

La série d'essais annoncée dans l'*Avant-Propos* de ce texte n'a jamais vu le jour. Une tentative voisine fut accomplie par PB avec *Le serpent de mer des primitives sémantiques* demeura, elle aussi, inachevée. Mais ces recherches, conjuguées avec d'autres, trouveront une certaine forme d'aboutissement dans  $\alpha\kappa\epsilon\rho\emptyset$ .

WH, 2003

## **BROUWER, CHWISTEK, CECCATO** et l'univers des "dichômes"

### **A v a n t - P r o p o s**

Le texte qu'on va lire fait partie d'une série d'essais destinés à être publiés dans des conditions variées. Leur présentation n'est donc pas rigoureusement homogène et leur contenu manifeste inévitablement une certaine redondance. Le premier de ces essais, intitulé "Grisa Blues" [0], était un rappel historique essentiellement consacré à l'effort consenti, au début des années 60, par les organismes scientifiques de la Communauté Européenne, dans le domaine de l'"intelligence artificielle". J'y constatais que nos travaux d'alors avaient, sur plusieurs points, anticipé les développements actuels. Mon propos, cependant, n'était nullement nostalgique et les divers essais de cette série s'efforceront d'en administrer la preuve.

Ce second texte propose tout d'abord un retour dans le passé de bien plus grande ampleur que le précédent. On verra pourtant que le projet qu'il sous-tend est demeuré le même : d'ailleurs deux de mes héros participèrent à l'un des séminaires de Blaricum, en 1961. Mais surtout il m'a semblé - à la suite de recherches très postérieures à la période "Grisa" - que l'explicitation d'une problématique commune aux trois pionniers de ce que je propose d'appeler la "dichomologie", n'était pas dépourvue d'intérêt au stade actuel de la recherche.

P.B. Mars 1986

(révisé – sans modification essentielle – en janvier 2003)

## **I. Introduction**

Après un bon quart de siècle de promesses insatisfaites, la pression de l'environnement socio-culturel pour une maîtrise informatique du langage naturel se fait plus vive. En Europe, et plus particulièrement en France, on voit se dessiner les contours d'une industrie ou d'un groupe d'industries nouvelles : les industries du langage [1].

Dans cette perspective on peut constater que si un certain nombre de problèmes techniques importants peuvent être considérés comme résolus, quelques "points durs" demeurent obstinément : ce sont évidemment ceux qui ont trait aux problèmes de la signification. Certes des techniques d'organisation d'une "représentation des connaissances" et de son exploitation pour l'interprétation de textes simples (des questions, notamment), existent, mais elles sont essentiellement développées dans le cadre d'univers très spécialisés (de nature technologique). On sent confusément qu'une véritable percée, nous permettant d'accéder réellement aux significations de textes "naturels" demande la mise en œuvre de concepts spécifiques et de techniques originales. Certains vont d'ailleurs jusqu'à mettre en doute la possibilité même d'un tel développement. Ceux qui ont conservé la foi en la "faisabilité de cette recherche sont engagés dans des voies en apparence très éloignées les unes des autres, mais dont on conserve l'espoir qu'elles vont pouvoir converger:

- représentation connaissances par des organisations structurées : "frames", "scripts", etc..
- approfondissement des relations entre le fonctionnement de l'activité linguistique et celui de la "pensée" .
- élaboration d'hypothèses sur l'activité cérébrale sous-jacente aux processus cognitifs
- simulation de cette de la signification par des formules utilisant des "primitives sémantiques"
- représentation des activités à l'aide de programmes d'automates "neuronaux".

Plutôt que d'une absence d'idées, on se trouve ici en présence d'un véritable embarras de richesses". On manque apparemment d'un fil directeur qui permette d'organiser efficacement la rencontre et la consolidation des travaux en cours. Or, pour identifier un tel fil directeur, nous disposons d'un capital dont la valeur n'a peut-être pas été clairement perçue par les chercheurs. Ce capital, c'est l'existence d'un noyau commun à des recherches déjà anciennes, et célèbres pour d'autres raisons :

- celles de BROUWER sur l'intuitionnisme,
- celles de CHWISTEK sur la métamathématique,
- celles de CECCATO sur l'activité cognitive.

Je me propose de dégager ici les points communs à ces trois approches et d'en évoquer les applications possibles au fameux problème "esprit-langage". Les conséquences à en tirer pour le traitement informatique du langage apparaîtront alors très naturellement.

Brouwer était un géomètre devenu logicien en passant par des préoccupations philosophiques et même métaphysiques. Sa première publication, datée de 1905 est intitulée : "*Life, art and mysticism*" (elle est reproduite dans le volume de ses œuvres [2]). Sans commune mesure avec les thèses de doctorat traditionnelles, elle fit scandale.

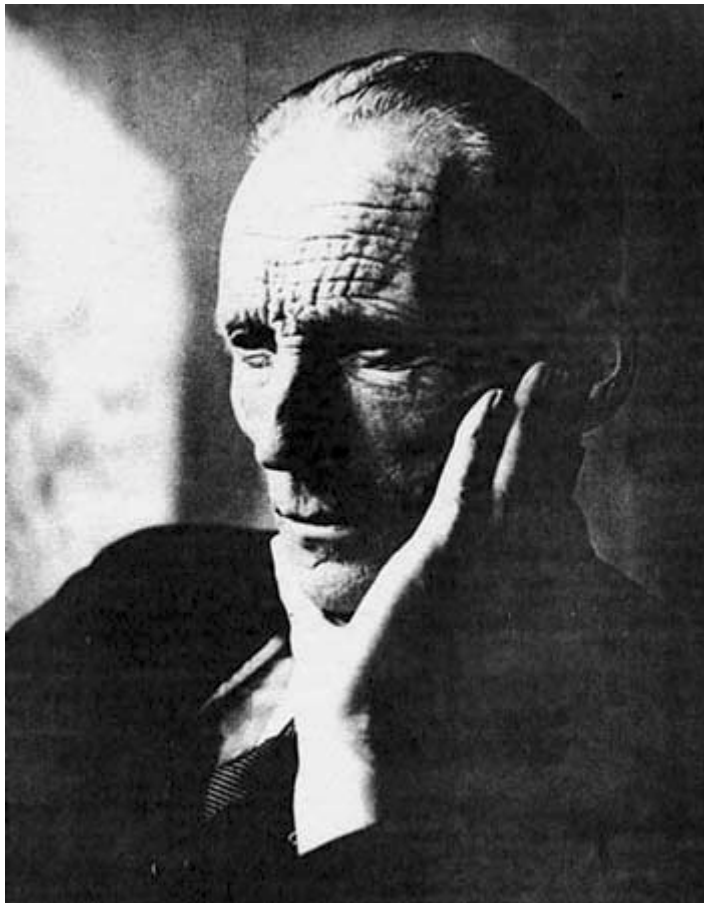
De son côté Chwistek, vigoureux rationaliste, dont les positions en logique sont celles d'un constructivisme et d'un nominalisme sans failles, est conscient de l'existence d'une "Pluralité des réalités". Son activité de critique et d'artiste l'atteste sans la moindre équivoque. On pourra s'en convaincre en lisant la biographie due à Karol

Estreicher (cf. [3]). Oeintre et romancier, il n'en est pas moins un critique averti de la tradition philosophique et surtout de l'héritage idéaliste que nous a laissé Platon.

Parti de préoccupations très différentes, et nourri d'une toute autre culture, Silvio Ceccato ouvre une problématique très voisine de celles que ses aînés explorèrent. Ici on se reportera à [4] qui reprend l'essentiel des œuvres de l'auteur antérieures à 1961.

On ne peut donc qu'être frappé du fait que ces trois auteurs ont construit leurs systèmes très différents) en utilisant des langages formels presque identiques : les arbres binaires (ou dichotomiques). La notion même de "two-icity", est explicitement présentée par BROUWER comme un concept de base dès 1907. Il ne s'agit pas là d'une simple coïncidence, mais, me semble-t-il, d'une indication précieuse qui pourrait ouvrir la voie à des progrès significatifs en recherche cognitive.

## II. Brouwer



Luitzen E.J. Brouwer (1881-1966)

Brouwer est universellement connu comme le fondateur du mouvement "intuitionniste" qui, dès le début du siècle, entreprit d'apporter une solution originale aux problèmes "de fondement" que posait aux mathématiciens la découverte des paradoxes de la théorie des ensembles.

L'entreprise de Brouwer et des quelques disciples qui l'accompagnèrent était réellement révolutionnaire : elle entraînait, en particulier l'abandon d'un principe essentiel de la logique classique, celui du "tiers exclu". Malgré l'importance des

découvertes proprement mathématiques de Brouwer en Topologie (dont le fameux théorème du "point fixe") ce furent les intuitionnistes qui furent exclus le plus souvent du courant principal de la logique mathématique (on dit que Hilbert, au début des années 20, traita Brouwer de "bolchevik" ce qui, à l'époque, n'était pas une mince accusation !).

Après des années de suspicion, et malgré les efforts de Heyting (qui, selon certains, s'était peut-être laissé prendre aux pièges de la formalisation), l'intuitionnisme vint reprendre une place importante dans la communauté scientifique, grâce aux efforts de logiciens tels que Kleene, Kreisel, Troelstra, etc... Kolmogorov lui-même y apporta une importante contribution. On doit pourtant remarquer que la vogue actuelle de l'intuitionnisme, liée aux préoccupations "constructivistes" qu'impose aux logiciens contemporains l'essor de la pensée et de la pratique algorithmiques (et Kolmogorov y a joué un grand rôle), a peut-être détourné l'attention de ce qui était le véritable projet de Brouwer et qui pourrait, aujourd'hui, nous éclairer.

Les historiens de la logique présentent, en effet, l'intuitionnisme comme la succession naturelle de Kronecker, puis des "pré-intuitionnistes" français : Poincaré, Borel, etc... Ils l'identifient souvent aux systèmes formels particuliers développés par Heyting, mais partiellement désavoués par Brouwer lui-même. En réalité il convient de situer Brouwer dans le milieu culturel très particulier qui fut le sien : celui d'un "Amsterdam - fin de siècle" certes moins célèbre que le "Vienne - fin de siècle", mais tout aussi original.

