DEUXIÈME PARTIE : ENTITÉS, REPRÉSENTATIONS, RÉFÉRENCE

Chapitre V

LA RÉFÉRENCE : THÉORIES, MODÈLES, SYSTÈMES

Nous consacrerons cette deuxième partie de notre thèse à un problème central de la langue, incontournable dans une modélisation à l'aide d'agents communicants, et jusqu'à présent relativement peu abordé dans ce cadre. De façon générale, il s'agit d'étudier par quels moyens un code de communication entre agents permet à ceux-ci d'échanger de l'information *à propos* de leur environnement, autrement dit comment un tel code peut *dénoter* des entités et leurs propriétés.

Dans notre seconde expérience, cette question recevait une réponse simplifiée, au profit de l'étude des propriétés syntaxiques du code. Les agents manipulaient des représentations conceptuelles des situations observées qui relevaient d'un système de catégorisation précis, partagé à l'identique par tous les agents. L'établissement de conventions de nommage des entités revenait, dans ce cadre, à l'accord sur un ensemble d'étiquettes. Toutefois, la « pragmatique simulée » propose, en vue d'une simulation conforme aux réalités de la langue et de son support individuel, que les « mots » soient fondés sur un système conceptuel individuel dépendant des valeurs de chaque agent. Il faut alors expliquer par quels moyens ces « mots » parviennent à *désigner* quelque chose pour plusieurs agents à la fois.

Néanmoins, ce n'est pas au *sens* ou à la sémantique, dans toute sa généralité, que cette deuxième partie sera consacrée, mais à un problème plus précis, qui lui est dans une certaine mesure relié : la *référence*. En effet, alors que le sens est une notion plus abstraite, qui peut s'appliquer à plusieurs niveaux (mot, syntagme, phrase, discours), la référence est par définition un lien instancié entre des entités, des expressions linguistiques et des représentations. Ce lien est visible à dans l'*utilisation* du langage, que l'on peut schématiser (*a fortiori* dans la perspective d'une modélisation par agents) comme la prédication de propriétés à propos d'entités, dans le cadre d'actes de langage (par exemple l'information).

Nous avons donc considéré indispensable une étude détaillée du problème de la référence, qui soit orientée vers la réalisation d'un modèle conforme à la « pragmatique simulée ». C'est pourquoi, dans ce chapitre, nous cherchons à comprendre les différents aspects du problème, qui peuvent varier notablement avec la discipline qui les aborde : philosophie du langage (V.A), linguistique (V.B), pragmatique (V.C). Nous verrons ensuite le point de vue du traitement automatique des langues (V.D), ainsi que les systèmes de « traitement de la référence » dans les textes ou le dialogue homme-machine

(V.E). La progression dans ce chapitre correspond donc à des considérations de plus en plus computationnelles, sur la voie de l'implémentation de systèmes de traitement.

Toutefois, la « pragmatique simulée » se détache assez nettement de ces systèmes, qui ne prennent pas suffisamment en compte le caractère situé du langage. C'est pourquoi nous exploiterons au chapitre suivant (Chapitre VI) des considérations plus rarement utilisées dans les études de la référence, notamment son ancrage neuropsychologique et son fonctionnement concret, afin de proposer un modèle d'utilisation de la référence pour un agent linguistique. Nous proposons ensuite (Chapitre VII) un modèle représentationnel de la référence, vue comme un lien entre entités du monde et représentations par la médiation d'expressions référentielles ; le modèle permet l'implémentation robuste d'un système construisant ces représentations à partir de textes. Notre système fait l'objet d'évaluations, analysées dans la troisième partie de la thèse.

V.A Ontologies de la référence

Dans cette section, nous allons situer les différentes définitions et propositions d'explication du phénomène de la référence, pour simplifier, de part et d'autre de la ligne de clivage entre les partisans d'une théorie de la référence « avec sens » et ceux de la « référence directe ». Nous tenterons de mettre en évidence les points consensuels (surtout les définitions) et les différentes critiques.

Il faut bien sûr prévenir le lecteur que notre tâche, qui peut paraître téméraire dans le cadre d'une thèse en informatique, se sait d'emblée limitée par la grande étendue et difficulté du domaine abordé, auquel sont consacrés déjà de nombreux ouvrages théoriques ou de synthèse – voir par exemple les indications bibliographiques dans (Moeschler et Reboul 1994, Récanati 1993, Sterelny 1994). En effet, le domaine a été non seulement traité dès la philosophie antique, mais il a constitué un sujet important pour la philosophie scolastique médiévale, avant de connaître un renouveau significatif grâce au courant logico-positiviste, et se poser comme problème de choix à la philosophie analytique. Il s'agira donc plus ici d'un passage en revue des principales théories de notre siècle, sans essayer de discuter celles-ci en détail, ce qui nous conduirait trop loin.

V.A.1 Problématique

Pour commencer, nous allons mettre en place les différentes entités entrant en jeu : les mots et syntagmes de la langue d'une part, les objets ou entités du monde de l'autre part, sans oublier les états mentaux qui leur correspondent, et que l'on a pu appeler « concept », « représentation », « idée », « sens », et dont le support sont les individus humains. Les liens entre les différentes entités ont reçu des noms et interprétations variées (ils sont notés pour l'instant 1, 2, 3 sur la Figure V-1). La nature exacte des « concepts » ou « représentations » et leur dimension individuelle ou

collective ont fait l'objet de nombreux débats. Il nous semble cependant que ce sont les différents liens et leur mode de fonctionnement qui ont posé le plus de problèmes.

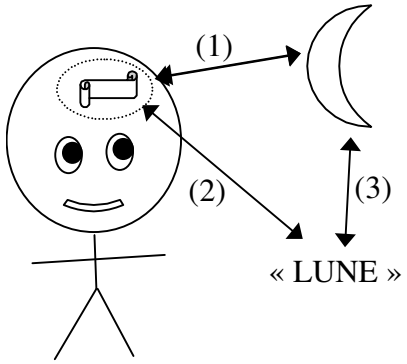

Figure V-1. Les entités participant au phénomène de la référence.

Nous pouvons déjà donner une idée approximative de la différence entre « sens » et « référence » : de façon générale, la « référence » d'une expression linguistique est l'objet auquel celle-ci se lie ou se rapporte, ou, pour reprendre une idée scolastique, l'objet qu'elle « remplace » dans l'énoncé en langue. Pour se limiter aux situations les plus simples, dans :

(1) Éteins le feu ! (dit par Ugh à Ugha après une succulente cuisse rôtie)

l'expression ou les sons *löfö* réfèrent au feu qui palpite encore devant les convives. De même, si l'on parle de *Shakespeare*, on dira que ces sons ou ce nom réfèrent, par exemple, à l'ancien propriétaire du squelette se trouvant à Stratford-upon-Avon sous l'avertissement « …cast be he who spreads my bones ».

V.A.2 La question du « sens »

Dans l'exemple (1), on peut certes parler du sens de la phrase (qui est d'ordonner…, etc.), mais on peut plus généralement affirmer que le sens du mot « feu » est un certain état incandescent de la matière, et s'applique à tous les feux que notre Ugh a pu voir. Le schéma donné sur la Figure V-1 ressemble beaucoup aux différents schémas que l'on a pu proposer pour le fonctionnement des *signes* en général, et en particulier au « triangle sémiotique », dont l'histoire remonte à l'antiquité – cf. l'aperçu de (Bordron 1991), auquel nous empruntons la terminologie donnée, qui évoque bien sûr notre schéma de la Figure V-1.

Saint Augustin :	*Verbum vocis*	*Verbum mentis*	*Res*
J. Locke :	Nom	Essence nominale	Chose
R. Carnap :	Signe	Objet intensionnel	Extension
Ch. Morris :	Signe	*Significatum*	*Denotatum*
« Structuralisme » :	Signifiant	Signifié	–

Table V-1. Noms donnés aux acteurs intervenant dans le « signe », d'après (Bordron 1991) ; nous ajoutons la dernière ligne.

Que peut-on donc conclure, à notre niveau, sur la différence entre le *sens* et la *référence* ? Il semble, de façon générale, que les études du *sens* aient privilégié le lien noté (2), entre le signe et le concept. Ainsi, pour prendre un exemple extrême, la réduction du signe au couple *signifiant/signifié* dans le structuralisme, c'est-à-dire au couple « assemblage de sons » / « concept ou représentation mentale », semble éliminer la réalité matérielle du champ de la linguistique structurale. Cette conception isole, pour l'étude du langage, deux systèmes reliés et autosuffisants, le système des signes phonologique (ou graphique) et le système des concepts.

À l'inverse, dans la conception commune de la référence, l'objet est fortement présent, et par conséquent ce sont, selon les écoles, le lien (1) ou le lien (3) qui sont étudiés. Il n'est pas moins vrai que dans les études approfondies de la référence, c'est l'ensemble des liens (1)-(2)-(3) qui est pris en compte ; dans ce cas, la *référence* reste le rapport entre un mot ou un concept, et un *objet*, le plus souvent concret, alors que le *sens* devient soit le nom d'un des liens précédents, soit le rapport entre un mot et la catégorie ou la classe d'objets auxquels ce mot se rapporte.

Enfin, pour prolonger cette idée, on remarquera qu'il est admis que la plupart des mots ont un *sens* (et en tout cas les noms, adjectifs et verbes), mais seule une catégorie plus restreinte de mots sont capables d'évoquer ou de *faire référence* à des objets ou entités, et cela seulement dans des syntagmes *et* des contextes appropriés. Le *sens* est donc comme une potentialité d'un système de signes à se rapporter à un univers extérieur, la *référence* est la capacité concrète d'un signe ou d'une combinaison à évoquer une entité.

V.A.3 La théorie descriptionniste classique et ses limites

Afin de donner un bref aperçu du débat philosophique autour de la référence, nous allons opposer, à la suite de (Alario 1995, Sterelny 1994), deux mouvances : la première est celle initiée par Frege, qui a proposé la première théorie logique moderne de la référence, et elle a subi des critiques variées, organisées dans notre exposé autour de la théorie de Kripke.

V.A.3.i La « référence avec sens » : Frege

Frege est à l'origine d'une théorie dite *descriptionniste*[1] de la référence (1892/1952, 1892/1971), à laquelle nous rattachons aussi les études logiques de Russell (1905/1956) sur le fonctionnement des descriptions définies. Dans cette approche, les expressions linguistiques spécifient un certain nombre de propriétés (justement leur *sens*) que doit satisfaire un objet du monde pour que le lien de référence s'établisse, ces propriétés constituant le « sens » de l'expression. Ainsi, pour prendre un exemple classique, lorsque l'on affirme :

(2) Le roi de France est chauve.

la description définie initiale serait logiquement équivalente à « il existe un unique X tel que RoiDeFrance(X) », avec décomposition du prédicat RoiDeFrance() en certaines propriétés élémentaires. Les expressions référentielles évoqueraient donc des objets parce qu'elles ont un sens (*connotation*) ; lorsque l'ensemble des objets (la *dénotation*) est vide, la phrase totale n'est ni vraie ni fausse, elle est dépourvue de sens, mais d'une façon moins critique que lorsque l'expression est mal formée (Wittgenstein commente ces différents cas de non sens). Certaines critiques ont visé le postulat d'existence et d'unicité de l'objet, qui devrait être infléchi pour tenir compte des informations ou des présuppositions communes des locuteurs, dans une optique plus pragmatique (Strawson 1950).

Dans son ensemble, la théorie réductionniste frégéenne fut largement adoptée, entre autres parce qu'elle répondait à un paradoxe classique de la référence, fondé sur la substitution des expressions coréférentes. Dans ses grandes lignes, il s'agit de savoir pourquoi les exemples suivants ne sont pas équivalents :

(3) Vénus est Vénus.

(4) L'étoile du soir est l'étoile du matin.

L'assertion (3) est une vérité *a priori* alors que l'assertion (4) possède un contenu informatif non nul, bien que les trois expressions référentielles désignent exactement le même objet. On dit que celles-ci ne sont pas substituables *salva veritate* (en préservant la valeur de vérité), contrairement à ce qui devrait arriver s'il s'agissait simplement d'étiquettes. On l'aura compris, la réponse de Frege est que ces « noms » possèdent en réalité un mécanisme référentiel (un « mode de donation de la référence ») différent, qui est le *sens* ou la *connotation*, et qui est susceptible d'une description compositionnelle logique[2].

---

[1] En anglais, elle est souvent qualifiée de « descriptive » ; nous préférons le mot *descriptionniste*, afin de souligner qu'elle n'est pas *descriptive* dans le sens où elle décrirait la référence, mais qu'elle considère les expressions référentielles comme des *descriptions* (plus ou moins abrégées) des entités réelles.

[2] De surcroît, la conception d'une référence avec sens permet de répondre à un deuxième paradoxe, dit de l'« opacité référentielle ». Ainsi, dans les exemples suivants :

Cécile croit que Vénus est une planète.
Cécile croit que l'étoile du soir est une planète.

V.A.3.ii Théorie frégéenne et exactitude des descriptions

Il faut comprendre que la théorie de la référence avec sens défendue par Frege s'étend jusqu'aux noms propres, qui sont eux aussi des raccourcis remplaçant une définition intensionnelle du référent. Ainsi, pour reprendre un autre exemple classique, Aristote est défini par le « philosophe grec », « maître d'Alexandre », « élève de Platon », etc. – on accumule autant de prédicats que nécessaires à l'identification.

Or, on a souvent remarqué que l'usage d'une référence pouvait être tout à fait heureux même lorsque le locuteur ignorait une bonne partie des propriétés définissant l'entité, ou se trompait grossièrement sur certaines d'entre elles. Pour répondre à cette critique, Searle (1958) a proposé d'élargir la définition du référent à un faisceau (ou *cluster*) de définitions, aucune d'entre elles n'étant strictement nécessaire ; on s'ouvrait ainsi à la notion de correspondance partielle entre le sens d'un terme et l'objet à qui il était appliqué.

Un autre défi pour les théories descriptives a été la « distinction de Donnellan » entre usage référentiel et usage attributif d'une expression référentielle (Donnellan 1966). Sans entrer dans les détails, nous citerons l'exemple classique :

(5) Pierre dit : « L'assassin de Jean doit être complètement fou ».

dans lequel Pierre peut avoir à l'esprit deux idées différentes. D'une part, il peut vouloir affirmer que l'assassin, quel qu'il soit (il ne le connaît pas), doit être fou pour avoir agi de la sorte – et c'est l'usage *attributif*. Mais Pierre peut aussi avoir aperçu l'accusé, Jules, au tribunal, et s'exprimer ainsi en parlant de Jules, alors même que sa culpabilité n'est pas certaine et donc que l'expression employée peut être vraie ou fausse – et c'est l'usage *référentiel*, dont la théorie descriptive s'accommode plus difficilement[3]. Une autre objection sera soulevée au sujet des noms propres.

V.A.3.iii La référence directe et les noms propres : Mill, Kripke, Putnam

Kripke (1982), s'inspirant de la théorie plus ancienne de J.S. Mill (1867), supprime le sens dans le lien entre le nom propre et l'entité singulière qu'il désigne, en utilisant les arguments précédents sur le caractère imparfait et non nécessaire des descriptions caractéristiques d'un objet. Kripke s'appuie entre autres sur la validité des énoncés contrefactuels comme :

(6) Aristote aurait pu ne pas devenir philosophe.

et utilise la théorie des mondes possibles pour introduire sa conception des noms propres comme *désignateurs rigides*, i.e. attachés aux mêmes individus dans tous les mondes

---

les deux phrases peuvent aussi avoir une valeur de vérité différente. Les expressions référentielles ne sont pas substituables *salva veritate* après un prédicat modal, car elles possèdent un sens responsable de leur référence, qui se manifeste dans ce contexte modal (ou oblique, ou opaque).

[3] Certains formalismes proposés pour tenir compte de cette distinction s'inspirent soit de la théorie des actes de langage (cf. V.D.4 et (Kronfeld 1986)), soit de la théorie de la pertinence (V.D.5 et (Moeschler et Reboul 1994)).

possibles, et ce sans passage par un *sens*, donc *dénotant* sans *connoter*. Le nom fonctionnant par référence directe est assimilé à une « étiquette » collée une fois pour toutes sur chaque objet singulier, objet supposé garder une identité suffisante à travers les mondes possibles.

Cette théorie qui ne manque pas de force présenterait peut-être un intérêt moins grand si elle concernait seulement les « noms propres » grammaticaux. Il se trouve cependant que la classe des « entités singulières » qu'elle concerne a une étendue non négligeable. Ainsi, Putnam (1973, 1975) étend la théorie à l'ensemble des espèces naturelles (animaux, éléments chimiques, etc.), et d'autres l'appliquent aux notions scientifiques comme « photon » ou « homme de Neandertal ». Pour tous ces aspects, la théorie descriptive d'origine frégéenne semble mise en défaut, au profit des désignateurs rigides.

Il reste néanmoins que l'on peut souhaiter, de la part des deux approches, un meilleur fondement « naturaliste », psychologique ou neurologique. D'une part, la décomposition frégéenne du sens à l'aide d'unités plus élémentaires doit éviter une régression infinie, en fournissant une sémantique des concepts élémentaires ; outre les difficultés de telles décompositions, le résultat est que la théorie de la référence devrait reposer sur une théorie sémantique pour l'heure inexistante. D'autre part, les partisans de la référence directe n'expliquent pas toujours en vertu de quoi les désignateurs rigides sont établis et transmis parmi les locuteurs, et le risque d'une régression infinie existe aussi. Face à cette situation, des solutions de remplacement ont été proposées plus récemment.

V.A.4 En quête de nouvelles fondations

V.A.4.i L'intérêt des « théories causales »

Les « théories causales » sont, selon Devitt et Sterelny (1987), des tentatives d'explication de la référence qui ne relèguent pas le problème de la sémantique référentielle à une théorie fondatrice pour l'heure inexistante, mais visent à élaborer une réduction « naturaliste » de la référence fondée sur d'autres sciences, comme la psychologie, la neurophysiologie, la biologie. Le terme « causal » évoque ici la volonté de décrire la « cause », au sens physicaliste, du rapport entre les mots et les concepts vus comme objets mentaux. Devitt et Sterelny identifient trois directions prometteuses (V.A.4.i.a – V.A.4.i.c) auxquelles nous ajouterons le point de vue de la pragmatique (V.A.4.ii).

V.A.4.i.a Les chaînes causales

L'idée d'une transmission d'individu à individu des conventions de nommage est en fait déjà présente dans une remarque de Kripke (1982 : 84-86) sur le fonctionnement des désignateurs rigides. Typiquement, lorsqu'un objet singulier ou un individu apparaissent pour la première fois à une communauté linguistique, un nom est choisi par convention et attaché par une « cérémonie de baptême » à l'objet. Cet attachement peut

se transmettre ensuite de proche en proche (chaîne causale), sans que des propriétés caractéristiques de l'objet soient nécessairement transmises, comme l'exigerait la théorie descriptive (Devitt et Sterelny 1987).

On peut opposer à ce point de vue le fameux exemple du « *gavagai* » de Quine (1960) : si le locuteur d'une langue inconnue pointe vers un lapin et prononce ce mot, comment savoir si le mot désigne le lapin avec ou sans pattes, les morceaux de lapin, le lapin gris, etc. ? Si, pour répondre, on admet que la transmission des conventions de nommage requiert déjà, de la part des individus, des facultés assez élaborées de perception, conceptualisation et structuration linguistique, cela ne fait que repousser le problème. Il faudrait alors pouvoir décrire aussi un mécanisme d'amorçage ou d'acquisition de ces facultés. Ainsi, la réduction opérée aux sciences naturelles est encore incomplète.

V.A.4.i.b La théorie de l'indication

La théorie de la fonction indicative, défendue par Dretske (1981), réduit le lien entre concept et symbole à leur corrélation stable pour les individus d'une communauté donnée. Le mécanisme perceptif sous-jacent serait donc responsable de la bonne reconnaissance, en moyenne, de l'appartenance de différents objets à une même catégorie, et donc de l'attachement d'un symbole constant à celle-ci. On s'éloigne ici du strict problème de la référence, pour aborder le problème plus général de la sémantique des mots, ou liens mots/concepts et concepts/percepts.

V.A.4.i.c Théories fonctionnalistes et téléologiques

Enfin, le mécanisme de formation des concepts a fait aussi l'objet d'explication « darwiniennes », qui tentent d'utiliser le mécanisme de la sélection naturelle pour expliquer la genèse (la cause) des liens entre concepts et percepts (Millikan 1984, Papineau 1987). Un point de vue analogue, bien que moins reconnu en philosophie, mais sur lequel nous aurons l'occasion de revenir, est celui de Derek Bickerton (1990). Dans l'ensemble, si le mécanisme proposé par ces théories peut être accepté pour l'émergence de concepts élémentaires – tels les « concepts » ou les représentations que se font les animaux de leurs ennemis, de leurs congénères ou leurs petits – il reste que le système conceptuel humain, complexe et surtout compositionnel, ne reçoit pas une description à la mesure de sa finesse de représentation.

V.A.4.ii Théories de la référence située

Un autre ensemble de théories, à visée moins ontologique que les précédentes, tente de prendre en compte l'importance du contexte dans l'usage de la référence. Ces théories reprennent en substance le *credo* de la théorie descriptive frégéenne, et tentent de le moduler en y introduisant des contraintes dépendant des intentions des utilisateurs, et/ou du contexte illocutoire.

V.A.4.ii.a Actes de langage et référence

La théorie des actes de langage développée par Austin (1962) et surtout par Searle (1969) a été reprise dans de très nombreux travaux philosophiques, linguistiques et informatiques, comme nous l'avons montré au II.D.1.i.a. Aux conceptions logicistes, véri-conditionnelles, ou informationnelles, la théorie des actes de langage vient ajouter une dimension supplémentaire, celle des prédicats *performatifs* qui précisent le mode sur lequel on doit considérer l'information transmise : promesse, ordre, requête, déclaration, assertion.

Cette théorie novatrice modélise des caractéristiques essentielles du contexte d'énonciation, mais à ses origines le problème de la référence n'y est pas central. Même des formalisations logiques récentes de la théorie elle-même (Vandervecken 1990a, 1990b) manquent de s'intéresser au problème de la référence, comme le fait remarquer Gérard Sabah (1997). Nous allons décrire ci-après une étude liant les choix des expressions référentielles aux intentions des locuteurs (Appelt et Kronfeld 1987). Par ailleurs, Kronfeld (1986) a analysé la distinction de Donnellan entre usage référentiel et attributif dans un cadre inspiré des actes de langage. Nous présenterons aussi plus bas (V.D.4) quelques systèmes de génération d'expression référentielles prenant en compte les actes de langage.

V.A.4.ii.b Référence, actes de langage et action rationnelle

Appelt et Kronfeld (1987) présentent une formalisation intéressante de l'usage de la référence, dans le cadre de la théorie de actes de langage et celle de l'action rationnelle (Cohen et Levesque 1985). Il s'agit d'une théorie représentationniste, qui postule l'existence chez l'individu ou agent de *représentations intensionnelles d'objets (RIO)*, pouvant être perceptives, discursives ou fonctionnelles ; chaque RIO peut dénoter une entité du monde. Un ensemble de RIO maximales dénotant la même entité, pour un agent donné, forment un *ensemble individualisant (EI)*, qui est donc une représentation complexe, subjective, d'un référent dans le cas où l'ensemble de RIO est « cohérent » (différentes aberrations sont évoquées : absence de référent réel, mélange de RIO dénotant deux entités réelles, deux EI pour la même entité).

Le fonctionnement des expressions référentielles dans une situation de communication est décrit par le *schéma référentiel* suivant (décrit d'abord en termes formels) :

> *Si le locuteur et l'interlocuteur croient mutuellement que U est un énoncé propositionnel et que GN est une partie de U et une expression référentielle potentielle, alors le locuteur a pour but que l'interlocuteur croie deux choses : (1) que le locuteur possède un EI représentant l'entité correspondant au GN, et (2) que le locuteur désire que l'interlocuteur croie que le contenu descriptif du GN est vrai de l'entité représentée par son EI.* (Appelt et Kronfeld 1987:644)

La référence serait donc un acte de langage, ayant un but et des conditions de succès. D'autres définitions et propositions viennent étoffer la théorie, qui rend compte de

certains propriétés de la référence. Tout d'abord, elle délimite les situations où la définition intuitive suivant est valable : *un agent réfère lorsqu'il possède une représentation mentale d'un objet, et désire qu'un autre agent en possède une aussi*. La théorie montre aussi comment une référence peut réussir alors que la description utilisée est crue fausse par l'interlocuteur (usage *référentiel* selon Donnellan), et montre enfin que l'hypothèse des noms standards (désignateurs rigides dans tous les mondes possibles) peut être remplacée par les RIO et les EI.

La théorie d'Appelt et Kronfeld a le mérite de décrire la question de la référence dans le cadre plus général de la théorie de l'action rationnelle. D'un point de vue ontologique, la question de la « naturalisation » est donc renvoyée vers cette dernière théorie, qui devrait être justifiée sur des bases physicalistes. En ce qui concerne la modélisation informatique, cette théorie convient aux modélisations à base d'agents rationnels, qui sont, nous l'avons vu (II.D.1), assez peu linguistiques. En revanche, son application à des animats (comme au II.C.1 ou II.D.2) suppose d'avoir implémenté la théorie de l'action rationnelle.

V.A.4.ii.c Pertinence et référence

La récente théorie de la pertinence (Sperber et Wilson 1986) a donné lieu aux propositions d'Anne Reboul (Moeschler et Reboul 1994:148-153), sur lesquelles nous reviendrons plus en détail au V.D.5.

Selon Sperber et Wilson, il existe un véritable « organe linguistique périphérique », qui produit une « forme logique » pour chaque énoncé perçu. Cette forme logique, que les auteurs distinguent de la « forme propositionnelle », est élaborée par le module périphérique à l'aide de capacités phonologiques, syntaxiques, sémantiques. Mais la forme logique n'est pas contextualisée, en particulier les différents référents ne sont pas intégrés dans la représentation, qui utilise à leur place des sortes de variables. Le rôle de l'« unité centrale », intégrative, de l'esprit sera d'utiliser les informations liées à la situation d'énonciation pour lier les variables de la forme logique aux représentations des référents, afin de rendre celle-ci utilisable[4].

Ce processus d'attribution des référents se fait par l'élaboration et l'infirmation d'hypothèses faisant intervenir la forme logique et l'*environnement cognitif mutuel*, afin de satisfaire aux différentes conditions du *principe de pertinence* proposé par Sperber et Wilson[5]. Nous retrouvons donc un écho de la conception frégéenne de la référence avec

[4] Cette conception modulaire s'inspire des travaux de Fodor (1983), et fait écho à l'organisation proposée par Chomsky (cf. notre **II.A.1.iv**). Ce schéma sera aussi utilisé pour notre modélisation de la référence dans un agent (au VI.D.3).

[5] Résumons brièvement les bases du principe de pertinence.

(1) Un acte de communication est toujours *optimalement pertinent* (à savoir 2 et 3).

(2) Le contenu communiqué est suffisamment *pertinent* (à savoir 4 et 5) pour être traité.

(3) La forme utilisée pour la communication est la plus *pertinente* (à savoir 4 et 5) possible.

(4) Plus un énoncé produit d'effets contextuels, plus il est dit *pertinent*.

