'indexation qui cherche à maintenir, par une même forme linguistique, une même stabilité référentielle¹⁸⁴.

C- Notion de désignateur rigide

Ce principe de rigidité, illustré de façon exemplaire par le nom propre, est également à l'œuvre dans le fonctionnement référentiel d'autres types d'expressions linguistiques. En effet, dans la théorie de Kripke, certaines descriptions définies peuvent également fonctionner comme des désignateurs rigides (c'est d'ailleurs sur ce point que Kripke se sépare de Mill).

Par ailleurs, il faut également préciser que, si la démonstration de la rigidité est particulièrement visible appliquée aux contextes modaux, Kripke ne la soutient pas moins dans tous les autres types de contexte¹⁸⁵. L'exemple suivant illustrera la position de Kripke sur ces deux derniers aspects (description définie comme désignateur rigide et absence d'opérateur modal) :

(4) π est censé être le rapport de la circonférence d'un cercle à son diamètre,

où « π » et « le rapport de la circonférence d'un cercle à son diamètre » sont des désignateurs rigides même s'ils ne le sont pas au même titre¹⁸⁶.

En effet, Kripke¹⁸⁷ distingue la rigidité *de jure* (par stipulation), caractéristique des noms propres, et la rigidité *de facto*, qui couvre le cas où les descriptions définies ne sont vraies que d'un seul objet, quel que soit l'univers envisagé. Kripke pense

nazie ait été doux comme un agneau ».

¹⁸⁴. Notons, cependant, comme le fait remarquer Kripke, que la notion de rigidité ne suppose pas de critère d'identité à travers les mondes ; l'identité référentielle est une conséquence de la rigidité, voir l'enjeu de la distinction étudiée dans le chapitre IV.

¹⁸⁵. Kripke 1982 [1972], p. 162-166.

Hors de contextes modaux, les deux exemples (1) et (2) s'opposent tout autant, mais l'opposition s'exprime ici en termes de valeur de vérité et non en termes de potentialité : « Le président de la République est un homme de gauche » (si c'est Chirac, c'est faux ; si c'est Mitterrand, c'est vrai) et « Chirac est un homme de gauche » (c'est faux, quel que soit le monde de référence).

¹⁸⁶. Kripke 1982 [1972], p. 48-49.

¹⁸⁷. *Ibid.*, p. 173.

surtout aux descriptions mathématiques : ainsi, le référent de la description définie « le plus petit nombre premier » est le nombre 2, dans tous les mondes possibles. La rigidité vient donc dans ce cas de la nature du référent désigné. La rigidité étant dans ce cas *de facto*, l'identification du référent se fait hors du contexte de l'énoncé, alors que, pour les désignateurs rigides *de jure*, le référent est directement identifiable dans l'énoncé. Nous revenons sur la portée de cette distinction dans le chapitre V.

D- Dimension discursive du nom propre

La théorie de la rigidité énoncée par Kripke montre que la langue peut envisager des cas de « référence directe [...] qui court-circuite la machine intensionnelle »¹⁸⁸, c'est-à-dire des cas où les unités linguistiques établissent une relation référentielle stable sans que l'on ait à se préoccuper de leur « sens ». Les désignateurs rigides illustreraient donc un cas d'autonomie de la relation référentielle, ce qui ne signifie pas que cette autonomie existe en soi ; elle est, en effet, issue non du réel mais de situations discursives¹⁸⁹. Selon Kripke, la fixation des référents des désignateurs rigides *de jure* se fait en effet par le biais d'une « chaîne causale » :

La référence semble finalement déterminée par le fait que le locuteur fait partie d'une communauté de locuteurs qui utilisent le nom. Le nom lui a été transmis grâce à une tradition, de maillon en maillon.
Kripke, 1982 [1972], p. 95

Après un « baptême initial » au cours duquel une description définie est associée à un individu et à un nom propre, se met en place une « chaîne causale » qui correspond à la transmission du nom propre d'un locuteur à l'autre. Mais, comme le précise Kripke, « la manière dont la référence d'un terme est fixée n'a rien à voir avec sa signification »¹⁹⁰.