Contrairement à Vienne où l'impulsion provient surtout des milieux artistiques et littéraires, Amsterdam voit se développer un mouvement de réflexion et de contestation radicale dirigé par des mathématiciens professionnels, dont certains - tel Pannekoek - seront, après 1917, les fondateurs du Parti Communiste Néerlandais (ce qui explique peut-être les attaques "ad hominem" de Hilbert). L'un d'eux, Gerrit Mannoury, s'inspirant de Lady Welby, fonde le mouvement "Signifik" dès le début du siècle. Il s'agit là d'une orientation résolument "psycho-linguistique", mais dont les applications sont orientées vers l'activité des mathématiciens. Mannoury écrit d'ailleurs, dans *"Les deux pôles de l'esprit (Etude de psychologie linguistique du point de vue communiste)"* :

« *Les mathématiques et la mystique sont les limites extrêmes entre lesquelles se meuvent le "savoir" et le "vouloir" humains.* » ([5]) :

La thèse de Brouwer est rédigée en 1907, en plein démarrage du mouvement "Signifik". Intitulée : *Over de grondslagen der wiskunde* (Sur le fondement des mathématiques), elle donne un rôle exceptionnel, pour un ouvrage de ce genre, à des considérations de psychologie et de linguistique. Le texte ci-après en témoigne :

"SUMMARIZING :

1) *Mathematics is created by a free action independant of experience; it develops from a single aprioristic basic intuition, which may be called invariance in change as well as unity in multitude.*

2) *The first act of construction has two distinct things thought together (also according to Cantor, Vortrag auf...); F.Meyer (verharadl...) says that one thing is sufficient, because the circumstance that I think of it may be added as a second thing; this is false, for exactly this adding (i.e. setting it while the former is retained) presupposes the intuition of twoicity; only afterwards this simplest mathematical system is projected on the first thing and the ego which thinks the thing." ([2], p.179)*

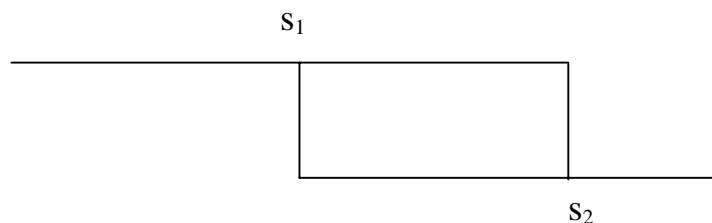
Quarante ans plus tard, à l'occasion du X<sup>ème</sup> Congrès International de Philosophie, Brouwer n'a pas modifié son point de vue :

*In the world of sensation experienced by mind, the free-will phenomenon of causal attention occurs. It performs identification of different sensations, and in this way, in a downing atmosphere of forethought, creates iterative complexes of sensations. (p.480) [...] Mathematics comes into being, when the two-ity created by a move of time is diverted of all quality by the subject, and when the remaining empty form of the common substratum of all two-ities, as basic intuition of mathematics, is e to an unlimited unfolding... (p.482)*

Dans son dernier article important, "Points and spaces", Brouwer est encore plus clair :

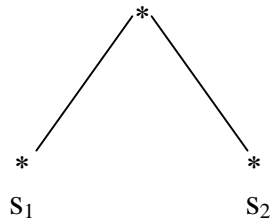
*The first act of intuitionism completely separates mathematics from mathematical language, in particular from the phenomena of language which are described by theoretical logic. It recognizes that mathematics is a languageless activity of the mind having its origin in the basic phenomenon of a move of time, which is the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained by memory. If the two-ity thus born is diverted of all quality there remains the common substratum of all two-ities, the mental creation of the empty two-ity. This empty two-ity and the two unities of which it is composed, constitute the basic mathematical systems. And the basic operation of mathematical construction is the mental creation of the two-icity of two mathematical systems previously acquired and the consideration of this two-icity as a new mathematical system. (p. 523).*

On est donc en présence ici d'une sorte de "phénoménologie de la pensée vide" qui met en évidence un "procès dichotomique élémentaire" considéré comme "atome de pensée". On pourrait proposer une représentation graphique du procès, telle que :



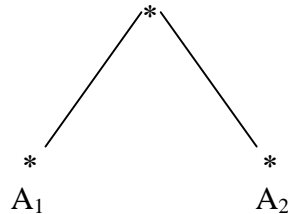
- où S<sub>1</sub> et S<sub>2</sub> sont les deux "unités de pensée abstraite" dont la composition, prise en compte dans un temps historique abstrait constitue le "système mathématique de base.
- l'axe des abscisses est celui du temps dans lequel se déroule l'événement "état de conscience éveillée".
- l'axe des ordonnées évoque un espace (dont la géométrie reste indéterminée) nécessaire à l'actualisation d'états de conscience simultanés.

Une représentation équivalente, mais plus traditionnelle du schéma ci-dessus était l'arbre dichotomique élémentaire :



Si maintenant on affecte à des arbres déjà définis des identificateurs (soit  $A_1$  et  $A_2$ ), l'action que Brouwer appelle : « *the mental creation of two-icity of two mathematical systems previously acquired* »

se traduira par la constitution du "sur-arbre" :



On notera qu'une représentation linéaire des arbres dichotomiques est possible, et ce ceci de trois manières qui correspondent aux représentations usuelles des lois de composition interne en algèbre :

Si on représente symboliquement "*the first act of intuitionism*" par le signe  $*$ , Le schéma et l'arbre précédent pourront être représentés par l'une des expressions :

$* A_1 A_2$  (représentation préfixée – dite "polonaise")

$A_1 * A_2$  (représentation infixée – usuelle en algèbre)

$A_1 A_2 *$  (représentation postfixée – familière aux informaticiens)

Il ne semble pas que Brouwer ait jamais explicité son modèle de construction dynamique des concepts. Son œuvre est essentiellement consacrée à l'illustration des conséquences de son point de vue sur la technique mathématique proprement dite : théorie des ensembles, analyse, topologie.

Pour cela il se borne à exposer le mécanisme d'engendrement de la suite (potentiellement) infinie des entiers naturels auxquels on pourrait associer les arbres (linéarisés ici dans le format "préfixé") :

- $*SS$  pour 1
- $**SSS$  pour 2

- ...
- \*\*...\*SSS...S (avec n \* et n+1 s) pour n

En ce qui concerne les concepts de nature plus directement "logique", tels que les connecteurs "et" et "ou", Brouwer ne suggère pas de représentation formelle spécifique. Au contraire, les représentations proposées et utilisées par Heyting, le continuateur officiel de Brouwer (quoique finalement désavoué par lui), et par d'autres, s'inscrivent essentiellement dans le cadre notationnel de la logique mathématique usuelle et Brouwer ne cessera de se montrer très réservé sur ces tentatives.

Pourtant, en 1948, il introduit une nouvelle idée révolutionnaire: celle du "sujet créateur", qui introduit des considérations "historiques" dans le processus de la démonstration. Cette idée qui ouvre des débats encore actuels - ne présente pas de rapport attesté avec le modèle "primitif" proposé par Brouwer. Elle s'y rattache pourtant, ou plutôt elle se rattache à certains prolongements de ce modèle, prolongements que suggérera, justement, la comparaison avec les deux autres systèmes que nous étudions ici.

Elle se rattache aussi à l'œuvre d'un mathématicien, M Pasch (1843-1930) qui, entre 1909 et 1927, présente à son tour une présentation de l'activité mathématique comme déploiement dans le temps [6] et que, curieusement, Brouwer et ses disciples semblent ignorer. Ce mathématicien, célèbre pour son travail d'axiomatisation de la géométrie, décrit l'activité mathématique comme un "jaillissement de concepts à partir d'éléments primitifs, et ce ci dans le cadre d'une combinatoire qui s'inscrit explicitement dans la durée.

Il distingue ainsi :

- les choses
- les événements
- les données (sorte particulière d'événements)
- les noms propres (sorte particulière de choses )
- les noms collectifs (noms génériques pour des choses données)

Ces entités conceptuelles peuvent entrer dans des relations (telles que – lorsqu'il s'agit d'événements – celles d'antériorité ou de postériorité). De même certaines combinaisons de concepts, ou constructions d'entités de niveau plus élevé sont spécifiées :

- chaînes d'événements
- "troupe" de choses

Ici encore, l'objectif recherché est une véritable "simulation" de l'activité mathématique. On devine cependant qu'une extension du modèle à un univers sémantique plus vaste est concevable. Mais une telle extension n'a pas été tentée à ce jour.

### III. Chwistek



Léon Chwistek (1884-1944),

Au moment même où Vienne et Amsterdam voient se développer des mouvements culturels révolutionnaires, la Pologne est également le théâtre d'une activité créatrice exceptionnelle. A Zakopane, ville de plaisance située dans les Tatras, à la frontière de la Bohême, trois jeunes gens se préparent à changer le monde :

Bronislaw Malinowski (1884-1942) deviendra le fondateur de l'ethnologie scientifique.

Ignacy Witkiewicz (1885-1939) peintre, romancier, philosophe, sera l'inventeur, bien avant Ionesco, du "Théâtre de l'absurde".

Léon Chwistek (1884-1944), créateur du mouvement "formiste" avec Witkiewicz, sera également écrivain, peintre, mais surtout logicien et inventeur d'un système formel "dichotomique dont nous allons rappeler les grandes lignes.

Avant le coup d'éclat de 1907, Brouwer était déjà un mathématicien connu. Son "théorème du point fixe" avait fait de lui un des plus célèbres géomètres de son temps. Chwistek, lui, pris par de multiples intérêts, ne va cesser de se mouvoir et d'alterner des activités artistiques, scientifiques elles-mêmes très diverses :

- sa première œuvre est un drame : *Coburg*, composé en 1901. - il s'intéresse ensuite à la philosophie de la perception et publie en 1909 un article *Sur les variations périodiques du contenu des images vues dans un contour donné*.