(5) Moins un énoncé demande d'efforts de compréhension, plus il est dit *pertinent*.

sens : les variables de la forme logique peuvent dénoter un certain nombre de référents (descriptions incomplètes) – mais le choix entre les différentes options doit maximiser un principe non logique, le principe de pertinence. Naturellement, la prise en compte par Sperber et Wilson de la communication « ostensive-inférentielle » dans toute sa généralité (incluant par exemples les désignations gestuelles ou *déictiques*), s'éloigne considérablement du paradigme frégéen.

V.A.4.ii.d Autres travaux : Evans, Récanati

Le livre de François Récanati (1993) s'inscrit lui aussi dans le débat classique entre référence avec sens et référence directe, sous la distinction plus complexe *de dicto / de re*. L'auteur se place dans un cadre clairement représentationniste, car il prend en compte les représentations des entités du discours pour chaque locuteur ; il distingue plusieurs niveaux d'information dans ces représentations, notamment les informations descriptives (de facture encyclopédique, connaissances sur le monde) et les informations perceptives (les deux dans la modalité référentielle *de re*). D'autre part, l'auteur revendique l'apport des théories pragmatiques, puisqu'il met en évidence trois niveaux de sens qui concourent à la détermination des référents : le sens de la phrase (« meaning »), ce qui est dit par la phrase, et ce qui est communiqué. Le côté computationnel de l'ouvrage se manifeste dans l'utilisation du nom de *dossiers d'objets* pour les représentations des référents, dont nous verrons l'origine plus bas (V.B.2.iv, V.B.3.ii).

L'utilisation de représentations conceptuelles pour les référents, contenant entre autres des informations perceptives, est aussi suggérée par Gareth Evans (1985). Ces représentations peuvent avoir une existence purement mentale, dans un espace construit par le locuteur, et bien distinct de la réalité, ce qui évoque la théorie des espaces mentaux de Fauconnier (cf. V.B.3.iii.a ci-dessous), où toutefois il n'est fait aucune mention d'un ancrage perceptif des représentations des référents.

V.A.4.iii Conclusion sur les théories philosophiques

On le voit, les dernières approches exposées semblent plus concernées par la résolution de la référence (le « comment ») que par la nature même du phénomène (le « qu'est-ce que c'est », ou l'ontologie). Nous allons continuer à nous rapprocher des études de la résolution de la référence en nous tournant d'abord vers les théories linguistiques, dans la section suivante. Mais avant cela, que peut-on conclure de notre passage en revue des théories philosophiques de la référence ?

Les théories précédentes ont le mérite de situer, au delà des controverses d'école, les principaux éléments qui ancrent le langage à la réalité matérielle environnante, notamment le statut de la référence par rapport au sens. Dans notre quête d'une implémentation informatique qui rende justice au phénomène du sens, nous constatons que la faculté de faire référence est *la manifestation concrète* du système conceptuel ou

---

On peut se demander si l'optimalité évoquée peut être exprimable mathématiquement ; l'enchaînement des principes entraînerait une certaine complexité formelle.

sémantique du locuteur. Il est donc essentiel de pouvoir modéliser le phénomène de la référence dans une société d'agents communicants. Pour ce faire, l'utilisation des théories plus situées, pragmatiques, est certes à privilégier, alors que les théories « descriptives » seules conviennent plutôt à l'approche symbolique du langage naturel.

Notre examen nous a permis aussi de circonscrire la gamme des problèmes à résoudre, afin précisément de nous assurer qu'aucun élément essentiel à notre appréhension de la référence n'a été oublié – et cela justifie sa longueur. La tendance qui s'en dégage, pour le domaine, est un appel plus fort aux sciences de la vie, qui semblent prendre le relais de la logique invoquée il y a un siècle. À ce titre donc, notre modélisation d'inspiration neurobiologique (VI.D.3) s'inscrit dans la problématique actuelle de « naturalisation de la référence » (Devitt et Sterelny 1987).

V.B Théories linguistiques de la référence

Nous regroupons sous le terme « linguistique » des théories de la référence qui, s'intéressant moins à la nature exacte du phénomène, étudient le comportement des expressions référentielles dans le langage. Ces théories se restreignent souvent aux rapports induits par la référence entre différentes expressions linguistiques.

Dégager des lois linguistiques par l'étude d'assertions détachées de leur contexte peut suffire en morphologie ou syntaxe, mais déjà moins en sémantique. Or, nous venons de voir que la question de la référence dépasse le seul système de la langue. Elle peut cependant se « projeter » au niveau linguistique sous forme de *coréférence* – relation entre deux expressions ayant le même référent. Les théories exposées ci-après tentent donc d'exprimer des critères linguistiques concrets de coréférence entre mots ou syntagmes. Leurs limites nous amèneront à l'étude de modèles tenant compte du contexte.

V.B.1 Chomsky : de la notation indicielle au liage

L'exemple extrême d'approche purement linguistique se trouve dans les travaux novateurs de Chomsky. Plus encore que Saussure avec les systèmes autonomes de signifiants et de signifiés, Chomsky détache l'objet de la linguistique de l'univers matériel (le monde des référents) et en fait un système formel dont il s'agit de dégager les règles internes dans un formalisme quasi mathématique (Chomsky 1995).

V.B.1.i Grammaire génératives et anaphore

Les travaux de Chomsky se limitent aux aspects les plus formels de la syntaxe du langage isolé du monde (phrases bien formées dépourvues de contexte). Les relations de coréférence étaient représentées à l'aide d'indices dans les premiers travaux de la grammaire générative (Chomsky 1957, 1965). Les mots de même indice sont coréférents, (* désigne une phrase inacceptable, à cause des coréférences) :

(7) *Cécile*$_i$ est arrivée en retard parce qu'*elle*$_i$ avait égaré sa montre.

(8) *$Elle_i$ est arrivée en retard parce que $Cécile_i$ avait égaré sa montre.

(9) $Elle_i$ est arrivée en retard parce que $Cécile_j$ avait égaré sa montre.

(10) Parce qu'$elle_i$ avait égaré sa montre, $Cécile_i$ est arrivée en retard.

La notation indicielle, critiquée plus tard pour la simplicité à laquelle était réduite la notion de référence, permettait néanmoins à Chomsky d'aborder un problème central pour la syntaxe, et qui apparemment ne pouvait déjà pas s'y réduire : le problème de l'anaphore pronominale, à savoir les diverses restrictions de l'usage des pronoms, dont l'exemple (10) donne un faible aperçu. Une théorie de l'anaphore pronominale devait donc expliquer pourquoi certaines co-indexations entre un pronom et un *antécédent* étaient possibles et d'autres non[6].

Le projet chomskyen d'une théorie purement syntaxique de l'anaphore pronominale était naturel dans le cadre de la grammaire générative, mais aussi quelque peu illusoire. En effet, Tesnière (1965) déjà dans son ouvrage sur la syntaxe structurale appelle l'antécédent « source sémantique » de l'anaphore, en faisant remarquer le peu de contraintes syntaxiques liant les deux termes de l'anaphore. Selon Tesnière, l'anaphore correspond à une relation extra-syntaxique, i.e. une connexion supplémentaire, d'ordre sémantique. L'anaphorique par lui-même est vide de sens (ce point sera repris dans la théorie linguistique de Milner (1989)) et c'est son antécédent ou « source sémantique » qui le lui donne.

V.B.1.ii Théorie du gouvernement et du liage

Les travaux générativistes ont très justement évolué vers une restriction de la portée des contraintes syntaxiques sur les pronoms. Dans le cadre de la théorie du gouvernement et du liage (Chomsky 1981), intégrée plus récemment à la théorie « principes et paramètres » (Chomsky 1995), la définition même d'un anaphorique est beaucoup restreinte (voir aussi (Haegeman 1991)). Ainsi :

> *Les GN exprimés se divisent en trois catégories : anaphoriques (réflexifs, réciproques), pronoms, et r-expressions (John, la racine carrée rationnelle de 2, et autres expressions 'quasi-référentielles'…)* – (Chomsky 1995:40)

La dimension syntaxique de la coréférence prend désormais des proportions relativement réduites[7]. Cette révision rend justice à la complexité réelle de la référence et

[6] L'antécédent désigne l'expression du texte avec laquelle le pronom est coréférent. Dans la terminologie utilisée, le pronom (dit anaphorique, ou en anglais « anaphor ») devait toujours être lié à un antécédent par un lien d'anaphore (en anglais « anaphora »).

[7] Les principes de liage (Chomsky 1995:93-96) qui correspondent à ces trois catégories font appel à la notion de *c-commande*, fondée sur la représentation de la phrase sous forme d'arbre syntaxique.

A. Un anaphorique doit être lié dans son domaine local.

B. Un pronom doit être libre (non lié) dans son domaine local.

C. Une r-expression doit être libre.

avec les définitions auxiliaires :

à son ancrage essentiel dans un monde extra-linguistique. Pourtant, dans les années '70, un grand nombre de travaux logiques ont été développés pour répondre au problème des coréférences tel qu'il était posé par Chomsky, comme nous allons le voir.

V.B.2 Approches « logicistes » de la coréférence

Le paradigme chomskyen a inspiré des travaux linguistiques partageant certains points communs : l'utilisation de la grammaire générative *et* de la logique, l'étude d'exemples linguistiques en général détachés de leur contexte et des locuteurs, un rôle privilégié accordé aux pronoms, et donc de façon générale une adoption tacite du formalisme « descriptiviste » frégéen – d'où notre qualification de « logicistes ».

Trois contributions illustrent de façon remarquable l'évolution ultérieure de la question : les travaux de Lasnik (1976) assurant la continuité du point de vue chomskyen, les travaux de Fauconnier (1974), précurseurs de son orientation vers la grammaire cognitive et les « espaces mentaux », et les travaux de Karttunen (1969-76), posant des problèmes qui trouvent leur écho dans l'approche pragmatique du projet CERVICAL (cf. V.D.5 ci-dessous)[8]. Mais auparavant, il nous faudra décrire les deux approches de la coréférence qui se dessinaient au début des années 1970.

V.B.2.i Théorie transformationnelle et théorie interprétative

Le début des années 1970 voit l'approche générativiste de la (co)référence mise en question par un certain nombre de contre-exemples ou paradoxes, auxquels tente de répondre une approche plus sémantique et logique.

La théorie transformationnelle ou classique s'intéresse particulièrement à la « coréférence » entre les pronoms et leurs « antécédents », de façon presque exclusive à l'intérieur d'une même phrase. La conception qui prévaut est que dans la *structure profonde* (l'actuelle D-structure), il n'y a pas de pronoms : chaque feuille ou sous-arbre correspondant à un « référent » contient une description adéquate et suffisante du référent, par exemple un même groupe nominal (GN). Lors de la génération de la *structure de surface*, certains de ces sites sont substitués par des pronoms, qui se rattachent par des liens d'anaphore aux GN exprimés qu'ils désignent. Les GN possèdent les traits [± spécifique] et [± défini], qui ont une influence sur le choix et les transformations des pronoms. Enfin, dans la *phrase*, les coréférences sont marquées partout à l'aide d'indices (cf. exemples 9–12).

Cette approche a été contestée sur un plan grammatical (et non de principe) par Emmon Bach et James McCawley, entre autres auteurs (Bach 1968, 1970, McCawley 1968). Ceux-ci montrent que la théorie classique n'est pas capable de dériver certaines

---

• α « lie » β si α c-commande β, et α, β ont le même indice.

• α c-commande β si le nœud α ne domine pas le nœud β, et tout γ qui domine α domine aussi β –

Voir aussi le livre récent, en français, d'Anne Zribi-Hertz (1996).

[8] Les travaux plus récents de Lauri Karttunen, par exemple sur l'analyse syntaxique à l'aide d'automates, ne semblent pas être une extension de ces travaux plus anciens sur la coréférence et les présuppositions.

phrases bien formées, dès que celles-ci font intervenir à côté des pronoms des phrases relatives quelque peu complexes :

(11) L'homme qui la recherche aura la vie qu'il désire.

Nous ne rentrerons pas dans le détail de cette argumentation, qui porte le nom de paradoxe de Bach-Peters.

La solution de remplacement proposée considère les expressions référentielles (pronominales ou non) comme des variables, reprenant en cela les indices de coréférence introduits par Chomsky. Le point novateur est la soumission de ces variables à des quantificateurs logiques, qui rendent superflus les traits [± *spécifique*] et [± *défini*] (Karttunen 1976:378). Il s'agit donc d'une théorie « logiciste », qui permet de rendre compte de coréférences croisées ou enchâssées parfois assez subtiles. Cette théorie « interprétative », connue sous le nom de « sémantique générative », compta parmi ses défenseurs G. Lakoff, R. Langacker et G. Fauconnier (cf. aussi V.B.3.iii ci-dessous), et imprima sa marque sur la DRT (V.B.3.i).

V.B.2.ii Lasnik (1976) et « l'orthodoxie générativiste »

Le point de vue de Howard Lasnik illustre bien la réaction du courant génerativiste aux critiques précédentes, qui consista à focaliser l'étude de la référence sur la question de la pronominalisation. Dans son article « Remarks on Coreference » (1976), Lasnik part de l'analyse d'un exemple simple :

(12) Jean prétend qu'il est malade.

Ici, *Jean* peut être coréférent avec *il*, et la syntaxe aura à rendre compte des contraintes de cette relation, ou bien ne pas l'être. Dans ce dernier cas, l'approche interprétative affirme que le référent aura été mentionné antérieurement dans le discours, ou désigné dans le contexte (deixis). Le point de vue de Lasnik, suivant en cela (Postal 1966), c'est que dans ce dernier cas la source du référent (le mode de donation, en termes frégéens) n'est pas du tout du ressort de la « grammaire », mais relève de principes plus généraux de coopération dans la communication ; de nombreux travaux plus récents rendent justice à cette distinction (V.D.4 et V.D.5).

Lasnik restreint ainsi le champ de ses études, en excluant le problème de l'identification des expressions coréférentes, et en critiquant aussi la théorie interprétative. Celle-ci serait incapable d'expliquer la différence entre (13) et (14) ci-dessous, et le terme même de coréférence serait abusif lorsque les pronoms ne réfèrent pas , comme en (15) ou (16) (nous traduisons) :

(13) *Le fait qu'il_i avait perdu chagrinait *chaque candidat*$_i$.

(14) *Chaque candidat*$_i$ était chagriné par le fait qu'il_i avait perdu.

(15) *Qui*$_i$ croit qu'il_i peut accomplir cette tâche ?

(16) *Personne*$_i$ ne pense sérieusement qu'il_i est stupide.

Le point essentiel et novateur de cet article de Lasnik est que certains pronoms doivent être présents dans la structure profonde (D-structure), leur usage étant régi par les principes susmentionnés. Ainsi, le GN coréférent doit être récupérable ailleurs dans la D-structure ; cela représente un changement notable au niveau conceptuel de la théorie chomskyenne.

V.B.2.iii Fauconnier (1974) : « syntaxe ou sémantique ? »

Le titre de ce livre de Gilles Fauconnier (1974) marque déjà une distance par rapport à l'approche générativiste. Le livre débute par une comparaison entre la théorie transformationnelle et la théorie interprétative sur le problème de la pronominalisation. Fauconnier montre les avantages de la seconde théorie, capable de représenter correctement l'exemple (17) à l'aide de la forme logique suivante :

(17) Tous les concurrents espèrent qu'ils vont gagner
$\forall x$ (x espère (x gagner)) [et non bien sûr $\forall x$ (x espère ($\forall y$ (y gagner)))].

En revanche, l'interprétation par variables se révèle circulaire dans :

(18) Jean embrasse les filles qui *le* lui demandent.

Après avoir aussi analysé le problème des description définies, Fauconnier propose une théorie « allégée », qui prend en compte l'« histoire dérivationnelle » d'un arbre syntaxique dans la représentation des coréférences[9]. Selon l'auteur,

> *Pour devenir des pronoms* [en partant d'un arbre syntaxique complet ou saturé] *les SN vides doivent recevoir des traits à partir des syntagmes pleins du même réseau de coréférence ; ceci est réalisé par la règle de copiage* [proposée par l'auteur]. *De la même façon, un adjectif reçoit des traits à partir d'un syntagme coréférentiel avec celui qui déclenche l'accord.*

Sur un plan plus général, synthétisé dans son introduction, le livre étudie la possibilité de l'élimination de la coréférence du domaine de la syntaxe, pour l'affecter exclusivement au domaine de l'interprétation sémantique. La réponse de l'auteur est que cela est impossible, et sa conclusion est que la distinction traditionnelle entre syntaxe et sémantique n'est pas viable. À ces deux branches devrait se substituer une théorie cognitive plus intégrée. Fauconnier s'oppose donc à Lasnik, comme on le voit, en ce qu'il conteste l'autonomie du langage par rapport au « système interprétatif », ce qui donnera naissance notamment aux principes de la grammaire cognitive (Langacker 1991).

V.B.2.iv Karttunen (1969-76) : paradoxes de la logique des référents

Nous allons exploiter un important article de Lauri Karttunen (1976) pour exemplifier à la fois les limites de la théorie interprétative (V.B.2.iv.a) et montrer en quoi le problème abordé garde son actualité (V.B.2.iv.b). Dans son ensemble, l'article analyse

[9] Cette théorie permet à Fauconnier de décrire aussi les phénomènes d'accord du français dans un formalisme unifié, de nature logico-grammaticale, et non réductible à la seule syntaxe.

un certain nombre de cas où un GN indéfini introduit un nouveau référent, et discute aussi la question de la spécificité des référents.

V.B.2.iv.a Des paradoxes logiques poussés à l'extrême…

Afin de montrer la supériorité représentationnelle de la théorie interprétative ou « sémantique générative », Karttunen développe avec beaucoup de finesse un exemple que la théorie indicielle chomskyenne ne peut expliquer. Karttunen considère les ambiguïtés liées à la spécificité ou non du *musée* dans l'exemple :

(19) Bill intends to visit *a museum* every day.
[Bill désire visiter *un musée* tous les jours.]

La théorie transformationnelle utilise le trait [± *spécifique*] pour caractériser le statut du mot *musée*, et cela conditionne son éventuelle reprise par un pronom ; il n'y aurait donc que deux possibilités. Or, Karttunen montre que (19) possède trois interprétations possibles (du moins en anglais). Le trait [± *spécifique*] est donc insuffisant. En revanche, le formalisme logique de l'autre approche permet bien d'exprimer les trois sens, que nous paraphrasons ci-dessous.

(20) Il y a un musée que Bill désire visiter tous les jours.

(21) Bill désire qu'il y ait un musée qu'il visite tous les jours.

(22) Bill désire faire une visite de musée par jour.

Au delà de la clarté de l'argument, se pose bien sûr le problème de l'utilisation réelle d'une telle assertion, qui nous paraît très improbable, surtout avec les interprétations (22) et plus encore (21), assez difficile à percevoir. Cet exemple fonctionne parce que l'auteur *force les interprétations* de son énoncé ; la réaction habituelle à un tel énoncé serait probablement l'incompréhension, et il paraît improbable qu'un individu même averti puisse facilement trouver les trois interprétations. De plus, il est loin d'être certain que les trois interprétations seraient exprimées de la même façon dans des langues très différentes (même la traduction française semble moins convaincante que l'original anglais) – nous verrons plus loin que ce n'est probablement pas le cas (cf. surtout le VI.B.3). Nous allons donc nous tourner vers ce qui, dans l'article de Karttunen, paraît applicable à une analyse de la référence dans son *usage courant et indépendant de la langue étudiée.*

V.B.2.iv.b … mais une problématique qui garde son actualité

L'article de Karttunen (1976) pose le problème de la résolution de la référence en des termes très généraux. Dans le cadre du projet CERVICAL (V.D.5), Anne Reboul a attiré notre attention sur le début de l'article cité, dont voici un extrait :

> *Considérez un système conçu pour lire un texte dans une langue naturelle, l'interpréter, et stocker d'une certaine façon le contenu. [Un tel système] doit être capable de construire un fichier contenant des entrées [fiches] pour tous les individus, c'est-à-dire événements, objets, etc., mentionnés dans le texte,*

fichier qui, pour chaque individu, enregistre ce qui est dit de lui. Certes, pour l'heure, il semble qu'un tel système de compréhension ne soit pas à notre portée. […] Dans cet article, je souhaite discuter une des caractéristiques qu'un tel système doit avoir : il doit être capable de reconnaître quand une entité nouvelle est mentionnée dans le texte et la stocker, en même temps que sa caractérisation, en vue des références à venir.

Pour la première fois peut-être, un principe computationnel concret est énoncé pour la résolution de la référence dans la compréhension de textes[10]. Quoique général, ce principe préconise l'usage de représentations complexes pour les référents du discours, et en esquisse le fonctionnement. Le problème de la création d'une nouvelle « fiche » ou représentation est en effet essentiel, et il l'est resté depuis, même si les critères ainsi que le formalisme utilisés ont pu changer.

Plusieurs aspects du contexte de production sont distingués (en introduction) par Karttunen, notamment les contextes locaux des interlocuteurs[11] :

[Un tel] système d'interprétation de textes doit être capable d'assigner une valeur de vérité à la proposition […] dans laquelle le GN apparaît. Il doit être sensible aux propriétés sémantiques des verbes ayant pour compléments des phrases relatives ; distinguer entre l'assertion, l'implication et la présupposition ; et en final distinguer ce qui existe pour le locuteur de ce qui existe seulement pour les autres. […] Il faut distinguer entre le monde du locuteur et d'autres mondes et admettre qu'ils ne contiennent pas les mêmes référents.

Enfin, on trouve aussi la suggestion que la structuration du discours a une influence sur l'existence et l'accessibilité des référents (§1.3, *Short Term Referents*), idée qui sera développée dans de nombreuses théories postérieures (V.D.2). Ainsi, Karttunen montre à l'aide d'exemples comment un GN indéfini peut ne pas introduire un référent stable (ou existant réellement), mais un référent temporaire (ou hypothétique), présent tant que le discours continue dans la même modalité présuppositionnelle (mais aucune formalisation logique n'est donnée).

En conclusion, nous pouvons dire avec l'éditeur du recueil, James McCawley, que le travail de Karttunen annonce un véritable programme de recherche, qui est largement sorti du cadre générativiste ou logique auquel sa période semblait le limiter.

[10] Ce principe, de par sa généralité, se retrouve sous une forme ou sous une autre dans de nombreux programmes de compréhension de textes postérieurs, fonctionnant souvent sur des domaines restreints, tel le système réalisé par Gérard Sabah (1978).

[11] Une approche qui prolonge celle de Karttunen sur ce point précis se trouve dans (Clark et Marshall 1981), qui analysent l'influence des connaissances mutuelles des interlocuteurs sur l'emploi de la référence définie.

V.B.2.v Webber (1979) : formalisme logique et discours

Bonnie Lynn Webber (1979) propose une modélisation logique détaillée de certains phénomènes référentiels, exposée dans sa thèse de doctorat et quelques articles connexes. À nos yeux, sa modélisation logique emprunte beaucoup aux idées énoncées bien avant par Karttunen (peut-être insuffisamment cité par Webber), tout en apportant des analyses logiques originales. Le point plus novateur est en revanche la prise en compte du rôle du discours dans le formalisme logique lui-même.

Webber suppose que chaque phrase à analyser est convertie en une forme logique (avec des variables, prédicats, quantificateurs), à laquelle s'appliquent des règles logiques permettant de dégager l'ensemble des entités que la phrase évoque (auxquelles elle fait référence), ainsi que leurs propriétés. Les principaux apports de Webber sont ici les analyses fines de certains pronoms indéfinis, dans des constructions complexes comprenant des quantifications, comme :

(23) Each boy gave each girl a peach. [Chaque garçon donna à chaque fille une pêche]

Dans cet exemple, le but est de caractériser l'ensemble des pêches, et la forme logique résultante contient bien le fait que chacune a été donnée à une fille par un garçon.

Pour ce qui est de la prise en compte du contexte, Webber affirme que sa théorie vise la représentation linguistique des entités que les interlocuteurs possèdent dans leur « modèle du discours », ou représentation de la situation. Webber affirme, de façon assez discutable (Webber 1981:283) :

> *1. Un des objectifs du discours est de permettre à un locuteur de communiquer à un interlocuteur un modèle qu'il possède d'une situation.*
> *2. Un tel modèle du discours peut être vu comme une collection d'entités.*
> *3. Un locuteur utilise une anaphore définie pour se référer à une entité qui est déjà dans son modèle du discours. Ce faisant, le locuteur suppose sur la base du discours déjà exprimé qu'une entité semblable se trouve dans le modèle du discours de l'interlocuteur […]*
> *4. […] Les entités du discours peuvent avoir les propriétés des individus, ensembles, événements, actions, faits, etc.*

On est certes surpris par cette « définition cognitive » de l'anaphore, qui s'oppose aux vues génératives. En effet, si l'on distingue d'un côté les expressions qui se réfèrent aux entités existant déjà dans l'esprit de l'interlocuteur, et de l'autre celles qui introduisent de nouvelles entités, il reste celles évoquées par association d'idées. Si une expression ultérieure se réfère à une telle entité, sera-t-elle anaphorique pour ceux qui ont fait l'association, et non anaphorique pour les autres ?

Webber étudie aussi les entités « évoquées naturellement » (non anaphoriquement) par des expressions, i.e. introduites par référence. Le point original est le rattachement de la représentation logique des entités à l'assertion qui les a introduites ; accompagnée d'une « quantification restreinte » définie par Webber, cela permet de représenter sous forme logique des phrases comme :

(24) Chaque homme qui possède un âne le bat

en écrivant une formule logique qui peut être paraphrasée par :

(25) « Chaque homme qui possède un âne, bat l'âne qu'il possède et que l'on a mentionné. »

Plutôt que le cadre logique lui-même (voir aussi l'analyse critique de (Hirst 1981b)), nous retiendrons donc la tentative de modélisation des univers de discours, qui, à l'exception du point contesté plus haut, relie de façon novatrice et assez consensuelle les approches réductrices logiques ou « textuelles » à la situation concrète d'énonciation.

V.B.3 Théories formelles plus récentes

V.B.3.i La DRT de Kamp (1981, 1993)

La « Discourse Representation Theory » (Théorie de Représentation du Discours ou DRT) est un formalisme logico-sémantique (Kamp 1981a, Kamp 1981b, Kamp et Reyle 1993) qui permet de représenter les référents, parmi d'autres propriétés d'un discours. Le formalisme logique des variables et quantificateurs est enrichi par l'introduction de « Discourse Representation Structures (Structures de Représentation du Discours ou DRS) qui regroupent les référents et les propriétés prédiquées. Ces DRS sont soumises à des relations d'inclusion, ce qui permet de représenter l'accessibilité (ou non) des référents d'un univers de discours à un autre, par exemple dans le cas de contextes négatifs ou modaux. L'évolution des DRS est déduite de l'analyse logique (syntaxique et sémantique) des phrases, ce qui peut être soumis aux mêmes objections que la théorie frégéenne. Son mérite est, à nos yeux, d'avoir organisé et formalisé l'appareil logique afin de le relier de façon fiable aux différents ensembles parmi lesquels le mécanisme de compréhension doit choisir les référents des expressions – sur la base des considérations traditionnelles sur les articles défini et indéfini, par exemple.

La prise en compte de facteurs plus élaborés dans l'évolution des DRS, comme le contexte illocutoire et les croyances/intentions serait déjà un pas vers plus de réalisme, en attendant une représentation plus riche des référents, englobant des connaissances sémantiques préexistantes (dites « encyclopédiques »), et leur importance pour chacun des locuteurs. Mais de telles extensions n'ont pas encore été réalisées, malgré les développements autour de la DRT. Parmi ceux-ci, il faut citer ceux de Nicolas Asher (1993, 1997), qui étudient la représentation dans le cadre de la DRT des entités abstraites ou des entités qui évoluent dans le temps ; cela se traduit, dans ses travaux, vers une prise en compte plus importante du discours, sans pour autant aller jusqu'aux facteurs situationnels proprement dits.