C'est cette particularité de la fixation de la référence, réalisée une fois pour toutes après un baptême initial, qui distingue les pronoms des noms propres : si, comme le note Milner, la référence virtuelle des deux types d'unité « contient déjà la mention de l'unité en cause »¹⁹¹, cette référence virtuelle est circulaire dans le cas du pronom, elle ne l'est pas dans le cas du nom propre ; autrement dit, la référence du nom propre n'est pas suspendue à l'énonciation, elle existe en dehors d'elle

¹⁸⁸. Nef 1991, p. 100.

¹⁸⁹. Kripke 1982 [1972], p. 82-83 : « Il est faux que nous déterminons l'objet auquel nous faisons référence grâce à des propriétés qualitatives qui seraient à notre disposition et qui permettraient de singulariser l'objet en question. [...] En général, ce à quoi nous faisons référence dépend non seulement de ce que nous pensons nous-mêmes, mais des autres gens de la communauté, de l'histoire du chemin suivi par le nom pour nous atteindre, et ainsi de suite. C'est en suivant cette histoire qu'on parvient à la référence ».

¹⁹⁰. Kripke 1982 [1972], p. 124.

¹⁹¹. Milner 1978, p. 333.

précisément en vertu de la chaîne causale qui constitue sa référence :

[Les noms propres] donnent apparemment lieu à la circularité : la définition de *Jean*, c'est d'être dit *Jean* mais en fait c'est une fois pour toutes et sans qu'il faille répéter l'énonciation qu'un sujet reçoit un nom propre. Ainsi la classe des êtres appelés *Jean* est objectivement déterminable, sans intervention à chaque énoncé singulier d'un sujet énonciateur
Milner, 1978, p. 334, n. 1.

Cette propriété de non-circularité dégagée par Milner appuie la position de Kripke précédemment rapportée sur la distinction entre fixation référentielle et signification du nom propre.

Le principe de « rigidité » établi par Kripke nous paraît éclairer avantageusement le mécanisme de l'indexation. En effet, la notion de « désignateur rigide » incarne la possibilité d'établir dans la langue une permanence référentielle transcendant la diversité des mondes de référence, et cela grâce à l'insensibilité que témoigne le nom propre à l'égard des propriétés de l'objet qu'il désigne. Une telle possibilité rend viable le pari de l'indexation, à quelques nuances près, que nous développerons dans le chapitre V consacré au descripteur : en effet, la notion de désignateur rigide ne recouvre qu'un type d'unité bien précis (essentiellement les noms propres et quelques cas de description définie) ; or l'indexation, si elle manipule bon nombre de noms propres, comme nous le montrerons ci-après, utilise surtout ce que les logiciens nomment les « désignateurs accidentels », soumis à la variabilité des univers de référence.

III.2.2- Pratiques professionnelles et usages du nom propre

Au regard de ce que fait apparaître le principe de rigidité (un principe de désignation stable car indépendant des propriétés de l'objet que désigne un nom propre), il ne nous semble pas indifférent que nombre de pratiques recourent massivement à l'emploi de noms propres, la pratique d'indexation au premier chef.

A- Sur-représentativité du nom propre en indexation

Nous rendons compte ici d'observations tirées de l'expérimentation que nous avons menée¹⁹².

Comparant les indexations réalisées par dix organismes documentaires sur un même numéro du journal *Le Monde*, nous avons relevé que la part des noms propres utilisés comme descripteurs oscille entre un tiers et deux tiers de l'ensemble

¹⁹². Le cadre de l'expérimentation est précisé dans l'annexe 2.

des descripteurs utilisés¹⁹³. On trouve principalement des noms de personnes, mais aussi des noms géographiques (pays, villes, régions) et des noms d'entreprises¹⁹⁴. Outre ces données quantitatives qui indiquent déjà, comme le note Marandin et comme nous le développerons ultérieurement, une certaine « immédiateté » du choix lexical en faveur du nom propre¹⁹⁵, nous avons noté, dans les pratiques mêmes de l'indexation, une organisation privilégiant les accès par noms propres. Un cas exemplaire est illustré par le type d'indexation réalisé au centre de documentation de la *Fondation nationale des sciences politiques*. La classification utilisée donne la priorité à la localisation géographique, l'indication « thématique » étant seconde : « chaque fois que cela est possible, on classe d'abord sous le nom d'un pays, ou sous celui d'un ensemble géographique et seulement si aucune localisation n'est possible dans les rubriques générales »¹⁹⁶.