- son premier essai sur le fondement des mathématiques date de 1912 : *The principle of contradiction in the light of recent investigations of Bertrand Russell*.

- après une période consacrée à la création du mouvement "formiste", Chwistek écrit un article essentiel *Antinomies of formal logic* où il propose le système qui sera connu - et adopté - par la plupart des logiciens sous le nom de "Théorie simple des



types" (qu'on attribue souvent, à tort, à Ramsay).

- mais Chwistek n'est pas satisfait du compromis ainsi réalisé. Des 1924, il propose une *Theory of constructive types* : théorie "ramifiée", sans l'"axiome de reductibilité" que Russell avait proposé, puis abandonné.

- de 1928 à 1932 (c'est l'époque où Curry et Church développent leurs systèmes), il se lance dans l'édification d'un système formel original dans lequel il voit la possibilité d'assurer de façon rigoureuse le fondement de toute activité rationnelle : logique, mathématique, physique, etc.

- en 1935, il publie un traité : *Granice Nauki* (les limites de la Science) qui rassemble les résultats obtenus. Mais deux élèves se sont joints à lui : Hetper et Herzberg, qui vont lui permettre d'accomplir des progrès techniques décisifs qui sont présentés dans plusieurs textes, dont deux paraissent dans le célèbre *Journal of Symbolic Logic*.

- la guerre empêchera Chwistek d'achever sa tâche : Hetper et Herzberg sont assassinés par les nazis. Lui-même se réfugie en URSS et publie encore deux textes dans les Comptes-rendus de l'Académie des Sciences de Géorgie.

Il meurt à Moscou en 1944 non sans laisser deux ouvrages posthumes : *La méthode générale des sciences positives* [7], rédigé en français, et paru chez Hermann en 1946, puis *The Limits of Science* [8], rédigé en anglais et paru chez Kegan Paul en 1948. Ce dernier ouvrage est une adaptation du livre de 1935 enrichie des résultats nouveaux obtenus avec ses collaborateurs. Les œuvres que Chwistek avait publiées en polonais seront rassemblées et commentées par K.Pasenkiewicz en 1961 [9].

La personnalité de Chwistek et le développement peu orthodoxe de sa carrière l'ont empêché d'être vraiment connu et apprécié de la communauté scientifique. En 1948, pourtant le logicien américain John Myhill offre une première analyse sérieuse de cette œuvre souvent obscure. En Pologne même, il faudra attendre 1967 pour qu'un chercheur, Z. Zwinogrodski, poussé par Pasenkiewicz, en fasse une analyse satisfaisante [10].

J'avais été frappé, en relisant Chwistek que j'avais découvert dès la fin de la guerre, grâce au petit livre publié chez Hermann, par l'actualité de ses préoccupations, au moment même où, à Orsay, je m'efforçais de développer des activités de recherche dans le domaine de la manipulation assistée par ordinateurs des systèmes formels de la mathématique et de la Logique. À l'occasion du Congrès "*Mathematical foundations of Computer Science*", tenu à Varsovie en 1973, j'ai présenté une communication dont le résumé contenait le passage suivant :

*Chwistek's complex of systems ("fundamental", "auxiliary", "proper", "meta", "generalised meta") is intended to provide a consistent foundation to the whole of mathematics, and therefore to science in general and possibly beyond. At least "constructive" and even "predicative" mathematics are obtained. The systems are built up from a two element alphabet into an unlimited hierarchy through a list of carefully selected "directives". A modern reader will notice that*

- a) *the "fundamental system" is an object language which can be used as a syntactic device as well as an encoder or an inter-preter, through "unstratified control" techniques (in Gorn's terminology);*
- b) *the "auxiliary system" is built up from the "fundamental" one in a way which is very similar to the way "macroassemblers" are developed from machine languages;*
- c) *hierarchies of "proper" and "meta" systems also bear some resemblance to high level programming languages and systems. [11]*

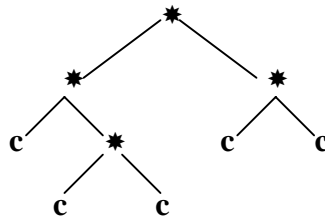
Il n'est évidemment pas question d'analyser ici en détail les techniques formelles que Chwistek et ses élèves ont progressivement mises au point. L'essentiel est ici d'identifier les mécanismes qu'il a utilisés, et d'en étudier le comportement. Ce mécanisme peut se décrire ainsi :

- Il s'agit d'un code n'utilisant que les signes typographiques \* et c
- La formation des mots du code se définit ainsi (règles **RP**):
  - c est un mot du code
  - si M et N sont des mots du code, la chaîne \*MN est un mot du code.

C'est ainsi que **\*c\*cc\*cc** est un mot du code, tandis que **\*\*\*c\*cc\*\*** ne l'est pas.

Il s'agit donc d'une "présentation préfixée" de l'ensemble des "arbres dichotomiques" :

C'est ainsi qu'au mot **\*c\*cc\*cc** correspondra l'arbre :



Sur cette base, Chwistek construit une hiérarchie infinie de systèmes formels en procédant par étapes. Il propose successivement :

- un système de codage (celui que nous venons de décrire) ; en termes modernes, le langage engendré par ce code est un sous-ensemble du monoïde libre engendré par l'alphabet { \*, c }, sous-ensemble spécifié par les règles de production **RP**.
- un système de base qui donne le codage de symboles primitifs :
 

<b>*cc</b>		est le code pour <b>0</b>
<b>**cc*cc</b>	(donc <b>*00</b> )	est le code pour <b>1</b>
etc.		
<b>**nn*nn</b>		est le code pour <b>n+1</b>

Les nombres entiers sont donc codés par des arbres dichotomiques équilibrés. Deux mots "asymétriques" sont alors définis :

**\*\*cc\*\*cc\*\*cc** est le code pour **I**  
**\*\*\*cc\*\*cc\*\*cc** est le code pour **II**

Ces mots "élémentaires" servent à définir de nouveaux symboles et en premier lieu les connecteurs du calcul des propositions, en particulier le "symbole de Scheffer" (souvent désigné par / et prononcé "non-et"). C'est ainsi que :

**\* $\overline{I E}$   $\overline{II F}$**  est le code pour **/EF**

- un système auxiliaire ; c'est dans ce système que sont explicités les codes
  - des entiers de "type" **L** (où **L** est un mot du code)
  - des fonctions primitives propositionnelles
  - des variables propositionnelles
  - des variables "sémantiques"
  - des quantificateurs
  - des axiomes
  - de la règle de substitution
- une hiérarchie de "systèmes propres" qui dépendent de deux indices numériques (eux-mêmes codifiables dans le système de base). On peut alors écrire, dans ces systèmes, des énoncés qui expriment les propriétés des objets de type correspondant au système propre dans lequel on se situe.
- une hiérarchie ramifiée de "métasystèmes" dépendant de trois indices numériques. On peut, dans ces métasystèmes, démontrer les théorèmes qui ont été énoncés dans les systèmes propres qui leur correspondent.
- Une hiérarchie de "métasystèmes limites" et de "métasystèmes généralisés"; cette hiérarchie a été conçue par Hetper, mais nous n'en connaissons pas de description précise ([12], p.124).

On arrive ainsi à un système complexe dont l'articulation peut être visualisée comme suit :