V.B.3.ii Thèse de Heim (1982)

La thèse de Heim (1982) sur la sémantique des groupes nominaux définis et indéfinis représente une tentative plus isolée et de moindre portée que la DRT ou la grammaire cognitive (ci-après). Heim s'inspire de la métaphore du fichier contenant des fiches pour chaque entité du texte (introduite par Karttunen), et propose assez classiquement des critères de création et de renouvellement de ces fiches en fonction des

expressions référentielles. Son modèle est plus élaboré que celui de Karttunen en ce qu'il prend en compte l'évolution des référents par le biais d'une réactualisation de chaque fiche à chaque nouvelle expression référentielle qui lui est attachée, réactualisation déclenchée par une estimation de la cohérence de chaque fiche.

V.B.3.iii La grammaire cognitive

La grammaire cognitive s'est élaborée dans les années 1980, notamment à l'Université de Californie à San Diego. Sans prétendre résumer l'ensemble important de ces travaux, nous pouvons dire que leur but est en général la modélisation des *représentations cognitives des humains* à l'aide de différents formalismes, afin d'expliquer par ce biais certaines aptitudes cognitives comme le langage, voire plus récemment les mathématiques (Lakoff et Nunez 1997, à paraître). En ce qui concerne le lien entre représentations cognitives et langage, et son application à la question de la référence nous aborderons deux travaux importants : celui de Fauconnier (1984) sur les espaces mentaux, et la synthèse de Langacker (1991) sur la grammaire cognitive.

V.B.3.iii.a Les espaces mentaux de Fauconnier (1984)

Dix ans après l'ouvrage sur la coréférence déjà cité (Fauconnier 1974), Fauconnier propose une modélisation des processus cognitifs à l'aide des « espaces mentaux » (Fauconnier 1984). Ces espaces sont à concevoir comme autant de zones de l'esprit, reliées entre elles, peuplées de représentations d'objets du monde extérieur, avec leurs propriétés et leurs relations. Le langage permet non seulement de créer ou modifier les représentations, mais surtout d'introduire de nouveaux espaces associés au contexte de chaque discours. La question de la référence revient alors à décrire les moyens par lesquels une expression linguistique est attachée à une représentation dans l'espace mental approprié.

Sans entrer dans les détails, il nous semble que le principal mérite de cette théorie est l'explication de certains phénomènes référentiels complexes, comme certaines métonymies, ou certaines anaphores associatives, par exemple :

(26) *L'omelette au jambon* est partie sans payer. [le client ayant commandé…]
George Sand est sur l'étagère de gauche. [le(s) livre(s) écrit(s) par…]

L'explication est fondée sur le concept de *fonction pragmatique* ou *connecteur,* entre le *déclencheur* de la référence (par exemple la représentation de l'omelette au jambon) et la *cible*, i.e. le référent véritable (par exemple la représentation du client ayant commandé l'omelette et après lequel il eut fallu courir). Ces connecteurs peuvent relier entre eux des espaces différents ; ce sont précisément les relations entre espaces qui permettent de rendre compte des cas où la coréférence ou la pronominalisation sont impossibles. À un niveau encore plus élémentaire, ce seraient l'existence et l'unicité à l'intérieur d'espaces mentaux déterminés qui entraînent l'usage de l'article défini, alors que l'introduction de représentations d'entités dans un espace se ferait à l'aide d'articles indéfinis.

Nous renvoyons à (Moeschler et Reboul 1994:172-175) pour une critique du concept de fonction pragmatique, lui reprochant sa faible valeur explicative et prédictive.

Nous noterons seulement que dans son principe cette théorie semble reprendre une idée répandue, le calcul logique des désignations de Russell (V.A.3.ii), et lui ajouter seulement la restriction des espaces mentaux particuliers. Le mécanisme de la référence lui-même, sur lequel l'auteur fournit peu de détails, serait fondé sur le symbolisme logique. Par ailleurs, il existe des modèles computationnels d'« accessibilité » des référents bien plus détaillés que les espaces mentaux, et plus proches de l'implémentation (V.D.2 ci-dessous).

V.B.3.iii.b Langacker (1991) et le fonctionnement des nominaux

Langacker (1991) développe un ensemble de notations et représentations graphiques propres au courant de la « grammaire cognitive ». De son propre aveu, Langacker emprunte certains éléments au travail de Fauconnier pour sa propre théorie de la référence, notamment la représentation topologique des « espaces mentaux ».

Le principal problème de Langacker touchant la référence est le fonctionnement des *nominaux* (expressions utilisables dans un énoncé) par rapport aux simples *noms* ; son analyse apporte aussi, indirectement, une réponse au problème de la référence. Selon Langacker, l'opposition *nom/nominal* est décrite par l'opposition *type/instance*. Le passage de l'un à l'autre se fait en trois étapes, que nous résumons dans la Table V-2. Il faut noter le statut ambigu du processus d'*instanciation*, qui apparaît à certains endroits comme l'ensemble de la transformation du nom en nominal, et à d'autres comme la première étape de ce processus[12].

Fonction sémantique	**Réalisation lexicale**
(spécification de type ± détaillée)	NOM (ou combinaison de noms)
Instanciation (?)	?
Indication de la quantité ou de la proportion	NOM + QUANTIFICATEUR (première couche)
Ancrage dans le cadre du discours et par rapport aux locuteurs	NOM + QUANTIFICATEUR + PRÉDICAT ÉPISTÉMIQUE (deuxième couche)
(emploi dans un énoncé)	NOMINAL RÉSULTANT

Table V-2. Du nom au nominal :synthèse d'après (Langacker 1991:ch.2-3).

Un des points faibles de cette analyse est la lexicalisation des différentes étapes. En effet, Langacker montre sur le nominal suivant :

(27) those three black cats [ces trois chats noirs]

[12] Dans la section de (Langacker 1991) consacrée à l'instanciation (2.2), les entités examinées sont toutefois des nominaux (produits finaux du processus), avec quelques cas particuliers : noms propres, prédicats nominaux, fonctions (comme « pape »). Ensuite, l'auteur étudie par deux fois les considérations de nombre ou quantité (2.3 et 3.2-3.3) et les « prédications épistémiques » (2.4 et 3.1).

comment au noyau central, spécifiant le type (*chat noir*) on ajoute d'abord le quantificateur (*trois*), puis le prédicat d'ancrage dans la situation (*ces*). Cet exemple est en effet particulièrement clair, mais la plupart des exemples examinés par la suite ne s'encadrent pas, de l'avis même de l'auteur, dans cette analyse. Dans bien des cas, l'un ou l'autre des trois éléments (type/quantificateur/ancrage) manque, ou bien une lexicalisation unique permet d'exprimer plusieurs fonctions. On peut alors mettre en question le bien fondé même de ces distinctions, malgré leur clarté logique. Dans de nombreux cas, on peut employer en tant que nominal un syntagme qui est très proche du nom lui-même : nom propre, prédicat nominal, démonstratifs, déictiques, etc. Dans ces cas, l'analyse logique précédente rend compte assez mal du fonctionnement référentiel.

Cette analyse s'avère fortement tributaire du calcul russellien de la désignation, qu'elle amène cependant à un niveau de détail intéressant par la distinction des trois étapes dans la construction du sens des nominaux. Un des aspects les plus novateurs est l'introduction des prédicats épistémiques dans la construction du nominal, qui situent celui-ci dans un espace mental, à une position déterminée. Il est cependant regrettable que, suivant en cela la tradition logico-sémantique, la théorie soit détachée à la fois des exemples multilingues (nous en verrons l'utilité avec les théories fonctionnalistes au VI.B), ainsi que des situations concrètes d'emploi.

Nous retiendrons donc des travaux de la grammaire cognitive, ici ceux de Langacker et Fauconnier, la synthèse qu'ils opèrent entre le formalisme logique et une certaine prise en compte de la situation d'énonciation, qui se traduit par une représentation symbolique spécifique.

V.B.4 Linguistique descriptive de la langue française

Nous avons jusqu'ici considéré des travaux portant sur la nature et le fonctionnement de la référence en général, en général sans considération de langue (ou, de façon implicite limités à l'anglais). Il existe cependant des travaux liés plus spécifiquement à la langue française : d'abord, dans la veine des théories précédentes, la théorie linguistique de Milner (1982, 1989), puis les études sur des phénomènes linguistiques « fins » réalisées par Kleiber, Charolles, Corblin.

V.B.4.i La théorie de la référence de J.C. Milner (1982, 1989)

J.C. Milner (1989, 1982) propose une théorie ouvertement linguistique de la référence. Milner distingue la *référence actuelle* d'une expression dans une situation concrète d'emploi, de la *référence virtuelle* ou référence « lexicale » hors contexte. Les termes qui manquent de référence virtuelle, comme les pronoms, doivent recevoir leur référence actuelle, en contexte, par un processus de *saturation sémantique*. À savoir, le sujet interprétant ajoute à ces termes l'information sémantique qui manque pour caractériser un référent à partir d'autres informations, d'origine contextuelle. Ce mécanisme de saturation, lorsqu'il fait appel aux informations contenues dans l'environnement illocutoire, par exemple pour les déictiques, n'est plus du ressort de la

linguistique (Moeschler et Reboul 1994:351). Nous ne pouvons bien sûr que souscrire à l'idée de saturation sémantique que nous allons tenter de préciser dans notre système.

V.B.4.ii Travaux descriptifs de Kleiber, Charolles, Corblin

Nous ne pouvons pas donner ici des détails à propos des nombreux travaux sur différents point linguistiques liés à la référence, en français, et en particulier concernés par le fonctionnement des pronoms. Ces travaux constituent un précieux réservoir d'analyses. Leur application aux modèles informatiques n'est pas toujours aisée, dans la mesure où la rhétorique l'emporte parfois sur l'esprit de synthèse. Mais chaque langue possédant des marqueurs référentiels propres, ces études s'avèrent essentielles pour les détails de chaque marqueur français, en vue d'une future implémentation.

Georges Kleiber est l'auteur d'un grand nombre de travaux sur la référence, et le problème particulier de l'anaphore, dont pas moins de cinq livres (Kleiber 1981, 1989, 1994a, 1994b, 1995). Il s'agit souvent de la description de phénomènes référentiels extrêmement particuliers, et non de la proposition d'un modèle pour la référence. Ces exemples sont parfois le prétexte d'analyses plus générales, comme celle des noms propres à partir de la construction *CE + Nom propre* (Kleiber 1994b:ch.4). Autre exemple, ses nombreux travaux sur le pronom *il*, qui semblent défendre en fin de compte un mode de fonctionnement pragmatique, du moins dans certains cas ; nous présenterons cependant ci-dessous un modèle plus achevé de traitement pragmatique de la référence (V.D.5).

Enfin, des linguistes comme M. Charolles (1987)et F. Corblin se sont intéressés aux phénomènes référentiels en liaison avec la cohérence des discours, question sur laquelle nous reviendrons au V.D.2. Ainsi, F. Corblin (1987, 1995) analyse la notion de chaînes de référence : les mentions successives d'un même référent dans un discours, utilisant des expressions référentielles variables. La structure de ces chaînes est intimement liée à la cohérence du discours et la nature des expressions référentielles employées est liée à l'« accessibilité du référent », (Moeschler et Reboul 1994:462), cf. V.C.2 ci-dessous.

V.C Référence, contexte, pragmatique

On peut dire, en simplifiant beaucoup, que la *pragmatique* est l'étude de la communication située dans un contexte, par opposition à la syntaxe et à la sémantique qui se concentrent sur les propriétés intrinsèques des énoncés. Les rapports entre la pragmatique et la linguistique sont examinés dans (Moeschler et Reboul 1994), surtout au chapitre introductif : bien que le terme « linguistique » suggère l'étude du langage dans son ensemble, on a plutôt tendance à limiter la linguistique à l'étude de la *compétence* et rattacher l'étude de la *performance* (interprétation en contexte) à la pragmatique. Ce domaine s'est développé après les années 1980.

Nous avons présenté, parmi les théories ontologiques de la référence, quelques approches situées : les actes de langage au V.A.4.ii.a, la théorie de la pertinence au

V.A.4.ii.c. Notre analyse des fondements de la référence a montré qu'il s'agit d'un phénomène qui ne prend son sens que dans un *contexte d'énonciation*. Nous avons montré comment les approches logicistes s'orientent vers une prise en compte du contexte. La présente section illustre cette évolution avec deux points de vue « proto-pragmatiques » sur le contexte, et sur la mémoire qu'en ont les utilisateurs, ceux de Talmy Givón, et de Mira Ariel.

V.C.1 Talmy Givón (après 1979) : la mémoire des référents

Nous appellerons « fonctionnalistes » les travaux linguistiques de Talmy Givón, eu égard à l'importance accordée au fonctionnement ou usage réel de la langue. Nous exploiterons plus loin, de façon détaillée, son étude de l'expression linguistique de la référence, en vue de notre modélisation par agents communiquants (cf. VI.B.2). Pour l'heure, nous montrerons comment Givón entend enrichir le point de vue classique sur la référence par une analyse de la structure de la mémoire de travail dédiée à la compréhension de la référence.

V.C.1.i Vers une modélisation de l'ensemble des référents accessibles

Givón (1984-1990:388 sq.) critique de façon succincte l'approche frégéenne des descriptions définies en remarquant que dans l'exemple classique :

(28) Le roi de France est chauve.

le caractère référentiel du GN défini est traduit par une assertion existentielle (ici fausse). Selon Givón, ce GN seulement présuppose le partage de l'idée que le roi de France existe, et dans la mesure où ce partage repose sur une convention communicative entre *les deux* interlocuteurs, les extensions du formalisme logique à l'aide d'une sémantique présuppositionnelle ne résolvent pas le problème[13].

Dans une perspective que se veut « pragmatique », Givón tente de montrer que le *monde réel* de la logique doit être remplacé par un *univers du discours* ayant une structure et une évolution particulières. Givón détaille le fonctionnement des différents *fichiers* qui contribuent à l'univers du discours, regroupant les référents que les interlocuteurs s'accordent à supposer existants. Dans une situation de communication, afin d'interpréter le discours, chaque interlocuteur maintient un *fichier des entités actives*. Ce fichier est alimenté par trois sources : d'abord le *discours* lui-même, qui permet d'introduire de nouvelles entités, puis par le *contexte déictique immédiat*, et par le *fichier permanent* des entités connues de la communauté des locuteurs, activées par le discours (Figure V-2).

[13] Givón fournit une liste (peu structurée) d'arguments contre les théories logiques, dont nous reproduisons les entrées (pp. 388-390) : (a) le monde réel par rapport à l'univers du discours ; (b) quantificateur universel par rapport à la non-référentialité ; (c) degrés de non-référentialité ; (d) intention référentielle du locuteur ; (e) modalités véri-conditionnelles par rapport aux modalités référentielles ; (f) référence et pragmatique du discours.

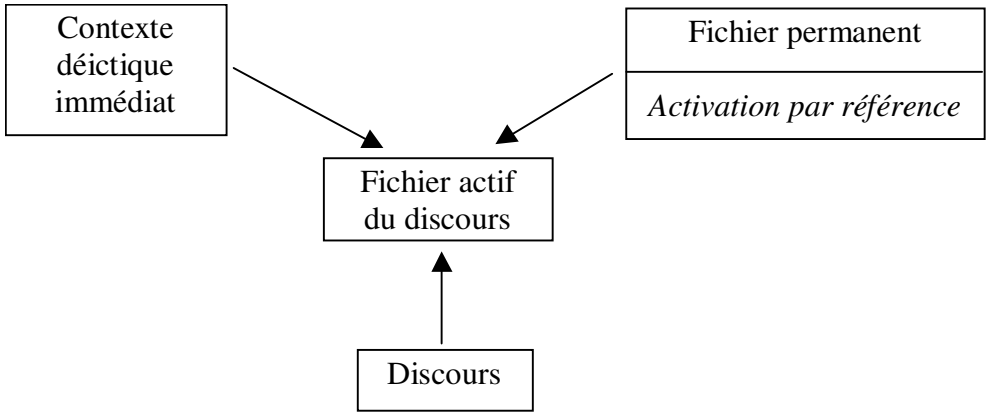

Figure V-2. Les trois sources d'information alimentant le fichier actif d'après Givón (1984-1990:405).

Nous remarquerons d'abord la reprise du terme « fichier » utilisé par Karttunen (V.B.2.iv). Cette tentative de structuration, qui tente de prendre en compte la *situation* de communication, est cependant discutable dans les détails. Par exemple, dans le contexte déictique, outre les pronoms de première et deuxième personne, et les démonstratifs, on trouve des référents associés aux référents déjà existants (« sa mère », « sa tête », « sa maison ») : on voit mal ce que cela a de déictique, et pourquoi ces références fondées sur la connaissance du monde sont séparées du fichier permanent – en réalité, leur fonctionnement est plutôt hybride. D'autre part, la séparation entre référents *introduits* par le discours et référents permanents *activés* par le discours paraît discutable dans le détail.

Nous retiendrons toutefois le principe d'un fichier des référents actifs – il s'agit plutôt de *représentations* actives – ainsi que l'idée d'un partage des sources d'information qui créent ou activent des référents.

V.C.1.ii Évolution de la mémoire des référents

Dans un ouvrage plus récent, Givón (1993) détaille son modèle de la mémoire « de texte » ou « épisodique » et décrit les principales opérations qui régissent son évolution. Cette mémoire contient à chaque instant un seul *nœud textuel* actif, qui représente l'ensemble des informations attachées à un référent, provenant d'une série d'énoncés (qui ont ce référent pour sujet). La suite du texte détermine l'évolution du nœud actif, selon le schéma donné ci-dessous (le terme *référent* remplace *nœud actif*).

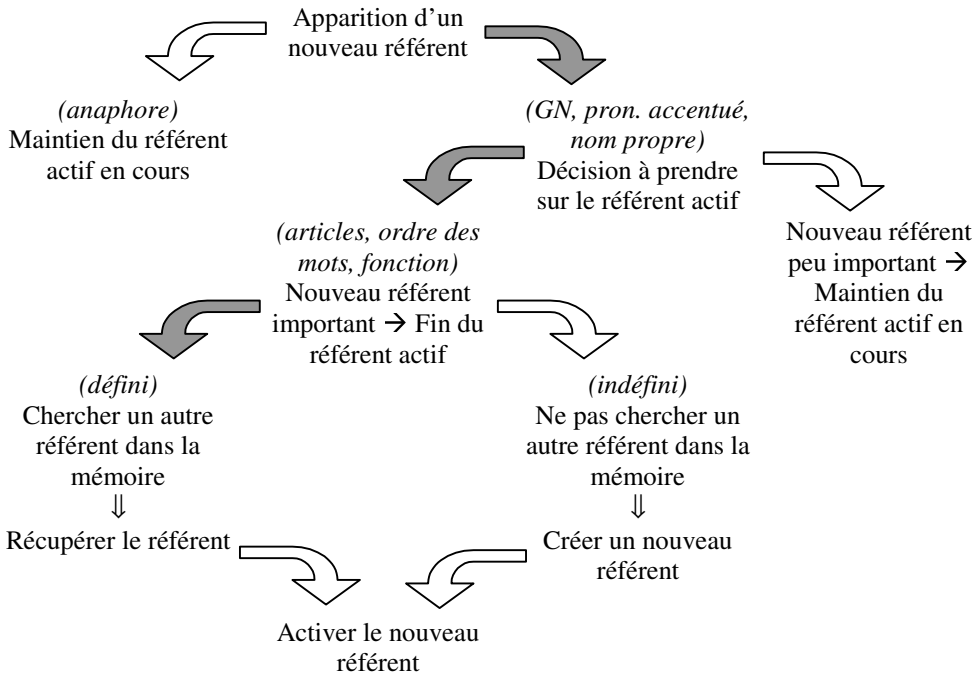

Figure V-3. Opérations cognitives pour le suivi de la référence. Les flèches grisées signalent les cas linguistiquement marqués. Adapté de (Givón 1993:245)..

Une telle modélisation est à nos yeux bienvenue, et nous en trouverons un exemple plus élaboré avec la théorie du centrage ; elle rappelle aussi l'évolution des espaces mentaux vue ci-dessus (V.B.3.iii.a), et participe plus généralement du mouvement de prise en considération du contexte mental de chaque locuteur, facteur déterminant dans l'attribution des référents. Certains éléments sont toutefois discutables, surtout l'unicité du référent actif (ou *nœud textuel actif*), qui est une simplification d'une échelle plus complexe d'activation. Aussi, les critères de choix aux branchements sont peu explicités. Pourtant, des idées plus anciennes de Givón (1984-1990:ch.11) sont plus nuancées, et montrent que le couplage entre l'expression linguistique et le caractère défini est en réalité plus flou, comme nous le verrons au VI.B.2.

V.C.2 Pragmatique et référence : la question de l'accessibilité

Les études réalisées par Mira Ariel (1990, 1994) sur l'accessibilité des référents se placent dans une optique proche de la précédente. Les interlocuteurs conservent en mémoire un ensemble de référents (représentations), et les expressions référentielles d'un texte doivent activer chacune le référent correspondant. Ariel étudie le lien entre la nature de l'expression référentielle utilisée et le niveau d'activation autorisé pour le référent qu'elle veut réactiver – en d'autres termes les marqueurs linguistique d'accessibilité (du référent) pour les expressions référentielles. Il s'agit donc d'une propriété en quelque sorte duale de l'activation de la mémoire des référents que nous avons vue chez Givón ; l'hypothèse d'Ariel se trouve à l'interface entre la pragmatique, la linguistique et la référence.

Référent peu actif en mémoire	Nom propre complet
(faible accessibilité)	Description définie longue
	Description définie courte
	Nom de famille
	Prénom
	Adjectif démonstratif distal
	Adjectif démonstratif proximal
	Pronom démonstratif distal
	Pronom démonstratif proximal
	Pronom accentué
	Pronom non accentué
Référent très actif en mémoire	Pronom clitique
(forte accessibilité)	Pronom réfléchi, anaphore 'zéro'

Table V-3. Échelle d'accessibilité d'Ariel (simplifiée).

L'échelle d'accessibilité d'Ariel (Table V-3) *estime* le niveau d'activation du référent lorsqu'il est désigné par tel ou tel type d'expression référentielle. Elle régit en principe à la fois la génération et la compréhension des énoncés. On peut certes critiquer l'idée de considérer seulement le *type* de l'expression, indépendamment de son contenu, et aussi faire remarquer que cette échelle risque de se réduire tout simplement à des contraintes de distance textuelle entre expressions coréférentes (Reboul *et al.* 1997).

Il semble néanmoins que ces principes ne soient pas vraiment menés à bien dans la formulation détaillée de la théorie. Les critiques d'Anne Reboul (Reboul *et al.* 1997) lui reprochent à la fois une tendance à la simplification (la référence serait le lien entre la représentation et le texte seulement) et une visée de réduction de la référence à un phénomène exclusivement linguistique (à travers l'échelle donnée).

Nous retiendrons cependant l'originalité de cette proposition qui lie un facteur pragmatique (l'activation des référents) à l'expression linguistique ; même si elle n'exprime qu'une propriété graduelle, son bien fondé devrait pouvoir mis en évidence par une étude informatique, réalisant une moyenne sur de nombreux textes.

V.D Modèles computationnels de la référence

Nous passons maintenant aux contributions théoriques qui ont été effectivement implémentées dans des systèmes ; ce ne sont pas encore les *systèmes* en eux-mêmes qui nous intéressent (cf. V.E), mais les *théories computationnelles développées* – et qui sont implémentables, ou presque, comme l'illustrent les systèmes (ou parfois les prototypes) réalisés. Après un bref aperçu des approches discursives de la référence au début des

années 1980 (V.D.1), nous développerons deux approches novatrices de la focalisation (travaux de Grosz et de Alshawi, cf. V.D.2). Nous exposerons ensuite deux théories plus récentes du discours et de la référence (le centrage et la « théorie des veines », cf. V.D.3), puis des travaux sur la référence et les actes de langage, dans une perspective de génération de textes (V.D.4). Nous présenterons pour conclure un modèle computationnel inspiré par la pragmatique, élaboré dans le cadre du projet CERVICAL auquel nous avons participé (V.D.5).

V.D.1 Référence et contexte vers 1980

Plusieurs travaux ont essayé, vers 1980, de prendre en considération le contexte discursif dans des modèles implémentables ; il s'agit donc d'une coopération heureuse de la linguistique et de l'informatique pour la « compréhension » de la langue et des références. Une discussion de ces travaux se trouve dans (Hirst 1981a, Hirst 1981b), et aussi en français dans l'ouvrage de Gérard Sabah (1989:228-236). Afin d'éviter un troisième résumé de ces travaux, nous renvoyons le lecteur aux références précédentes, et présentons un aperçu fondé sur (Hirst 1981b).

Hirst résume et discute les travaux suivants, auxquels s'ajoutent ceux de B.L. Webber (vus au V.B.2.v ci-dessus) :

- La thèse de Kantor (1977) utilise la notion d'activation d'un concept à un point donné d'un discours pour expliquer pourquoi certaines reprises par des pronoms sont plus compréhensibles que d'autres. Le point intéressant est le remplacement d'un classement binaire des concepts en « thème » et « non thème », ou encore « focalisés » / « non focalisés », par une échelle graduelle ; cette idée n'est pas testée par l'auteur dans un système.

- La thèse de Lockman (1978) et l'article de Hobbs (1979) sur la structure du discours préconisent un traitement en plusieurs étapes : identification des relations de cohérence (ou relations rhétoriques) dans un texte, construction des ensembles des référents « focalisés », résolution des références par recherche dans cet ensemble. À ceux qui observent que la résolution des références est elle-même nécessaire à la détermination des relations rhétoriques, Hobbs répond que ce mécanisme utilise des connaissances différentes à chaque étape, donc qu'il n'agit pas d'un cercle vicieux mais d'une « spirale ». Un point intéressant est que ces travaux abordent la construction des relations de cohésion (rhétoriques) dans le cas général, et n'utilisent pas, comme les travaux de Grosz (ci-dessous), des schémas de dialogue prédéfinis – mais ils semblent assez loin de l'implémentation.

- Barbara Grosz, dans sa thèse (Grosz 1977b) et ses travaux ultérieurs (Grosz 1977a, Grosz 1981), étudie la résolution de la référence dans des dialogues orientés par la tâche. Elle montre comment l'utilisation de connaissances élaborées sur la structure

de la tâche permet de résoudre les références aux objets qui y interviennent. Grosz prend l'exemple de dialogues dans lesquels un expert enseigne à un novice à assembler un compresseur, à l'aide d'outils et pièces détachées. L'« espace de focalisation » à un moment donné contient les référents possibles, à savoir les objets et outils correspondant à l'étape ou sous-étape courante de la tâche. À l'intérieur de cet espace, l'appariement entre l'expression référentielle à résoudre et le référent (i.e. sa représentation) se fait à l'aide des critères de compatibilité classiques, en comptant sur le nombre restreint de référents accessibles pour rendre ces critères discriminants.