Introduite sous forme de facettes et non plus d'indices de classification, l'entrée privilégiée par le nom propre est également à l'œuvre au centre de documentation du *Monde*. En effet, les descripteurs retenus doivent nécessairement figurer dans l'une des quatre rubriques suivantes : « territoires étrangers », « territoires français », « personnes physiques ou morales » et « auteurs cités ».

On notera également le caractère systématique de la mention de la localisation géographique dans les indexations réalisées à la *Documentation française* : le champ « descripteurs géographiques », obligatoire, est distingué des champs d'indexation dits thématiques, où, par ailleurs, l'on trouve bon nombre de noms propres (de personnes, de partis politiques, etc.).

Cette sur-représentativité du nom propre en lieu et place de descripteurs « thématiques » nous paraît doublement significative : outre qu'elle confirme le projet de l'indexation de réaliser, par tous les moyens linguistiques possibles, une permanence référentielle, elle pointe également la faillite, ou plutôt, l'inadéquation des langages documentaires, qui tiennent qu'il est possible de traiter toutes les unités de langue comme des symboles sans signification. En effet, ce n'est pas parce que le descripteur doit, prioritairement, permettre de désigner, qu'il peut se passer de signifier.

Ce que souligne également la présence massive du nom propre en indexation, du

¹⁹³. Les résultats plus précis de cette étude sont présentés dans l'annexe 4.

¹⁹⁴. Pour des raisons de commodité, nous avons assimilé, dans notre typologie, les noms de manifestations culturelles (comme Le Festival d'Avignon, La Fureur de lire), les noms de partis politiques et les noms de syndicats dans une même catégorie, celle des « noms d'entreprises ».

¹⁹⁵. Marandin 1988, p. 79.

¹⁹⁶. Pour la journée d'indexation examinée, on a comptabilisé 5 cas sur 85 (environ 6 pour cent) de classifications échappant à ce primat de la localisation géographique.

moins dans le cadre de l'expérience réalisée, c'est le caractère discursif, énonciatif, de l'indexation : en effet, on peut rapprocher le phénomène observé d'un paramètre pragmatique dit « hétéro-facilitatif » à l'origine de comportements linguistiques destinés à éviter les ambiguïtés interprétatives¹⁹⁷ ; or, précisément, le nom propre, par opposition aux descriptions définies, neutralise ce type d'ambiguïtés.

Sur ce dernier point, notre expérience fait apparaître des cas d'indexation par nom propre qui ne semblent s'expliquer que par une volonté d'éviter toute ambiguïté interprétative.

Ainsi de l'indexation de deux interviews, l'une de Simone Veil, alors ministre des Affaires sociales, de la santé et de la ville¹⁹⁸, et l'autre de Jean Chrétien, premier ministre canadien¹⁹⁹. Les deux interviews sont prioritairement, voire exclusivement²⁰⁰, indexées par le nom propre de la personne interviewée, alors que Simone Veil est interrogée sur la question précise de la prévention du sida et non sur sa carrière politique ou personnelle. De même, Jean Chrétien s'exprime, dans cet article, non sur lui-même mais sur la présence canadienne dans l'ex-Yougoslavie.

Ces faits d'indexation nous semblent souligner deux points :

- d'une part que l'indexation ne consiste pas, dans la majeure partie des cas, à dire le contenu d'un texte : il s'agit avant tout d'essayer de le situer. En cela, les noms propres présentent des points d'ancrage forts, en termes spatio-temporels pour les noms géographiques, en termes de foyers énonciatifs pour les noms de personnes ;
- d'autre part, et la remarque d'un indexeur est sur ce point éclairante²⁰¹, le recours au nom propre apparaît comme le moyen de ne pas prendre de « risque interprétatif » ; l'indexation par nom

197. Apothéloz et Reichler-Béguelin 1995, p. 238 : « Il s'agit du contrôle par le locuteur du bon déroulement de l'acte référentiel, qui s'exprime en particulier dans l'anticipation des ambiguïtés référentielles ».

198. *Le Monde* du 1/12/1994, p. 12.

199. *Id.*, p. 6.