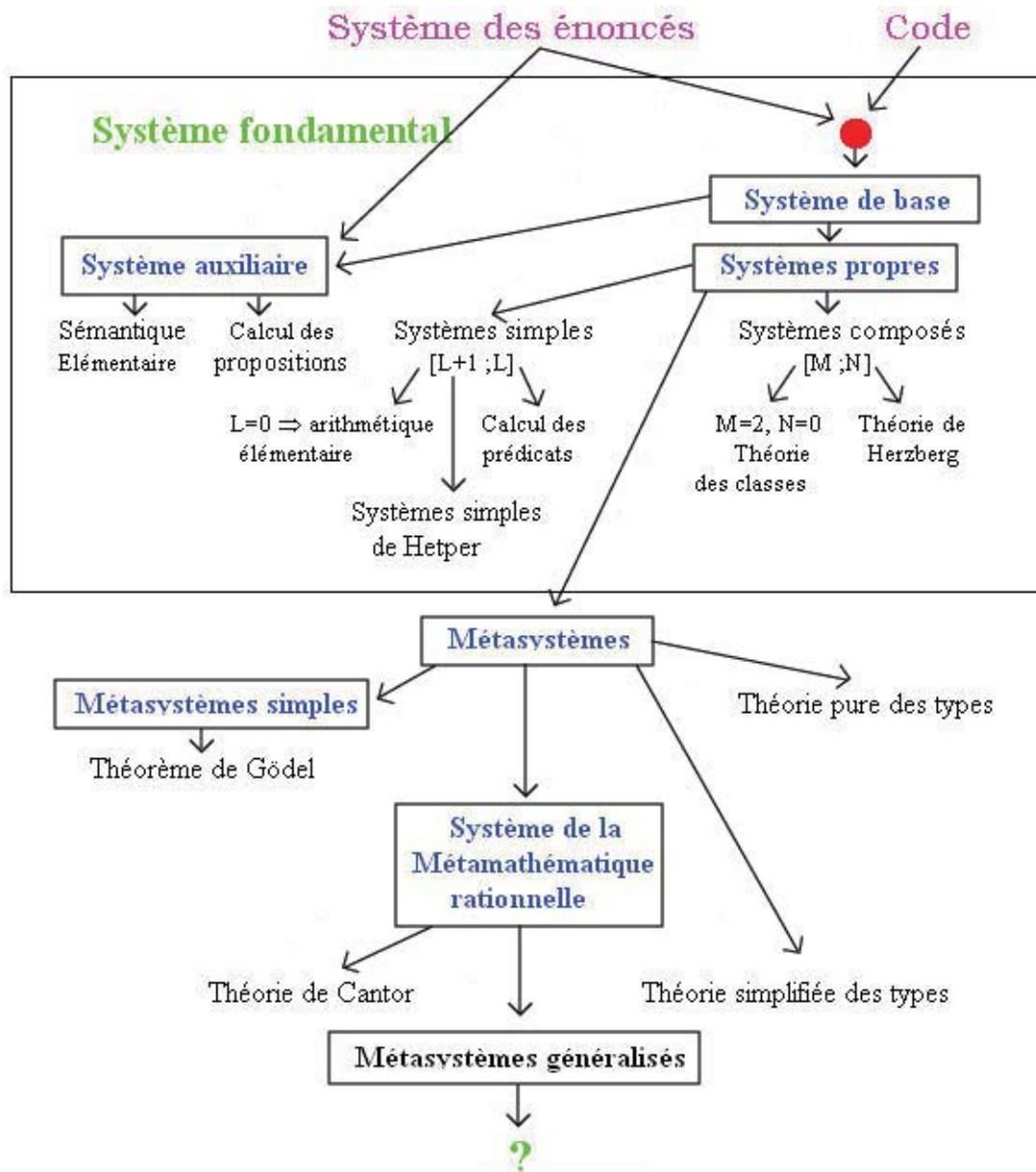


Tableau d'ensemble des systèmes de Chwistek

Bien que la tâche que s'était fixée Chwistek n'ait pas été réellement achevée, ce que nous connaissons de son travail donne à penser qu'une reconstruction d'un système formel de la mathématique constructive à la fois cohérent et complet (donc se situant en dehors du champ d'application du théorème de Gödel) serait possible sur cette base. Une telle possibilité – qui avait échappé à Russell mais que Myhill avait bien perçue – repose sur une "astuce technique" de la codification : la mise en œuvre systématique d'une sorte d'effet de "zoom" (ou de "fractalisation" inverse).

Lorsqu'on dispose d'un mot du système de base, on peut, en effet remplacer toutes les occurrences de  $c$  dans ce mot par un autre mot du système afin d'obtenir une construction qui combine les propriétés "syntaxiques" du premier avec les propriétés "sémantiques" du second. On obtient sans peine un langage formel qui inclut son propre métalangage sans donner prise aux paradoxes ni sacrifier la capacité d'expression du système. Par contre le mécanisme de zoom entraîne une croissance

exponentielle de la longueur des expressions (en **c** et **\***) lorsqu'il devient nécessaire de les expliciter – et c'est la forme ultime de toute démonstration.

Prenons, à titre d'exemple, la proposition

$$\mathbf{p} \wedge \mathbf{q} \quad (\mathbf{p} \text{ et } \mathbf{q})$$

où **p** et **q** seraient des variables propositionnelles du système propre (54) – c'est un niveau que l'on atteint pour des énoncés "métamathématiques tels que le théorème d'incomplétude de Gödel – se traduit par

**\*\*\*NN\*\*NN\*NN\*\*\*NN\*NN\*NN**

où **N** est

**\*\*\*PP\*\*PP\*PP\*\*\*QQ\*QQ\*QQ**

**P** étant une variable propositionnelle du système (54), peut alors s'écrire :

**\*\*\*CC\*\*CC\*CC\*\*\*SS\*\*SS\*SS**

**C** et **S** étant les codes des entiers 5 et 6. Ce sont, rappelons-le, des arbres dichotomiques complets qui présentent donc respectivement 32 et 64 occurrences de **c**. Un calcul simple montre alors que la forme développée de la proposition  $\mathbf{p} \wedge \mathbf{q}$  contient exactement 138.240 occurrences de **c**.

Comme l'indique Chwistek lui-même :

*..we have to do with very long symbolic expressions and as there is not yet any machine to build up such expressions... ([12], p.1) [...]*

*Only when these abbreviations are eliminated will correct theorems of the system be obtained. It is obvious that this could be done by a special machine ([9], p.166).*

Une telle vision était évidemment prémonitoire. Mon objectif, lors de ma communication de 1973, était d'inciter les chercheurs à reprendre les idées de Chwistek en utilisant les ressources des techniques de l'informatique. Mais celles-ci n'étaient peut-être pas suffisamment développées, à l'époque, pour que cela soit effectivement réalisable.

#### IV. Ceccato



Silvio Ceccato (1914-1995)

Silvio Ceccato, issu d'une riche famille de Vicenza (il s'agit des *Pompes Ceccato*) décida très tôt de rompre avec son milieu familial pour entreprendre une carrière artistique, plus précisément une carrière musicale : étude de piano et de composition. Il se passionna bientôt pour les problèmes de l'esthétique puis de l'épistémologie. Son premier travail important s'intitule de façon très significative :

*Un contributo alle ricerche sul fare dell'uomo* ([4] I, p.53).

Dès lors Ceccato se livre à une critique extrêmement aiguë de ce qu'il appelle la "tradition cognicienne". Porté par un mouvement qui fait parfois songer à Wittgenstein, il tente, en 1946, une "axiomatisation du connaître" ([4] I, p.167). Des amis provenant d'horizons variés : physique, ingénierie, psychiatrie, se joignent à lui : la "*Scuola operativa italiana*" est en gestation. Elle réunira Enzo Morpurgo, Entico Maretti, Vittorio Somenzi et Giuseppe Vaccarino.

A l'occasion des "entretiens de Zurich", en 1947 un texte essentiel est rédigé : *Quelques espèces d'opérations*. Ce texte (écrit en français) présente pour la première fois les travaux de ce qui s'appelle encore "Centre Italien de Méthodologie et Analyse du Langage" ([4] I, p.261). ces recherches se définissent progressivement par opposition à la tradition philosophique, tout en se situant dans la mouvance de Bridgman et de Dingler. Après "*Il Teocono*" ([4] II, p.135), texte à l'humour dévastateur qui présente l'histoire de la philosophie comme un jeu de société, c'est, en 1949 la parution chez Hermann de "*Il linguaggio con la tabella di Ceccatieff*" ([10] II, p.241) qui voit se préciser l'orientation plus spécifiquement linguistique de la "scuola".

Mais c'est à Londres, en 1953, qu'un choix fondamental est annoncé : celui d'une tentative de vérification de l'analyse par la simulation :

*The Italian Operational School has lately been formed in order to realize a plan which comprises :*

- a) a particular sort of operational analysis of thinking;  
 b) the construction of an apparatus that thinks in the manner shown by the analysis.

....

*The construction of the Homoculus cogitans, for our help and delight, will show whether our results are right.*"([4] II, pp.646,647)

Pour organiser cette simulation il devient donc nécessaire de formaliser et, pour cela d'identifier des opérations "élémentaires" qu'il sera possible de composer. Cette formalisation se précise peu à peu entre 1953 et 1956, sous l'influence d'Enrico Maretti, un ingénieur qui prend en charge la construction d'un "modèle mécanique des opérations mentales, modèle baptisé (sans trop de modestie), "*Adamo II*" ([13]).

Bien entendu *Adamo II* n'est qu'une ébauche. La réalisation d'un tel projet réclamerait, en effet, la mise en jeu de moyens puissants, sans aucune comparaison avec le gadget électro-mécanique présenté en avril 1956 à la *Mostra dell'Automatismo*, à Milan. Il faudrait, pour réaliser une expérience "en vraie grandeur", disposer de calculatrices électroniques, de programmeurs, etc.. Il faut donc obtenir des crédits, ce qui est rendu possible par l'engouement soudain – et parfois encombrant – que manifestent les médias pour les domaines de la documentation et de la traduction automatiques.

L'équipe de Ceccato se renforce et se rassemble à Milan, au *Centro di Cibernetica e di attività linguistiche* et est hébergé par la prestigieuse *Università degli studi*. Un contrat de recherche est passé avec l'*US Air Force* (!) et fera l'objet d'un rapport publié en 1960 [14]. Un autre est signé avec EURATOM où je dirige alors une équipe qui anime le développement de la recherche européenne en Intelligence Artificielle [15].

Des linguistes (des documentalistes, psycho-linguistes et spécialistes du russe et de l'allemand) ont rejoint l'équipe de Milan et ont mis au point une variante de la "Table de Ceccatieff : il s'agissait là d'une technique de catégorisation des mots d'une langue naturelle (appliquée au russe, à l'anglais et à l'italien. [16].

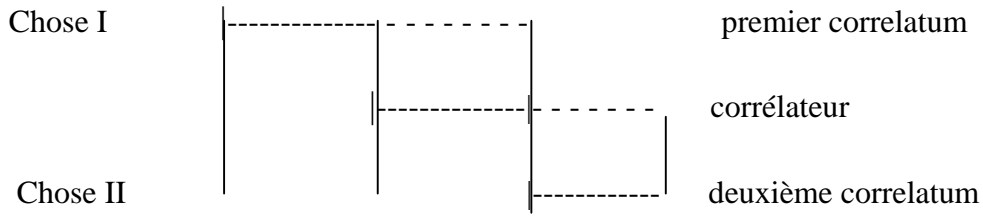
Mais la pression de la demande (américaine, surtout) empêche qu'un recul soit pris à temps et qu'une méthodologie unifiée se dégage. Bientôt les conséquences du fameux "rapport ALPAC" sont là les crédits sont coupés et l'équipe se disperse. Après mon départ, EURATOM renonce à poursuivre son soutien.

Ceccato n'abandonne pourtant pas et développe sa recherche, dans un cadre purement italien, dans la direction de la psychologie de l'enfant et des applications pédagogiques (cf. [17] où il examine quelques unes des difficultés qui sont mentionnées ci-après). Mais il reste ignoré des chercheurs, à la notable exception de Sowa qui en fait, à juste titre, un pionnier, dans son récent ouvrage de synthèse [18].

Au cours de ces années, l'analyse Cecatienne s'est développée simultanément (mais aussi inégalement sur quatre niveaux que je vais présenter tour à tour à l'aide d'exemples, avant d'en proposer un commentaire.

- le premier niveau, celui des "opérations", qui sont les constituants fondamentaux de la pensée et du langage (et donnent son nom à la *Scuola operativa*), est essentiellement celui de la corrélation, un concept fondé sur l'image, évidemment chère à Ceccato, de la polyphonie.