- Les travaux de Candace Sidner, sur laquelle nous reviendrons ci-dessous, et notamment sa thèse (Sidner 1979a), présentent en grandes lignes une conception analogue à celle de Grosz. La différence principale réside dans le refus de Sidner d'utiliser des schémas prédéfinis de dialogue, et dans sa tentative d'élargir l'application de l'« espace de focalisation » à des dialogues plus généraux. Pour ce faire, elle utilise des « frame » ou schémas situationnels, qui correspondent à des situations prototypiques données au système (Charniak 1978, 1981). Sidner applique ses idées à un système d'organisation d'emplois du temps, pour lequel le nombre de « frame » à fournir reste raisonnable.

L'aperçu de ces théories montre donc un intérêt croissant pour la limitation de l'ensemble des référents accessibles, principale innovation apportée aux critères logiques classiques. Les méthodes pour gérer cet ensemble sont quant à elles très variées. Cet intérêt est lié aussi à la possibilité d'applications informatiques.

V.D.2 Théories de la focalisation

Pour mieux comprendre les concepts liés à la focalisation, exprimés en anglais par « focus », « focusing » ou « focus space », nous allons donner quelques détails de la théorie initiale proposée par Candace Sidner (1979b), en particulier sa conception de la référence, puis sa théorie proprement dite. Nous commenterons ensuite un travail plus récent qui reprend et développe la focalisation, celui de Hiyan Alshawi (1987).

V.D.2.i Calcul de la focalisation à l'aide de schémas : thèse de Sidner (1979)

V.D.2.i.a Conception de la référence

Afin d'asseoir sa théorie sur une base conceptuelle solide, et de prévenir les critiques des linguistes, Candace Sidner brosse un tableau simple mais réaliste du phénomène de la référence ; nous le résumons d'après (Sidner 1981, 1983).

Sidner accorde en effet que l'anaphore ne peut être simplement vue comme le lien d'un pronom (par exemple) à un antécédent situé dans le texte, parce que les expressions ne sont pas supposées référer à d'autres expressions, mais à des entités du monde. Il reste qu'un pronom ne peut en général être compris en l'absence d'un « antécédent », qui peut d'ailleurs ne pas référer exactement à l'entité désignée par le pronom.

Sidner considère que les locuteurs possèdent chacun une base de données constituée de *représentations cognitives* d'entités. Une expression référentielle sera dite *spécifier* un tel élément cognitif (appelé sa *spécification*). Ces éléments cognitifs, d'autre part, *représentent* pour chaque individu des entités du monde réel. On voit donc que le terme « référence » a été occulté, étant jugé problématique et chargé de connotation diverses. Deux expressions référentielles qui spécifient le même élément cognitif sont dites *co-spécifier* dans le texte, indépendamment de leur nature (noms ou pronoms). Cette conception permet alors à Sidner de se démarquer d'une vue trop simpliste de l'anaphore, et de définir celle-ci comme « le processus computationnel qui utilise la spécification existante d'une expression pour trouver la spécification d'un anaphorique » (Sidner 1981:218). En reformulant, seraient donc des anaphoriques les expressions dont la spécification a déjà été réalisée par un expression antérieure.

Cette conception présente à nos yeux un progrès notable vers une conception réaliste de la référence ; l'introduction des représentations participe d'une tendance mentaliste plus large. On peut toutefois s'interroger sur l'utilisation effective que Sidner fait de cette description : sur un plan purement logique, sa distinction des anaphoriques et non anaphoriques amène son modèle exactement au stade antécédent/anaphorique, puisqu'à chaque « antécédent » on associe un élément cognitif, et chaque anaphorique doit être rattaché à l'élément cognitif antécédent. Si l'on peut apprécier la généralité du modèle initial (expression ↔ représentation ↔ entité réelle), on ne peut que regretter son immédiate réduction. Nous retrouverons le côté représentationniste du modèle dans les propositions du projet CERVICAL (cf. § V.D.5 ci-dessous), ainsi que dans notre propre version (VII.B).

V.D.2.i.b Algorithme de Sidner pour la détermination des éléments focaux

Sidner utilise plusieurs structures pour tenir compte de la focalisation de l'attention référentielle dans un discours, lors de chaque énoncé. L'élément le plus important est l'élément focal du discours (« discourse focus »), et il y a aussi un acteur focal (« actor focus »), qui sera une entité ayant une importance plus locale, ou transitoire[14]. Enfin, les éléments qui ont été focaux sont stockés séparément, dans une pile, alors que pour la phrase courante ont considère une liste d'éléments focaux potentiels.

L'utilisation de ces données discursives pour la résolution peut être schématisée dans le principe suivant (Sidner 1981:222) :

> *Si un pronom apparaît comme argument verbal de type autre que agent, choisir l'*élément focal du discours *comme co-spécifieur du pronom, sauf si cela est exclu par les contraintes syntaxiques, sémantiques ou pragmatiques. Si le pronom apparaît comme argument de type agent, choisir l'*acteur focal *comme co-spécifieur, sauf si cela est exclu par les contraintes précédentes.*

[14] La traduction des termes n'est pas aisée. Gérard Sabah (1989:233-6) utilise par exemple le terme de *foyer* pour « discourse focus » et *acteur central* pour « actor focus ».

Une bonne partie des considérations de Sidner traitent de la détermination de ces éléments et de leur évolution dans le texte, et aussi, en particulier, de leur amorçage lors de la première phrase. Les algorithmes proposés prennent en considération la façon dont les anaphores précédentes ont été résolues.

V.D.2.ii Les facteurs de contexte d'Alshawi (1987)

Le livre de Hiyan Alshawi (1987) présente un modèle générique de mémoire et contexte pour la compréhension de la langue par une machine. Sans entrer dans le détail, la mémoire du système contient des *entités* représentant des objets, classes, prédicats, relations. Ces entités peuvent être sujettes à d'autres relations. Selon l'auteur, un partie centrale du problème de la compréhension d'un texte consiste à activer des entités appropriées dans la mémoire, l'activation étant une propriété numérique de chaque entité de la mémoire. Le contexte joue un rôle essentiel dans la mise à jour de l'activation.

Alshawi propose un ensemble générique de *facteurs de contexte*, à savoir des critères qui renforcent l'activation des entités de la mémoire auxquels ils s'appliquent. Étant donné la proximité avec l'idée de focalisation, et plus encore avec l'activation des concepts introduite par Kantor (V.D.1), nous résumons les types de facteurs de contexte proposés (Alshawi 1987:16-20) :

- 'Récence' : trois facteurs renforcent l'activation des entités mentionnées dans la *phrase*, resp. le *paragraphe*, resp. le *texte*, courants. Le facteur 'deixis' accroît encore cette activation si les trois facteurs de 'récence' dépassent un certain seuil.
- 'Emphase' : une seule entité peut être renforcée, si elle se trouve dans une construction emphatique (sujet d'un verbe au passif, ou sujet d'un prédicat nominal).
- 'Historique du traitement' : une entité ayant pris part (même comme intermédiaire) dans le traitement précédent est renforcée.
- 'Sujet du domaine' : les entités qui appartiennent au domaine des textes traités (donné à l'avance) sont renforcées.
- 'Association' : les entités conceptuellement proches d'une entité dont l'activité dépasse un certain seuil sont renforcées.

Ces facteurs ont été réellement utilisés dans le système dont traite la majeure partie de l'ouvrage, avec des poids numériques donnés (Alshawi 1987:94). Le système réalisé illustre les considérations génériques sur la compréhension de la langue ; il a pour tâche la construction d'une base de données relationnelle à l'aide de textes anglais traitant des constructeurs de produits informatiques.

Les facteurs précédents sont utilisés effectivement. Alshawi en ajoute d'autres, à titre d'exemple :

- Le temps et l'espace : activer les entités de la mémoire ayant une proximité temporelle et spatiale par rapport aux concepts décrits présentement.

- Les buts : augmenter l'activation des entités proches des buts des locuteurs.
- Association aux agents : augmenter l'activation des entités (objets, événements) reliés aux agents figurant dans les textes, p.ex. les personnages principaux.
- Séquences d'événements : activer les entités faisant parties de séquences prévisibles (analogues aux scénarios de Schank).
- Structure prédéfinie de dialogue : activer les (autres) entités qui en font partie.

L'utilisation des facteurs de contexte pour la résolution de la référence est certes sans surprise. Lorsqu'une expression référentielle est détectée, un ensemble de contraintes (syntaxiques et sémantiques) est propagé dans la mémoire, qui marque toutes les entités compatibles. Le système examine l'*activation* des entités référençables, qui est calculée à l'aide des facteurs. Puis, *grosso modo*, l'entité compatible ayant le niveau d'activation le plus élevé est choisie comme référent. Curieusement, le problème de la création de nouvelles entités lorsque toute résolution de référence définie échoue n'est pas explicité.

Le mérite d'Alshawi est d'avoir conçu un mécanisme générique, où la structuration de la mémoire des référents est réalisée simplement par un coefficient numérique d'activation. Cette idée a peut-être plus de pertinence psychologique par rapport aux mémoires trop structurées (comme chez Sidner, ou dans le « centrage » ci-dessous), et son implémentation est aisée. D'importants travaux postérieurs adoptent le modèle proposé par Alshawi, comme ceux de Lappin et Leass (1994) pour la résolution de l'anaphore dans les textes, ou de Huls, Claassen et Bos (1995) dans le dialogue multi-modal. Notre propre système s'en inspire aussi (cf. VII.C.3.v), et nous décrirons une technique de réglage automatique des poids des facteurs de contexte (au VII.F.2).

V.D.2.iii Collaboration de Grosz avec Sidner (1985-86)

En conclusion de cet aperçu de la focalisation au sens large, nous rappelons que Barbara Grosz et Candace Sidner ont réalisé en collaboration un important travail mettant en commun une partie de leurs idées précédentes (Grosz et Sidner 1985, 1986). La présentation de ce travail dépasse le but de notre exposé, malgré ses conséquences pour la référence. L'article étudie en effet les rapports, dans le dialogue orienté par la tâche, entre les buts des interlocuteurs, le dialogue lui-même, et les « espaces focaux » des participants ou leur « état attentionnel » ; ce dernier représente l'ensemble des référents accessibles, à l'instar des travaux précédents.

V.D.3 Cohérence du discours et référence

V.D.3.i La théorie du centrage : Grosz, Joshi, Weinstein (1983-1995)

La théorie du centrage, ou « centering theory » est issue de travaux présentés vers 1983 (Grosz, Joshi et Weinstein 1983), mais formalisés de façon détaillée seulement en

1995 (Grosz, Joshi et Weinstein 1995). Il ne s'agit pas d'une ontologie de la référence, ni d'une méthode de résolution, mais d'une théorie descriptive reliant la cohérence des assertions d'un discours aux différentes expressions référentielles utilisées.

Chaque *assertion* (*U*, « utterance ») d'un discours se voit associer un ensemble de *centres d'intérêt à venir* ($C_f(U)$, « forward-looking centers ») et un unique *centre d'intérêt passé* ($C_b(U)$, « backward-looking center »). De l'aveu des auteurs, ces notions reprennent et précisent, respectivement, les notions d'*éléments focaux potentiels* et d'*élément focal du discours* introduites par Sidner (V.D.2.i.b)[15]. Les différentes entités ou centres sont susceptibles d'être *réalisées* par une assertion ou fragment d'assertion : ce terme désigne soit la dénotation directe (le NP_n réalise le centre c_i) soit l'évocation après inférences, ou application d'autres facteurs ; les auteurs donnent en fait peu de détails sur la « réalisation » d'un centre.

Une proposition centrale est la suivante :

- L'élément le plus actif de $C_f(U_n)$, parmi les centres d'intérêt à venir de U_n, qui est réalisé dans U_{n+1} est le $C_b(U_{n+1})$, i.e. le centre d'intérêt passé, ou plutôt courant, de U_{n+1}.

La théorie défend un certain nombre de contraintes pour qu'un discours soit cohérent. La cohérence n'implique pas que le centre d'intérêt soit constant, mais plutôt que les transitions soient effectuées selon certaines règles. Les auteurs définissent trois types de transitions :

- continuité du centre : $C_b(U_n) = C_b(U_{n+1})$ et $C_b(U_{n+1})$ est l'élément le plus activé de $C_f(U_{n+1})$;
- rétention du centre : $C_b(U_n) = C_b(U_{n+1})$ mais $C_b(U_{n+1})$ n'est pas l'élément le plus activé de $C_f(U_{n+1})$;
- changement du centre : $C_b(U_n) \neq C_b(U_{n+1})$.

Parmi les contraintes postulées, nous mentionnons les suivantes :

- Chaque assertion U_n possède un unique centre d'intérêt courant $C_b(U_n)$.
- Les éléments de $C_f(U_n)$ sont partiellement ordonnés, ce qui guide (comme on l'a vu) la détermination de $C_b(U_{n+1})$. On peut remarquer que les auteurs discutent assez peu les facteurs d'ordonnancement, et l'existence même d'un « plus grand » élément (ou « plus actif »).
- Contrainte sur la réalisation : si un quelconque élément de $C_f(U_n)$ est réalisé (i.e., désigné ou repris) par un pronom dans l'assertion U_{n+1}, alors le centre $C_b(U_{n+1})$ doit être aussi réalisé par un pronom. Cela formalise l'idée, semble-t-il, que le premier élément à être pronominalisé est le thème.

[15] Ce qui montre que le terme « backward-looking center» (notre « centre d'intérêt passé ») est à concevoir plutôt comme « centre d'intérêt courant, provenant du passé ».

- Préférence entre les transitions : la continuation du centre est préférable (i.e. demande moins d'effort de compréhension, assure une plus grande cohérence) à la rétention du centre, elle même préférable au changement.

- Le choix de $C_b(U_n)$ est strictement local, ne dépendant que de $C_f(U_{n-1})$.

Une thèse qui, à nos yeux, semble assez problématique est l'idée que le centrage n'est pas gouverné seulement par la syntaxe, la sémantique, ou la pragmatique, mais par une « combinaison de facteurs discursifs » (rappelons aussi l'analyse donnée par Joshi (1998), I.B.3). Cette thèse est assez peu détaillée, et on comprend mal si la théorie du centrage elle-même est le « facteur discursif » en question. Par ailleurs, le calcul des centres fait appel à un certain nombre de facteurs syntaxiques ou sémantiques qui sont à peine esquissés dans l'article : la pragmatique semble quant à elle totalement absente.

De façon plus applicative, l'article cité dresse la liste des utilisations réalisées (avant 1995) du centrage : nous en retiendrons l'application à la résolution des anaphores, par Breck Baldwin (1993). La théorie du centrage ne fournit pas de *mécanisme* de résolution de la référence, et Baldwin utilise les contraintes du centrage pour calculer d'abord les différents centres, puis en déduire, par satisfaction de contraintes, les référents des pronoms. La distinction entre expressions pronominales ou non joue donc un rôle important dans la théorie.

Parmi les réserves que l'on peut émettre, la première concerne le formalisme même de la théorie, qui ne semble pas fondé sur un modèle cognitif, et tient peu compte de la position du locuteur dans la situation de communication. Les ensembles de centres $C_f(U_n)$ supposent une organisation assez élaborée de la mémoire du locuteur. Le caractère abstrait du formalisme, et sa limitation au discours, est visible dans le concept de « réalisation », qui est une solution de remplacement assez peu satisfaisante au concept de « référence ». Le centrage semble donc plus pertinent dans le domaine de la cohérence discursive que dans celui de la référence proprement dite[16].

Se pose enfin le problème de l'implémentation du centrage ; malgré certaines tentatives (Baldwin 1993, Hahn et Strube 1997), il semble que les connaissances nécessaires pour la construction des centres, peu explicitées dans la théorie, ne soient pas encore à la portée des programmes.

V.D.3.ii La théorie des veines : Cristea, Romary, Ide (1998)

Une généralisation récente et originale du centrage est la « théorie des veines », qui propose des règles analogues au centrage, valables pour la structure plus globale du discours (Cristea, Ide et Romary 1998). Très schématiquement, l'idée est de construire une structure arborescente fondée sur la structure rhétorique du discours, les « veines » étant les chemins dans cet arbre, et les feuilles étant des (fragments de) phrases. Cette

[16] Il est vrai aussi que le centrage avoue des ambitions « locales », i.e. limitées aux relations entre phrases consécutives. En ce sens, il s'agirait d'un point de vue complémentaire, en ce qui concerne la structure du discours, du travail de Grosz et Sidner (1986) – leurs rapports sont discutés dans (Grosz et Sidner 1998).

structure permet de décrire un certain *domaine d'accessibilité* référentielle pour chaque phrase (ensemble de référents). Les différents ensembles de centres constituent des parties du domaine d'accessibilité, mais il n'y a plus, apparemment, d'ordre selon l'activation des centres. Les auteurs fournissent une vérification expérimentale trilingue des principes énoncés, et calculent aussi une mesure de cohérence ou de transition pour les textes.

V.D.4 Actes de langage et génération d'expressions référentielles

La théorie des actes de langage, et plus encore les travaux de Grice sur les maximes conversationnelles, ont eu une influence théorique sur l'idée de référence, mais aussi une portée pratique. Nous évoquons ici des modèles computationnels implémentés qui utilisent ces théories, pour la *génération* d'expressions référentielles. On peut certes se demander pourquoi les travaux en génération ont pris une tournure différente de ce que nous avons vu précédemment pour l'analyse. Il faut supposer que la contrainte de pertinence est essentielle à la génération d'énoncés utilisables, alors que la compréhension d'énoncés supposés déjà pertinents nécessite des connaissances plus fines, de nature différente.

V.D.4.i Planification d'expressions référentielles : Appelt (1985)

Douglas Appelt (1985) propose une théorie et un système pour générer des expressions référentielles dans une tâche d'enseignement, par la machine, d'une suite d'actions permettant d'accomplir un but donné (par exemple démonter une pièce). La théorie prend en compte la différence de connaissance entre les agents (y compris la perception de la situation) ainsi que les différents buts des actes de langage produits par la machine. En particulier, sont distingués le cas où on se réfère à un objet mutuellement connu, et le cas où on demande à l'interlocuteur d'identifier un référent et, éventuellement, on l'informe du nom de celui-ci.

Le système réalisé affiche des ambitions assez réduites. Toutes les pièces qui interviennent dans la tâche ont des noms standard, contrairement à la théorie plus étendue d'Appelt et Kronfeld (1987), décrite au V.A.4.ii.b. De plus, il n'y a apparemment pas de dialogue : la machine engendre un plan d'action et un ensemble d'instructions pour que l'homme puisse exécuter la tâche. L'exemple donné dans l'article cité consiste à générer la phrase « remove the pump with the wrench in the toolbox », où le GN final contient en complément, de façon judicieusement informative, l'emplacement de l'outil[17].

V.D.4.ii Travaux sur la génération : Dale (1992-1995)

Les travaux de Dale et Reiter (Dale 1992, Dale et Reiter 1995) se placent à un niveau différent. À la différence de la théorie d'Appelt et Kronfeld (V.A.4.ii.b), les auteurs ne s'intéressent pas aux actes de langage ; les expression référentielles

[17] Traduction : « enlevez la pompe avec la clé *de la* boîte à outils » - en français, l'incorporation de l'emplacement paraît moins naturelle (on dirait « la clé *qui se trouve dans* la boîte »).

considérées, des groupes nominaux pleins, doivent simplement permettre l'identification du référent par l'interlocuteur, en dehors d'autres actes communicatifs. À la différence du système d'Appelt, les référents n'ont pas de noms standard (sinon l'identification serait triviale). Il s'agit de construire des expressions référentielles en combinant des propriétés des objets, telles leur catégorie sémantique et leurs caractéristiques.

Les auteurs fournissent pour cela un algorithme produisant des GN qui satisfont une version simplifiée des maximes de Grice (1975:65). Brièvement, un tel GN doit constituer une description juste, contentant exactement la quantité d'information caractéristique nécessaire, et la plus courte possible. L'algorithme proposé satisfait une forme faible de ces maximes, mais s'avère très rapide.

Signalons une extension des travaux de Dale, dans le cas des descriptions incomplètes, appliquée à un système d'enseignement des descriptions correctes en français, destiné aux enfants autistes (Mouret et Rolbert 1998).

L'ensemble de ces travaux sur la génération des expressions référentielles montrent que les problèmes rencontrés sont certes différents de ceux de la compréhension, mais acceptent jusqu'à un certain point un cadre de description commun. Les approches adoptées ici sont en réalité des variantes des théories à faisceaux de descriptions, et à ce titre présentent les mêmes limitations contextuelles.

V.D.5 Un modèle représentationnel inspiré par la pragmatique : le projet CERVICAL

Le projet CERVICAL développe des propositions concrètes de modélisation des représentations des référents, et propose des règles pour leur évolution[18]. L'application envisagée est plutôt le dialogue homme-machine (DHM), mais les réalisations du projet sont surtout théoriques. En effet, et suivant en cela un certain nombre de travaux antérieurs, le projet CERVICAL définit la notion de *représentation mentale (RM)* d'un référent comme une structure de données munie de plusieurs entrées : logique (relations avec d'autres RM), encyclopédique (informations sur la catégorie du référent et sur son individualité), visuelle (aspect du référent), spatiale (position, orientation), lexicale (expressions utilisées pour désigner la RM), d'identification (pointeur sur l'objet informatique dans une application de DHM). Plusieurs opérations sont définies, abstraitement, sur les RM : création, fusion, duplication, groupement, extraction. Le point de vue des RM s'oppose donc à une conception exclusivement linguistique de la référence, et de nombreuses critiques ont d'abord été apportées par le projet à ce point de vue (Reboul *et al.* 1997).

[18] L'acronyme signifie : « Communication et référence : vers une informatique collaborant avec la linguistique ». Il s'agit d'un projet ayant regroupé des chercheurs du LIMSI / Groupe Langage et Cognition (notamment Gérard Sabah, Isabelle Robba et nous-même) et du LORIA / Projet Langue et Dialogue (notamment Anne Reboul, Laurent Romary et Bertrand Gaiffe). Le projet a été financé par le GIS - Sciences de la Cognition, entre 1996 et 1998.

Une partie importante des travaux consiste à préciser les conditions des opérations citées, et leurs pré- et post-conditions sur le *domaine de référence*. Cette dernière notion reprend l'idée, que nous avons déjà rencontrée, de modéliser l'ensemble des référents accessibles à la désignation, à un instant donné. Les études menées dans le cadre du projet abordent le fonctionnement des déterminants *le N* et *ce N* (cf. aussi (Gaiffe 1992)), des numéraux ordinaux (ainsi que *suivant*, *dernier*, *autre*), des anaphores associatives, des collections d'objets. À cela s'ajoute l'étude de la référence temporelle et des RM-événement.

Ces considérations théoriques restent assez éloignées d'une possible implémentation : en effet, les règles concrètes de déclenchement des opérations sont encore à préciser, et les connaissances nécessaires au déclenchement (souvent implicites dans les descriptions) font encore défaut. Par exemple, le remplissage des différents champs des RM pose déjà des problèmes difficiles à l'analyse syntaxique et sémantique des textes, et les conditions de création (ou non) des RM ne sont pas toujours plus explicites que dans les travaux antérieurs (comme ceux de Karttunen)[19].

Outre nos propres apports aux discussions théoriques, notre position dans le cadre du projet a été résolument orientée vers une application concrète, dont les performances pussent être évaluées quantitativement. Nous avons donc simplifié la notion de RM, en particulier pour que ses entrées puissent être instanciées même avec les données assez élémentaires qu'un analyseur fournit sur des textes non restreints. Le résultats est un modèle simplifié, son implémentation dans un résolveur de références, et les outils qui l'accompagnent, décrits en détail au Chapitre VII.

V.E Systèmes de résolution de la référence

Nous arrivons enfin au niveau le plus applicatif de notre analyse, celui des systèmes ayant pour mission principale ou secondaire le traitement des références. De très nombreux systèmes tombent bien sûr sous cette caractérisation, preuve que la référence est un problème central et incontournable du traitement élaboré de la langue. Il n'existe pas, à notre connaissance, de synthèse récente de ces réalisations, contrairement aux travaux plus anciens sur l'anaphore de Hirst (1981a) et de Bosch (1983), ou la thèse de Vidalenc-Sabourin (1989). On peut toutefois trouver dans les actes des campagnes d'évaluation MUC[20] (MUC-6 1995, MUC-7 1998), de nombreux systèmes évalués sur la résolution des coréférences (cf. VIII.A.3.iii).

[19] Cependant que des critiques de nombreux travaux antérieurs montrent sur des exemples l'insuffisance des critères purement linguistiques, et la nécessité de prendre en compte la pragmatique, i.e. la situation de communication, les gestes, les objets visibles.

[20] Message Understanding Conferences, sur lesquelles nous reviendrons plus en détail.

V.E.1 Classement des systèmes

Très souvent, les systèmes réalisés s'inspirent d'une ou plusieurs des théories décrites plus haut, en particulier des modèles computationnels existants. C'est pourquoi nous n'allons pas fatiguer le lecteur avec de longues analyses ; nous avons choisi de synthétiser dans un tableau les informations que nous possédons sur les principaux systèmes ayant trait à la référence. Nous donnerons en particulier l'évaluation qui a pu être faite de ces systèmes (résultats numériques[21] ou estimations), ainsi que les techniques et connaissances utilisées, afin de dégager pour les différentes théories employées leur pertinence pratique. Pour les besoins de l'exposé, nous séparons les systèmes en trois catégories :

- résolution de la référence dans les systèmes de compréhension de textes (a1), ou résolution de l'anaphore seulement (a2) ;
- résolution de la référence dans les systèmes de dialogue homme-machine (DHM). Il est rare que l'on se limite dans ce cas à la résolution des seules anaphores. Cette catégorie n'étant pas centrale pour nos préoccupations ultérieures, focalisées sur les textes, nous allons être moins exhaustifs, et renvoyer par exemple à la thèse de Bertrand Gaiffe (1992) sur la référence dans le DHM ;
- traitement de la référence dans les systèmes de génération. Nous en avons déjà parlé, et nous n'y reviendrons pas ;
- systèmes génériques/généraux de compréhension de la langue, catégorie dans laquelle nous évoquerons quelques systèmes récents travaillant sur le français, rassemblés dans le rapport de l'ARC A4 du réseau FRANCIL (Sabatier *et al.* 1997) (hors tableau, V.E.3).

V.E.2 Synthèse sous forme de tableau

Le tableau suivant synthétise dans l'ordre les trois catégories : DHM, compréhension de la référence dans les textes, puis de l'anaphore.

[21] L'évaluation de la résolution de la référence sera étudiée en détail au Chapitre VIII. Ici, il suffira de savoir que la mesure numérique du « taux d'erreur » se décompose en deux scores, le *rappel* (références non résolues) et la *précision* (références mal résolues), le succès étant égal à 100% moins l'erreur. La plupart des systèmes ainsi testés l'ont été dans les campagnes MUC (VIII.A.3.iii).