200. Par exemple, concernant l'entretien de Simone Veil : sur les six organismes documentaires qui ont sélectionné cette source, deux ne l'indexent que par le nom propre « Veil Simone » ; deux l'indexent par le nom propre et y ajoutent des noms communs ; un des organismes documentaires qui ne dispose pas, dans la liste des descripteurs autorisés, de noms propres de personnes (accessibles à la recherche par d'autres biais) effectue une indexation qui pointe sur la personnalité de l'interviewée : « structure du gouvernement / personnage » ; enfin, le dernier organisme documentaire à avoir sélectionné cet article hésitait entre une indexation « biographique » [par nom propre] jugée non satisfaisante et une indexation « thématique » qui semblait trop floue (indexation par le terme « sida »).

201. Un commentaire était joint à l'indexation de l'article de Simone Veil : « Avant, on aurait indexé en biographie, mais maintenant on essaie de privilégier le contenu » : l'indexation par nom propre apparaît ici comme ce qui permet de ne pas s'exprimer sur le « contenu ».

propre permet à coup sûr de retrouver un document mais elle dispose ce faisant d'accès très sélectif : il faut **connaître** préalablement le nom propre, on ne peut pas le « trouver ».

Les noms propres en indexation, par leur nombre d'une part et par leurs utilisations - souvent inattendues - d'autre part, ne peuvent manquer d'attirer l'attention de l'analyste ; cependant, nous n'avons pas trouvé, dans la littérature consultée, de remarques ni d'explications de ce point.

Dans le cadre de notre analyse, la présence massive des noms propres en indexation, qui mériterait une réflexion plus approfondie, se comprend comme un moyen, pour une pratique, de stabiliser l'« instabilité constitutive des objets de discours »²⁰² afin de présenter une version stable du monde.

D'autres pratiques recourent à un semblable procédé de stabilisation.

B- Attractivité du nom propre dans les pratiques

L'immédiateté du choix lexical en faveur du nom propre se marque, de façon générale, dans les stratégies de désignation adoptées par les sujets parlants, qu'ils soient ou non en situation professionnelle.

Sur ce point, Marandin relève, dans son analyse du thème de discours, que la détermination du nom d'un thème n'est pas insensible à l'attractivité du nom propre²⁰³. Il propose une explication en termes de « cause ultime », qu'il reprend de Ricœur [1977] : le nom propre « révèle le caractère fini de l'explication », le nom propre sature une interprétation, pourrait-on dire.

En matière de pratique professionnelle, Beaulieu [1995] remarque :

Les mass-media et, en particulier, les concepteurs de périodiques culturels ont abondamment recours au nom propre pour stimuler la reconnaissance d'une information a priori, première phase du processus de la connaissance. Mais que se passe-t-il ensuite ? De quelle manière participe-t-il à l'élaboration des idées en s'insérant dans la consécution des inférences ? En approfondissant une information déjà contenu dans les prémisses d'un raisonnement ou en produisant quelque chose qui ne s'y trouve aucunement impliqué ? Il y a là un choix qui, d'après nous, explique une partie des écarts entre le savoir des individus.
Beaulieu, 1995, p. 41.

Nous reprenons l'intégralité des interrogations de Beaulieu sur le sur-emploi des noms propres dans certaines pratiques dans la mesure où elles soulignent une limite

202. Dubois et Mondada 1995, p. 273.

203. Marandin 1988 p. 79 (et note 17) : « /Lancelot/ constitue un choix plus immédiat que /la vengeance des licomes/ ».

spécifique à l'utilisation du nom propre : au-delà de la reconnaissance d'un référent, que permet le nom propre d'un point de vue interprétatif ? S'il est saturé de ce point de vue, peut-il permettre de construire plusieurs objets de discours ?

Nous examinerons de près ce point dans le chapitre V : si le nom propre semble se présenter comme le candidat descripteur idéal, en ce qu'il permet d'assurer une stabilité de la référence, il apparaît bien vite comme manquant cruellement de « plasticité » discursive ; or c'est plus cette plasticité-là que l'idéal d'une stabilité référentielle qui se révèle importante en indexation.

Nous avons proposé de voir, dans l'utilisation des noms propres en indexation, un moyen qu'utilise la pratique pour assurer une stabilité référentielle : nous avons dégager ce processus de stabilisation en nous appuyant sur le modèle de la rigidité établi par Kripke.