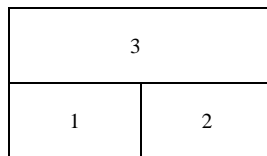
C'est ainsi qu'on écrira :



avec les conventions suivantes :

- déroulement du temps —————>
- temps pendant lequel la "chose" est faite
- - - - - temps pendant lequel la "chose" est maintenue

Une écriture équivalente mais plus commode à manipuler pour la gestion d'une combinatoire est la suivante (la plus utilisée par Ceccato et ses collaborateurs) :



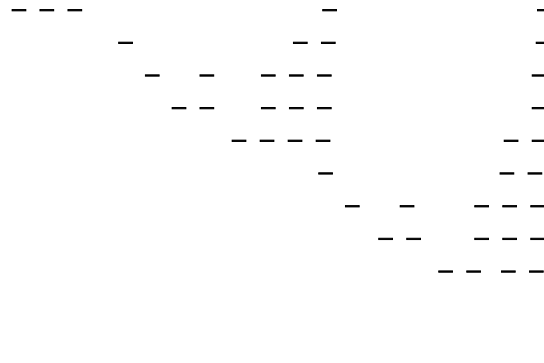
1 : premier correlatum;                      2 : deuxième correlatum;                      3 : corrélateur

- au second niveau le modèle général de la corrélation est utilisé pour développer une combinatoire des "états d'attention" avec laquelle on s'efforcera d'opérer une analyse des concepts et des mots qui leur sont associés dans les langues usuelles. Ici aussi, plusieurs formalismes – équivalents – sont utilisés, que nous illustrerons en présentant, sous différentes formes, l'analyse Ceccatienne de "punto" (point) :

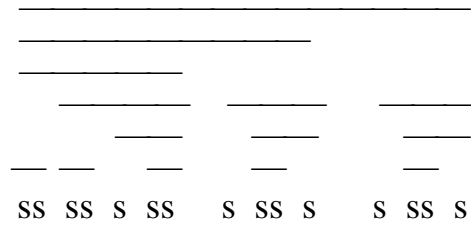
a) notation "musicale"







b) notation "Hilbertienne" (avec utilisation du "vinculum" : SS) :



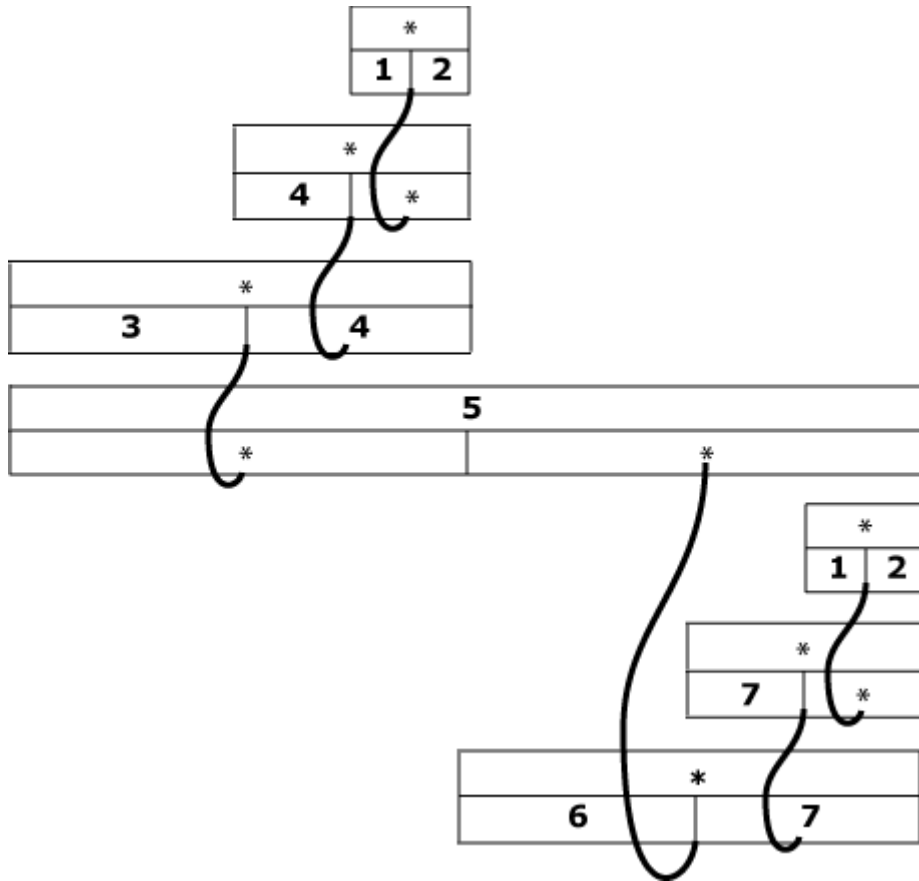
c) notation "polonaise" (évoquée succinctement par Ceccato, [17]) :

D<sub>3</sub> D<sub>3</sub> D<sub>2</sub> D<sub>1</sub> SSD<sub>3</sub> D<sub>1</sub> SS D<sub>2</sub> S D<sub>1</sub> SS D<sub>2</sub> S D<sub>1</sub> D<sub>2</sub> SSS D<sub>2</sub> S D<sub>2</sub> D<sub>1</sub> SSS

que l'on pourrait d'ailleurs traduire "à la Chwistek" par

\*\*\*SS\*\*SS\*S\*\*SS\*S\*\*\*SS\*S\*\*\*SS

- le troisième niveau est celui où les concepts s'articulent pour former des mots fléchis, des syntagmes, des phrases, des textes. Ici l'analyse "fine" devient évidemment difficile à maîtriser. On utilisera essentiellement des "réseaux corrélacionnels tels que :



Qui est une représentation graphique de la structure syntaxique de la phrase

(1) The (2) book (3) Mary (4) writes (5) and (6) John (7) reads

- le quatrième niveau est celui où s'inscrit l'organisation globale des connaissances : c'est, en fait, l'encyclopédie, concept souvent évoqués, aujourd'hui, par les chercheurs. Sous le nom de "sphère notionnelle", Ceccato propose une représentation des connaissances proche de celle définie à la même époque par Jean-Claude Gardin ainsi que par André Leroy et moi-même.

Pour profonde et anticipatrice que soit l'œuvre de Ceccato, elle n'en laisse pas moins subsister ici et là des zones d'incertitude que l'interruption des recherches proprement algorithmiques n'a pas permis d'éliminer et qui se situent précisément aux points d'articulation des divers niveaux de l'analyse : *concaténation*, *arborisation*, interprétation des assemblages organisés en *macrostructures* et *mégastructures*

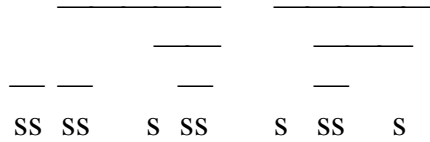
### **concaténation**

La concaténation de deux schémas s'effectue évidemment à l'aide du processus de corrélation (au sens ceccatien du terme). La construction progressive de schémas complexes à partir de schémas simples s'effectue toujours ainsi, chez lui.

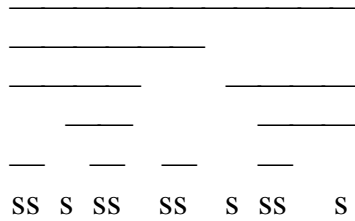
Considérons par exemple les concepts place (postto) dont le schéma est

\_\_\_\_\_

\_\_\_\_\_



et moment (momento) dont le schéma est



L'examen des schémas déjà définis permet de reconstruire progressivement ces deux concepts en vue d'une analyse sémantique en profondeur. Posons en effet :

chose =  $\overline{SS} = C$                       et                      objet =  $\overline{S SS}$                       puis

singulier =  $\overline{S}$  objet      pluriel =  $\overline{chose}$  objet

espace =  $\overline{chose}$  pluriel      temps =  $\overline{pluriel}$  chose = T

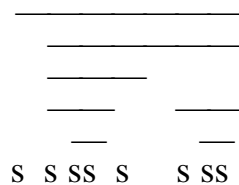
On peut en déduire :      place =  $\overline{espace}$  singulier                      et      moment =  $\overline{temps}$  singulier

On voit ainsi comment peuvent se propager, dans l'élaboration progressive des schémas, des symétries et dissymétries sémantiques fondamentales.

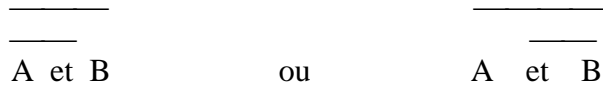
**arborisation et linéarisation**

Dès que le projet de simulation du langage par des schémas aborde le niveau du discours, il devient nécessaire de fournir une représentation de l'articulation syntagmatique qui découle naturellement du principe général de l'analyse. Une difficulté apparaît alors :

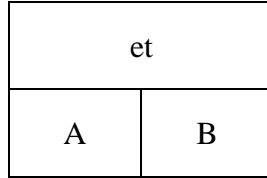
-                      au niveau "atomique" les corrélateurs usuels (au sens de la logique) se voient bien attribuer des schémas spécifiques. C'est ainsi que le schéma pour et est



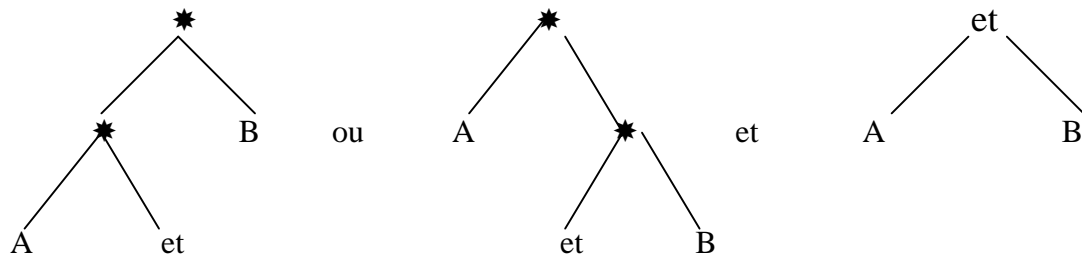
de telle sorte que le schéma correspondant au syntagme "A et B" peut être



- au niveau "moléculaire", on écrira simplement, au contraire,



On a donc deux types d'arborisation possibles :

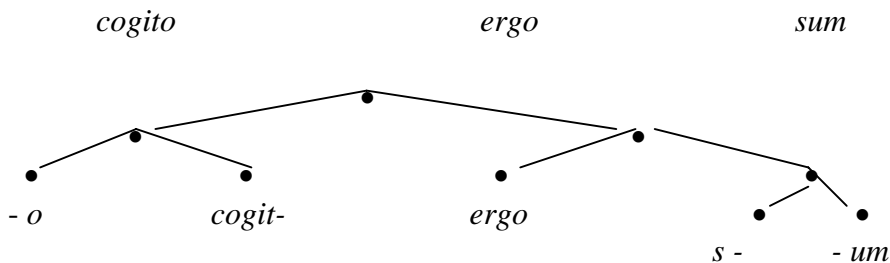


C'est la deuxième version qui prédomine à partir de 1961. Mais du coup on perd les bénéfices de l'analyse "fine".