Nom du système - Institution	Auteur(s) - Année	Fonctionnalités - Capacités	Connaissances utilisées pour le traitement des références	Évaluation communiquée	Commentaires
SHRDLU	(Winograd 1972)	DHM pour la manipulation de blocs. Compréhension des références aux blocs	Identification de formes logiques	Exemples	Un des premiers systèmes réalisés Exemple de performances avec des techniques élémentaire
Projet HITS	(Luperfoy 1992)	Interface de DHM appliquée à trois systèmes (éditeur pour CYC, conception d'interfaces et de présentations)	Trois niveaux : base de conn., modèle du discours, analyse linguistique/déictique Utilisation d'« ancres discursives » accumulant les expressions référentielles par appariement syntxico-sémantique	Utilisation satisfaisante dans des systèmes réels	Représentation intéressante des référents. Cf. notre VII.B.4 Utilise pour l'anaphore le système décrit plus bas (Rich et LuperFoy 1988)
CRIN, Nancy	(Pierrel 1987)	DHM en langue parlée, compréhension de la parole	Système multi-expert : le module « sémantique » assigne les référents qui sont des noms propres en général	Application à une messagerie vocale d'un sys. simplifié	Intégration de la reconnaissance de la parole avec l'analyse syntaxico-sémantique des messages
LORIA, Nancy	(Gaiffe 1992)	Manipulation de fenêtres par le DHM (commandes)	Analyse linguistique élaborée (« axiologies ») d'un certain nombre de marqueurs référentiels	Non implémenté	Thèse de doctorat Considérations reprises dans le projet CERVICAL (V.D.5)
EDWARD	(Huls, Claassen et Bos 1995)	Dialogue multi-modal. Interprétation des anaphoriques et déictiques, mono et multimodales	Lexique, réseau sémantique à la KL-ONE, facteurs de saillance (Alshawi)	125 expressions testées, toutes traitées correctement	
SRI	(Kehler, Martin, Cheyer, Julia, Hobbs et Bear 1998)	Dialogue multimodal ; informations touristiques (cartes, annuaires)	Saillance utilisée pour désambiguïser les pointages et les expressions référentielles. Domaine limité aux objets affichés		Il existe une autre réalisation en cours par J.-C. Martin (LIMSI)

COMPRÉHENSION DE TEXTES					
LUNAR	(Woods, Kaplan et Nash-Webber 1972)	Recherche d'information	Ex. : une règle par pronom	Utilisation réelle (?)	Cf. SHRDLU
PAL	(Sidner 1979b)	Planification d'activités (assistant personnel). Résolution des anaphores définies (pronoms, GN – associations)	Utilisation de « frames » pour délimiter les ensembles de référents autorisés	Exemples	Détails au V.D.2.i
CAPTURE	(Alshawi 1987)	Extraction d'informations à partir de textes. Nombreux noms propres ou GN définis	Modélisation très générale de la saillance, prétraitement linguistique limité	Exemples	Cf. notre V.D.2.ii
New York University	(Grishman 1995)	Tâche MUC-6			
La Hack 2 + CogNIAC - University of Pennsylvania	(Baldwin *et al.* 1995)	Tâche MUC-6 : coréférence + résolution des anaphores			
PIE - University of Manitoba	(Lin 1995)	Tâche MUC-6	Résolution de contraintes référentielles entre ER (avec différents poids), une fois le texte entier examiné		Méthode originale, « holiste » et efficace
FASTUS - SRI	(Appelt *et al.* 1995)	Tâche MUC-6			
LaSIE - University of Sheffield	(Gaizauskas, Wakao, Humphreys, Cunningham et Wilks 1995)	Tâche MUC-6	Utilisation d'un réseau sémantique élaboré, adapté à la tâche		
RESOLVE - University of Massachusetts at	(McCarthy et Lehnert 1995)	Intégré à un système pour MUC-6	Utilisation d'arbres de décision pour résoudre la coréférence		

Amherst					
Apple Corp.	(Kennedy et Boguraev 1996b)	Résolution des anaphores et « coréférences évidentes » sur des corpus non restreints	Comparaison des formes de surface des expressions ; expansion d'acronymes		Passage de traitement de l'anaphore à celui de la coréférence
RÉSOLUTION DES ANAPHORES					
–	(Hobbs 1978)	Algorithme de résolution des anaphores pronominales		Etudes numériques dans l'article	
–	(Carbonell et Brown 1988)	Système de résolution des anaphores	Approche multi-stratégies ; des règles complexes utilisant une analyse syntaxicosémantique complète		
–	(Rich et LuperFoy 1988)	Système de résolution des anaphores	Architecture tableau noir pour intégrer des méthodes de résolution partielle ; règles complexes utilisant une analyse syntaxicos-sémantique complète		
–	(Dagan et Itai 1990)	Méthodes statistiques d'apprentissage du rattachement des pronoms			Utile surtout en vue d'une coopération avec d'autres méthodes
–	(Cardie 1992)	Acquisition automatique d'heuristiques de désambiguisation	Peu de connaissances, car limité aux pronoms relatifs		
« Algorithme de Lappin et Leass »	(Lappin et Leass 1994)	Résolution des anaphores pronominales, algorithme implémenté	Intégration des contraintes sélectionnelles avec un mécanisme d'activation inpiré d'Alshawi	Cf. études numériques de l'article	Algorithme « fondateur » ou « paradigmatique »
	(Dagan, Justeson, Lappin, Leass et Ribak 1995)	Un algorithme fondé sur les statistiques		Cf. article	Version statistique de l'algorithme de Lappin et Leass
Apple Corp.	(Kennedy et Boguraev 1996a)		Robuste, textes non restreints		

–	(Mitkov 1998, Mitkov et Belguith 1998)	Résolution des anaphores sur des manuels techniques (anglais, français, arabe, polonais)

V.E.3 Systèmes génériques de compréhension

Aux systèmes étudiant spécifiquement le problème de la référence, il faut bien sûr ajouter les parties traitant la référence dans les systèmes plus généraux. On pourrait ainsi, seulement en France, étudier un à un les systèmes ayant répondu à l'ARC A4 (Action de recherche concertée) du réseau FRANCIL de l'Aupelf-Uref, systèmes répertoriés dans le rapport final (Sabatier *et al.* 1997).

À titre d'exemple, nous citerons le système CARAMEL développé au LIMSI (Sabah 1990a, Sabah 1990b, Sabah et Briffault 1993), qui préconise une division de l'expertise linguistique entre agents, dans une structure de tableau noir. Un des agents est chargé de la résolution des anaphores, à l'aide de règles assez classiques ; toutefois, ses capacités de communication avec d'autres agents lui permettront de faire appel à d'autres connaissances en cas de nécessité. Il faut citer aussi le système TALISMAN, développé à l'IMAG (Stéfanini et Demazeau 1995, Stéfanini et Warren 1996), qui préconise aussi un mécanisme multi-agent, dans lequel la division de l'expertise, et aussi la résolution des anaphores, est relativement semblable à celles de CARAMEL.

V.E.4 Remarques de synthèse sur les systèmes

Afin de conclure sur les systèmes que nous avons mentionnés (essentiellement ceux du tableau V.E.2), nous allons tenter de dégager quelques points communs aux systèmes, qui les opposent assez nettement aux modèles computationnels (V.D) dont les visées sont plus ambitieuses.

Ces systèmes construisent rarement une représentation élaborée des référents, étant le plus souvent limités à la coréférence linguistique, voire à l'anaphore. Le plus souvent, ils sont conçus comme un « module » (plus ou moins intégré à d'autres), et sont déclenchés de façon séquentielle après les analyses syntaxique et sémantique de chaque phrase, suivie éventuellement d'une analyse plus « pragmatique » ou discursive (calcul de la saillance ou de la focalisation). Ces différentes analyses font rarement l'objet d'interactions plus élaborées, telle par exemple celles suggérées par CARAMEL.

Les connaissances utilisées par ces systèmes sont assez variables, ainsi que le niveau d'analyse préalable exigée pour le texte. De façon générale, ces connaissances sont inversement proportionnelles à la robustesse et généralité du système. Parmi les connaissances utilisées, on peut citer les contraintes sélectionnelles : différents accords morpho-syntaxiques et sémantiques, logiques (quantificateurs, négation), anaphoriques (intra-phrastiques) ; puis, des heuristiques plus graduelles : parallélisme syntaxique, calcul de l'activation (facteurs de contexte), calcul de la focalisation (locale ou globale). Par rapport à ces méthodes, nous verrons que notre système (Chapitre VII) intègre des techniques appartenant aux deux catégories, tout en ayant une bonne robustesse.

V.F Conclusion d'ensemble

Notre revue des travaux existants sur la référence obéissait, dans son ensemble, au souci d'être aussi exhaustive que possible dans le cadre d'une thèse. En effet, le problème a souvent été abordé : sa modélisation dans la « pragmatique simulée » ne doit pas contredire la réalité. Nous avons montré comment la conception de la référence a évolué, notamment dans la seconde moitié de notre siècle. Nous avons ainsi mis en évidence un cheminement allant d'une conception logiciste de la référence, vers une prise en compte croissante du contexte d'énonciation, par exemple avec la théorie de la pertinence et la pragmatique.

Plus récemment, les tentatives de réduction de la référence aux sciences de la vie ont surtout montré les limites des formalismes logiques, et proposé des directions de recherche nouvelles. Les résultats restent en grande partie à venir, et il est probable qu'ils devront s'intégrer dans une théorie naturaliste plus générale du langage – un défi pour les sciences cognitives actuelles.

Sur un plan computationnel, si les modèles ont rencontré et ont dû résoudre les problèmes liés au contexte d'énonciation, force est de reconnaître que les réalisations se sont limitées à une prise en compte très réduite de ce contexte. Un élément assez répandu est la notion de saillance ou activation des référents pour un locuteur, qui vient compléter les contraintes de sélection entre expressions référentielles, d'inspiration logiciste. Remarquons cependant que relativement peu de modèles, et encore moins de systèmes, manipulent une véritable représentation des référents, beaucoup se limitant à la résolution des « coréférences » entre expressions.

Quant à la prise en compte du contexte d'énonciation extérieur aux locuteurs, les rares modèles qui tentent de le faire utilisent des *connaissances* pragmatiques fournies explicitement au systèmes (les schémas de tâches, par exemple), et sont très loin de *situer* le système lui-même comme agent dans un environnement – ce qui impliquerait naturellement une autonomie bien plus importante.

Réciproquement, les agents communicants doués d'une certaine autonomie, que nous avons vus au II.D, ne possèdent pas à l'heure actuelle de capacité référentielle. Pour tenter de combler le vide entre ces deux pôles, nous allons exploiter au chapitre suivant des études de neuropsychologie et de linguistique fonctionnelle qui éclaireront le problème de la référence sous un jour nouveau, et nous permettront de proposer un modèle d'utilisation de la référence par un agent.

Chapitre VI

DU FONCTIONNEMENT DE LA RÉFÉRENCE À LA CAPACITÉ RÉFÉRENTIELLE D'UN AGENT

Nous avons examiné au Chapitre V un certain nombre de théories ontologiques, linguistiques et pragmatiques de la référence, ce qui nous a permis de bien mesurer l'importance du phénomène pour toute modélisation du langage. Toutefois, deux aspects importants sont peu abordés dans les études traditionnelles de la référence, deux aspects sur lesquels nous fonderons un modèle de la faculté référentielle. Il s'agit des analyses des bases neuropsychologiques individuelles de la capacité à faire référence, et son fonctionnement dans les situations de communication concrètes.

Nous allons tenter de rassembler les connaissances pertinentes pour ces deux aspects (VI.A puis VI.B), et de les enrichir (VI.D.1) grâce à deux études linguistiques personnelles, montrant la variabilité linguistique des phénomènes référentielles (VI.C). Nous pourrons alors proposer une modélisation de la référence pour un agent (VI.D.3), dans la lignée cadre de la « pragmatique simulée » et des deux études du Chapitre III.

VI.A Bases de la référence dans les langages émergents

Les bases individuelles de la faculté référentielles, peu étudiées en elles-mêmes, seront plus claires si nous les étudions d'abord dans deux des trois cas d'émergence d'une langue (cf. II.B ci-dessus) : l'acquisition du langage – pour les enfants normaux, VI.A.1.ii, et sourds-muets, VI.A.1.iii – et la créolisation (VI.A.2). En effet, ces cas présentent des versions simplifiées du langage ou des « langues-limite », et les études de phénomènes référentiels y sont plus nombreuses. Nous évoquerons en final des études pouvant se rapporter à la faculté référentielle de l'adulte (VI.A.3) ; il s'agira le plus souvent d'études psycholinguistiques, car la neurophysiologie en propose peu. Il faut noter que les études concernant la référence en elle-même sont quasiment inexistantes dans ces domaines, peut-être parce que le sujet est réputé trop abstrait, ardu, ou « philosophique ». Ce sont donc des études périphériques que nous tentons de rassembler, afin de dégager un tableau d'ensemble.

VI.A.1 Acquisition du langage et référence

VI.A.1.i Capacités cognitives pré-linguistiques

VI.A.1.i.a L'identification des objets

Une des facultés qui conditionnent, à un niveau élémentaire, la fonction référentielle est la capacité de l'enfant à *isoler*, dans le flot des perceptions provenant du monde, des objets ou entités, et à leur *attribuer une certaine catégorie* – éléments auxquels l'enfant pourra par la suite apprendre à associer des mots. Pour ce qui est des objets, le cas le plus étudié est la perception visuelle de l'unité des objets dans des scènes visuelles potentiellement complexes, où foisonnent les contours et les ombres souvent superposés ou en mouvement[1].

La perception des objets est depuis longtemps une propriété étudiée expérimentalement en psychologie (rappelons Helmholtz puis les gestaltistes). Cette est d'abord apparue comme inhérente au système perceptif, mais en réalité elle émerge graduellement, entre quatre mois et deux ans (Spelke 1990). Les premiers critères de séparation en objets seraient la séparation selon la profondeur dans le champ visuel, et le mouvement indépendant. Les contours ou les propriétés des surfaces, bien que différenciés par l'enfant, ne sont utilisés que tardivement pour l'identification des objets, après l'apparition des capacités préhensiles et locomotrices.

Les objets sont perçus aussi comme ayant une certaine continuité dans le temps, en analysant la continuité des mouvements même lorsque les objets sont temporairement occultés. Par ailleurs, des études sur la perception des objets dans les modalités visuelle et tactile montrent que ces deux modalités opèrent un découpage homologue en objets. Cela laisserait penser que le mécanisme d'identification d'objets est relativement abstrait, indépendant de l'expérience, et se trouve à l'interface du *système perceptif* et du *système conceptuel* ; l'enfant dispose très tôt de ce mécanisme, ou en tout cas d'une partie suffisante pour identifier des objets standard.

La compréhension et simulation de cette faculté est abordée par exemple dans le modèle RCI proposé par Edelman pour la vision (II.C.2.ii), dans une perspective gestaltiste. Pour ce qui nous intéresse ici, la mise en évidence de ces capacités permet de répondre de manière naturaliste à une partie de l'interrogation philosophique sur l'acquisition du sens des mots par imitation (l'exemple du « *gavagai* » de Quine, V.A.4.i.a). Lorsqu'un lapin (ou tout autre « objet ») apparaît dans le champ visuel d'un adulte ou d'un enfant, celui-ci forme spontanément la représentation classique d'*un objet*, qui inclut toutes les parties solidaires et exclut l'environnement et les propriétés transitoires. Le système perceptif, fruit d'une sélection naturelle d'un intérêt visible, ne privilégie pas, voire exclut, des représentations du type « pattes avant et queue de lapin rentrant dans le terrier depuis la droite, sous le soleil de midi ».

[1] Toutefois, les enfants aveugles acquièrent aussi la notion d'objet, et plus généralement le langage, au même rythme apparemment que les voyants (Landau et Gleitman 1985).

VI.A.1.i.b L'identification des catégories

Une autre partie de l'interrogation quinéenne se retrouve dans le problème de l'acquisition des catégories : comment l'enfant sait-il que « *gavagai* » s'applique au *lapin* et non au lapin gris, ou à ce lapin particulier ? En d'autres termes, quelle est la capacité des enfants à former des catégories ou concepts, à partir des instances observées, et sur quelles bases ? Et aussi, quel est le niveau de généralité de ces catégories ?

Il s'agit à nouveau d'un domaine étudié par de nombreuses expériences psychologiques – voir notamment (Carey 1985). L'acquisition des premiers mots montre un certain flou dans la catégoricité : les noms sont en général utilisés de façon trop restrictive, et les verbes de façon trop générale. Il reste que, par rapport au grand nombre de pseudo-catégories possibles, les enfants regroupent très tôt les objets selon les « genres naturels », par exemple *objet inanimé*, *animal*, *humain*, avant un an. Des expériences ont montré que des enfants plus âgés sélectionnent spontanément, en général, les traits corrects déterminant les catégories usuelles, aux dépens de traits plus visibles. Cette aptitude paraît tellement naturelle que la question de son éventuelle innéité fait l'objet de nombreux débats. Il faut en tout cas constater que les enfants adoptent très tôt une structuration en catégories compatible avec la structure par « genres naturels » connue chez les adultes.

VI.A.1.i.c Esquisse d'explication computationnelle

Il ne nous appartient pas de nous prononcer sur la question des « genres naturels »[2], mais nous pouvons faire l'hypothèse suivante. On peut concevoir la catégorisation comme le regroupement d'objets ayant un certain nombre de traits en commun[3]. Dans ce cas, deux niveaux de choix sont possibles : le nombre de traits discriminants, puis les traits eux-mêmes. Si les traits peuvent être classés par ordre d'importance (forme > couleur > poids, par exemple) alors la finesse d'un système de catégories est donnée par le nombre de traits qui définissent les catégories. Ainsi, la prise en compte d'un très grand nombre de traits produit un système trop fin, où les classes sont égales aux individus, alors qu'un nombre réduit de traits produit un système catégoriel trop grossier.

Ce qui serait alors commun aux hommes (et peut-être inné) serait un certain niveau de granularité du système catégoriel, fondé peut-être sur des capacités neuro-physiologiques communes aux hommes, et émergeant dès le plus jeune âge. On obtiendrait ainsi une explication naturaliste des « genres naturels », vérifiable à l'aide

---

[2] Les « genres naturels » sont les catégories dans lesquelles diverses cultures divisent les entités présentes dans la nature (par exemple les espèces d'animaux). Ces catégories présentent une uniformité assez étonnante dans le monde, notamment dans le principe de fonctionnement, et sont souvent compatibles avec les classifications scientifiques (génétiques ou phylogénétiques, pour les animaux) – voir notamment (Atran 1990, Carey 1985).

[3] Cette conception est très discutée, notamment parce que les objets d'une catégorie ne partagent pas tous les traits minimaux nécessaires. D'autres formalisations utilisent la notion de prototype, ou la logique floue – voir par exemple (Rosch 1977).

d'un modèle de catégorisation biologiquement réaliste (par exemple ceux d'Edelman). À ces catégories correspondrait alors, de façon explicablement uniforme, un système de noms communs que l'enfant peut acquérir très rapidement, parfois en une seule présentation (« fast mapping », cf. (Carey 1985)). Ces noms désigneraient ainsi les *niveaux de base* du système catégoriel (Rosch 1978) (« chien » par rapport à « quadrupède »).

Quoi qu'il en soit, cette section VI.A.1.i montre grâce aux travaux psychologiques l'existence précoce, préalable au langage, de deux capacités complémentaires : la faculté d'identifier les entités correctes dans l'environnement et la faculté de catégoriser ces entités en un système de classes d'une granularité assez universelle. Ces propriétés justifieront notre modèle des *noms propres* / *noms communs* (cf. VI.D.3.i), et nous allons maintenant les relier à l'acquisition du langage.

VI.A.1.ii Émergence de l'expression linguistique de la référence

VI.A.1.ii.a Les noms d'objets

La rapidité avec laquelle les enfants acquièrent le lexique pourrait donc s'expliquer par la puissance du système conceptuel, qui fournirait les concepts auxquels l'enfant attache les mots entendus. Naturellement, il y a un élément essentiel qui est la possibilité même de saisir les mots dans le flux de parole, fruit d'une maturation phonétique, puis phonologique, qui démarre très tôt (cf. II.B.3.i.a) ; le mot à apprendre est rarement présenté de façon isolée, surtout dans le cas fréquent où on ne parle pas « mamanais » à l'enfant.

En ce qui concerne l'apprentissage d'un nouveau mot, il s'accompagnerait d'un processus de recherche active, dans l'environnement proche, des objets pouvant correspondre aux mots inconnus prononcés par l'adulte. Le suivi du regard de l'adulte (synchronisation déictique) permetrait à l'enfant de trouver l'objet désigné (Baldwin 1991).

On a montré que les noms d'objets ainsi acquis ne possèdent pas un sens « complexif », i.e. ne désignent pas l'objet *plus* son contexte, mais seulement l'objet, indépendamment du contexte ou de l'action accomplie (Huttenlocher et Smiley 1987). Il apparaît aussi dans cette étude que les usages « plus général » et « moins général » sont rares ; le second ne s'y produit pas (peut-être parce que les objets ne sont pas assez variés) et les occurrences du premier sont interprétées comme ayant une fonction communicative. Ces expériences montrent que le sens acquis pour les noms des objets reproduit le sens de l'âge adulte : les mots désignent des catégories d'objets, dont l'extension variera avec l'âge, et dont l'intension est initialement absente ; ils ne désignent pas un certain contexte indifférencié, comme on a pu le penser.

VI.A.1.ii.b La distinction objet spécifique/non spécifique

L'acquisition du système d'articles et du pluriel a été étudiée dans les expériences de Michael Maratsos (1976)[4]. Celles-ci montrent qu'à l'âge de trois ans les enfants sont capables d'exprimer linguistiquement la différence entre un usage *spécifique* d'un nom, pour désigner un individu d'une catégorie, et un usage *non-spécifique*, pour désigner n'importe quel individu de la catégorie[5]. L'acquisition d'un marqueur du pluriel suit de peu celle des articles, et permet d'exprimer autrement l'usage non-spécifique (du moins en anglais).

Dans ces expériences, on demandait aux enfants de compléter des phrases du type suivant (la partie à compléter est en italiques, nous traduisons) :

(1) Bill a acheté un chien et un chat, mais ses enfants n'aiment que *le* chien.
Bill voulait acheter un chien et un chat, mais il n'a pas pu trouver *un* chien assez beau.

On remarquera que ces phrases n'exploitent pas la distinction article défini/ indéfini en elle-même, mais la distinction entre un sens spécifique et non spécifique, qui s'exprime dans la deuxième partie des phrases, et non dans la première (cf. mots soulignés). Ce serait cette dernière distinction que l'enfant maîtrise dès trois ans, et non celle des articles (du moins en anglais) – ce qui ne peut manquer d'étonner, puisqu'elle est n'est pas marquée linguistiquement en tant que telle dans ce que l'enfant entend.

Cette capacité de l'enfant à différencier très tôt, grâce à des modifieurs, l'usage spécifique de l'usage non-spécifique des noms communs, est à rapprocher de la distinction entre l'identification des individus (prolongement de la faculté de perception des objets) et de leurs catégories (fondée sur le système catégoriel évoqué plus haut). L'apprentissage purement linguistique de cette distinction semble mal aisé, et c'est pourquoi il faut prendre en compte son ancrage cognitif.

VI.A.1.ii.c Les pronoms anaphoriques

L'acquisition des pronoms est un élément important dans l'étude de l'émergence de la référence, mais elle présuppose déjà des capacités complexes, parmi lesquelles celles exposées plus haut. Nous la citerons donc par souci d'exhaustivité. L'étude des pronoms est en général abordée par le côté linguistique, assez détaché de la situation d'emploi[6].

À titre d'exemple, les études de Jakubowicz (1991) en langue française se placent dans le cadre de la grammaire universelle chomskyenne, et cherchent à

> *…déterminer si les comportements attestés [i.e. l'usage des pronoms] résultent de l'élaboration progressive du lexique ou sont plutôt une*

[4] N'ayant pas eu accès à cet ouvrage, nous reprenons l'aperçu donné dans (Bickerton 1981:146 sq.).

[5] Ce dernier usage pourrait être considéré comme une référence « à la catégorie elle-même », mais cela doit probablement être distingué de la réification de la catégorie, comme dans « le lion est un exemple de mammifère ».

[6] On peut mentionner les deux volumes édités par Barbara Lust (1987), et des travaux au LPE, Paris, autour de M.-F. Ehrlich (concernant un âge plus avancé, capable de lecture).

conséquence de la maturation tardive d'un des Principes du Liage au cours du développement du langage chez l'enfant. (Jakubowicz 1991:235)

Un des résultats les plus significatifs est que l'usage des anaphores (au sens génératif : pronoms réfléchis et réciproques) est maîtrisé à un niveau proche de l'adulte dès l'âge de trois ans, alors que l'usage des pronoms personnels sujet et objet est en lente amélioration entre trois et six ans. Ces considérations renseignent peu sur le fonctionnement de la référence en elle-même, mais tendent à accréditer l'idée que seul un nombre restreint de pronoms ont un fonctionnement prédit par la grammaire vue comme système linguistique autonome, et que les autres suivent des critères variables, plutôt contextuels ou discursifs.

VI.A.1.iii La distinction entre noms et verbes dans les gestes spontanés

Nous avons déjà rencontré au II.B.3.iii la situation des enfants sourds de parents entendants, qui prenaient l'initiative d'un code de communication gestuel, présentant des propriétés semblables au langage. Il faut citer ici l'usage des signes représentant des *noms*, par rapport à ceux représentants des *verbes* ou *adjectifs* (Goldin-Meadow, Butcher, Mylander et Dodge 1994). Les expériences confirment le caractère central de l'opposition entre noms et verbes, que nous retrouverons dans le langage adulte avec les travaux de Hopper et Thompson (1984) au VI.B.3, cités aussi par Goldin-Meadow *et al*.

Les résultats montrent que les sujets ou *thèmes* des assertions sont en général des objets, auxquels correspondent des « gestes nominaux », et les prédicats ou *rhèmes* sont des actions et attributs, auxquels correspondent des « gestes verbaux/adjectivaux ». Ces deux catégories de gestes manifestent des différences de marquage morphologique – abréviation des premiers, inflexion des seconds – et syntaxique – gestes nominaux placés après les déictiques (dans des phrases nommant des objets), gestes verbaux placés avant les déictiques. Aussi, lorsque des gestes nominaux et verbaux sont produits ensemble, les premiers prennent la place qu'occupaient les « gestes déictiques » (pointages).

Les auteurs rappellent en conclusion que la distinction entre objets et actions est une capacité conceptuelle fondamentale, et à ce titre assez indépendante du langage. L'étude d'enfants soumis à des stimuli linguistiques très limités montre que ceux-ci parviennent toutefois à spécifier cette distinction dans un code de communication, à l'aide de moyens qui se retrouvent dans de nombreuses langues.

Le caractère référentiel des gestes nominaux semble assez imprécis, au contraire des pointages directs. Aussi, la détermination ultérieure de ces gestes nominaux, afin d'exprimer le caractère spécifique/non-spécifique, semble ici absente, contrairement à un autre cas où les enfants contribuent à l'émergence d'une langue : la créolisation.

VI.A.2 Expression de la référence dans les langues créoles

Nous exploitons ici un autre travail de Derek Bickerton, son modèle d'emploi des articles dans les langues créoles (Bickerton 1981). Le lexique ou les catégories sémantiques des langues créoles, prérequis pour la référence, ne semblent pas se différencier de ceux des autres langues. En revanche, selon Bickerton, l'expression du

caractère défini ou indéfini des entités nominales présente une remarquable analogie avec celle mise en évidence par Maratsos (cf. VI.A.1.ii.b ci-dessus) chez les enfants.

Plus précisément, Bickerton tente de dégager la logique d'emploi des articles dans quelques exemples de langues créoles. Il reprend ainsi la distinction entre un sens spécifique d'un GN, référence à un exemplaire précis, et un sens non-spécifique, désignant n'importe lequel des exemplaires de la catégorie. À ce trait que nous pouvons qualifier de sémantique, Bickerton ajoute le caractère « présupposé » (connu) de l'entité évoquée par le GN, qui dépend de la situation de communication du discours. Ces deux traits (**s**pécifique (+/–), **p**résupposé (+/–)) diviseraient l'espace « sémantique » en quatre régions auxquelles on rattache, dans chaque langue, l'usage d'un ou plusieurs articles. Les quatre régions ou cas reçoivent aussi des noms plus traditionnels (Figure VI-1).

