

Langage naturel et intelligence artificielle

Gilles Bernard, Jym Feat

► To cite this version:

Gilles Bernard, Jym Feat. Langage naturel et intelligence artificielle. Histoire Epistémologie Langage, SHESL/EDP Sciences, 1988, Stratégies théoriques, 10 (1), pp.151-170. <www.persee.fr/doc/hel_0750-8069_1988_num_10_1_2255>. <10.3406/hel.1988.2255>. <hal-01294596>

HAL Id: hal-01294596

<https://hal.archives-ouvertes.fr/hal-01294596>

Submitted on 30 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives| 4.0 International License

LANGAGE NATUREL, INTELLIGENCE ARTIFICIELLE

G.Bernard, J.Feat

Mots-clés : *intelligence artificielle, linguistique formelle, informatique linguistique, histoire*

Résumé *Cet article présente un panorama d'ensemble des activités de recherche consacrées au traitement du langage naturel en intelligence artificielle (IA). Après un rapide survol des conditions socio-économiques de ces recherches, l'article aborde : i) l'influence des théories langagières sur la recherche en IA, en particulier neuropsychologie, théorie psychocognitive, grammaires transformationnelles et grammaires de cas ; ii) l'histoire des concepts et outils techniques, élaborés pendant ce qu'on peut caractériser comme une lutte entre deux grands courants, sérialiste (Von Neumann) / paralléliste (connexionnistes, réseau d'objets).*

Date de rédaction : février 1988

Caractéristiques de la publication originale

Statut : *article publié en revue nationale avec comité de lecture*

Lieu de publication : *in Stratégies théoriques, Collection Histoire et Théories Linguistiques 10/1*

Pages : 151-170 (20 pages)

Date de publication : 1988

Caractéristiques de la reproduction

Les notes de fin de document ont été mises en bas de page.

Préface ajoutée par l'auteur en 1999

Cet article est une présentation historique, qui, en tant que telle, n'appelle pas beaucoup de commentaires. Un chercheur a critiqué l'emploi du terme « *langue naturelle* », et je reconnais bien volontiers qu'il s'agit là d'un anglicisme malvenu, justifiable dans « *langage naturel* » (en face de « *langage formel* ») – bien que je lui préférerais aujourd'hui *langage humain* –, mais pas ici.

On y remarquera tout de même l'importance que nous donnons aux technologies neuro-mimétiques, inhabituelle dans ce type de contexte. Il faut savoir que les années 1988-1990 marquent l'apogée dans l'application de notre démarche à la mécanisation des opérations linguistiques. Mais ce succès même nous a amené dans une impasse technologique fondamentale ; alors que nous nous dotions des concepts et des outils nécessaires à dépasser le stade du traitement de phrases isolées pour traiter de textes réels, la quantité de connaissances à utiliser devenait trop grande et leur qualité supposait un trop grand travail de préparation. Cet obstacle, typique de l'intelligence artificielle de l'époque et que d'autres avaient rencontré avant nous, nous posait de surcroît un problème spécifique, lié à notre conception de la relation linguistique - informatique : nous ne pouvions plus faire évoluer notre théorie linguistique par l'épreuve de la mécanisation.

Il est donc certain, avec le recul, qu'à l'époque de cet article, avant même d'avoir achevé la mise au point de nos mécanismes, nous commençons déjà à prendre conscience de cette difficulté, et que nous nous tournions vers une autre branche de l'intelligence artificielle, qui permettait d'envisager l'acquisition automatique des connaissances et une meilleure modélisation des mécanismes de la mémoire.

On trouvera un autre bilan dans « *La linguistique et l'intelligence artificielle* » (document 33), réalisé onze ans plus tard. Bien qu'il n'y ait pas continuité de l'un à l'autre, il peut être intéressant de voir comment mon point de vue sur la relation entre ces disciplines a évolué, grâce aux avancées décisives en cours dans le domaine du traitement de corpus, et aux travaux de la seconde période de ma recherche.

LANGAGE NATUREL ET INTELLIGENCE ARTIFICIELLE

Gilles BERNARD et Jym FEAT

ABSTRACT : This paper presents an overall assessment of the whole range of research activities devoted to Natural-Language Processing within the field of Artificial Intelligence (AI). After a quick survey of the social and economic conditions of research, the paper addresses itself to : i) the influence of language theories on AI research : neuropsychology, psychocognitive sciences, transformational grammar and case grammars ; ii) the history of technical instruments and concepts, worked out during what may be described as a struggle between two basic trends : serialistic (Von Neumann) / parallelistic (connexionists, networks).

RÉSUMÉ : Cet article présente un panorama d'ensemble des activités de recherche consacrées au traitement du langage naturel en intelligence artificielle (IA). Après un rapide survol des conditions socio-économiques de ces recherches, l'article aborde : i) l'influence des théories langagières sur la recherche en IA, en particulier neuropsychologie, théorie psychocognitive, grammaires transformationnelles et grammaires de cas ; ii) l'histoire des concepts et outils techniques, élaborés pendant ce qui peut être vu comme une lutte entre deux grands courants, sérialiste (Von Neumann) / paralléliste (connexionnistes, réseaux).

Sommaire

1. LE CONTEXTE SOCIO-ÉCONOMIQUE

2. THÉORIES ET MODÉLISATIONS

2.1 Modèles psycho-cognitifs

2.2 Modèles linguistiques

2.3 Théories logiques

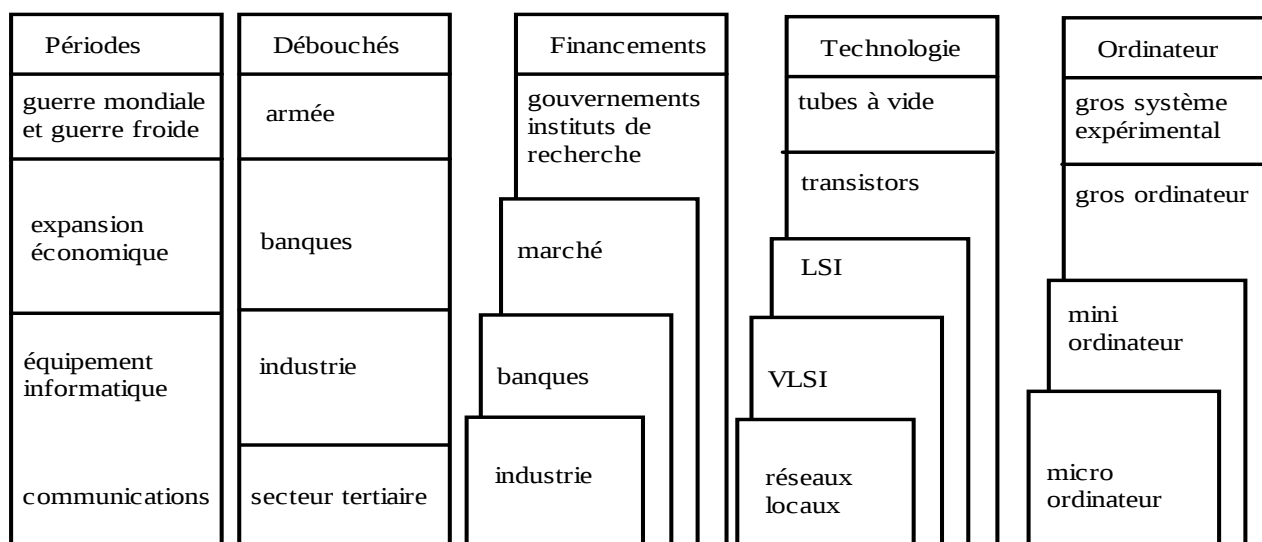
3. CONCEPTS ET MISE EN OEUVRE

4. PERSPECTIVES

RÉFÉRENCES

Du fait que le développement de l'informatique est récent, que la stabilisation de la profession est encore loin d'être atteinte, et que de nombreuses institutions se disputent une part du gâteau, il est difficile de dégager des écoles bien nettes et bien distinctes, les filiations étant incertaines ou multiples. Il nous faudra donc examiner le domaine du traitement du langage naturel (TLN) sous divers angles ; nous prendrons successivement trois points de vue : le contexte social et économique, les théories & modélisations, les concepts & techniques de mise en oeuvre¹. Les TLN seront ainsi appréhendés au milieu des différentes conditions qui les ont fait naître.