**macrostructures**

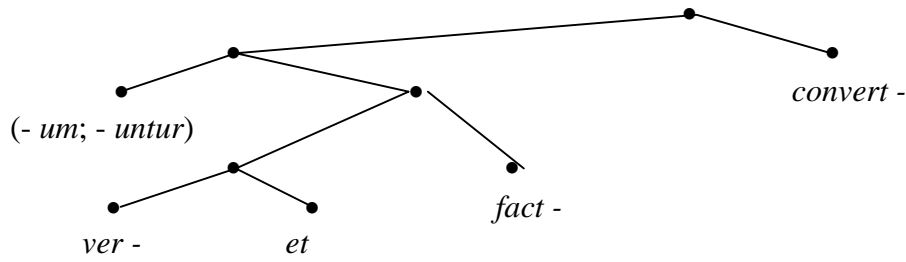
Deux phrases seulement ont fait l'objet d'une analyse complète au sens ou Ceccato l'entendait. En voici une représentation structurale qui utilise une technique voisine de celle des "stemmas" de Tesnière :

- de **Descartes**

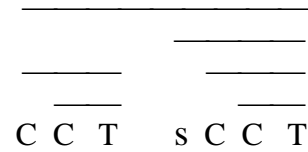


- de **Vico**





On remarquera que, dans le schéma détaillé qui est donné en [13], les sous-arbres associés à "o" et "um" sont identiques et ont la forme



De même, les deux "- um" et le "- untur" sont associés au même sous-arbre



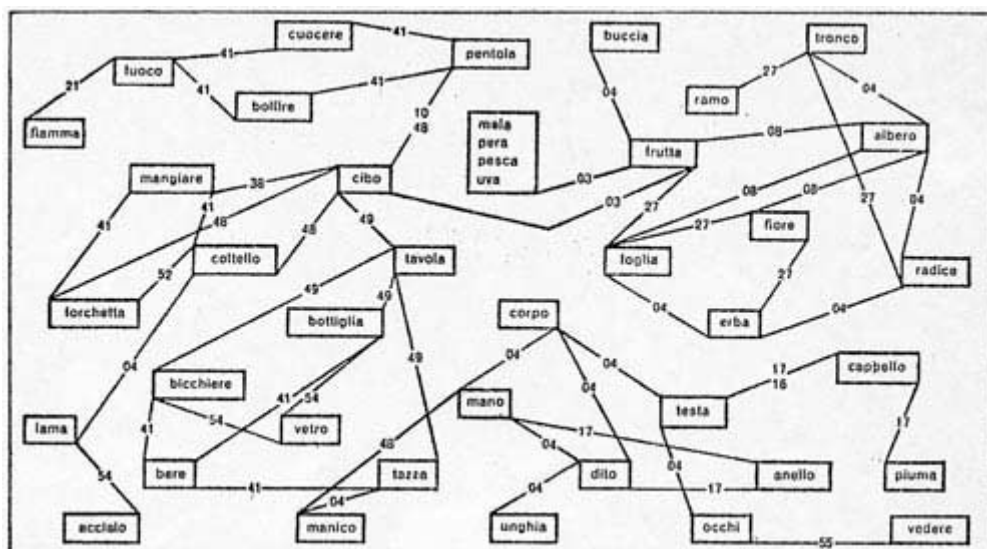
C et T désignant les sous-arbres définis plus haut. Comparant ces deux expressions on pourrait alors en déduire que

"première personne du singulier de l'indicatif présent"	=	nominatif singulier neutre	s	nominatif singulier neutre
---------------------------------------------------------------	---	----------------------------------	---	----------------------------------

Mais c'est évidemment là une identité difficile à justifier sur un plan linguistique, psychologique ou cognitif. De plus on n'aperçoit pas clairement pourquoi la mise en facteur d'un sous-arbre est légitime dans un cas, mais pas dans l'autre.

### mégastructures

Ceccato a présenté à plusieurs reprises des indications sur l'outil encyclopédique auquel il donne parfois le nom de "sphère notionnelle". Voici un exemple (d'un fragment de cette sphère, évidemment, proposé dans [16], vol. 2, p.236).



### SFERA NOZIONALE

#### Tipo di rapporto

01	elemento - collezione
02	esemplare - classe
03	specie - genere
04	parte - tutto
05	componente - composto
06	caratteristiche costitutive - cosa caratterizzata
07	caratteristiche consecutive - cosa caratterizzata
08	che è prodotto - che produce
09	che è prodotto - dove si produce
10	che è contenuto - che contiene
11	che è sostenuto - che sostiene
12	che è spinto - che spinge
13	che è diretto - che dirige
14	che proviene - da dove proviene
15	che precede - che segue
16	che è coperto o chiuso - che copre o chiude
17	che è ornato - che orna
18	la cosa di cui si prende cura - chi si prende cura
19	la cosa principale - l'accessorio
20	mezzo di protezione - la cosa da cui ci si protegge
21	materia - forma
22	materia - suo stato fisico
23	stadio di sviluppo che precede - stadio che segue
24	parentela ascendente - parentela discendente
25	parentela collaterale
26	appartenenza storica
27	contiguità
28	opposizione
29	relazione economica
30	relazione semantica
31	oggetto - attività
32	oggetto - oggetto dell'attività
33	oggetto - risultato dell'attività
34	oggetto - materiale dell'attività
35	oggetto - strumento dell'attività
36	oggetto - dove è attivo di solito
37	oggetto - quando è attivo di solito
38	attività - oggetto
39	attività - risultato
40	attività - materiale
41	attività - strumento
42	attività - dove si svolge di solito
43	attività - quando si svolge di solito
44	risultato - materiale della stessa attività
45	risultato - strumento della stessa attività
46	oggetto - risultato della stessa attività
47	oggetto - materiale della stessa attività
48	oggetto - strumento della stessa attività
49	oggetto - dove si trova di solito
50	oggetto - quando si trova di solito
51	materiale - strumento della stessa attività
52	complementarità di strumenti
53	complementarità di oggetti
54	cosa - materiale
55	funzione - organo

Une telle structure, caractéristique de l'organisation du système au niveau 4, est fondée sur les organisations de niveau 3, mots et relations qui en sont les constituants. On ne comprend cependant pas aisément comment les analyses en constituants plus élémentaires (de niveau 2) de ces concepts justifient le "graphe conceptuel" qui nous

est proposé.

Lorsqu'on s'efforce de faire le bilan de ces recherches, on se trouve en présence d'une situation assez complexe :

- on dispose d'un modèle d'analyse "atomique" des concepts et de l'analyse effective de quelques dizaines de concepts.
- L'équipe a mis au point une méthode d'analyse des "mots-outils" du langage naturel (articles, conjonctions, prépositions, etc...)
- des linguistes qui ont rejoint l'équipe de Milan ont préparé un système de classification des verbes (appliqué tout spécialement au russe et à l'anglais).
- à l'échelon le plus élevé le concept de "sphère notionnelle" a été dégagé, concept qui est équivalente au concept actuellement à la mode de "frame"

Il serait donc intéressant, ici aussi, de reprendre le flambeau en prenant appui – ce que Ceccato ne fit jamais – sur les travaux similaires d'autres équipes.

## **V. Un commun dénominateur**

Il est maintenant possible de dresser un tableau des similitudes et des différences qui existent entre les points de vue de nos trois auteurs avant d'envisager une possible synthèse et de nouveaux développements.

- Brouwer et Chwistek

ont le désir naturel pour des mathématiciens qui se sont formés au moment de la découverte des "paradoxes" de la théorie des ensembles, de donner un fondement solide à leur discipline. Contemporains de Russell et de Hilbert, leur œuvre est avant tout œuvre de logiciens, ce que n'est pas Ceccato.

- Chwistek et Ceccato

proposent, pour formaliser les systèmes qu'ils décrivent, plusieurs mécanismes de codification équivalents entre eux : ce sont en fait des arbres dichotomiques. Par contre, ils élaborent, à l'aide de ce code commun, des "sémantiques" entièrement distinctes :

La sémantique de Chwistek explicite la sémantique d'un univers bien particulier, celui de la typographie : choix d'un alphabet et de règles d'assemblage des caractères, puis de définitions abrégées. Mathématique et méta-mathématique sont alors essentiellement des activités d'écriture.

Celle de Ceccato s'attache à décrire, depuis le niveau "atomique" jusqu'au plus général, les opérations de la pensée, y compris celles qui caractérisent un langage.. Il s'agit donc de deux "simulations", mais qui s'attachent à des univers bien différents.



Quelques uns des participants au séminaire organisé en commun par EURATOM et IBM Europe à Blaricum (Pays-Bas). Ceccato est à l'extrême gauche et Brouwer à l'extrême droite. On peut reconnaître aussi Claude Berge, Hao Wang, Jean-Claude Gardin, ainsi que les organisateurs : David Hirschberg et l'auteur.