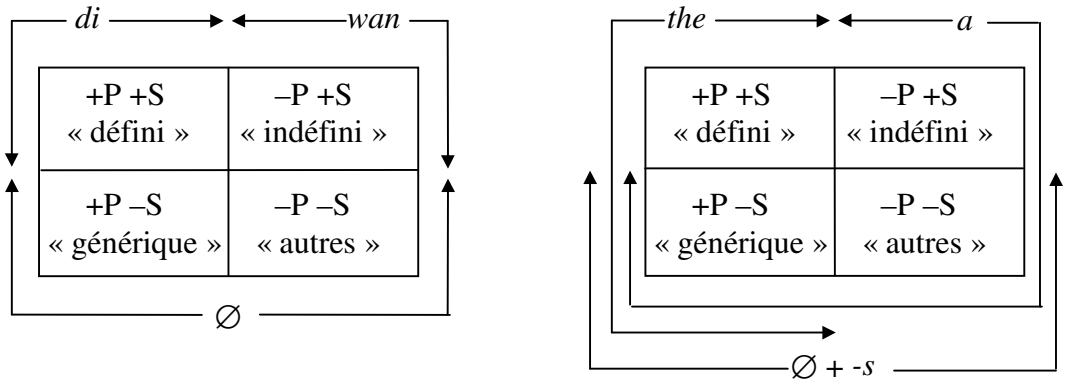

Figure VI-1. Système des articles en créole du Guyana (à gauche) et en anglais (à droite) selon (Bickerton 1981).

La thèse de Bickerton est que les articles des langues créoles suivent en grande majorité le découpage logique selon les deux traits P et S, montré sur la Figure VI-1 pour l'exemple du créole guyanais, à base anglaise (les items lexicaux varient bien sûr d'une langue créole à une autre, en fonction de la base). En revanche, une langue comme l'anglais présente une certaine liberté dans le choix des articles. Ces résultats sont à relier au processus de créolisation et aussi à l'acquisition du langage : si l'espace <P, S> est ancré cognitivement, la solution la plus « naturelle » est d'exprimer les différentes régions par des marqueurs différents. C'est ce qui se produirait lors de la créolisation, en l'absence de pression de la part des parents, et c'est aussi une tendance observée chez les enfants normaux (distinction spécifique +/–, cf. VI.A.1.ii.b). À titre de conjecture, Bickerton affirme la continuité du codage linguistique des quatre régions : jamais deux région diagonales ne recevront à elles seules la même expression. On comprend enfin que cet exemple, parmi d'autres, serve à Bickerton (1984) pour étayer son hypothèse du bioprogramme linguistique (II.B.3.ii.b).

Ce modèle a le mérite d'étudier l'ancrage cognitif des articles, et à ce titre nous l'utiliserons au VI.D. Il permet aussi de mettre en évidence la liberté d'emploi des articles en anglais (nous proposerons un schéma pour le français au VI.D.1), ce qui peut

expliquer l'incapacité des théories logiques à les caractériser, et nos observations expérimentales du VI.C. Ce modèle tente aussi de caractériser les *situations* d'emploi des articles, ce qui permet justement de saisir la spécificité des systèmes créoles par rapport aux articles défini/indéfini du français[7].

On peut toutefois critiquer le manque d'exemples, et surtout l'inexistence d'un vrai protocole d'analyse des données permettant de bien préciser le statut réel de l'objet désigné. Cela fait qu'il y a un risque non négligeable de privilégier la synthèse théorique aux dépens des faits réels, et cela a parfois été reproché à Bickerton (cf. le commentaire de P. Seuren dans (Bickerton 1984:208-9)).

Une analyse plus approfondie met en évidence d'autres faiblesses du modèle. On peut s'étonner du croisement d'un trait sémantique avec un trait discursif ; on s'interroge aussi sur la possibilité de fournir un véritable sémantique formelle, logique, qui définisse précise ces deux traits. Ces lacunes apparaissent aussi si l'on tente d'interpréter les étiquettes attachées aux quatre cas, dont la symétrie est discutable. En particulier, le cas « autres » se comprend assez mal (Bickerton évoque les GN dans des contextes modaux, par exemple). Il manque aussi dans (Bickerton 1981) plus d'exemples pour situer ces distinctions (en l'absence d'un formalisme), car le rattachement à l'un des quatre cas précédents n'est pas toujours clair. Nous verrons par ailleurs une variante de ce modèle, accompagnée d'exemples, élaborée par Bickerton et Talmy Givón (VI.B.2).

VI.A.3 Ancrage de la faculté référentielle chez l'adulte : pistes de recherche

Étant donné que l'on ne peut pas dire grand-chose de la référence dans le processus d'évolution du langage humain, il nous reste à évoquer quelques études qui abordent son ancrage chez l'adulte[8]. Nous avons examiné *in extenso* les points de vue de la philosophie et de la linguistique (Chapitre V), toutefois ces approches ne fournissent pas de réduction naturaliste exploitable pour notre paradigme de la « pragmatique simulée ». Afin de combler ce vide, nous allons indiquer quelques directions de recherche neurophysiologiques qui tentent d'explorer le système représentationnel. En effet, et nous venons de le voir pour l'acquisition, plusieurs capacités jouent un rôle déterminant pour la référence : tout d'abord la perception de l'objet en son unité, puis son identification et sa catégorisation, et enfin le lien avec les aspects linguistiques.

Sur cette longue voie, seuls les tout premiers pas ont été franchis. Nous avons déjà évoqué les travaux sur les « détecteurs de traits » groupés en cartes dans les premiers

[7] On sait par exemple que *l'usage* des articles français ne correspond pas exactement à celui des articles anglais (par exemple, les noms abstraits sont souvent privés d'article en anglais). On ne renseigne donc pas suffisamment sur le système des articles quand on dit qu'en créole à base française (antillais, haïtien ou mauricien) l'article indéfini est *enn* (ex. : *enn chat*) et que l'article défini est *la* (ex. : *chat la*), comme on voit dans (Chaudenson 1979:91) ; il faudrait préciser aussi dans quelles conditions on *utilise* ces articles.

[8] Si l'on adopte la description de l'évolution du langage donnée par Bickerton (II.B.2.ii), on peut dire que la protolangue pouvait déjà utiliser un équivalent des noms communs, puisque le système représentationnel était déjà élaboré.

étages du système visuel (Hubel et Wiesel 1979). Le problème qui se pose maintenant est l'intégration de ces traits, leur groupement selon les objets auxquels ils appartiennent (et non selon leur adjecence) : il s'agit du « binding problem », ou question du liage des stimuli perceptifs. Au delà des travaux psychologiques, les neurologues tentent d'en élucider les mécanismes.

À titre d'exemple, voici quelques moyens d'étudier ce problème. Une série d'études (Fries, Roelfsema, Engel, König et Singer 1997) utilise des chats amblyopes, c'est-à-dire dont le fort strabisme fait qu'un œil devient dominant alors que l'autre, bien que fonctionnel, perd tout pouvoir. Un objet pourra donc être *perçu pleinement* avec l'œil dominant, mais jamais avec l'autre, bien que l'information optique soit transmise. Neuronalement, on constate une *synchronisation* en fréquence des différents neurones sensibles aux traits lorsque l'objet est présenté à l'œil dominant, synchronisation qui est absente pour l'œil dominé, bien que les neurones correspondants (distincts des premiers) soient activés. D'autres expériences étudient les bases neuronales de l'identification d'un objet *pertinent pour une action*, autrement dit la formation d'une « représentation de l'objet » dans le champ attentionnel (Roelfsema, Engel, König et Singer 1996). Des singes sont entraînés à réaliser une certaine action lorsqu'ils identifient un certain objet ; les mesures de l'activité neuronale montrent qu'à l'identification des objets *pertinents* pour l'action correspond aussi une synchronisation des détecteurs locaux.

Les études de cette même équipe concernent d'un certaine façon la *conscience* d'un objet (Engel, Fries, Roelfsema, König et Singer 1998), (très) préliminaire à la possibilité d'y faire référence. Il s'agit à nos yeux plutôt d'une *conscience primaire*, dans le sens décrit par Edelman (II.A.2.iii), et non de la conscience d'ordre supérieur. Le mécanisme de synchronisation paraît une proposition plausible, et l'utilisation de certains créneaux dans les signaux oscillants permettrait d'expliquer la présence de plusieurs objets dans le champ attentionnel. Les études expérimentales citées sont compatibles avec les considérations plus théoriques de Moshe Abeles (1982) sur un codage temporel de l'information neuronale – direction de recherche très prometteuse à l'heure actuel.

Ces études sont certes très préliminaires, et même les modèles de catégorisation, comme ceux présentés par Edelman, n'ont pas encore été complètement expliqués neuronalement, bien qu'ils se fondent sur de nombreux éléments vérifiables. Nous allons donc nous tourner vers la linguistique pour des considérations sur *l'emploi* de la référence en situation.

VI.B Analyses fonctionnalistes de la référence

Les travaux des linguistes fonctionnels ne semblent pas être traditionnellement comptés parmi les « travaux sur la référence ». Cette relative exclusion montre bien l'emprise qu'exercent sur le sujet les approches traditionnelles, logicistes ; les études plus récentes de la référence, notamment les théories naturalistes et la pragmatique ne semblent pas tirer un parti suffisant de la linguistique fonctionnaliste.

Nous souhaitons montrer ici l'importance de cette approche pour la « pragmatique simulée », et tirer de ces travaux quelques principes essentiels à la modélisation par agents informatiques des phénomènes référentiels.

VI.B.1 L'approche fonctionnaliste

Les approches que nous qualifions de fonctionnalistes partagent un souci important du rôle de chaque énoncé dans son contexte d'emploi (*fonction*), et exploitent de nombreux exemples réalistes dans une grande variété de langues. Le cadre théorique est réduit au profit de l'analyse et du classement de phénomènes linguistiques concrets.

Nous exploitons ici, parmi les représentants de ce « courant », les travaux de Talmy Givón (1979a, 1979b, 1984-1990, 1993, 1995), Derek Bickerton (1981), Paul Hopper et Sandra Thompson (1984, 1985). De façon plus générale, les *fonctionnalistes* citent souvent les travaux sur les universaux linguistiques dans la lignée d'Edward Sapir (1921), tels ceux de (Greenberg, Ferguson et Moravcsik 1978, Hawkins 1988) – pour plus de détails, voir les ouvrages cités et leurs bibliographies. Nous utilisons ici deux contributions très représentatives, qui appliquent les principes fonctionnalistes à l'étude de la référence, et sont en rupture avec l'approche « logiciste », afin d'éclaircir la question de *l'usage* de la référence.

VI.B.2 « Caractère défini et référentialité » : Talmy Givón (1984)

Un chapitre de Givón (Givón 1984-1990:406-435) fournit l'intitulé de cette section ; ce chapitre reprend de travaux plus anciens (Givón 1978), et se prolonge, pour l'anglais, dans (Givón 1993 :ch.5)[9].

VI.B.2.i Usage référentiel et caractère défini

Givón distingue d'abord l'usage référentiel d'un GN de son usage non-référentiel, et met en évidence sur les exemples (29) un continuum entre les deux cas (les premiers exemples sont plutôt référentiels, contrairement aux derniers – nous traduisons) :

(29) Si vous voyez *l'homme* au chapeau vert, dites-*lui*…
Si vous voyez *un homme* au chapeau vert, dites-*lui*…
Si vous voyez *quelqu'un* avec un chapeau vert, dites-*lui*…
Si vous voyez *n'importe qui* avec un chapeau vert , dites-*lui*…

L'hétérogénéité du cas non-référentiel est illustrée sur les exemples suivants :

[9] Cette étude aboutit à un « modèle cognitif » de résolution de la référence fondé sur la cohérence du discours, moins original que les analyses qui nous intéressent ici, puisque des travaux importants existent déjà sur la cohérence et la référence, comme nous l'avons vu au V.D. Nous avons déjà exposé (**V.C.1**) les propositions de Givón pour modéliser un univers du discours, contenant un fichier des référents actifs alimenté par trois sources : le discours, le contexte déictique immédiat, et le fichier des référents permanents.

(30) *Tous les professeurs* sont chauves. [non référentiel générique]
Je cherche *un professeur*. [non référentiel individuel]
Jean est *un professeur* célèbre. [non référentiel attributif]

Givón propose ce qui serait une « relation nécessaire » :

> *Sous la portée de la modalité propositionnelle* FACTUELLE, *les arguments nominaux peuvent être seulement référentiels. Sous la portée de la modalité* NON FACTUELLE, *les arguments nominaux peuvent aussi être non référentiels.*

Une deuxième distinction importante est opérée à l'intérieur de l'usage référentiel. Il s'agit d'une définition pragmatique des usages *défini* et *indéfini*. L'usage est défini si le locuteur pense que son interlocuteur peut assigner au nominal une référence unique. Il est indéfini s'il pense que celui-ci ne peut pas assigner une référence unique ; cela utilise donc le fichier des référents actifs dans la situation de communication. S'ajoute à cela le cas défini et non-référentiel, appelé *générique*. On a donc une intéressante classification pragmatique des fonctions des groupes nominaux :

DÉFINI[10] *↔ RÉFÉRENTIEL INDÉFINI ↔ NON RÉFÉRENTIEL ↔ GÉNÉRIQUE*

Les deux cas extrêmes doivent être vus aussi comme adjacents, bien qu'ils appartiennent à des plans conceptuels différents, celui des *individus* vs. celui des *catégories*. Givón situe les quatre situations sur une boucle hélicoïdale, analogue aux quatre cases du rectangle de Bickerton (Figure VI-1) – il s'agit en effet d'un travail réalisé en collaboration.

Ensuite, Givón tente de trouver un répertoire général de mécanismes linguistiques permettant d'exprimer ces situations d'emploi, avec des exemples dans différentes langues. Par rapport au schéma quaternaire ci-dessus, l'hypothèse de Givón et Bickerton est que seules des situations voisines sur la boucle hélicoïdale ou le rectangle peuvent partager un même mécanisme d'expression (par exemple l'article défini).

VI.B.2.ii Catégories sémantiques et catégories pragmatiques

Givón fait un pas supplémentaire vers la prise en compte de la pragmatique lorsqu'il remet en question la classification précédente, taxée de « sémantique » :

> *Jusqu'ici, nous avons traité le contraste entre le référentiel et le non référentiel comme sémantique, faisant intervenir l'intention référentielle du locuteur indépendamment du contexte de communication. Ce traitement, bien que divergeant de la tradition logique, préserve néanmoins un trait de cette tradition, à savoir l'approche de la référence comme un attribut inhérent à la sémantique propositionnelle. Dans cette section, nous allons voir que […]*

[10] Le cas « défini » est forcément référentiel. Nous avons remplacé les '>' d'origine par '↔'. Il y a dans les signes '>' une idée de particularisation ou de marquage, en allant de la gauche vers la droite, mais la notion de marquage n'étant pas toujours consistante dans le chapitre exploité, nous la laisserons de côté (par exemple, il n'est pas clair si l'incorporation de l'objet non référentiel au verbe reflète un marquage ou non).

dans le langage humain la référence est une question d'importance communicative. (Givón 1984-1990:423-435)

La prise en compte de l'importance communicative fait intervenir la « saillance des référents », qui n'est pas un trait sémantique. Givón montre qu'il y a des situations où le choix de la modification linguistique du nominal ne dépend pas du caractère référentiel (entité identifiable ou non) ou défini (entité connue ou non), mais de *l'importance de l'entité* pour la suite du discours. Cet exemple et quelques autres entraînent une modification du classement initial (Table VI-1).

Les quatre catégories (A, B, C, D) de la Table VI-1 sont déterminées par des ensembles de manifestations linguistiques. Le résultat le plus remarquable est la non-concordance, au centre, des lignes 3 et 4 (ce que Givón appelle facteurs « sémantiques » et « pragmatiques ») : on constate qu'*un usage référentiel peut recevoir les marques d'un usage non référentiel si l'entité n'a pas une importance suffisante dans le discours*. On peut noter la disparition des génériques, et aussi remarquer que les deux dernières lignes rappellent la théorie du centrage (V.D.3.i) – mais leur exploitation est ici assez réduite.

Facteurs	A	B	C	D
Marquage morpho-syntaxique	$\{m_{A1}, m_{A2}, \ldots, m_{An}\}$	$\{m_{B1}, m_{B2}, \ldots, m_{Bn}\}$	$\{m_{C1}, m_{C2}, \ldots, m_{Cn}\}$	$\{m_{D1}, m_{D2}, \ldots, m_{Dn}\}$
± présent dans « fichier actif » ± entité précise	DÉFINI USAGE RÉFÉRENTIEL	INDÉFINI USAGE RÉFÉRENTIEL	USAGE NON RÉFÉRENTIEL	USAGE NON RÉFÉRENTIEL
± saillant dans « ficher actif »	DÉFINI	INDÉFINI + IMPORTANT RÉFÉRENTIEL	INDÉFINI − IMPORTANT RÉFÉRENTIEL	USAGE NON RÉFÉRENTIEL
Continuité discursive (% au passé)	Haute	Basse	Basse	Basse
Persistance discursive (% à l'avenir)	Haute	Haute	Basse	Basse

Table VI-1. Classement des phénomènes référentiels (adapté de Givón (1984:426))

Plusieurs points sont donc à retenir. D'abord la méthodologie : l'utilisation de nombreux exemples réalistes, multilingues et situés. Puis, les principes théoriques : la remise en question du cadre logique, la prise en compte du contexte à travers le modèle « pragmatique » des fichiers, l'importance communicative (nous dirons « saillance »). Enfin, certains traits importants du modèle : la distinction du caractère (non) référentiel et du caractère (in)défini, et les classifications des cas référentiels.

VI.B.3 La référence et la catégorie grammaticale des noms : Hopper et Thompson (1984)

VI.B.3.i Fondements pragmatiques de l'expression linguistique

VI.B.3.i.a Les noms et les verbes

Le but de l'article de Hopper et Thompson (1984) est de montrer que le fondements des catégories grammaticales essentielles du *nom* et du *verbe* ne sont ni morpho-syntaxiques, ni sémantiques, mais discursives et en un certain sens pragmatiques. De nombreux exemples seront cités à l'appui, certains empruntés à Givón.

Pour Hopper et Thompson, les fonctions discursives prototypiques des noms sont les entités, sujets du discours, ou encore les « percepts ayant une stabilité temporelle » (Givón 1979b:320-21) ; les verbes expriment en revanche l'évolution de ces percepts, donc typiquement une action ou un événement. Les auteurs montrent que plus la fonction d'un mot se rapproche de l'un de ces deux pôles, plus celui-ci possède de caractéristiques attribuées en général aux noms respectivement aux verbes, et ce indépendamment de son statut morpho-syntaxique ou sémantique. La fonction est donc nécessaire et suffisante pour déterminer la catégorie morpho-syntaxique.

VI.B.3.i.b Corrélation entre syntaxe, sémantique et discours pour les noms

Que peut donc apporter ce travail assez général à notre étude de la référence ? En ce qui concerne les *entités*, ce serait leur statut référentiel *pour les locuteurs* qui est exprimé dans la langue. Les auteurs citent Du Bois (1980) et Givón (1981), pour qui le codage linguistique des entités d'un discours est assez indépendant de leur sémantique référentielle (l'existence ou l'unicité dans un univers). Hopper et Thompson montrent que l'on peut exprimer dans le langage deux modes d'introduction des entités : le mode manipulable et le mode non manipulable (manipulable *dans le discours*). La catégorie des *noms* est par associée aux *entités manipulables*. Cette dichotomie peut être qualifiée de pragmatique.

Nous pouvons alors reformuler la thèse de l'article cité : le langage n'exprime pas l'existence objective (vériconditionnelle) des entités, mais leur existence dans l'univers du discours. Afin de prouver cela, les auteurs examinent trois dichotomies : manipulable / non manipulable (plan « pragmatique »), référentiel / non référentiel (plan logique ou sémantique), et typique / non typique de la catégorie des noms (plan langagier, marqueurs morpho-syntaxiques). L'exposé montre que le deuxième plan est assez indépendant des deux autres, qui sont en revanche assez proches.

VI.B.3.ii Expression de la non-référentialité

Hopper et Thompson illustrent d'abord le cas où les trois plans précédents s'accordent : les expressions non référentielles sont en général peu saillantes, et souvent

dépourvues de certains traits nominaux[11]. Nous allons examiner ici les cas des noms sous la portée d'une négation, ce qui nous permettra au passage de mieux comprendre un exemple donné par Karttunen (1976:366-67) et V.B.2.iv. Les auteurs citent d'abord Givón :

> *Givón (1979b:93-103) montre que sous la portée de la négation, l'opposition « usage référentiel / usage non référentiel » est neutralisée pour les noms non anaphoriques. Les noms indéfinis sous la portée de la négation ne peuvent bien sûr pas se référer à des entités concrètes…*

Le point délicat est qu'il faut interpréter le caractère indéfini des noms comme une indéfinitude conceptuelle, logique, et certainement pas grammaticale (article indéfini des langues latines ou germaniques). Les auteurs fournissent des exemples en plusieurs langues montrant comment les marqueurs référentiels de certaines langues sont perdus sous la portée de la négation ; en français, on aurait :

(31) (a) J'ai acheté *un* livre.
(b) Je n'ai pas acheté *de* livre.
(c) ?Je n'ai pas acheté *un* livre. [acceptable dans un certain cadre discursif]

Puisque l'argumentation des auteurs finit sur ces exemples, il nous a paru nécessaire de mieux l'expliquer. Ce qu'il s'agit de prouver, c'est la corrélation entre un certain marquage morpho-syntaxique caractéristique des noms et la saillance dans le discours ; sous la portée de la négation la saillance est faible, et le marquage aussi.

Que dire des exemples de Karttunen (1976:366-67) :

(32) Bill has *a car*. *It* is black. [Bill a *une voiture*. *Elle* est noire.]

(33) Bill doesn't have *a car*. **It* is black. [Bill n'a pas *de voiture*. **Elle* est noire.]

(34) Bill didn't see *a misprint*. [≈ Bill n'a pas vu *une erreur*/Bill n'a pas vu *d'erreur*.]

On constate que la traduction des exemples en français montre déjà que selon le sens, le marquage linguistique est différent, en français. La présence du marquage nominal (ici l'article) signale une plus forte saillance dans le discours, qui se traduit concrètement dans ces courts exemples par une reprise plus *naturelle* du nom par la suite. Cela est visible sur la série d'exemples suivants :

(35) ?Jean n'a pas *une voiture*.
[en français, ce marquage ne semble pas compatible avec la négation]

(36) Jean n'a pas *de voiture*. **Elle* est rouge.

(37) Jean n'a pas *de voiture*. *~Elle* lui a été volée.
[inexistence logique, mais saillance non nulle, donc reprise possible]

[11] La liste exacte varie bien sûr avec la langue. Les auteurs en examinent plusieurs : incorporation de l'entité patient au verbe, incorporation de l'entité auxiliaire (« oblique »), composition nominale, prédicats nominaux, nominaux sous la portée de la négation – tous cas où les mots désignant des entités présentent peu de traits nominaux, ont une faible saillance dans le discours, et une valeur référentielle quasi nulle.

(38) Jean n'a plus *de voiture*. *Elle* lui a été volée.
[la portée temporelle de la négation diminue]

(39) Jean n'a pas *sa voiture*. *Elle* est au garage.
[même négation, mais marqueur référentiel fort]

Le point faible de l'argumentation de Karttunen est donc d'utiliser un cas où l'anglais ne possède pas, comme ici le français, des marqueurs référentiels plus différenciés. Son but est de modifier le formalisme logique pour rendre compte, par la logique, des deux sens de « *a misprint* » dans (34). Or, ces deux sens correspondent en réalité à deux situations différentes dans l'univers du discours[12] ; on doit donc bien étudier le lien entre le plan linguistique et l'univers du discours, et non directement le lien entre le plan logique et le plan linguistique.

Certes, le contenu logique des énoncés influence l'univers du discours, mais de façon partielle : des analyses sémantiques fines comme celle de Karttunen sont nécessaires . Elles doivent cependant être effectuées pour chaque langue, car comme on le voit sur les exemples (38) et (39), à une même détermination logique du prédicat, des marqueurs linguistiques de l'entité peuvent faire correspondre deux situations opposées (voiture existante ou non).

VI.B.3.iii Faible manipulabilité, faible typicalité… mais usage référentiel

La suite de l'article de Hopper et Thompson détache encore plus le plan logique du plan de l'expression. Jusqu'ici, les exemples montraient l'expression linguistique de la faible saillance, dans le cas non référentiel ; nous avons cependant vu comment sur l'anglais le marquage linguistique n'était pas assez caractéristique, contrairement au français. Les exemples suivants montrent des situations où des noms qui réfèrent d'un point de vue logique à un référent peu saillant (non manipulable) sont néanmoins dépourvus des marques les plus typiques des noms, et se comportent d'un point de vue linguistique comme des noms non référentiels.

Il y a trois catégories d'exemples dans l'article cité : marquage des noms qui seront ou non repris par la suite (toutes choses égales par ailleurs), marquage des noms anaphoriques, marquage des anaphores associatives de type « partie du corps ». Pour la première catégorie, on donne entre autres deux versions d'un texte en hébreu :

(40) Après le travail, je suis allé à la bibliothèque, je me suis assis, j'ai lu *un livre [sefer-exad = livre+un]*, et c'était un excellent livre.
Après le travail, je suis allé à la bibliothèque, je n'avais rien à faire, donc j'ai lu *un livre [sefer = livre 'tout court']*, et j'ai lu deux journaux, et je suis rentré.

Pour la deuxième catégorie, les auteurs citent un exemple en nahuatl :

(41) Avec quoi a-t-il coupé *le pain* ?
Il a *pain*-coupé avec un couteau. [pain-coupé traduit l'incorporation au verbe]

Enfin, pour les parties du corps, nous reproduisons l'exemple donné en français :

[12] Avec la terminologie de Hopper et Thompson, manipulable / non manipulable, ou plus généralement saillant / non saillant.

(42) Il s'est blessé *la main* / *Il s'est blessé *sa main*. / *Il a blessé *sa main*.

VI.B.3.iv Conclusion sur les trois plans : logique, langage, représentation

Pour résumer la théorie dégagée de la partie concernant les noms dans (Hopper et Thompson 1984), on peut d'abord citer les auteurs :

> *Le fait que le véritable critère qui sous-tend l'assignation des articles (manifestant le caractère typiquement nominal) soit en réalité la saillance ou la PROÉMINENCE dans le discours est fortement suggéré par des faits concernant la relation entre la morphologie nominale et le discours dans [diverses] langues.*

Nous retiendrons donc les rapports entre les trois plans : plan logique, expression linguistique, univers des locuteurs (du discours). Un lien fort unit les deux derniers, alors que le premier, proche de la sémantique des phrases, n'est qu'un des facteurs qui régit l'évolution des univers représentationnels des locuteurs. Il ne faut donc pas chercher – à la manière des logicistes – à relier directement l'expression linguistique aux situations logiques, mais passer par le monde des représentations, et analyser par quels moyens linguistiques propres à chaque langue s'exprime le statut pragmatique des entités.

VI.B.3.v Le changement catégoriel d'un lexème

Pour conclure sur l'ensemble de l'article exploité, i.e. sur les catégories nominale et verbale, nous résumons dans la Table VI-2 l'interprétation des transitions entre catégories. La question est de savoir comment une racine sémantiquement nominale peut être utilisée comme verbe, et inversement.