1. LE CONTEXTE SOCIO-ÉCONOMIQUE



Pendant quelques années, les recherches sont essentiellement consacrées à la reconnaissance des formes et de la parole (pour robots), guidées par les impératifs militaires de la deuxième guerre mondiale puis de la guerre froide. Lors de l'expansion économique des années 1950, l'informatique connaît sa première expansion quand elle équipe les banques en systèmes de gestion informatisés (ex. IBM, à partir de 1949).

Avec les techniques d'intégration LSI (Large Scale Integration) et VLSI (Very Large Scale Integration), et la baisse des coûts de production qui s'ensuit et aboutit aux micro-ordinateurs, une deuxième expansion se fait : l'informatique devient un élément d'équipement ; l'État et les collectivités s'équipent, et les banques se lancent dans l'aventure en finançant des projets (joint venture). L'ordinateur personnel se répand de plus en plus, le secteur tertiaire, dont l'importance s'accroît, augmente sa rentabilité avec les premiers outils de bureautique (Xerox) ; autre facteur non négligeable, les pays en voie de développement fournissent une main d'œuvre à bon marché.

Autour des réseaux locaux, et des nouvelles techniques afférentes, une troisième révolution est en marche, qui profitera, on le verra plus loin, aux architectures de neuro-ordinateurs.

L'intérêt des chercheurs pour le TLN est fonction de ce contexte : les recherches sur la reconnaissance et la génération de la parole ; la représentation des connaissances, qui, en vertu de contraintes économiques fortes, ne sera liée au TLN que dans les instituts de recherche ; le domaine de l'intelligence artificielle (IA) ; l'interfaçage des bases de données avec les langues naturelles. Actuellement, les réalisations en TLN se répartissent dans 6 grands secteurs d'activités, cités par ordre d'importance des budgets de recherche (d'après le Clearinghouse for Natural Language Processing, Battelle Columbus Laboratories) :

- interface en LN pour bases de données (40 %)
- fabrication d'outils spécifiques pour systèmes intégrés (20 %)
- traduction automatique d'un LN en un autre (20 %)
- parcours de texte pour indexation automatique
- génération de texte pour documents standardisés
- reconnaissance de la parole

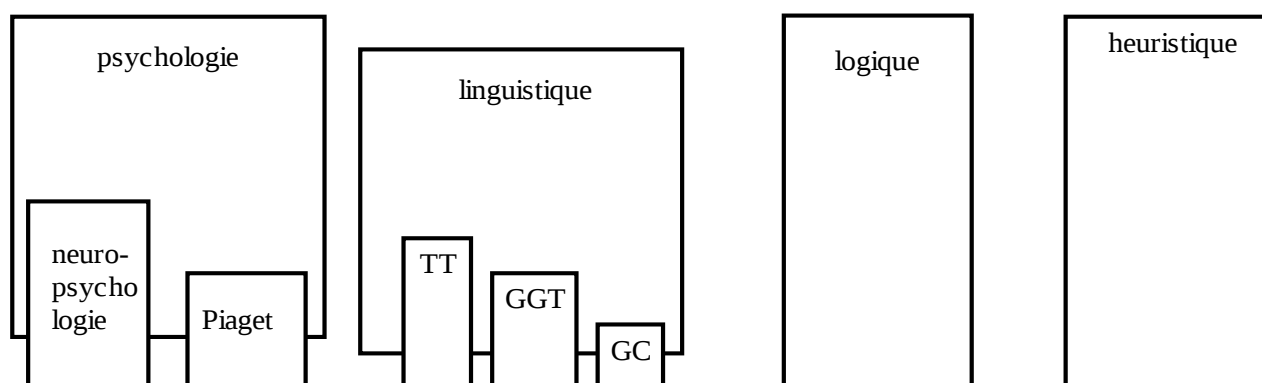
1. Ces points de vue ont probablement à voir avec les trois composants que dégage Sylvain Auroux pour les disciplines scientifiques : théorique, pratique, sociologique..

Les trois premiers secteurs drainent 80 % du budget de la recherche, aussi bien publique qu'industrielle, parce que leurs réalisations ont un débouché immédiat. Ce fait ne devrait pourtant pas nous faire oublier les autres secteurs, actuellement négligés du fait de leur peu de rentabilité. A terme, ils sont cependant plus prometteurs que les secteurs florissants, pour deux raisons :

- la faveur dont jouissent certains développements n'est pas due à leur qualité, mais à leur efficacité relative. Les interfaces de bases de données n'ont pas vraiment besoin de comprendre les nuances des LN, et leur vocabulaire peut être limité au domaine couvert par la base. Ces applications sont donc tout à fait adéquates, mais leurs méthodes ne sont pas généralisables et n'intéressent pas vraiment le linguiste.
- les difficultés que connaissent les autres secteurs sont dues au manque d'une bonne théorie linguistique qui puisse être mise en oeuvre avec des performances correctes compte tenu de la technologie matérielle et logicielle actuelle.

2. THÉORIES & MODÉLISATIONS

Trois courants théoriques influencent le TLN : l'approche psychocognitive, l'approche linguistique, et l'approche logico-mathématique. Nous ferons figurer les nombreux informaticiens qui n'utilisent pas de théorie du TLN sous l'étiquette d'approche heuristique.



2.1 MODÈLES PSYCHOCOGNITIFS

Autour du mythe du « robot pensant », les premiers chercheurs en IA pensent que l'étude biologique du cerveau doit servir de base aux développements des facultés électroniques des machines. Le problème de la communication des machines avec le monde extérieur, autrement dit, de leurs moyens de perception, a très tôt attiré un grand nombre de chercheurs.

Les premiers modèles de la perception sont fondés sur des théories neurophysiologiques. Ceux qu'on appellera les connexionnistes (cf partie 3) élaborent et implémentent, dans les neuro-ordinateurs, les concepts de réseau de neurones (Rosenblatt, 1949), de seuil (Widrow, 1959), de hiérarchie (HearSay, 1973). L'attention accordée aux processus d'apprentissage dirige une bonne partie de la recherche.

En 1956, McCarthy développe un langage (Lisp) qui permet de mettre en oeuvre sa conception des processus de pensée, présentée dans un article (1958) qui sera réédité par Minsky en 1968. A partir de ce langage, Papert, un disciple de Piaget, élaborera plus tard le langage LOGO, à destination pédagogique.