Ce modèle montre les enjeux, pour la construction de la référence en indexation, de l'emploi de telle ou telle unité. Le recours aux noms propres et/ou aux noms communs n'est pas indifférent parce que leur pouvoir référentiel n'est pas le même : si le nom commun réfère en discours à des objets différents, c'est sur la base de sa signification lexicale ; le nom propre réfère, lui, à un individu particulier de façon régulière sur la base d'un acte d'énonciation. À confondre les deux types d'unité, on en vient à confondre aussi les deux types de fonctionnement et à postuler ainsi que l'on peut « rebaptiser » les mots de la langue naturelle par des descripteurs de tout type linguistique.

IV- Conclusion du chapitre

La question de la référence nous est apparue centrale en indexation, notamment sous deux aspects :

- celui de la stabilité référentielle d'une part, puisque sa recherche constitue l'objectif premier et explicite de l'indexation ;
- celui des objets référentiels visés d'autre part, dans la mesure où il concerne la cible matérielle de l'indexation.

Sur ces deux points, le modèle réaliste sous-jacent à l'approche classique de la référence ne nous fournit guère d'éléments : la stabilité référentielle y apparaît sous la forme d'un axiome, en cela peu apte à rendre compte des faits de variation observables dans les pratiques. Les objets de l'indexation, le document notamment, semblent être exclus du champ des interrogations, en raison même du présupposé de leur pré-existence.

Ces deux aspects de la référence en indexation restent difficiles à appréhender dans le cadre du modèle réaliste. On peut montrer en effet qu'ils engagent une réflexion plus globale, d'une part sur le rapport entre le sens et la référence, et d'autre part sur la relation entre objets mondains et objets textuels, qui ne se comprend que dans un modèle non réaliste de la référence tel qu'il se donne dans le champ linguistique. Il nous est apparu, à ce titre, nécessaire de rappeler les grandes lignes de la problématique de la référence en linguistique : les distinctions entre types de référent comme entre acte de référence et propriété référentielle, mais aussi entre signification lexicale et construction référentielle.

Sous le nouvel éclairage retenu, nous avons repris les deux principaux problèmes de référence que nous avons identifiés :

- l'examen concernant la construction du référent-document en indexation nous a amené à concevoir l'opération d'indexation comme un double processus de contextualisation, comprenant la phase dite de sélection documentaire, assimilée alors à une étape de décontextualisation, à laquelle répond une étape de recontextualisation, au terme de laquelle une source devient un document ;
- l'étude de l'effet de stabilité référentielle recherché en indexation nous a conduit à disposer, au cœur de la construction de la référence documentaire, le principe de rigidité tel que Kripke a pu le concevoir. Sous cet angle, il apparaît que le descripteur révèle un

comportement proche de celui du désignateur rigide, la première manifestation notable en étant l'utilisation massive du nom propre en indexation.

Chemin faisant, est apparue une dimension - celle du discours - qui, pour être nécessaire à la compréhension de l'acte de référenciation en indexation, n'en est pas moins absente des approches classiques. C'est ainsi que si le chapitre II, consacré au lexique, nous amenait à étudier la question de la référence, ce chapitre III nous conduit à examiner la dimension du discours en indexation : nous effectuons là un saut, car notre approche de l'indexation s'éloigne alors radicalement du discours classique. Il nous semble nécessaire de franchir le pas. En effet, la constitution des fondements théoriques de l'indexation, telle que nous l'entendons dans cette recherche, risque sinon de rester incomplète. Certes, nous avons proposé, dans le paragraphe III de ce chapitre, un modèle d'utilisation de l'approche non réaliste de la référence en indexation, mais il nous manque encore beaucoup d'éléments pour que les remarques proposées puissent constituer des fondements théoriques : en quoi un document constitue-t-il un « énoncé » ? En quoi l'indexation réalise-t-elle une « énonciation » ? Comment se définissent, dans ce cadre, le rôle et la morphologie du descripteur ?

En faisant le bilan des problèmes théoriques de l'indexation que nous avons pu dégagés par notre étude du lexique et de la référence en indexation, nous pourrions proposer un modèle d'utilisation de la langue qui repose sur un modèle de fonctionnement de la langue valide du point de vue de la théorie linguistique : en cela, les réponses que nous proposerons, notamment aux trois questions précédemment posées, pourront prétendre à une certaine généralité.

C'est à la mise au point d'un modèle d'utilisation de la langue en indexation que nous consacrons la seconde partie de cette étude.