	Racine « nominale » (sémantiquement)	Racine « verbale » (sémantiquement)
Usage nominal : marqueurs nominaux : m_{N1}, m_{N2}, …, m_{Nn}.	• normal	• ajouts morphologiques élaborés • événement → entité • entité référençable • sémantique : entité ET événement
Usage verbal : marqueurs verbaux : m_{V1}, m_{V2}, …, m_{Vp}.	• pas d'ajout morphologique • entité ↛ événement • usage associatif ou métaphorique • sémantique : événement	• normal

Table VI-2. Sémantique et usage des catégories nominale et verbale : synthèse d'après (Hopper et Thompson 1984:744-46).

Nous pensons donc que les entités sont des concepts du premier ordre, ou des percepts stables, alors que les actions sont des concepts du second degré, qui signifient les changements d'état des premiers. Si ceux-ci à leur tour sont réifiés (figés), alors ils

peuvent acquérir le statut de concept de premier ordre ; le contraire semble en revanche dépourvu de sens.

Cette conception fournit des bases exploitables pour *l'implémentation des catégories de nom et verbe dans les modélisations par agents communicants.* Il s'agit bien sûr d'une simplification, mais le mérite de l'article est de montrer qu'elle est cohérente avec le fonctionnement du langage, même si chaque langue possède ensuite ses propres conventions d'expression. Une implémentation d'un tel cadre donnerait à un programme des bases linguistiques, tout en assurant que ses éventuelles erreurs obéiront à une certaine logique, et ne seraient pas des erreurs dans toutes les langues.

VI.C Classification d'exemples : propositions pour une théorie de l'emploi de la référence en français

VI.C.1 Une expérience de remplacement d'articles

Nous avons analysé les références dans un article politique extrait du *Monde Diplomatique*, afin de mettre en évidence la variabilité de l'expression linguistique de la référence dans un texte « naturel », et non dans des exemples construits[13].

VI.C.1.i Données de l'expérience

L'extrait utilisé compte 348 mots, parmi lesquels il y a 82 occurrences de noms communs, et 16 occurrences de noms propres (Table VI-3 – texte en annexe XI.A). Sur les 82 occurrences de noms communs, on peut estimer qu'environ 30 correspondent à des « formes figées » du point de vue des articles, c'est-à-dire des constructions relativement invariables : prépositions ou adverbes construits avec des noms ('au début de', 'en raison de', 'sans excès'), compléments quasi incorporés ('couronnée de succès'), noms de mois ('en mars'), etc.

Parmi les référents du texte, on distingue d'abord les noms propres, supposés connus du lecteur (Bosnie, OTAN). À partir de ceux-ci, le texte peut référer à certains attributs canoniques, ou à des entités déterminées par ces noms propres : '[l]es ruines de Stari Rasadnik, un faubourg de Brcko, dans le nord-est de la Bosnie' ou 'les troupes de l'OTAN'. D'autres entités sont introduites par l'auteur : 'un frigo pour les bières', 'trois tables sous une bâche' ; toutes cependant n'ont pas une identité aussi précise, par exemple 'la condition de réfugié'. On est en présence d'un continuum référentiel, qui s'étend jusqu'aux emplois où l'entité désignée n'est ni spécifique, ni générique : 'accords de *paix*', '*l'éventualité* d'*une reprise* des hostilités', '*l'indifférence* générale', '*l'optimisme*'. On voit que ces noms désignent souvent des notions abstraites, des états ou propriétés, des sentiments, bref des « entités non manipulables ».

[13] « La paix américaine en panne en Bosnie », par Thomas Hoffnung, *Le Monde Diplomatique/Manière de voir*, n°43, 01-02/1999, pp. 33-38 (paru dans *Le Monde Diplomatique* de 09/1998 sous le titre « La Bosnie à l'heure du "ni guerre ni paix" »).

VI.C.1.ii Étude des remplacements définis ↔ indéfinis

Notre expérience consiste à remplacer le plus possible d'articles définis par des articles indéfinis, puis l'inverse. Les deux textes résultants sont donnés en annexe XI.A pour que le lecteur puisse se convaincre de l'acceptabilité des remplacements effectués ; *acceptabilité* signifie que le sens du texte n'est pas changé (même s'il y a un changement local) et que l'utilisation des deux variantes paraît équivalente.

On constate que l'on peut remplacer environ 12 articles définis ou noms sans articles (sur les 37 occurrences « non-figées ») par des *articles indéfinis*, avec un changement de sens négligeable (noms soulignés sur la Table VI-3, ou figurent aussi les chiffres). Sur les 24 « formes figées », on peut en fait transformer la première dans la liste d'exemples ci-dessous :

(43) dans *l'*indifférence générale → dans *une* indifférence générale

(44) dans le cadre *d'*élections générales → dans le cadre *des* élections générales

(45) partagent *la* même infortune → partagent *une* même infortune

(46) chassés par *les* extrémistes serbes → chassés par *des* extrémistes serbes

(47) considérée comme *le* miroir de la Bosnie → considérée comme *un* miroir de la Bosnie

(48) soumis *aux* pressions des États-Unis → soumis à *des* pressions des États-Unis

Pas d'article = 6 =	• (la condition de) réfugié, (accords de) paix, (négociations de) paix, (une… menace de) guerre, (accords de) paix, (le cadre) d'élections (générales).
Formes figées = 13 =	• (un bout de) café, (en) mars, (à) peine, (sans) excès, (en) décembre, (de) rigueur, (en) septembre, (en bonne et due) forme, (pas de) date (butoir), (n'est plus d')actualité, (en) raison (d'une), (mise en) œuvre, (couronnée de) succès.
Art. déf. sg. (le, la, de+le, de la, à+le, à la) = 19 =	• la (même) infortune, la condition (de refugié), la capitale, l'autorité (du gouvernement de M.), (l'autorité) du gouvernement (de M.), la guerre, (à l'abri) du soleil, le miroir (de la Bosnie), la signature, l'optimisme, la naïveté, le règlement (de la question…), la question (bosniaque), l'état-major (de l'Alliance), (de Stabilisation de) la paix, leur retrait, l'éventualité (d'une reprise des), la mise (en œuvre), du volet (militaire des accords).
Formes figées = 7 =	• au milieu (des ruines de), le nord-est (de), au début (de la guerre), (à) l'abri (du soleil), le cadre (d'élections), (dans) l'indifférence (générale), la (première) fois.
Art. déf. pl. (les, de+les, à+les) = 12 =	• (au milieu) des ruines (de Stari…), les bières, les extrémistes (serbes), des accords (de… Dayton), les parrains (…des négociations), (les parrains…) des négociations (de paix), (soumis) aux pressions (des États-Unis), ses troupes, les (premières) élections, les Bosniaques, les troupes (de l'OTAN), (volet militaire) des accords (de paix).
Formes figées = 4 =	• des (60 000) hommes (de l'IFOR), (se sont rendus…) aux urnes, les 12 et 13 septembre 1998, (une reprise) des hostilités.
Art. indéf. sg., et numéraux = 11 =	• un bout (de café), un faubourg, un frigo, une bâche, une scène, une affaire (de quelques mois), un retrait, un (nouveau) mandat, un mandat (qui…), une (quelconque) menace (de guerre), une reprise (des hostilités).
Formes figées-3	• (il y a seulement) un an, (pendant) un an, une (nouvelle) fois.

Indéf. pl. (art., adj., num.) = 5 =	• trois tables, quelques « clients », des Serbes, des Musulmans, quelques mois.
Formes figées = 2	• deux ans (plus tard), (près de) 30 000 hommes.
Noms propres = 16 =	• Stari Rasadnik, Brcko, Bosnie, Sarajevo, M. Alija Izetbegovic, Brcko, Brcko, Bosnie, Paris, Dayton, l'Alliance atlantique, Bosnie, l'Implementation Force (IFOR), l'OTAN, la Force de Stabilisation (de la paix) (SFOR), Bosnie.
«Pronominaux» = 5 =	• les premiers, les seconds, certains, (proclamait-)on, personne.

Table VI-3. Noms du texte étudié classés selon leurs articles. On souligne les cas où on peut remplacer par l'indéfini, ou par le défini (le premier GN accepte les deux).

Parmi les noms que l'on ne peut pas changer figurent principalement les éléments reliés de façon unique aux noms propres ('*la* signature… des accords… négociés à Dayton', 'l'état-major de l'Alliance') ; cette unicité se propage de proche en proche à partir des noms propres. Il s'agit en général d'instances uniques ou bien déterminées de leur catégorie, reprises plusieurs fois dans le texte, et supposées connues du lecteur.

Dans le sens inverse, le remplacement des indéfinis par des articles définis est encore plus aisé. On peut effectuer entre 7 et 10 remplacements (soulignés sur la Table VI-3), parmi les 16 cas (et 5 « formes figées »). On a ainsi, par exemple :

(49) *Un* frigo pour les bières, trois tables sous *une* bâche et *quelques* « clients » : *des* Serbes et *des* Musulmans qui partagent… → *Le* frigo pour les bières, trois tables sous *la* bâche et *les* « clients » : *les* Serbes et *les* Musulmans qui partagent…

(50) était *une* affaire de quelques mois → était *l*'affaire de quelques mois

(51) en raison d'*une quelconque* menace de guerre → en raison de *la* menace de (*la*) guerre

(52) l'éventualité d'*une* reprise des hostilités → l'éventualité de *la* reprise des hostilités

On remarque donc que très peu d'indéfinis semblent ne pas pouvoir être changés ; les véritables numéraux ('un an', 'deux ans') en font partie. Les autres cas introduisent de nouveaux objets : 'un bout de café', 'une scène inimaginable' et 'une nouvelle fois', mais on peut se demander si parfois leur construction avec l'article défini ne serait pas acceptable à l'oral.

Nous nous sommes surtout attaché à l'opposition défini/indéfini, aux dépens du cas sans article. Ce cas comporte essentiellement des formes figées (13 sur 19), ce qui doit être relié à une tendance diachronique du français à supprimer les formes sans article, beaucoup plus fréquentes dans la langue ancienne (avant le XV^{e} siècle surtout). Ici, l'article est omis dans 'la condition de *réfugié'*, 'les accords/négociations de *paix'*, 'une menace de *guerre'*, cas où les noms sans article désignent une propriété plutôt qu'une entité.

La séparation des deux sens de l'expérience (défini ↔ indéfini) ne doit pas faire oublier que cela illustre le *même phénomène* : la polyvalence d'emploi des articles. On peut certes assumer que les choix de l'auteur sont les plus naturels, ou bien on peut essayer d'évaluer, selon plusieurs personnes, le choix le plus naturel. La différence entre les deux formes est souvent une différence de degré dans l'importance de l'objet, ou dans

la certitude que le lecteur doit le connaître (ou *devrait* le connaître), ou dans la probabilité de l'existence de l'objet lui-même.

VI.C.1.iii Conclusions de l'expérience

Cette expérience, en dépit de ses faibles proportions, tend à remettre en question le caractère absolu de certains principes « logicistes ». Ainsi, les nombreux cas d'interchangeabilité des articles définis et indéfinis en français suggèrent que :

- on ne peut affirmer exclusivement que l'article défini s'applique aux entités déjà introduites dans le discours, et l'article indéfini introduit des entités nouvelles. On observe en revanche que les entités supposées uniques et connues (les noms propres et leurs « attributs ») nécessitent l'article défini, alors que les entités secondaires manifestent une grande variabilité. Lorsque le choix est possible, il semble que l'article défini confère une importance accrue au nom déterminé ('*les* extrémistes serbes' vs. '*des* extrémistes serbes').

- il semble difficile d'affirmer que le caractère référentiel, ou bien la spécificité, soit un trait binaire. Tel semble être le cas pour les entités saillantes du discours, mais pour de nombreuses autres le choix n'est pas exigé par le texte, ni probablement opéré par le lecteur. En fait, les noms dans ce cas semblent caractériser des états ou propriétés plutôt que des entités, selon un continuum d'interprétations ('accords *de* paix', 'négociation [*de*|*d'une*|*de la*] paix', 'stabilisation *de la* paix').

- le schéma quadripartite de Bickerton semble devoir être considéré avec prudence, lorsque l'on prend en compte des textes réels. Certes, ce schéma rend compte d'une certaine latitude dans l'emploi des articles plus que des cadres formels plus rigides. Toutefois, il semble que la distinction des quatre cas soit floue, et doive composer avec d'autres facteurs.

En somme, cette expérience tend à confirmer la différence entre un usage « fortement nominal de noms » (entités saillantes, connues, individualisées) et un usage « moins nominal » (entités peu importantes, peu individualisées, propriétés, états mentaux, sentiments, etc.) – comme relevé par Hopper et Thompson (VI.B.3.). Naturellement, l'analyse de textes plus variés (éventuellement semi-automatique), permettrait de mieux justifier nos observations ; cette expérience préliminaire indique la démarche à suivre.

VI.C.2 Une méthode de classement d'exemples réels

Notre seconde expérience est motivée par la volonté d'éprouver la répartition des différents cas de référence en quatre classes, donnée par Bickerton ou Givón : soit DÉFINI ↔ RÉFÉRENTIEL INDÉFINI ↔ NON RÉFÉRENTIEL ↔ GÉNÉRIQUE (Givón), ou bien +S+P ↔ +S–P ↔ –S–P ↔ –S+P, avec les traits 'Spécifique' et 'Présupposé connu' (Bickerton).

On peut déjà constater que l'appariement entre ces classes, données ici dans le même ordre, ne semble pas aller de soi (en particulier pour la troisième classe).

Nous avons donc tenté de recenser un nombre suffisant d'usages de noms communs, afin de les classer selon les cas précédents. Il nous est alors apparu que ces quatre cas étaient sous-spécifiés, et nous avons construit une hiérarchie de cas plus fine, que nous présentons d'abord (VI.C.2.i) ; nous donnons ensuite le classement des exemples (VI.C.2.ii), puis la conclusion de l'expérience (VI.C.2.iii), sur la pertinence des catégories référentielles proposées.

VI.C.2.i Taxonomie référentielle de l'usage des noms

Sur la base des nombreux cas étudiés, nous avons élaboré un classement possible des usages nominaux par rapport à la référence. Nous donnons ici les caractérisations de ces classes en langue, puisque ce n'est pas vers un modèle formel discret que nous nous orientons. Les frontières des classes voisines sont « floues » (au sens de la « logique floue »), ce qui était déjà le cas chez Givón. Une décomposition en classes floues du continuum référentiel s'accorde avec le modèle structurel de la capacité référentielle esquissé ci-après (VI.D.3), et avec les thèses fonctionnalistes précédentes.

Le critère de classement utilisé est l'extension de la référence ou dénotation de chaque nom, à savoir le nombre et la nature des objets désignés. Nous nous limitons pour l'instant au nombre singulier grammatical, et nous partons du principe qu'un nom commun exprime une certaine catégorie, regroupant un certain nombre d'instances. Nous avons ainsi (exemples Table VI-4) les emplois nominaux suivants :

- [PROP] – le nom n'évoque pas une instance de la catégorie, mais plutôt une propriété caractéristique de la catégorie ;
- [ZÉRO] – le nom désigne une instance inexistante : l'énoncé peut exprimer une propriété qui n'est vraie d'aucune instance de la catégorie, ou bien introduire une instance qui, explicitement, n'existe pas (représentation ayant le trait 'inexistant')[14] ;
- [UNIQ] – le nom désigne l'unique instance d'une classe à un seul élément, connue de la communauté linguistique, fut-elle concrète (phénomène naturel) ou abstraite (phénomène psychologique) ;
- [1-CONN] – le nom désigne une seule instance, déjà introduite dans le texte ;
- [1-REL] – le nom désigne une seule instance, qui se rattache par association ou par une propriété fréquente à une instance déjà introduite dans le texte (celle-ci figurant comme complément du nom ou dans une phrase relative) ;
- [1-NOUV] – le nom introduit une seule instance (dont il nous apprend l'existence) ;

[14] On pourrait donc séparer en deux ce cas. Le problème des objets inexistants, dans des contextes modaux, est qu'ils peuvent acquérir une existence en tant que représentation, si on s'y réfère plusieurs fois.

- [1-INDÉT] – le nom désigne une seule instance, encore indéterminée (c'est-à-dire sur laquelle on ne pourrait littéralement coller une étiquette), mais dont on connaît certaines propriétés ;
- [1-ALÉA] – le nom évoque une et une seule instance prise au hasard dans la catégorie (et dont on ne connaît en général pas d'autre propriété) ;
- [N-ALÉA] – le nom évoque *une ou plusieurs* instances prises au hasard dans la catégorie ;
- [TOUT] – le nom évoque toute instance ou toutes les instances de la catégories (sauf peut-être certains membres marginaux)[15] ;
- [CATÉG] – le nom désigne la catégorie en tant qu'objet, ou instance d'une méta-catégorie.

VI.C.2.ii Données de l'expérience

Étant donné la variété des situations, on se limitera à l'emploi en français, au singulier, des articles définis (*le*, *la*, *l'*), indéfinis (*un*, *une*), et du nom sans article. Cela paraît bien sûr très restrictif, mais nous verrons que la variété des situations demande de commencer à un niveau modeste.

Il y a certes de nombreux autres cas intéressants. Le partitif serait sans doute le suivant à considérer, proche qu'il est de l'article défini (*du* ≈ *de* + *le*)[16]. Le pluriel est déjà assez fortement lié au singulier du moment où des collections d'objets peuvent être désignées par un nom au singulier (usage générique, par exemple). Il y a enfin les différents articles et adjectifs indéfinis, sur lesquels de longues et nombreuses études existent déjà, par exemple celles de Georges Kleiber (cf. V.B.4.ii). On le voit, notre analyse ne représente qu'un pas très préliminaire.

Le tableau ci-dessous donne 75 exemples[17] classés selon les cas d'emploi donnés plus haut. Pour chaque cas, nous donnons dans l'ordre des exemples avec article défini, puis avec article indéfini, puis sans article. Le choix des exemples est personnel (avec l'aide de dictionnaires) en vue de couvrir la plus large gamme d'emplois. Les exemples n'ont été classés qu'après leur recueil en vrac. Nous avons extrait certains exemples de textes réels, et nous en avons conçu d'autres. Il y a bien sûr un risque de sortir les phrases de leur contexte (situation ou texte), c'est pourquoi nous avons employé des noms propres et pronoms déictiques, pour permettre au lecteur d'ancrer chaque exemple dans une situation déterminant son sens.

---

[15] On remarquera l'équivalence logique de *tout* et de *n'importe lequel* : ce qui est vrai de n'importe quel individu est vrai de tous les individus.

[16] Selon Grévisse (n°655), le partitif n'était pas usité en ancien français (*mangier pain*). L'emploi de *del* comme partitif se rencontre après le XIII^{e} s. (aussi comme *dou*), par exemple *dou pain* (Littré, « *du* »).

[17] Les exemples sont numérotés de (53) à (127), à la suite des autres exemples de la thèse.

Toutefois, même dans ces conditions, le classement d'un exemple dans tel ou tel cas peut être incertain : c'est une preuve de plus de l'existence du continuum référentiel, et nous pensons que l'on peut surtout hésiter entre des classes voisines.

[PROP]	(53)	Le chien est depuis longtemps *l'ami* de l'homme.
	(54)	L'occasion fait *le larron*.
	(55)	« Un chien errant entra dans notre vie. … Nous l'appelions Bobby, comme il convient à *un chien* chéri. »
	(56)	Paul est *un docteur* qui connaît son métier.
	(57)	Marco Polo arrivait à *pied* depuis la Chine.
	(58)	Le prévenu affirma sous *serment* qu'il ne savait pas tirer au pistolet.
	(59)	Il fait un temps de *chien*.
	(60)	Paul est *docteur*.
	(61)	*Trêve* de philosophie ! Parlons concrètement.
	(62)	Je l'obtiendrai de *gré* ou de *force*.
[ZÉRO]	(63)	Jean n'a pas vu *le moindre lion* en Afrique.
	(64)	Vraiment, on ne mettrait pas *un chien* dehors !
	(65)	Il n'y avait plus *un chien* dehors.
	(66)	Jean n'a pas vu *un lion* en Afrique.
	(67)	Si Jean avait *un âne*, il le battrait.
	(68)	Où veux-tu trouver *une licorne* ?
	(69)	Jean n'a pas vu de *lion* en Afrique.
[UNIQ]	(70)	*Le soleil* se couche.
	(71)	Ses yeux étaient verts comme *la mer*.
	(72)	« Les animaux malades de *la peste*. »
	(73)	Je vais d'abord imposer *le silence*.
	(74)	« *L'amour* n'a plus de goût, non plus que *la dispute*. »
	(75)	« Ces idées peuvent vous venir aussi à *la table* familiale, quand on enfonce la fourchette dans le rôti. »
		–
	(76)	Contentement passe richesse. Noblesse oblige. Prudence est mère de sûreté.
	(77)	L'orateur imposa d'abord *silence*.
	(78)	Trêve de *philosophie* ! Parlons concrètement.
[1-CONN]	(79)	J'ai acheté un chien et un chat. *Le chien* est très amical.
		–
		–

[1-REL]	(80)	Pierre s'est cassé *la jambe*.
	(81)	Ulysse donna *l'ordre* de fermer les portes.
	(82)	« Après *l'apprentissage* sur le segment, nous avons transféré Ève sur une droite. »
	(83)	Jean est parti chasser *le renard* qui mangeait ses poules.
	(84)	Pierre s'est cassé *une jambe*.
	(85)	« Après *un apprentissage* sur un segment, nous avons transféré Ève sur une droite. »
	(86)	Pierre n'avait pas vu *une erreur* qui s'était glissée.
	(87)	…Le procès est fini. *Accusateur* et *accusé* ont été renvoyés dos à dos par le juge.
	(88)	Ulysse donna *ordre* de fermer les portes.
	(89)	« Après *apprentissage* sur un segment, nous avons transféré Ève sur une droite. »
[1-NOUV]	(90)	Ma radio ne marchait plus. J'ai dû changer *la batterie*.
	(91)	Je suis allé hier voir *le docteur* : tout va bien. Il m'a dit…
	(92)	« Nul ne le vit débarquer dans la nuit unanime, nul ne vit *le canot* de bambou s'enfoncer dans *la fange* sacrée… » [première phrase d'un récit]
	(93)	J'ai acheté *un chien* et *un chat*. Le chien est très amical.
	(94)	Jean a chassé hier *un sanglier*.
	(95)	« *Un chien* errant entra dans notre vie. … Nous l'appelions Bobby, comme il convient à un chien chéri. »
	(96)	Paul lit *un journal*.
	(97)	Ces idées peuvent vous venir aussi à la table familiale, quand on enfonce *une fourchette* dans *un rôti*.
	(98)	Nous nous sommes mis à *table*.
	(99)	Pierre lui a demandé *pardon*.
[1-INDÉT]	(100)	« Ces idées peuvent vous venir aussi à la table familiale, quand on enfonce *la fourchette* dans *le rôti*. »
	(101)	Paul lit *le journal*.
	(102)	Il s'est assis *au volant* et a démarré.
	(103)	Tu devrais voir *un docteur*.
	(104)	Bill cherche *une femme*.
	(105)	Jean est parti chasser *un sanglier* pour le repas de Noël.
	(106)	Que serions-nous sans *une femme* ?
	(107)	Que serions-nous sans *femme* ?
[1-ALÉA]	(108)	Le chien est depuis longtemps l'ami de *l'homme*.

	(109) « *Le médecin* peut réussir là où *le prêtre* échoue. »
	(110) Qui se sert de *l'épée* périra par *l'épée*.
	(111) *L'occasion* fait le larron.
	(112) Anne ne supporte pas de voir *un chien* dans sa maison.
	(113) Vous m'avez reçu comme *un chien* dans *un jeu* de quilles.
	(114) « La théorie de la fonction indicative défendue par Dretske réduit le lien entre *concept* et *symbole* à leur corrélation stable pour l'individu. »
	(115) La mère esquimau enveloppe son enfant dans une peau de *renard*.
[N-ALÉA]	(116) *Le chien* est depuis longtemps l'ami de l'homme.
	(117) Jacques ne sait pas tirer *au pistolet*, et encore moins à *l'arc*.
	(118) Jean est parti chasser *le renard*.
	(119) « Vous abandonnerez *au chien* » le fruit de votre chasse.
	(120) « Je suis épais comme *un sandwich SNCF*. »
	(121) Je sens une odeur de *chien* dans cette pièce.
[TOUT]	(122) *Le lion* est carnivore, contrairement à *la gazelle*.
	(123) *Le chien* est un quadrupède.
	(124) Que serions-nous sans *la femme* ?
	(125) *Une mère* peut-elle ne pas aimer son enfant ?
	(126) Un chien est un mammifère.
	–
[CATÉG]	(127) *Le chien* fait partie de l'embranchement des vertébrés.
	–
	–

Table VI-4. Taxonomie référentielle des usages des noms.

La première constatation est, nous le pensons, la difficulté à classer un certain nombre d'exemples. Toutefois, il faut remarquer que ce choix d'exemples ne respecte probablement pas la distribution statistique « naturelle », les « cas simples » étant en réalité beaucoup plus fréquents. Il serait utile de mesurer les différentes fréquences sur un corpus, mais c'est une tâche très longue. Par ailleurs, il serait significatif de trouver aussi des *exemples négatifs* (i.e. incorrects) pour chaque cas, pour éviter de conclure à une *trop* grande liberté d'emploi des articles ; cela fait partie des travaux à venir.

VI.C.2.iii Discussion

La principale conclusion qui se dégage de cette expérience est l'extrême diversité d'emploi des articles en français. Il ne semble pas aisé de réduire le nombre de catégories utilisées dans la taxonomie, et on tendrait plutôt à l'augmenter pour cerner plus finement

les différents cas. À cette tendance s'oppose toutefois le fait que l'interprétation référentielle de nombreux exemples est mal aisée, chacun pouvant probablement occuper une région du continuum référentiel, et non un point.

Il ne faudrait pas conclure à une interchangeabilité trop aisée des articles en français : de nombreux usages sont probablement historiquement fixés, alors que d'autres obéissent à des critères qui ne sont pas saisis par notre taxonomie. On remarquera à cet égard que celle-ci fait bien intervenir des propriétés de deux natures : (1) « sémantiques », i.e. le nombre de référents, leur détermination ou non, et (2) « pragmatiques », par exemple si ces référents sont connus du locuteur ou non.

Cette expérience assez préliminaire indique une démarche à suivre plutôt qu'une étude aboutie. Elle est cependant assez significative pour permettre des avancées sur plusieurs points. D'abord, ces analyses nous permettront de donner un schéma d'utilisation des articles en français analogue à ceux de Bickerton, sur lequel nous ferons figurer aussi les situations d'emploi précédentes. Ensuite, notre mise en évidence du continuum référentiel sera reliée à notre modèle représentationnel d'utilisation de la référence par un agent, et les situations d'emploi seront reliées aux niveaux d'activations des différentes représentations.

VI.D La référence dans la « pragmatique simulée »

VI.D.1 Application du modèle de Givón/Bickerton au français

Une des motivations de notre taxonomie était d'être plus fines que les quatre cas des modèles proposés (conjointement) par Givón et Bickerton. Se pose toutefois la question de la compatibilité des deux descriptions ; c'est pourquoi nous avons tenté de situer nos catégories dans le plan <S, P> (s*pécifique*, *présupposé*), en nous aidant bien sûr des exemples et de considérations théoriques.