Schank, Minsky, entre autres, utilisent des modèles psychologiques issus des connexionnistes et de Piaget. Schank (1973) ramène le sens d'un énoncé à une combinaison de prédicats (relateurs) primitifs supposés universels. Ces primitives auraient permis de normaliser les représentations internes à la machine, sous forme de graphes de dépendance conceptuelle. Mais il semble que le conceptuel ne se laisse pas découper de façon aussi rigide, et que seuls quelques énoncés simples se soumettent à la transformation. Minsky (1975) élabore les

frames sur le modèle des schémas piagétien, et Schank un peu plus tard, les scénarios sur le même modèle. Les scénarios permettent de reconstruire les présupposés nécessaires à la compréhension d'un texte : des connaissances standard concernant les situations évoquées, l'environnement dans lequel elles se déroulent, et les facteurs qui peuvent influencer le déroulement normal. Mais les scénarios s'avèrent, à l'usage, très lourds à gérer dès que plusieurs situations se superposent. Schank complète son modèle en 1977 en intégrant aux scripts la dépendance conceptuelle ; Riesbeck (1975, 1979) met ce modèle en oeuvre dans son analyseur ELI.

Dans ce courant, des chercheurs comme Schank et Wilks refusent l'idée de devoir engendrer une structure intermédiaire entre la structure linéaire de l'énoncé et sa représentation interne en termes de concepts.

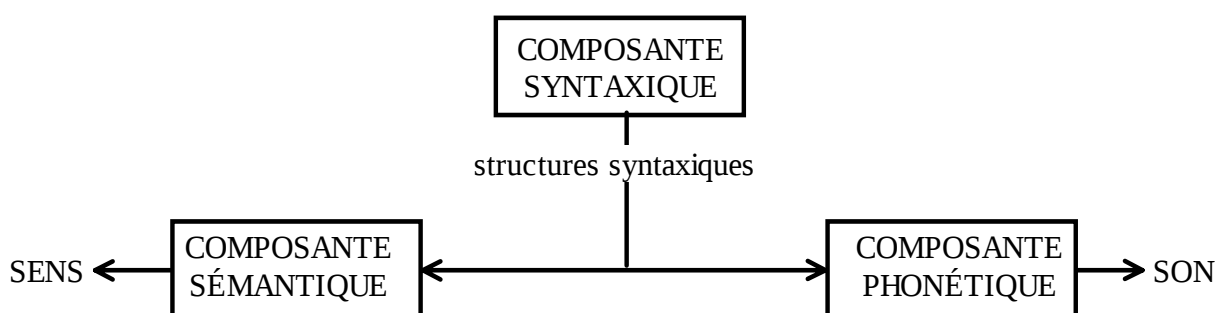
2.2 MODÈLES LINGUISTIQUES

A partir des années soixante, on constate un intérêt pour les théories linguistiques ; trois grands courants se dégagent : les théories transformationnelles (TT), la grammaire générative et transformationnelle (GGT) et les grammaires de cas (GC).

ÉLIMINATION DU SENS. Les théories transformationnelles (dont le théoricien le plus connu est Harris) relèguent la sémantique à une position tout à fait secondaire : les analyseurs syntaxiques ne comportent pas de sémantique, sauf sous la forme de contraintes limitant l'application des règles syntaxiques. Les modèles de Harris construisent les observables qui sont les familles d'énoncés reliées par transformations. Le Linguistic String Project de Sager (1967) est fondé sur la technique d'analyse en chaînes de Harris (1962). L'analyseur de Sager, réécrit par Salkoff (1973) pour le français, à l'aide d'un modèle de Gross, n'utilise de contraintes sémantiques que pour filtrer les différentes analyses syntaxiques possibles. La structure engendrée par l'analyseur ne prétend d'ailleurs pas représenter plus que la structure superficielle de l'énoncé.

RÉAPPARITION DU SENS. La GGT (représentée d'abord par Katz-Postal 1964 et par Chomsky 1965, puis par d'autres ouvrages de ce dernier) permet de construire des modèles où l'analyse syntaxique précède l'analyse sémantique. Elle pose la syntaxe au centre des relations entre la structure de la phrase d'une part, son sens et sa forme phonétique d'autre part. Ces modèles se caractérisent par :

- une composante syntaxique de base contenant virtuellement toutes les phrases possibles de la langue sous leur forme structurée dite « profonde »
- des règles de transformation permettant de passer des structures générées par la base aux formes dites « de surface »



Ce diagramme fameux montre trois processeurs, dont l'un n'a d'ailleurs pas d'entrée, mais engendre de l'information qui sera traitée en parallèle par les deux autres. Aucune intégration n'est apparemment prévue. De cet organigramme, les informaticiens n'ont retenu que le processeur syntaxique, puisque ni l'interface avec le processeur sémantique, ni le processeur lui-même, n'ont jamais été formalisés. Les informaticiens ont donc dû décider seuls de la façon dont le flot d'information circule, et est interprété.

L'ATN (Augmented Transition Network) de Woods (1970), dont l'une des premières applications est le système LUNAR (1972), est élaboré sur un modèle de la GGT pour servir à l'analyse syntaxique, considérée comme un processus séparé, préliminaire à l'analyse sémantique. Il met en oeuvre des techniques puissantes et présente des avantages considérables par rapport aux analyseurs précédents. C'est la raison pour laquelle il a influencé de très nombreux analyseurs de langage naturel. D'autres applications, comme le système déterministe Parsifal de Marcus (1980), ont suivi l'évolution de la GGT jusqu'à aujourd'hui.

Dans la GGT, ne sont pas pris en compte les phénomènes de temps, d'aspect et de modalité, non plus que, pendant longtemps, le calcul des références. De plus elle pose souvent comme universaux du langage des structures de l'anglais, à la rigueur adaptables aux autres langues SAE [Standard Average European]. D'où la difficulté, par exemple, à faire un modèle pour des langues VSO comme l'arabe.

INTÉGRATION DU SENS. Les grammaires de cas (Fillmore, Anderson, Halliday) supposent que les relations entre les lexèmes peuvent être ramenées à un nombre fini de cas primitifs, supposés universels, représentant la structure profonde liée à un énoncé, assimilée à son sens (ceci les distingue de la GGT). Ces systèmes de cas et de traits sémantiques ont une histoire déjà longue, puisque les derniers au moins résultent de la sophistication des techniques de définition qu'on trouve déjà chez Aristote. Winograd (1972) implémente, avec le programme SHRDLU, une grammaire systémique, i.e. une grammaire de cas.

Ces modèles ne sont valables que pour de petits univers particuliers où les relations sont homogènes, et peuvent être ramenées à quelques types. Dès que les univers sont complexes et dynamiques, il faut choisir entre multiplier les types, ou perdre le pouvoir de les discriminer. Les cas sont donc arbitraires, dépendants de l'univers choisi et du point de vue implicite du concepteur. Une grammaire de cas ne peut être une solution générale que si elle justifie par la théorie le nombre de cas du modèle.

DISPARITION DE LA SYNTAXE. Des modèles casuels ont été utilisés par des chercheurs de la « branche heuristique », qui ne se préoccupaient pas de théorie. La tendance est à minimiser, voire même négliger l'importance du rôle joué par la syntaxe dans l'organisation de la phrase. Cette approche se retrouve dans les travaux de Brown, Burton & Bell (1974), ceux de Codd (1974), de Waltz & Goodman (1977), ou de Hendrix (1977), qui tous utilisent exclusivement un ATN sémantique pour extraire directement l'information à partir de la phrase.

L'analyseur ELI de Riesbeck (ci-dessus) occupe une position intermédiaire : il est conçu de telle sorte que le composant sémantique dirige le processus de compréhension, alors que la syntaxe est d'importance secondaire (ELI n'effectue d'analyse syntaxique que si le composant sémantique le demande).