- Brouwer et Ceccato

se proposent de construire un modèle d'activité psycho-linguistique à l'aide d'une combinatoire de schémas dichotomiques qui représentent des "atomes de pensée". Après avoir esquissé une analyse phénoménologique de la pensée "naissante", Brouwer s'intéresse (et ce sera sa dernière contribution dans ce domaine) à l'expression d'une dynamique spécifique : celle du "sujet créateur". Il se situe ainsi à un niveau très élevé de complexité dans la hiérarchie des "opérations mentales". Chwistek n'est pas insensible à cette problématique mais ne s'y attache pas explicitement.

Ce commun dénominateur que l'on peut découvrir à l'étude de trois auteurs se présente essentiellement comme une problématique de l'emploi du temps : si le mécanisme fondamental de la pensée se présente en effet comme un développement, ou même un déploiement dans le temps, il faut examiner notamment

- dans quelle mesure Brouwer, en se limitant à l'examen du processus d'engendrement de la suite des nombres entiers, ne laisse pas de côté les véritables difficultés de son projet.
- comment il faudrait prendre en compte, dans le(s) système(s) de Chwistek, le



contenu procédural des opérations qui le caractérisent : substitutions, abréviations, expansions, reconnaissance de sous-expressions, etc.

Dans un cas comme dans l'autre, on trouvera des suggestions intéressantes dans l'œuvre de Pasch (1843-1930), cité plus haut. Ce mathématicien, célèbre pour son travail de pionnier : l'axiomatisation de la géométrie, a consacré toute la fin de sa vie à une tentative, qui demeure, aujourd'hui encore, presque totalement ignorée d'analyse "dynamique" de l'activité mathématique (cf. [15] & [16]).

Mais c'est en examinant de façon réellement approfondie l'œuvre de Ceccato que notre problématique se dessine plus précisément, et que des solutions peuvent être ébauchées. En effet :

- Comme Chwistek, Ceccato propose un système d'analyse qui se développe à plusieurs niveaux :
  - états élémentaires d'attention et leurs corrélations.
  - interprétation de concepts "de base" (et, pour chaque langage naturel, des mots qui leur correspondent), comme combinaisons de termes du premier niveau
  - présentation de l'organisation syntagmatique de la chaîne parlée comme un système de corrélations d'entités du deuxième niveau.
  - constitution d'une (ou de) sphère(s) notionnelle(s) comme un système de corrélations d'entités du troisième niveau.
- la circulation des informations et la prise en compte des contraintes pour les processus "inter-niveaux" ne pose pas de problème dans le cas du système de Chwistek puisque les objets manipulés sont strictement "autonymes". Mais, du coup, le système ne se prête pas aisément à la représentation du langage naturel.
- réciproquement, l'articulation "inter-niveaux" qui est évidemment indispensable à la cohérence du système de Ceccato n'est pas - à mes yeux - suffisamment explicite. On en jugera grâce aux exemples ci-après.
  - a) Concaténation vs composition :
  - b) La linéarisation
  - c) Les macrostructures
  - d) La mégastructure :

C'est ici qu'il pourrait être opportun de mettre à l'épreuve les possibilités qu'offrent les progrès récents de la recherche cognitive, possibilités particulièrement évidentes lorsqu' on s'appuie plus spécifiquement sur les travaux de Yorrick Wilks (le premier exposé d'ensemble est celui de [18]. Cf. aussi [19] & [20]).

Le système de la "sémantique préférentielle" (preference semantics), se présente comme une grammaire qui ne pose pas de hiérarchie préalable entre règles syntaxiques et règles sémantiques. Il se développe, lui aussi, selon une hiérarchie de niveaux :

- les "primitives sémantiques", dont le nombre, qui n'est pas définitivement fixé, dépasse la centaine,
- les "formules", combinaisons arborescentes de primitives,
- les "schémas" (templates), combinaisons relationnelles (non nécessairement arborescentes) de formules,
- les "métaschémas" (metatemplates)

Toutes ces remarques permettent d'esquisser une synthèse qui respecterait les contraintes ci-après :

- but à atteindre :

Modéliser l'activité linguistique – et en amont l'activité mentale – dans les termes d'un groupe d'opérations organisées le long d'une hiérarchie combinatoire à partir d'une opération de base semblable à celle qu'envisagent Brouwer et Ceccato, et grâce à l'itération de cette opération au sein d'une machine abstraite susceptible d'en effectuer le travail.

- explicitation du modèle :

Cette machine abstraite devrait être munie de caractéristiques permettant d'exploiter des concepts semblables à ceux que l'informatique a rendu familiers : système "hôte", processus, pointeurs, etc.. de tels concepts permettraient de guider l'intuition tant pour la spécification des entités intermédiaires que pour l'expression des aspects proprement dynamiques.

Dans mon intervention, cités plus haut, au colloque de Varsovie, j'avais souligné que les analyses de Ceccato permettaient d'avancer assez loin dans la direction évoquée par Chwistek.

- mise à l'épreuve du modèle :

Un modèle de ce type fonctionnera, dans l'ensemble, comme un logiciel de CFAO. Il se comportera comme un automate à qui l'on présente des textes "naturels" et qui, lorsque ces textes ont été acceptés (donc se sont avérés "bien formés" syntaxiquement et sémantiquement) déclenche des actions spécifiques : production d'images, contrôle de robots, émission de sons, etc.. La représentation interne des mots du langage naturel exploité utilisera des techniques du type Chwistek/Ceccato. Mettre à l'épreuve le modèle, ce sera vérifier qu'au fur et à mesure de sa construction, c'est-à-dire de la spécification et de la mise en place des entités qui le constituent, le système aura un comportement cohérent lorsqu'on lui fournira des phrases "correctes" et réagira de façon adéquate à des phrases "incorrectes".

Il est possible, dès à présent, de se faire un idée des problèmes à résoudre :

- identification des attributs

L'informatique nous a habitués à considérer que les objets qu'on y manipule :

données, programmes, etc. sont pourvus d'attribuent qui en expriment certaines propriétés et permettent d'assurer la cohérence des assemblages.

Chez Ceccato, les objets sont des arbres binaires (souvent fort complexes) associés à des mots du langage naturel (ils jouent en quelque sorte le rôle d'identificateurs). On peut donc naturellement leur associer une structure et un type. Mais ce sont dans tous les cas de complexité, des objets d'ordre zéro (au sens des logiciens : ce sont des constantes).

Or il est clair que le mécanisme de l'édification progressive des réseaux implique la mise en œuvre d'objets fonctionnels. Au niveau atomique, on dispose bien de l'opération élémentaire de construction par corrélation, mais au niveau macroscopique il faudrait que l'on puisse assigner une fonctionnalité précises aux conjonctions, prépositions, et, bien entendu, aux verbes, si l'on veut reproduire de façon naturelle un mécanisme de reconstruction syntaxique (à la Tesnière) des phrases.

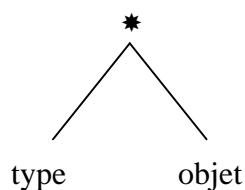
Pour que cela soit possible ici, il faudrait que, dans certains arbres, des nœuds soient rendus vacants et puissent se voir affecter les sommets d'autres arbres. On aurait ainsi à l'œuvre les mécanismes d'abstraction et d'application fonctionnelle du lambda-calcul. C'est ainsi qu'on pourrait résoudre le problème d'arborisation énoncé plus haut.

- types et catégories :

Il existe une certaine ressemblance entre le type des objets de la logique et les catégories grammaticales de la linguistique : ces concepts permettent en effet d'explicitier les mécanismes de contrôle de compatibilité qui régissent la structuration des syntagmes.

Il est donc important qu'un modèle destiné à réaliser une simulation psycholinguistique possède une représentation des types à la fois complète et efficace. Or, de ce point de vue, la technique imaginée par Chwistek n'offre une incontestable richesse d'expression qu'au détriment de l'efficacité : les objets d'un type donné T, où T est une expression bien formée du système, sont engendrés par la mise en œuvre d'une procédure analogue à RP, mais où le signe *c* serait remplacé par l'expression T.

Il est donc tentant de conserver l'idée de Chwistek mais d'éviter l'explosion combinatoire qui en résulte en opérant une mise en facteur du type, mise en facteur que l'on pourrait représenter, dans la notation Chwistek-Ceccato par



Mais c'est là précisément la technique utilisée par Ceccato dans son analyse des phrases de Vico et de Descartes : le "o" (du cogito) et le "um" dans "cogito

ergo sum" sont effectivement préfixés, dans sa représentation, tout comme mes deux "um" et le "untur" du "verum et factum convertuntur".

- Explicitation des processus :

La modélisation informatique – que Chwistek avait évoquée et Ceccato entamée – soulève évidemment des problèmes pratiques considérables mais la résolution de ces problèmes pourrait, en retour, apporter une contribution appréciable à la théorie, en particulier en ce qui concerne l'explicitation des aspects dynamiques du fonctionnement de la machine abstraite. Ainsi :

- Les systèmes formel se présentent comme des assemblages (réalisés sous contraintes) de signes typographiques. Ces signes appartiennent à un alphabet spécifique et sont reconnus comme tels par le logiciel qui les gère. Mais ce logiciel peut reconnaître aussi d'autres caractères dits "sans écho" qui ne contribuent pas directement à la présentation des entités du système, mais jouent un rôle (souvent essentiel) dans le déclenchement des actions auxquelles ces entités participent. C'est le cas, par exemple, du "retour chariot" (ou "entrée"), ou, en combinaison avec d'autres caractères, des touches "control" ou "escape" des claviers informatiques usuels. Les entités invisibles qui sont évoquées par ces caractères bien particulier sont en fait des acteurs du système hôte qui décident de l'acceptation ou du rejet des expressions présentées au système., gèrent la construction d'éléments complexes par corrélation ou composition, déclenchent l'analyse, etc.. ces entités présentent donc un caractère essentiellement dynamique.
- Les opérations de reconnaissance et d'assemblage des expressions ne peuvent pas, le plus souvent, se dérouler de façon linéaire. Mais l'échec d'une tentative d'assemblage, par exemple, n'entraînera pas, en général, le rejet des constituants mis en jeu mais provoquera leur empilement en vue d'essais ultérieurs et l'activation d'autres niveaux (inférieurs ou supérieurs) dans l'architecture du système.
- L'exploitation – indispensable à l'explicitation fine du sens – des informations de la sphère notionnelle implique aussi l'évocation "en parallèle" d'autres activités de reconnaissance et d'assemblage, comme le suggère d'ailleurs J. Pollack et D. Waltz [19].