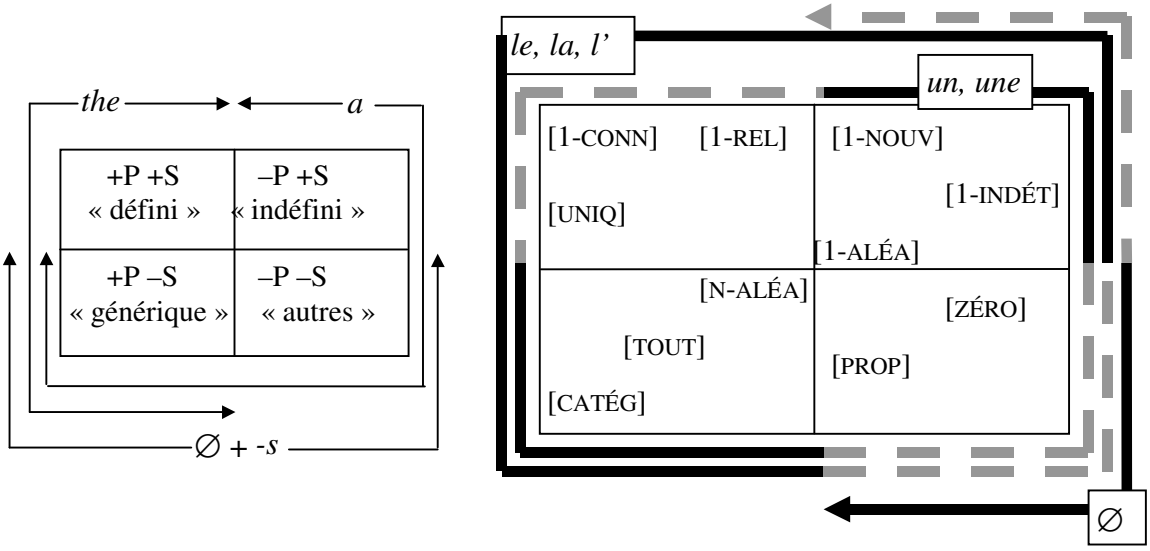

Figure VI-2. Application du modèle de Bickerton au français (anglais à gauche).

Le résultat est donné sur la Figure VI-2 (à droite) ; nous avons exploité les « coordonnées » S et P pour préciser la position de nos cas à l'intérieur de chacun des rectangles. On remarquera ainsi la proximité des cas [1-ALÉA], [N-ALÉA] et [TOUT], qui représentent la désignation de n'importe quel individu de la catégorie du nom (un seul quelconque, un ou plusieurs, tous). Nous avons situé les deux cas [1-NOUV] et [1-INDÉT] dans la même région, car ils désignent un individu de la classe, plus ou moins déterminé (Paul a mangé *un croissant* / Paul voudrait *un croissant*), alors que [1-INDÉT] devrait se trouver dans le région non-référentielle selon Givón. Celle-ci est à nos yeux occupée seulement par [PROP] et [ZÉRO], et encore ce dernier cas peut être vu comme la représentation d'un individu avec la propriété 'inexistant'. Il semble donc que le plan <S, P> ne soit pas la meilleure façon de rendre compte des variétés de référence.

En ce qui concerne l'expression linguistique, nous avons distingué, sur la base des exemples étudiés, un emploi 'courant' (trait plein) et un emploi 'rare' (trait grisé pointillé) – en nous fondant sur notre jugement de la fréquence des exemples considérés. Par rapport à l'anglais, on observe une plus grande extension de l'article indéfini, la possibilité d'un article absent (sur le schéma anglais on considérait le pluriel sans article), et enfin une plus grande ambiguïté (ou liberté) du choix des articles.

VI.D.2 Prise en compte de la situation de communication

VI.D.2.i Représentations manipulées par les locuteurs

Nous avons déjà fait remarquer que le modèle de Givón/Bickerton mélangeait un trait « sémantique » (S) avec un trait « pragmatique » (P). Nous voudrions distinguer un peu mieux, dans ce paragraphe, les deux traits, en introduisant la notion d'activation en mémoire et la reliant au « caractère d'objet » tel qu'il a été défendu par Hopper et Thompson (VI.B.3), pour préparer ainsi la voie à notre modèle représentationnel.

Un facteur déterminant dans l'expression de la référence (en fait l'expression du caractère nominal plus généralement) est la ressemblance de l'entité à laquelle on réfère avec un *objet*, ainsi que la *saillance* de cet objet dans le discours (sa « manipulabilité », cf. VI.B.3.iv). Nous parlerons ici *d'activation d'une représentation*, puisque ce sont en fait les représentations des entités qui sont traitées dans l'esprit des locuteurs.

Notre taxinomie de la référence (VI.C.2.i) avait en vue un critère « sémantique », que nous préférons appeler ici *représentation cognitive*, à savoir le nombre et la nature d'exemplaires évoqués par l'usage d'un nom commun. *L'activation* joue aussi un rôle dans cette taxinomie : par exemple, la différence entre [1-NOUV] et [1-CONN] réside en l'existence préalable ou non d'une représentation pour l'entité évoquée. Toutefois, nous allons considérer l'« activation » (A) et la « situation cognitive » (SC) comme deux propriétés relativement indépendantes ; il ne faut donc pas s'étonner si l'une des deux ne permet pas de déterminer à elle seule l'expression linguistique.

VI.D.2.ii Rôle asymétrique des locuteurs : émetteur et récepteur

Plus important, les SC et A propriétés ne sont pas symétriques vis-à-vis de la situation de communication, car on a :

- pour l'émetteur d'une assertion : la situation est entièrement connue, l'activation (« importance ») des référents aussi, et celui-ci doit choisir l'expression linguistique (référentielle : ER), en tenant compte aussi des connaissances de l'interlocuteur ;
- pour le récepteur de l'assertion : la situation n'est que partiellement connue (« contexte »), l'activation des référents aussi (« importance jusqu'à ce point »), et celui-ci reçoit l'expression linguistique (ER) ; son but est de remettre à jour à l'aide de celle-ci sa représentation de la situation et les activations.

On pourrait donc penser à la représentation simpliste :

- Locuteur : $SC + A \rightarrow ER$
- Interlocuteur : $ER \rightarrow SC + A$

ce qui par trop schématique. L'interlocuteur doit comprendre, à travers une expression référentielle ER, non seulement la nature des entités évoquées, mais aussi leur importance dans le discours.

On comprend maintenant un autre problème du modèle de Givón/Bickerton : l'asymétrie de cette situation ne semble pas du tout prise en compte, ni même commentée. On peut donc se demander lequel des deux usages le modèle caractérise. Plus généralement, faut-il passer nécessairement par deux modèles distincts (émission/réception), ou bien peut-on trouver un cadre de description commun ?

VI.D.2.iii Une mémoire commune de l'interaction

Nous pensons qu'il faut introduire la notion de mémoire de l'interaction (MI) pour chaque locuteur, qui enregistre ce qui s'est dit au cours de l'échange ; par définition,

cette mémoire est identique dans chacun des locuteurs, puisqu'elle stocke ce qui est *mutuellement connu*. L'occurrence d'une assertion correspond donc à :

- Locuteur : $SC_{ém}^{pré} + A_{ém}^{pré} + MI^{pré} \rightarrow Expr + MI^{post}$

- Interlocuteur : $SC_{réc}^{pré} + A_{réc}^{pré} + MI^{pré} + Expr \rightarrow SC_{réc}^{post} + A_{réc}^{post} + MI^{post}$

On peut donc concevoir qu'une certaine synchronisation des représentations des locuteurs s'effectue dans la MI, grâce bien sûr à une « négociation » qui n'a rien de simple, et sans oublier que le locuteur a toujours « une assertion d'avance » sur son interlocuteur. En des termes intuitifs, l'émetteur tient *aussi* compte des connaissances du récepteur, et celui-ci est en *conscient*.

La description homogène de la situation, que nous allons adopter, consiste donc à considérer une seule *mémoire de l'interaction*, par rapport à laquelle on peut à la fois déterminer l'expression référentielle ER à utiliser pour émettre une assertion, et mettre à jour les représentations des référents lors de la réception d'une ER. Chaque agent possède une telle mémoire, et nous admettrons que son utilisation est « symétrique », i.e. il suffit de décrire son utilisation en termes d'appariements entre ER et représentations des référents (nombre, statut, activité), indépendamment du sens d'utilisation (des référents vers l'ER ou de l'ER vers les référents).

Cela permet donc de préciser un élément supplémentaire de notre modèle d'usage de la référence inspiré des fonctionnalistes. Nous allons maintenant nous tourner vers un point essentiel, qui est l'ancrage du modèle dans un agent doué de perceptions, conçu dans le cadre de la « pragmatique simulée ».

VI.D.3 Ancrage de la référence dans les agents communicants

VI.D.3.i Représentation des individus et des catégories

Comment traduire les considérations précédentes en une structure capable de s'intégrer aux mécanismes sensori-moteurs d'un agent situé dans un environnement simulé ? Peut-on déjà, à ce stade, proposer un mécanisme « réaliste » qui rende compte des propriétés de la référence qui sont apparues dans ce chapitre comme centrales ? La proposition que nous allons faire représente une hypothèse plausible dans ce sens.

Dans la partie initiale du chapitre, nous avons mis en évidence deux capacités fondamentales : reconnaître des individus et leur affecter une catégorie. Nous avions déjà rencontré ces capacités, de façon assez explicite, dans le modèle *Darwin II* élaboré par G. Edelman (II.C.2.i.b) ; ce système était précisément constitué de deux voies (« Darwin » et « Wallace ») responsables respectivement de la reconnaissance des individus et de celle des catégories, et étaient reliées de façon réentrante.

Nous proposons donc de réutiliser ce mécanisme, que ce soit de façon connexionniste ou symbolique, dans une structure à deux niveaux (Figure VI-3). Le premier niveau contient des détecteurs de traits présents dans les entrées perceptives

(essentiellement la vision). Le second contient, dans une voie (I), des catégoriseurs sensibles aux particularités des individus (détecteurs « fins » de conjonctions de traits), et dans l'autre (C) des catégoriseurs sensibles à la catégorie de l'individu (détecteurs « moins fins »). Ces seconds niveaux de chaque voie constituent deux « mémoires de travail » pouvant conserver les « représentations » des objets perçus récemment[18]. Ces mémoires seraient elles-mêmes reliées à des mémoires à long terme des individus et des catégories.

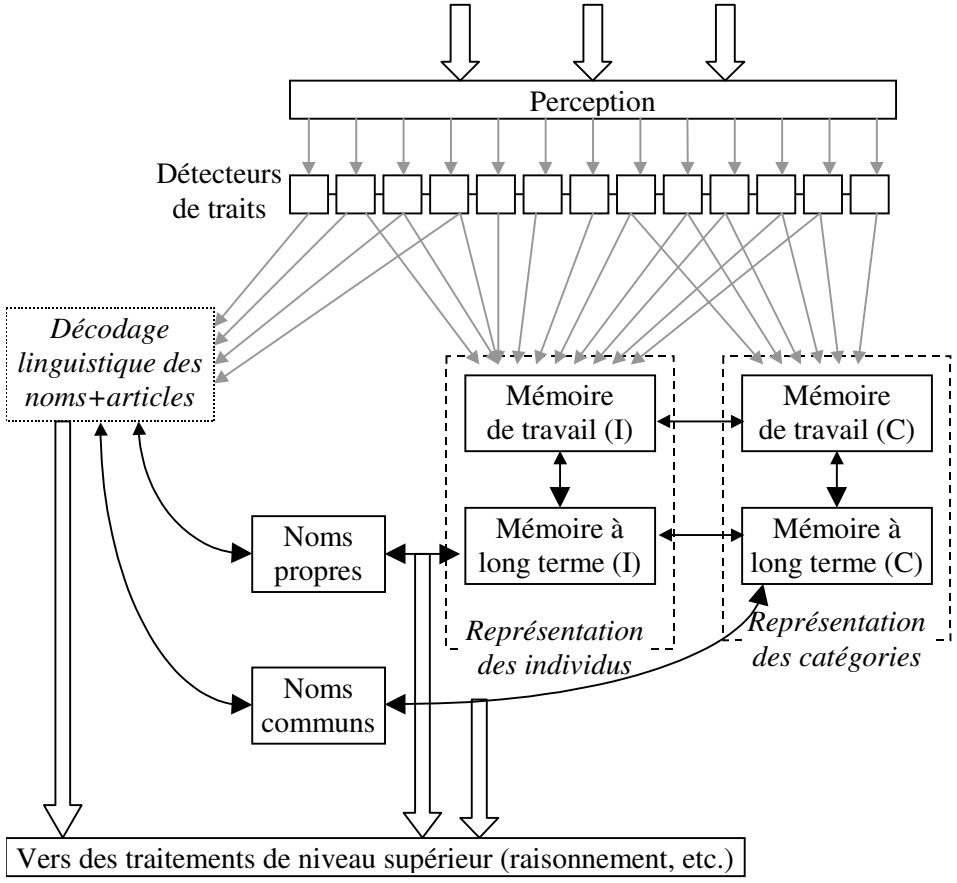

Figure VI-3. Proposition de structuration de la faculté référentielle pour un agent.

Ce sont les mémoires à long terme qui participent à la constitution des *symboles* ; ceux-ci sont des associations stables entre des individus ou catégories mémorisés et des « tokens » ou éléments particuliers (saillants) de chaque individu/catégorie (il s'agit d'une conception développée aussi par Edelman, cf. **II.A.2.iv**). Pour ce qui est de la parole, les « tokens » sont des percepts sonores, catégorisés de façon détaillée pour être distingués les uns des autres. Pour un agent, il est plus simple de réaliser une modalité de communication séparée de la perception visuelle, dans laquelle la catégorisation des « tokens » soit aisée (par exemple un détecteur de trait par token).

[18] Une façon de concevoir ces « mémoires », dans une perspective connexionniste/temporelle, ce serait d'imaginer des décharges synchronisées périodiques, chaque objet occupant un créneau de la période ; cela pourrait expliquer une capacité assez réduite constatée expérimentalement (« sept unités » ?).

Pour que ces « tokens » puissent jouer le rôle de mots, il faut qu'un mécanisme d'apprentissage (par cooccurrence) les relie aux entités représentées dans les mémoires à long terme, individuelle et catégorielle, et constitue ainsi les *symboles*. On comprend assez bien que les « tokens » associés aux individus seront l'analogue des *noms propres*, alors que ceux associés aux catégories formeront les *noms communs* (cf. Figure VI-3). Ce serait alors un pas qui nous rapprocherait beaucoup de l'état intermédiaire du *protolangage* proposé par Bickerton.

Ces associations doivent fonctionner de façon réentrante (réciproque) : la perception d'un exemplaire d'une catégorie doit évoquer le symbole associé à la catégorie, et à l'individu (y compris le « token »). Inversement, l'évocation du « token » entraîne l'activation du symbole, et donc l'activation d'une représentation de la catégorie et/ou individu dans la mémoire de travail à court terme[19]. Cela constitue donc un modèle très simple mais *bidirectionnel* du fondement des mots servant à référer aux objets ; le modèle peut donc être utilisé à la fois pour la production et pour la compréhension d'expressions référentielles, conforme en cela à l'idée d'une synchronisation des mémoires des interlocuteurs d'un dialogue (cf. VI.D.2.iii ci-dessus).

VI.D.3.ii Statut référentiel des représentations

La question qui se pose à ce stade est la capacité du modèle à représenter les différents usages de la référence observés dans ce chapitre. En d'autres termes, comment se projettent sur la structure proposée les onze cas décrits par notre taxonomie référentielle ? Nous avons vu que le modèle de Givón/Bickerton présentait certaines difficultés à ce propos ; notre modèle structurel et représentationnel pourra-t-il offrir une explication plus adéquate de ces propriétés ? Pour essayer de répondre affirmativement, nous faisons les hypothèses suivantes :

- chaque entité perçue possède simultanément une représentation RI_e dans la mémoire de travail individuelle et une autre RC_e dans la mémoire de travail catégorielle, ces représentations étant liées grâce aux connexions réentrantes entre ces deux mémoires (« binding ») ;
- chaque représentation possède une certaine activation dans la mémoire respective, i.e. $act(RI_e)$ et $act(RC_e)$. La somme des ces activations est bornée (cf. droite en pointillé sur la Figure VI-4), ce qui suppose un mécanisme d'inhibition, ou de distribution d'activation, entre les deux mémoires ;
- l'activation *totale* d'une entité correspond à sa saillance ou activation dans le discours et le contexte ;

[19] On voit ici que le modèle de la mémoire demanderait à être précisé : les représentations d'individus ou catégories évoquées par les symboles ne devraient pas avoir le même statut que celles effectivement perçues. On devrait donc distinguer une mémoire de travail « conceptuelle » de la mémoire perceptive à court terme.

- *le statut référentiel d'une entité est représenté par la combinaison des deux activations act(RI_e) et act(RC_e).* Les deux mémoires de travail sont elles-mêmes catégorisées à des niveaux plus élevés, en même temps que les mémoires de « tokens », pour obtenir les représentations conceptuelles des référents (flèches larges sur la Figure VI-3).

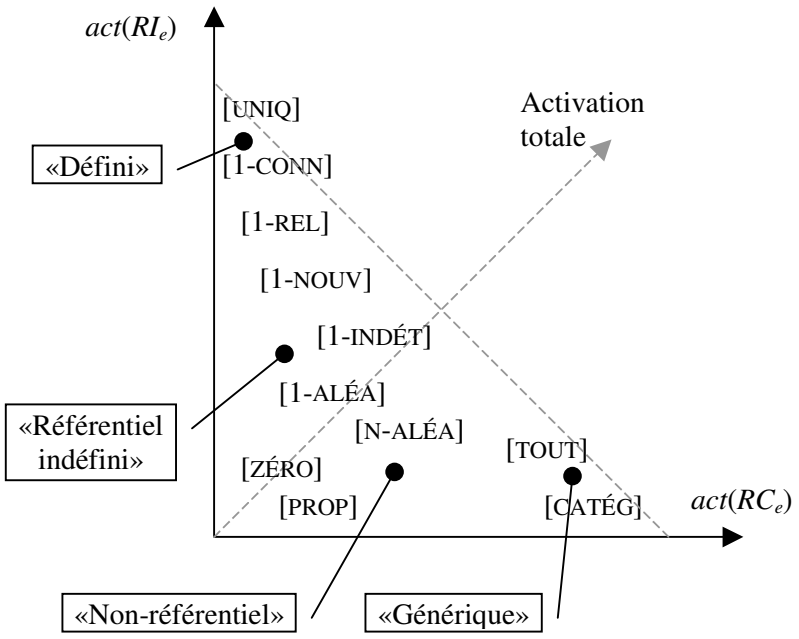

Figure VI-4. Interprétation des catégories linguistiques de la référence en termes d'activation des représentations individuelles (RI) et catégorielles (RC).

La dernière hypothèse est tout à fait centrale, puisqu'elle implique, sur des bases structurelles, que le statut référentiel d'une entité est représenté dans notre modèle par le couple < act(RI_e), act(RC_e) >, comme indiqué sur Figure VI-4. Chaque cas de référence correspond donc à une certaine région de ce plan, à l'intérieur du triangle des valeurs autorisées (somme bornée).

Les représentations d'entités manipulables en tant qu'individus (comme celles mises en avant par Hopper et Thompson, et nos cas [UNIQ], [1-CONN], [1-REL], [1-NOUV]) auront une RI_e beaucoup plus active que la RC_e ; par ailleurs, une différence entre act(RI_{e1}) et act(RI_{e2}) rendra compte de la différence de saillance de deux entités de type individu. Une situation analogue a lieu dans la région symétrique : si RC_e est beaucoup plus active que RI_e, c'est qu'il s'agit d'une entité de type classe, la zone des individus particuliers n'étant pas activée (cas de [CATÉG] et de [TOUT]).

Lorsque à la fois RC_e et RI_e sont très peu activées, il faut à notre sens interpréter l'évocation comme celle d'une propriété de la classe, i.e. le cas [PROP] (et peut-être [ZÉRO]). Il reste enfin la région centrale, ou le caractère individuel est moyennement marqué, et comparable au caractère générique. Nous y avons situé les cas restants – [1-INDÉT] [1-ALÉA] et [N-ALÉA] – sachant que la question de leur activation globale reste

à juger, i.e. leur position le long de la bissectrice : on peut supposer ou non que cette saillance est nécessairement plus faible que celle d'une entité de type [1-CONN] ou [TOUT].

Nous avons aussi indiqué sur la figure les régions correspondant aux quatre cas de Givón/Bickerton, qui ont bien sûr des frontières approximatives. Nous retrouvons dans notre modèle la contiguïté des quatre rectangles, sauf apparemment celle des cas « défini » et « générique ». Ces cas appartenaient déjà chez Givón à deux plan différents de sa « boucle hélicoïdale » (cf. VI.B.2.i) ; ici, leur point commun est de correspondre à une activation maximalement asymétrique des deux mémoires, et constituer peut-être les seuls cas d'activation globale maximale.

VI.D.3.iii Expression linguistique de la référence

L'expression linguistique des cas référentiels peut ou non utiliser, nous l'avons vu, des modifieurs ajoutés aux noms. À ce stade de notre modélisation, nous n'avons pas précisé sur la structure comment ceux-ci seraient employés (probablement à l'intérieur du module linguistique) ; une tâche préalable consiste à relier ces modifieurs à la représentation des cas référentiels par les activations des RI_e et RC_e.

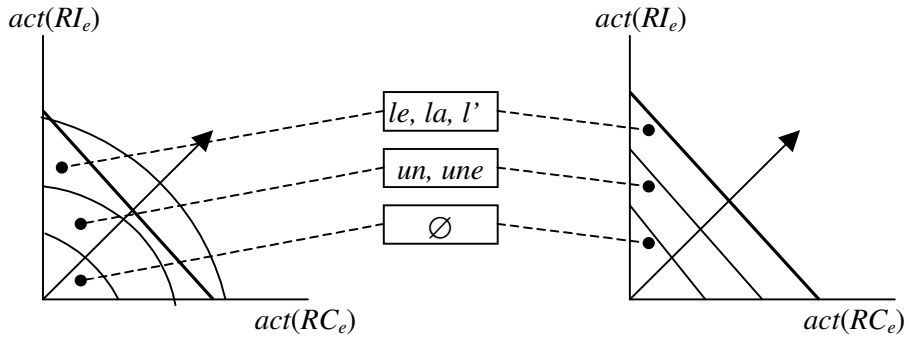

Figure VI-5. Deux hypothèses d'emploi des articles en fonction de la représentation individuelle et catégorielle (le trait gras signale l'activation globale maximale).

À titre d'hypothèse, nous donnons sur la Figure VI-5 deux critères possibles d'emploi des articles. À gauche, c'est la différence entre l'activation des deux mémoires qui détermine globalement l'article : si l'une des deux est beaucoup plus active que l'autre, on utilise le défini, si les deux sont proches, l'indéfini, et si de plus l'activation globale est faible alors on n'utilise pas d'article. Dans la partie droite, c'est simplement l'activation globale (la somme des représentations dans les deux mémoires) qui est le facteur déterminant, ce qui rappelle les thèses de Hopper et Thompson. Dans les deux cas, le résultat dépend bien sûr de la position des cas référentiels sur le schéma, ce qui fait que ces deux hypothèses ne sont pas tout à fait indépendantes.

Les oppositions entre *le* et *un* sont multiples en français, et dépassent probablement la simplicité de l'analyse précédente. Toutefois, l'opposition entre un référent nouveau et un référent connu (*un*/*le*) peut être décrite par notre modèle. En effet,

un référent qui n'était pas représenté en mémoire possède une activation nécessairement plus faible que celle d'un référent déjà représenté. De plus, l'éventuelle détermination du référent n'intervient qu'après une éventuelle reprise (ajout de propriétés), donc un changement d'activation.

VI.D.4 Discussion et conclusion

VI.D.4.i Discussion

Notre modèle est fondé en grande mesure sur les études du fonctionnement concret de la référence, tant celles, embryonnaires, de la neurophysiologie et de la psycholinguistique, que celles, plus détaillées, de la linguistique fonctionnaliste. L'implémentation proposée repose sur des méthodes que nous avons déjà utilisées, inspirées des théories d'Edelman, ainsi que sur une dualité individu/catégorie qui semble à la portée des réalisations actuelles, que ce soit par des moyens connexionnistes, ou bien symboliques, les conjonctions de détecteurs de traits.

Certains points sont à préciser en priorité, comme par exemple la représentation des articles (modifieurs) dans le module d'encodage/décodage linguistique, et l'utilisation des représentations élaborées au niveau référentiel par les niveaux d'ordre supérieur. Par ailleurs, la prise en compte de collections d'objets demanderait une extension spécifique de l'architecture. La validation pratique du modèle ne peut se faire qu'après son intégration dans une modélisation multi-agent plus vaste, esquissée au Chapitre III.

VI.D.4.ii Conclusion

Notre modèle d'ancrage représentationnel de la référence se détache assez nettement des considérations logicistes que nous avons décrites, ainsi que des modèles utilisant des représentations symboliques. Son implémentation semble pouvoir être réalisée également par des méthodes connexionnistes ou symboliques, à condition de respecter les capacités de catégorisation à l'aide de traits.

Toutefois, dans les deux cas, ce modèle présente un *caractère continu* tout à fait nécessaire pour expliquer la diversité d'emplois observés pour la référence. Notre taxinomie des cas référentiels délimite des régions d'un plan correspondant à l'activation de deux représentations en mémoire. Le modèle semble donc saisir une propriété spécifique de la référence, sa continuité irréductible au discret, visible dans ce passage (Langacker 1994:17) :

> *Un aspect important de la sémantique linguistique est notre capacité à interpréter une structure par rapport à un **fond** fourni par une autre. Il y a de nombreuses sortes de fonds, par exemple le discours préalable, les présuppositions et attentes pertinentes […] Il n'est pas difficile d'imaginer des discrétisations possibles dans ce domaine, mais je voudrais mettre en avant un paramètre important qui pourrait bien être continu : la saillance du fond […]*

> *Par exemple, une fois un référent discursif introduit, sa saillance tend à diminuer par la suite jusqu'à ce qu'il soit mentionné à nouveau, si cela se produit. Il y a bien sûr un ensemble discret de façons de le faire (par exemple avec un pronom, ou avec un article défini plus un nom), ce qui reflète des estimations discrètes du statut actuel du référent (Givón 1983, van Hoek 1992). Mais sa saillance* en soi *(et par là l'influence de sa présence implicite dans le fond) varie probablement de façon continue.*

Il est donc rassurant de constater, ici sur un exemple, que notre modèle est en accord avec des théories ou observations récentes. Il faut toutefois remarquer que celui-ci nécessite aussi, à côté du couple d'activations continues, le lien catégoriel des représentations à l'entité effectivement perçue ou conçue. Le modèle combine donc des éléments discrets et des éléments continus.

Afin de poursuivre plus avant notre étude de la référence, en l'absence d'un mécanisme complexe de catégorisation, nous nous sommes attaché à la réalisation d'un modèle symbolique, de facture plus proche des méthodes du TAL applicatif, qui met en œuvre un modèle représentationnel des référents, fondé sur les facteurs que nous venons d'évoquer : contraintes discrètes de « sélection » et « activation » des représentations. En attendant une résolution effective de la référence par des agents, comme décrit dans ce chapitre, nous allons étudier la dimension concrète, applicative, de cette résolution au chapitre suivant, puis le problème de son évaluation, dans la troisième partie de la thèse.