Tout récemment, des connexionnistes, McClelland et Kawamoto (1987), utilisant un modèle (lexi)casuel, ont développé un système capable de déterminer les différents « rôles sémantiques » joués par les éléments d'un énoncé. Les lexèmes sont décrits par ce qu'ils appellent des micro-traités sémantiques.

Par exemple, deux des micro-traités qui décrivent les substantifs seraient « humain » et « dureté », qui pourraient prendre comme valeurs respectives « humain/non-humain » et « dur/mou ». Les mots ne sont pas directement représentés dans le réseau du système, ils sont représentés en terme de l'activité des neurones qui représentent ces micro-traités. Le modèle dispose de groupes de neurones pour chacun des rôles majeurs qu'un substantif peut jouer dans une action exprimée par un verbe. Ces rôles sont « agent », « patient », « instrument » et « modifieur » (mot ou clause adverbiale). Ainsi, l'énoncé « the man ate a sandwich » déclencherait l'activation des micro-traités de « man » et « ate » dans l'ensemble des neurones correspondant à l'agent, et ceci représente le fait que l'agent du verbe « ate » est « man ».

EN RÉSUMÉ : le modèle d'Harris, très peu exploité, a eu le mérite d'être le premier à proposer un traitement formel ; pour les modèles génératifs, l'intégration du composant sémantique, et l'adéquation aux données linguistiques, sont sous la responsabilité de l'informaticien qui veut utiliser ce modèle ; les grammaires de cas sont un excellent outil pour décrire des univers homogènes, tels que des bases de données couvrant un domaine très restreint, où la liste de cas peut être finie.

2.3 THÉORIES LOGIQUES

Vu le grand succès des théories et modèles logico-mathématiques, nous ferons ici un simple survol des travaux pour en dégager les caractéristiques dominantes. Les modèles utilisés se basent en général sur le calcul propositionnel, le calcul des prédicats du premier ordre, la logique des mondes possibles, la logique floue, la logique modale, la logique procédurale...

Avec le principe de résolution de Robinson (1965) utilisé par Colmerauer (1978) pour créer le langage Prolog, semble s'ouvrir une voie prometteuse, grâce à la possibilité d'utiliser directement le calcul des prédicats du premier ordre, formalisme « universel ». Citons encore les travaux de Coles (1968), et de Sabatier (1984).

Des modèles logiques ont également été empruntés pour mettre en oeuvre des théories non logiques (les

connexionnistes utilisent la logique floue, Winograd intègre théorie casuelle et calcul des prédicats) ainsi que par des chercheurs n'utilisant pas de théorie linguistique ni psychologique (Zadeh 1978).

Tous ces travaux appliquent au TLN les modèles élaborés par Russell (1940), Carnap (1952), Šaumjan (1977), Colmerauer (1978), Montague (1970)... Mais ces modèles posent un problème théorique : on pense pouvoir retrouver à la racine du langage naturel, la logique formelle, alors que l'histoire nous enseigne que la logique formelle est (au moins pour une part) une formalisation de la logique naturelle qu'exprime le langage naturel.

Il s'ensuit un certain nombre d'inconvénients : s'il est vrai qu'on obtient de bons résultats quand le champ linguistique est relativement restreint (interrogation de bases de données spécifiques, contrôle de robots), il semble très difficile, comme l'a montré clairement le programme de Winograd, d'augmenter la puissance de ces systèmes sans provoquer en même temps une explosion combinatoire des temps de calcul et du volume des règles. En définitive, l'inconvénient majeur est l'arbitraire du choix de l'un ou l'autre modèle.

3. CONCEPTS ET MISE EN OEUVRE

En même temps que sont implémentés des modèles et/ou des théories du TLN, une discipline, l'intelligence artificielle, se constitue : d'une part, elle intègre un certain nombre de concepts, qu'elle considérera bientôt comme lui appartenant en propre, quelle qu'en soit l'origine ; d'autre part, elle connaît des incidents de parcours qui lui conféreront une histoire.

RÉSEAUX DE NEURONES. La reconnaissance des formes, domaine budgétairement favorisé à cause d'intérêts stratégiques militaires, a d'abord privilégié le traitement et la reconnaissance d'images.

En 1943, le neurophysiologue Warren S. McCulloch et le logicien Walter Pitts ont prouvé que n'importe quel modèle de réseau de neurones, dans lequel l'état d'un neurone quelconque pouvait être défini par une quantité finie d'informations, pouvait être simulé sur un ordinateur standard. L'hypothèse « quantité finie d'informations » est certainement une hypothèse forte : le nombre de bits nécessaires pour décrire l'état d'un neurone donné peut être si grand qu'il rendrait la simulation trop lente pour être praticable.

En 1949, Franck Rosenblatt (université de Cornell) met au point une machine constituée d'un réseau de 400 photorécepteurs, capable de reconnaître des formes simples. Cette réalisation, baptisée PERCEPTRON¹, suscite l'espoir et l'enthousiasme de nombreux chercheurs. Elle ouvre la voie à l'approche connexionniste, i.e. la technologie des réseaux de neurones artificiels, ordinateurs analogiques architecturés en réseau à multiples processeurs interconnectés.

Dix ans plus tard, Bernard Widrow (Université de Stanford) s'appuiera sur cette technologie pour construire une machine plus sophistiquée, utilisée en 1963 pour la reconnaissance de la parole.

La perception est certainement le problème jugé le plus difficile en IA. C'est aussi le secteur où les progrès sont les plus lents. Décrypter le monde environnant, l'analyser en éléments signifiants, est une activité humaine d'une fabuleuse complexité. On n'a pas encore construit d'ordinateur qui puisse apprendre une langue comme le fait un enfant, et c'est pourquoi les machines à percevoir ont toujours été l'objet d'un intérêt tout particulier.

L'ÉPISODE DE LA TRADUCTION AUTOMATIQUE. Dans les années 1960 aux États-Unis, quelques années plus tard en Europe, de grands projets de traduction automatique sont financés par les gouvernements, pour des raisons à la fois stratégiques et économiques. Les investissements sont comparables, toutes proportions gardées, à ceux consacrés à la course de l'espace, avec des résultats beaucoup moins probants. Le gouvernement des États-Unis demandera donc un rapport sur la question à Bar-Hillel, auteur d'articles critiques sur le sujet, rapport négatif qui mettra provisoirement un point final à la traduction automatique (Bar-Hillel 1964, ALPAC 1965).

1. Un perceptron est un réseau de neurones doté d'unités d'entrée qui peuvent chacune prendre la valeur 0 ou 1. On les appelle la rétine, par analogie à la vision des humains. Le perceptron est aussi doté d'unités appelées les prédicats. Chaque prédicat peut se connecter à n'importe quel sous-ensemble d'unités de la rétine, et sait calculer n'importe quelle fonction linéaire des valeurs de ces unités. Finalement, les prédicats sont connectés à une ou plusieurs unités de décision, qui retournent une seule réponse - oui, ou non - selon les valeurs des unités de la rétine. C'est ainsi qu'un perceptron est capable d'effectuer des tâches élémentaires de classification, c'est à dire qu'il peut classer les images de la rétine en fonction de leurs propriétés. Comme la perception est, fondamentalement, un problème de typologie (i.e. identifier des objets comme étant des chaises, des tables, etc...), on a bon espoir qu'un modèle plus élaboré du perceptron soit capable de perceptions complexes.

ÉCHEC AUX PERCEPTRONS. Parallèlement aux recherches sur les neuro-ordinateurs se développe une autre branche de l'IA, qui élabore les techniques de la programmation symbolique telle qu'elle est pratiquée aujourd'hui sur les ordinateurs standard, les deux branches étant en compétition en termes de budget de recherche. En 1969, Marvin Minsky (MIT) et Seymour Papert (MIT), qui appartiennent à la branche standard, attirent l'attention en publiant une analyse formelle des perceptrons à une couche (d'entrées). Ils prouvent mathématiquement que ce modèle de perceptron est incapable de calculer certaines fonctions d'entrée. Par exemple, la fonction de parité, qui détermine si le nombre de 1 en entrée est pair ou impair. Leur argumentation : si le perceptron n'est même pas capable de calculer une fonction aussi simple, il aurait bien du mal à effectuer les tâches requises pour la perception et l'intelligence.

A l'époque, les travaux de Minsky & Papert semblaient montrer que la technologie des perceptrons ne pouvait être viable comme ligne de recherche en IA. On ne s'aperçut pas, alors, que le modèle qu'ils avaient critiqué était le système le plus simple : le perceptron à une couche. Or, si on ajoute une autre couche d'unités entre les entrées et les prédicats, la puissance de calcul de la machine monte en flèche, et la critique de Minsky & Papert n'est plus pertinente. Et si on ajoute encore d'autres couches il devient alors très difficile de caractériser avec la logique formelle le comportement de ces réseaux. Cette différence ne fut pas bien comprise à l'époque, et les travaux de Minsky & Papert, conjugués aux problèmes de coûts des multiprocesseurs (cf plus haut), eurent pour effet d'étouffer presque complètement la recherche en ce domaine. Ils ne réussirent pas à décourager tout le monde, mais drainèrent cependant près de 80 % des chercheurs et des crédits de la branche connexionniste, au profit de la branche standard.

Depuis cette époque, des idées connexionnistes influencent les concepts utilisés en IA standard : on pourrait parler de faits de substrat ; en effet, les chercheurs qui ont abandonné la branche connexionniste n'en ont pas abandonné les idées. Inversement, les développements technologiques favorisés par l'IA standard mettent à la disposition des connexionnistes des matériels assez puissants pour redonner aujourd'hui à cette approche un souffle nouveau.

RÈGLES. L'idée se précise en 1972 grâce à Winograd et son Shrdlu, programme où chaque connaissance est un petit module procédural. La controverse procédural contre déclaratif (i.e. doit-on figer les connaissances dans le programme ou non)¹ autour de la technique de programmation de Winograd aboutit quelques années plus tard à Mycin, système expert en diagnostic médical, utilisant des connaissances déclaratives sous la forme de règles de production.

FRAMES. Minsky (1975), inspiré par les schémas de Piaget, élabore un modèle de représentation très structurée des connaissances, le frame, et, du même coup, pose comme cruciale l'adéquation entre notre vision du monde et sa représentation en machine. Un frame est un objet informatique qui incorpore dans une même structure, les connaissances caractéristiques du niveau de description où se place ce schéma dans la description globale. Les connaissances sont aussi bien déclaratives que procédurales : c'est l'apparition du concept de démon qui est une réaction automatique lors d'un accès au frame.

SCÉNARIOS. Schank (1977), suivant lui aussi Piaget, on l'a vu, élabore un mécanisme, le scénario (script), qui permet d'intégrer des successions standards d'événements dans la construction de plans, dans le but de restituer les parties sous-entendues d'une situation décrite en langage naturel. Les très grandes quantités d'informations nécessitées ont fait que ces scénarios n'ont pu servir qu'à des maquettes traitant de petits univers.

RÉSEAUX SÉMANTIQUES. Raphael (1964) et Quillian (1968) avaient introduit la notion de réseau sémantique ; l'ATN (réseau de transition augmentée) de Woods (1970), au départ implémentation d'un modèle syntaxique, est devenu un outil de représentation sémantique (cf plus haut) ; les réseaux sémantiques sont repris par Brachman (1977), Simmons et Slocum (1972), Carbonell (1978), Fahlman (1979), et Chouraqui (1981).

Un réseau sémantique est, à l'origine, un graphe composé de noeuds reliés par des arcs étiquetés. A chaque noeud correspond un concept, i.e. une classe d'entités [animal, éléphant, fleur] de l'univers qu'on veut modéliser. Les arcs représentent des relations particulières entre les noeuds : (éléphant est animal) (éléphant possède

1. Dans l'approche procédurale, les connaissances sont incorporées à la procédure qui gère l'ensemble des cas possibles. Dans l'approche déclarative, les connaissances sont «déclarées» indépendamment du programme, qui les consulte et gère leurs interactions.

trompe). Les relations peuvent, elles aussi, être représentées par des noeuds du graphe et intervenir à un autre niveau : (possède est relation).

Fahlman perfectionne le modèle en le transposant sur le plan matériel : chaque noeud deviendrait un processeur, et chaque arc une ligne de communication entre processeurs, l'information circulant simultanément sur un nombre quelconque de lignes. Quillian s'intéresse aux chemins possibles entre plusieurs concepts (ongle partie-de patte), (patte partie-de éléphant). Brachman introduit les valeurs par défaut, les contraintes de type sur les valeurs et les relations, la manière dont interagissent les connaissances. Chouraqui propose un système de raisonnement analogique travaillant sur un réseau sémantique. Tous ces travaux influencent à terme la conception de nouveaux langages de programmation (orientés objet).

LANGAGES ORIENTÉS OBJET. Le langage Simula est probablement l'ancêtre des langages orientés objet, et leur a légué les concepts de classe et de représentant. On retrouve ici, dans le courant standard (Goldberg 1972 chez Xerox avec SmallTalk-72, Winograd 1972 pour sa thèse avec MicroPlanner), une tentative pour récupérer des idées connexionnistes sur les structures de données et de contrôle. Plasma, SmallTalk et Loops, anciennement dénommés « *langages d'acteurs* », permettent une programmation originale en introduisant une nouvelle structure de contrôle : la transmission de message. Chaque objet, i.e. chaque élément du langage, est décrit sous deux aspects : un aspect statique, ensemble d'attributs pouvant prendre une ou plusieurs valeurs, et un aspect dynamique, ensemble de méthodes, comportements qui sont déclenchés à la réception d'un message (démons).

Ces langages succèdent aux langages relationnels, dont le plus connu est Prolog, qui devaient équiper la fameuse cinquième génération d'ordinateurs, prévue en particulier par les Japonais. Ces langages présentent des conditions nécessaires pour gérer des processeurs parallèles, ce qui paraît naturel, et des conditions de simplicité et de convivialité qui les ont mis à l'honneur auprès des utilisateurs non informaticiens.

Ce type de langage, mis à l'honneur dans plusieurs familles de produits industriels (Xerox, Apple MacIntosh), est de plus en plus utilisé pour le TLN, tant à cause de son confort en tant qu'environnement de développement, que de sa puissance pour la représentation d'objets évoluant dans des micro-univers dynamiques, donc de la mise en oeuvre de modélisations de l'activité cognitive humaine. Il permet, en particulier, de modulariser à l'extrême les connaissances du système. Ce point est important pour deux raisons : (1) chaque module (chaque objet) peut choisir la méthode la plus adaptée pour résoudre un problème (les objets les plus locaux ont les méthodes les plus locales), (2) en contrôlant la « profondeur » des points de vue, on peut enfin contrôler la profondeur des processus de perception, et simuler l'activité cognitive.

LA PROFONDEUR VARIABLE. Il y a là une telle évidence qu'on s'étonne de ne pas y avoir pensé avant les travaux de Kayser & Coulon (1982) sur la compréhension automatique de textes : un mot qui peut être compris avec plusieurs nuances différentes peut aussi être compris de façon sommaire, englobant les propriétés communes à toutes les acceptions. C'est l'auditeur qui décide, quand les circonstances l'exigent, de descendre dans les nuances, de spécifier plus ou moins. Il en suit qu'aucun énoncé n'a d'interprétation précisément définie, sauf peut-être pour son énonciateur. Un système automatique doit travailler sur l'interprétation, et celle-ci peut être très raffinée, ou au contraire rester « superficielle ».

PLURALITÉ. Il y a des analyseurs ayant de bonnes performances (encore que d'envergure limitée) qui travaillent avec des techniques de mots clé. D'autres emploient des méthodes déductives, divers modes d'inférence (montant, descendant, analogique, associatif, à plausibilité), pilotés par des stratégies et heuristiques diverses... Finie, l'utopie du mécanisme unique. La nouvelle génération d'analyseurs profite de l'expérience acquise en résolution de problèmes, où les programmes disposent des connaissances nécessaires pour choisir la méthode la plus efficace pour résoudre un cas particulier. Le programme de Brigitte Grau (1983) est, à cet égard, typique : pour suivre le fil conducteur d'une histoire ou d'une conversation, il utilise deux stratégies différentes. Montante (dirigée par les faits), quand il découvre, au cours de l'analyse d'un énoncé, l'indice d'un changement de thème, et cherche le lien avec le contexte. Descendante (dirigée par les hypothèses) lorsqu'au vu du contexte il fait des hypothèses sur le thème.

LE RETOUR DES CONNEXIONNISTES. La baisse des coûts de production, la présence d'un « substrat » connexionniste, ont favorisé, après une éclipse de vingt ans, un retour de l'approche connexionniste.

L'analyseur de LN sur neuro-ordinateur réalisé par McClelland et Kawamoto (cf partie 2.2), contient trois

niveaux de neurones. Le premier est le niveau lexical, qui sert de niveau d'entrée (l'équivalent de la rétine dans le modèle Perceptron). A ce niveau, chaque neurone correspond à un lexème particulier. Au niveau 2, le niveau sémantique, les entrées du niveau 1 activent les neurones qui représentent le sens des lexèmes. Le niveau 3 est le niveau logique, où les informations sémantiques se combinent pour former, pour chaque énoncé, les prédicats et les objets de ces prédicats.

Le regain d'intérêt pour l'approche connexionniste est aussi dû au fait qu'elle propose des solutions à un certain nombre de problèmes que se pose l'IA traditionnelle :

— Une solution aux problèmes de performance, grâce à des machines d'une architecture différente. Là où un seul processeur traiterait séquentiellement quelques bits de données de la zone mémoire, les unités de traitement interconnectées traitent toutes les données en parallèle. On sait que les "circuits" d'un cerveau humain sont beaucoup plus lents que les ordinateurs modernes. C'est le travail parallèle de plusieurs neurones qui lui confère sa performance. Par ailleurs, les machines traditionnelles sont conçues pour travailler sur des nombres et fournir des résultats exacts (avec des délais parfois prohibitifs), alors que les réseaux de neurones sont capables de trouver rapidement des solutions approchées pour des problèmes extrêmement complexes.

— Une solution aux problèmes de maintenance de l'intégrité de l'information (Truth Maintenance System en IA standard), puisque les neuro-ordinateurs, au lieu de travailler en logique booléenne (oui/non), prennent des décisions pondérées à partir de données floues, incomplètes, ou même contradictoires.

— Une solution aux problèmes de la capacité de la mémoire et de la vitesse d'accès à l'information : les machines conventionnelles enregistrent l'information dans une zone dédiée de leur mémoire, alors que les neuro-ordinateurs enregistrent l'information de manière répartie, par la modification des pondérations sur les lignes de communication entre les processeurs. Ce type de machine a également l'avantage d'être beaucoup plus tolérant quant aux pannes matérielles : les performances se dégradent progressivement avec chaque défaillance de composant (comme dans un cerveau biologique), alors que dans une machine conventionnelle, la défaillance d'un seul composant peut avoir des conséquences catastrophiques.

— Une solution aux problèmes de programmation : une machine conventionnelle doit être programmée par instructions. Les programmes, de plus en plus complexes, sont également de plus en plus difficiles à maintenir. Dans les neuro-ordinateurs, seuls les neurones sont programmés, leurs interactions ne le sont pas, et ils forment de manière spontanée leurs propres méthodes de traitement de l'information par auto-organisation lors de l'adaptation des connexions. Ils sont donc capables d'apprentissage, une faculté jusqu'ici difficile à mettre en oeuvre en programmation standard.

Ce dernier point est crucial. Il impliquerait que de telles machines n'ont pas plus besoin de modèles linguistiques tous faits, qu'un enfant n'en a besoin pour apprendre sa langue maternelle... Les neuro-ordinateurs, même en simulation, sont déjà pour certaines tâches largement plus performants que les machines conventionnelles. Les systèmes connexionnistes sont particulièrement adaptés aux problèmes de grandes dimensions où il s'agit de satisfaire simultanément des contraintes contradictoires, c'est à dire des problèmes où les règles de décision ne sont pas clairement définies. C'est justement le cas du TLN. Comme l'a montré expérimentalement Terrence Sejnowski (Univ. Johns Hopkins, Maryland), il faut 12 heures à un réseau de 309 neurones en trois couches (au total 18629 connexions variables) pour apprendre, à ce type de machine, à lire et à prononcer correctement 95% des mots d'un texte modèle (en anglais, pour un dictionnaire de mille mots) et 90% d'un nouveau texte. Et l'apprentissage d'un dictionnaire de vingt mille mots n'a nécessité qu'une semaine de calcul. Un programme de puissance similaire (DECtalk) avait demandé plusieurs années de travail à la firme Digital, et une somme considérable de recherches en linguistique et en traitement de la parole.

4. PERSPECTIVES

Il n'est pas facile de résumer quarante ans de recherches, qui sont également les quarante ans de constitution d'une discipline, encore moins quand il s'agit d'histoire récente. Les auteurs s'excusent si ils ont oublié un nom ici ou un fait là ; ils n'ont retenu, des événements et des hommes, que ce qui était nécessaire et suffisant pour montrer la constitution de cette discipline, l'intelligence artificielle, plus particulièrement sous un de ses angles les plus importants. Après tout, les ordinateurs se programment avec des langages ; et l'intérêt des informaticiens pour les langues naturelles n'est qu'un juste retour des choses. Ainsi, à travers cette histoire que nous venons de voir, il se dessine quelques mouvements généraux, quelques constantes, permettant de replacer en perspective le passé, et peut-être, par là, le futur.

DEUX COURANTS. Dès le début de l'IA, deux courants se sont opposés, à travers toutes les vicissitudes (restrictions de crédits, problèmes institutionnels, manque de contacts). Ces deux courants se fondent sur deux approches :

- pour la première, les processus sont ordonnés en *série*, en un ordre extrinsèque, imposé de l'extérieur ; il y a un référent, un contenant, unique, un seul programme maître, un seul processeur (qui en dirige éventuellement d'autres). C'est l'approche qui nous apparaît aujourd'hui comme traditionnelle, l'approche standard.
- pour la deuxième, les processus sont en *parallèle*, et leur ordre dépend de l'état de chaque processus : il s'agit d'un ordre intrinsèque, déterminé par les événements. Cette approche, l'approche connexionniste, est en ce moment de plus en plus en vogue.

Du conflit entre ces deux courants, on peut rapprocher la controverse procédural contre déclaratif. L'approche procédurale est une variante de l'approche sérielle, où les connaissances sont parcourues en série, dans un ordre défini par le programme. L'approche déclarative est une variante de l'approche paralléliste, où les connaissances sont consultées par le programme en fonction de leur définition propre et de leur pertinence à un moment donné. À la limite, le programme lui-même est simplement une connaissance comme les autres, qui gère les autres à un moment donné et peut passer la main à une autre connaissance.

Diviser ou intégrer, voilà l'opposition résumée en deux mots¹. Mais les deux approches sont également appréciables, pour qui peut en conjuguer les efforts. Et c'est bien la direction que semble prendre aujourd'hui la recherche sur le TLN.

LA THÉORIE ET LA PRATIQUE. Autre point qu'il vaut la peine de soulever, c'est celui du rapport — souvent épineux — entre les chercheurs en linguistique et en informatique. On dit que la recherche en IA est pratique, non théorique, alors que la recherche en linguistique produit des théories qu'il est difficile de mettre en pratique ; c'est en même temps le problème du rapport de la théorie à la pratique, normalement médiatisé par les modèles, qui est posé là.

On ne dispose pas encore aujourd'hui d'une bonne entente entre théories et pratiques. Si des liens sont tissés, ils sont aussitôt rompus. L'histoire des ATN est à cet égard exemplaire : au début, ils sont le versant informatique d'un modèle, ils sont réutilisés ensuite pour eux-mêmes, sans référence à un modèle ni à une théorie. De part et d'autre il reste encore du chemin à faire pour que soit fécondé le domaine du TLN.

La conjonction est souhaitable, à la fois pour les informaticiens, qui pourront ainsi généraliser plus facilement ce qui pour eux constitue des « astuces de programmation », et pour les linguistes, qui disposeront ainsi d'un nouveau champ expérimental ; sans omettre les psychologues, et toutes les disciplines qui peuvent apporter quelque chose à la question. Les scripts, aussi bien que la dépendance conceptuelle, auront eu le mérite de montrer que, pour élaborer un modèle de la compréhension du langage, il faut faire intervenir une réflexion interdisciplinaire où la psychologie et la linguistique, entre autres, ont leur mot à dire.

1. La controverse fort intéressante entre la phonologie générative standard et la phonologie générative naturelle est également une controverse de ce type ; l'un des points de la controverse est précisément de savoir si l'ordre des règles est extrinsèque (position standard de Chomsky-Halle), ou s'il est intrinsèque (position naturaliste de Hooper-Vennemann ; naturel a ici le sens d'intrinsèque).

RÉFÉRENCES

La bibliographie qui suit est restreinte, faute de place ; on consultera pour plus de détails Kayser et Coulon (1986).

Abbreviations utilisées :

ACM the Association for Computational Linguistics

BBN Bolt, Beranek & Newman, editors

IJCAI International Joint Conference on Artificial Intelligence

- ANDERSON JM [1973] « Maximi Planudis in Memoriam » in *Generative Grammar in Europe* [Kiefer & Ruwet eds], Reidel Publishing Co, Dordrecht
- AUROUX S [1982] « Linguistique et anthropologie en France » in *Collection Travaux d'Histoire des Théories Linguistiques* n°1, série 8, Paris
- BAR-HILLEL Y [1964] « A demonstration of the non-feasibility of fully automatic high quality translation » in *Language and information*, Addison-Wesley
- BOBROW DG, COLLINS AM [eds] [1975] *Representation and understanding studies in cognitive science*, Academic Press, New-York
- BRACHMAN RJ [1977] « What's in a concept : structural foundation for semantic networks » in *International Journal of Man Machine Studies*, n°9
- BRADY M, BERWICK R [1983] *Computational models of discourse* MIT Press, Cambridge Mass.
- BROWN RR [1976] *Semantic grammar : an engineering technique for constructing natural language understanding systems*, in BBN, Rapport 3453, ICAI Rapport 3, Cambridge Mass.
- BROWN JS, BURTON RR, BELL AG [1974] *Sophie : a sophisticated instructional environment for teaching electronic trouble-shooting*, in BBN, Rapport 2790, A1, Rapport 12, Cambridge Mass.
- CARBONELL J [1978] « Politics : automated ideological reasoning » in *Cognitive Science* n°2
- CARBONELL J, CULLINGFORD RE, GERSHMAN AV [1978] *Knowledge based machine translation*, Yale Univ., Dpt of Computer Science, Report 146
- CARNAP R [1952] *Meaning and necessity*, Chicago Univ. Press
- CHANGEUX JP [1984] « Les sciences du système nerveux vivent aujourd'hui une révolution » in *Le Courrier du CNRS* n°55/56
- CHOMSKY N [1965] *Aspects of the theory of syntax*, MIT Press, Cambridge Mass.
- CHOMSKY N [1980] *Rules and Representations*, Columbia Univ. Press, New York
- CHOURAQUI E [1981] *Contribution à l'étude théorique de la représentation des connaissances : le système symbolique ARCHES*, Thèse d'Etat, Ins. nat. poly. de Lorraine
- CODD EF [1974] « Seven steps to Rendezvous with a casual user » in *Database Management* [Klimbie & Koffema eds], North Holland Publishing Co, New York
- COLES LS [1968] « An on-line question-answering system with natural language and pictorial input » in *Proc. of the 23rd ACM National Conference*
- COLMEAUER A [1978] « Les grammaires de métamorphose » in *Natural language communications with computers* [Bolc ed], *Lecture Notes in Computer Sciences* n°63, Springer-Verlag
- COULON D [1982] « La compréhension, un processus à profondeur variable » in *Bulletin de Psychologie* XXXV (356)
- COULON D, KAYSER D [1986] « Informatique et langage naturel : présentation générale des méthodes d'interprétation des textes écrits » in *Technique et Science Informatique* Vol.5 n°2, Gauthiers-Villars
- CUTLER A, MEHLER J, NORRIS D, SEGUI J [1983] « A Language-specific Comprehension Strategy » in *Nature* n°304
- FAHLMAN SE [1979] *NETL : a system for representing & using real-world knowledge*, MIT Press, Cambridge Mass.
- FILLMORE ChJ [1968] « The case for case » in *Universals in Linguistic theory* [Bach & Harms eds], Holt, Rinehart & Winston Inc, New York
- FILLMORE ChJ [1971] *Some Problems for Case Grammar*, Academic Press, New York
- FODOR J [1983] *The Modularity of Mind*, MIT Press, Cambridge Mass.
- FORSTER K [1976] « Accessing the mental lexicon » in *Explorations in the Biology of Language* [Walker ed], Bradford Books, Vermont
- GRAU B [1983] *Analyse et représentation d'un texte d'après le thème du discours* Thèse de 3è cycle, Univ. Paris 6
- GROSS M [1975-1977] *Grammaire transformationnelle du français (1) Syntaxe du verbe (2) Syntaxe du nom*, Larousse, Paris
- HALLIDAY MAK [1961] « Categories of the theory of grammar », in *Word* n°17
- HARRIS ZS [1962] *String analysis of sentence structure*, Mouton, La Hague
- HARRIS ZS [1968] *Mathematical Structures of Language*, John Wiley & sons, New York
- HENDRIX GG [1977] « Human Engineering Techniques for applied Natural Language Processing » in *Proceedings of the 5th IJCAI*
- HENDRIX G, SACERDOTI ED, SAGALOWICZ D, SLOCUM J [1978] « Developing a natural language interface to complex data » in *ACM Transactions on database systems*
- HOOPEER JB [1976] *An introduction to Natural Generative Phonology*, Academic Press, New York
- JAYEZ JH [1982] *Compréhension automatique du langage naturel : le cas du groupe nominal en français*, Masson, Paris
- KATZ JJ, POSTAL PM [1964] *An Integrated Theory of Linguistic Description*, MIT Press, Cambridge Mass.
- KODRATOFF Y [1985] « Quand l'ordinateur apprend » in *La Recherche* n°170
- MARCUS MP [1980] *A theory of syntactic recognition for natural language*, MIT Press, Cambridge Mass.
- McCARTHY J [1958] « Programs With Common Sense » in *Mechanisation of Thought Processes*, Her Majesty's Stationnery Office, London, réédité [1968] in *Semantic information processing* [Minsky ed], MIT Press, Cambridge Mass.
- McCARTHY J [1980] « Circumscription : a form of non monotonic reasoning » in *Artificial Intelligence* n°13
- McCAWLEY JD [1968] « Lexical Insertion in a Transformational Grammar without Deep Structure » in *Papers from the 4th Regional Meeting of Chicago Linguistic Society* [Bailey, Darden & Davison eds], Dept of Linguistics, Chicago Univ.
- McCLELLAND JL, RUMELHART DE & alii [1986] *Parallel distributed processing : explorations in the microstructures of cognition* Vol.1-2, MIT

Press, Cambridge Mass.

- McCULLOCH/PITTS W [1943] « Analogical calculus of the ideas immanent in nervous activity » in *Bulletin of Mathematical Biophysics* n°5
- McDERMOTT DV [1974] « Assimilation of new information by a natural-language understanding system » in *Artificial Intelligence* n°291
- McDERMOTT J [1979] « Learning to use analogies » in *Proc. 6th IJCAI*
- MINSKI M, PAPERT S [1969] *Perceptrons : An Introduction to Computational Geometry*, MIT Press, Cambridge Mass.
- MINSKI M [1975] « A framework for representing knowledge » in *The Psychology of computer vision [Winston ed]*, McGrawhill
- MONTAGÜE R [1970] « Formal Philosophy » in *Selected Papers [Thomason ed]* Yale Univ. Press, New Haven
- NEF F [ed] [1984] *L'analyse logique des langues naturelles (1968-1978)*, Editions du CNRS, Paris
- PITRAT J [1980] *Un interprète de contraintes pour la compréhension du langage naturel*, Rapport 20 du GR 22, Univ. Paris 6
- QUILLIAN MR [1968] « Semantic memory » in *Semantic information processing [Minsky ed]*, MIT Press, Cambridge Mass.
- RAPHAEL B [1964] *Sir, a computer program for semantic information retrieval*, Thèse de PHD, Dpt of Mathematics, MIT
- RIESBECK CK, SCHANK RC [1976] *Comprehension by computer : expectation-based analysis of sentences in context*, Rapport de recherche 78, Yale Univ.
- RIESBECK CK [1975] « Conceptual analysis » in *Conceptual Information Processing [Schank ed]*, American Elsevier, New York
- RIESBECK CK [1979] « Representation to aid distributed understanding in a multiprogramming system » in *Associative networks : representation and use of knowledge by computers [Findler ed]*, Academic Press, New York
- ROBINSON JA [1965] « A machine-oriented logic based on the resolution principle » in *Journal of ACM* n°12
- RUSSELL B [1940] *Inquiry into meaning and truth*, Allen, London
- SABATIER P [1984] « Puzzle grammars » in *Actes du colloque Compréhension du langage naturel & programmation en logique*, Rennes
- SAGER N [1967] « Syntactic analysis of natural language » in *Advance in computers [Alt, Rubino ed]* Vol.8, Academic Press, New York
- SAGER N [1981] *Natural language information processing : a computer grammar of English and its applications*, Addison-Wesley
- SALKOFF M [1973] *Une grammaire en chaîne du français*, Dunod, Paris
- SALKOFF M [1979] *Analyse syntaxique du français : grammaire en chaîne*, John Benjamins, Amsterdam
- SCHANK RC [1973] « Identification of conceptualizations underlying natural language » in [Schank, Colby eds], Freeman, San Francisco
- SCHANK RC [1975] *Conceptual information processing*, American Elsevier, New York
- SCHANK RC [1977] « Representation of understanding of text » in *Machine Intelligence [Elcocks & Michie eds]* n°8, Holsted Press, New York
- SCHANK RC, ABELSON RP [1977] *Scripts, plans, goals, and understanding*, Lawrence Erlbaum Associates, Hillsdal, New Jersey
- SCHANK RC, COLBY KM [eds] [1973] *Computer models of thought and language*, Freeman, San Francisco
- SCHANK RC, RIESBECK CK [1981] *Inside computer understanding : five programs plus miniatures*, Lawrence Erlbaum Associates, Hillsdal, New Jersey
- SCHAUMJAN SK [1977] *Applicational grammar as a semantic theory of natural language*, Chicago Univ. Press
- SCHAUMJAN SK [1982] « The goals of linguistic theory and applicative grammar » in *Mathématiques et sciences humaines* n°77, EPHE, Paris
- SIMMONS RF, SLOCUM J [1972] « Generating English discourse from semantic networks » in *Communication of ACM* n°15(10)
- WALTZ DL, GOODMAN BA [1977] « Writing a natural language data base system » in *Proceedings of IJCAI*, MIT, Cambridge Mass.
- WILKS Y [1973] « An artificial intelligence approach to machine translation » in *Computer models of thought & language [Schank/Colby eds]*, Freeman, San Francisco
- WILKS Y [1975] « A preferential, pattern-seeking, semantics for natural language inference » in *Artificial intelligence* n°6
- WINOGRAD T [1973] « A procedural model of language understanding » in [Schank, Colby eds], Freeman, San Francisco
- WINOGRAD T [1975] « Frame representation & the declarative/procedural controversy » in [Bobrow, Collins eds]
- WINOGRAD T. [1975] *Understanding Natural Language*, Academic Press, New York
- WINOGRAD T [1983] *Language as a cognitive process Vol.1 : Syntax*, Addison-Wesley, Reading Mass.
- WINSTON PH [1978] « Learning by creating and justifying frames » in *Artificial Intelligence* n°10 :2
- WOODS WA [1970] « Transition Network Grammars for Natural Language Analysis » in *Communication of ACM* n°13(10)
- WOODS WA [1975] « What's in a link : foundations for semantic networks » in [Bobrow & Collins eds]
- WOODS WA, KAPLAN RM, NASH-WEBBER B [1972] *The LUNAR sciences natural language information system*, in BBN, Rapport final, Rapport 2378 Cambridge Mass.
- ZADEH LA [1978] « PRUF - a meaning representation language for natural languages » in *International Journal of Man Machine Studies* n°10
- ZADEH LA [1983] *A theory of commonsense knowledge* Memorandum M83/26, ERL, Los Angeles

Reçu février 1988

Universités Paris VII et Paris VIII

Adresse des auteurs :
Groupe Relpred
UA 1028 - 02 du CNRS
Dept. de Recherches Linguistiques

Université Paris VII
2 Place Jussieu 75251 Cedex 05