La réalisation de ces objectifs passe donc par l'élaboration d'un modèle disposant de capacités effectives – et efficaces – de gestion de l'espace et du temps dans un univers de signes : symboles, expressions, textes. Resteront à expliciter les rapports que cet espace et ce temps "internes" entretiendront avec l'espace et le temps "externes" auquel les signes de cet univers entendent faire référence.

Travailler dans cette direction serait – me semble-t-il – faire avancer les idées et les projets inachevés de Brouwer, Chwistek et Ceccato. Ce sont d'ailleurs là des préoccupations tout à fait actuelles comme on peut en juger par

l'observation de recherches récentes :

- Codage

Les arbres binaires ont été mis en avant à plusieurs reprises récemment : John Conway les utilise dans ses recherches combinatoires sur les nombres et les jeux (cf. [20] et aussi l'exposé qu'en donne Donald Knuth dans [21]). Jacques Roubaud les a mis à la base de ses algèbres non-associatives, conjuguant les intuitions de Schützenberger et celles de Benzecri en vue d'une théorie mathématique de la syntaxe. On trouvera une présentation de ces "nombres surnaturels" dans [22].

- Arbres fonctionnels

La recherche cognitive a toujours fait un grand usage de réseaux et de diagrammes variés. Yorrick Wilks, dans sa "preference semantics", utilise, lui, des arbres binaires. Rappelons l'architecture de son système (cf. [23] et [24]) : il se présente comme une grammaire générale (qui ne distingue pas entre règles syntaxiques et règles sémantiques et se développe – lui aussi- selon une hiérarchie de niveaux :

- les "primitives sémantiques" dont le nombre, qui n'est pas définitivement fixé, dépasse la centaine
- les "formules", combinaisons arborescentes de primitives
- les "schémas" (templates), combinaisons relationnelles (pas nécessairement arborescentes) de formules
- les "métaschémas" (blocs et réseaux)

C'est au niveau des formules qu'est utilisée la notion d'arbre binaire. Mais ces arbres peuvent être incomplets. Ils présentent des lacunes qui sont éventuellement remplies au fur et à mesure de la constitution des templates à partir des formules.

- Linguistique fonctionnelle

Les mécanismes linguistiques font un usage constant de la substitution des chaînes de caractères. Or les Logiques Combinatoires sont particulièrement bien adaptées à la description et à la mathématisation de ces procédures. On peut remarquer que les préoccupations de Chwistek rejoignent parfois celles de Curry et retrouver ainsi, dans le modèle que nous pourrions élaborer, une problématique proposée indépendamment par Saumjan et Harris [25] et clairement illustrée par Anne Daladier [26].

- Catégories et types

Mais les logiques combinatoire qui sont utilisées en linguistique théorique sont en fait des lambda-calculs typés. Il est donc naturel d'imaginer qu'ici aussi un phénomène du type Correspondance de Curry est à l'œuvre (correspondance entre les énoncés propositionnels positifs

et lambda-termes, ou les propositions valides correspondent à des lambda-termes.clos).

Or c'est précisément la préoccupation des chercheurs qui, à la suite de N.J. de Bruijn, se proposent de décrire formellement la sémantique de langages bien particuliers :

- les théories mathématiques (en vue de la vérification automatique des démonstrations)
- les langages de programmation (afin d'en valider la sémantique)

On trouvera une présentation générale des travaux de de Bruijn et de son équipe en [27]. Des développements ultérieurs ont été accomplis par Krivine ainsi que par Coquand et Huet [28].

Ainsi se referme une sorte de boucle méthodologique dans la mesure où

- De Bruijn et ses successeurs ouvrent une voie qui pourrait faire des rêves de Brouwer et de Pasch une réalité.
- A la métaphore "directe" de Jean Sammett [29], spécifiant un langage de programmation tel que COBOL en utilisant les concepts de "nom" et de "verbe" répond la métaphore "réciproque" d'Arbib et bien d'autres [30] qui suggèrent la possibilité d'une modélisation informatique des activités mentales.

Mais pour élaborer et évaluer de tels modèles avec une chance de succès il est essentiels de choisir, pour les construire, des univers qui soient à la fois limités en fonctionnalités et riches en potentialités. Des choix de cette nature caractérisent les systèmes de Chwistek et de Ceccato, notamment.

Il existe, en particulier, un univers qui pourrait peut-être fournir un agréable banc d'essais : c'est celui de la littérature – ou plus précisément de l'écriture sous contraintes (telle qu'elle a été définie par François Le Lionnais et Raymond Queneau et pratiquée avec succès par l'OuLiPo).

J'ajouterai que l'Atelier de Littérature Assistée par la Mathématique et les Ordinateurs (ALAMO) a été créé dans le but de fournir aux écrivains un certain nombre d'outils de travail qui vont bien au delà des classiques "traitements de texte" [31]. Pour l'analyse (PALAP) comme pour la synthèse (LAPAL), ces outils mettent en jeu des concepts qui se situent à la fois au niveau du langage naturel et à ceux des métalangages (de ce métalangage "naturel" qu'est la grammaire ou à des niveaux "artificiels" qui sont ceux de la manipulation formelle).

Un retour à Brouwer, à Chwistek, à Ceccato pourrait donc bien, si l'on approfondit les thèmes qui viennent d'être évoqués, donner un élan nouveau et fournir des perspectives inédites à des recherches et à des activités qui n'étaient que suspendues.

## VI. Bibliographie

- [0] Paul Braffort      *Grisa blues ou le système "ex-père"* . Rapport interne ALAMO, 1984.
- [1] B. Cassen          *Les industries du langage : les enjeux.*  
5ème Congrès "Reconnaissance des formes et intelligence artificielle". Thème 1 : journée de synthèse "Langage naturel". AFCET/ADI/IRIA, 1985, p.49.
- [2] L.E.J. Brouwer      *Collected works.*  
North-Holland,1975
- [3] Karol Estreicher    *Leon Chwistek. Biografia artysty (1884 — 1944)*  
Panstwowe Wydawnictwo Naukowe, Cracovie 1971
- [4] S. Ceccato          *Un tecnico fra i filosofi (2 vol.)*  
I. *Come filosofare* ; II. *Come non filosofare*  
Marsilio                      1964,1966
- [5] G. Mannoury        *Les deux pôles de l'esprit.*  
*Etudes de psychologie linguistique du point de vue communiste.*  
Editions Librairie du travail 1933
- [6] Pasch                *Der Ursprung des Zahlbegriffs*  
I.      Arch. Math. Phys. **28** (1920), p.17  
II.     Math. Z. **11** (1921)  
...  
*Mathematik am Ursprung*  
Leipzig, 1927
- [7] L. Chwistek         *La méthode générale des sciences positives.*  
*L'esprit de la sémantique.*  
Hermann, A.S.I. no 1014, 1946
- [8] L. Chwistek         *The Limits of Science.*  
Kegan Paul, 1948
- [9] L. Chwistek         *Pisma filozoficzne i logiczne (2 vol.).*  
PWN,1961,1963
- [10] Z. Zwinogrodski    *Z historii nominalizmu w filozofii matematyki.*  
*Systemy Chwisteka -J.R.Myhilla.*  
Rozprawy filozoficzne, TNT, 1969, p.449
- [11] P. Braffort         *Léon Chwistek and Computer Science.*  
International symposium on the mathematical foundation of  
computer science.(p.5), PWN 1973

- [12] L. Chwistek *New foundations of formal mathematics.*  
J. Symb. Logic 3 (p.1) 1938
- [13] E. Maretti *Adamo II*  
Civiltà delle macchine IV (p.25) 1956
- [14] S. Ceccato & alii *Linguistic analysis and programming for mechanical translation.* Feltrinelli 1960
- [15] E. Morpurgo, E. von Glasersfeld, S. Perschke  
*Compte-rendu des travaux effectués par le Centro di Cibernetica e di Attività Linguistiche dans le cadre du contrat avec Euratom*  
Rapport CETIS n° 24, 1963
- [16] S. Ceccato *Il punto* (2 vol.).  
IPSOA 1980
- [17] S. Ceccato *Il linguaggio con la tabella di Ceccatieff*  
Hermann 1950
- [18] J.F. Sowa *Conceptual structures*  
Addison-Wesley 1984
- [19] J. Pollack & D. Waltz *Interpretation of natural language*  
Byte, feb. 86, p. 189
- [20] J. H. Conway *On games and numbers*  
Addison-Wesley, 1976
- [21] D. Knuth *Surreal numbers*  
Addison-Wesley, 1974
- [22] Y. Wilks *Grammar, Meaning and the Machine Analysis of Language*  
Routledge & Kegan Paul, 1972
- [23] Y. Wilks *A Preferential, Pattern-seeking Semantics for Natural Language Inference*  
Artificial Intelligence, **6** (1975), p. 53
- [24] Z. Harris *A grammar of English on mathematical Principles*  
John Wiley, 1982
- [25] Anne Daladier *Problèmes de représentation d'une langue naturelle en  $\lambda$ -calcul. Approches formelles de la sémantique naturelle.*  
Université Paul Sabatier, Toulouse, 1982
- [26] N. J. de Bruijn *Automath, a language for mathematics*  
Les Presses de l'Université de Montréal, 1973
- [27] Th. Coquand & G. Huet *Constructions : a high order proof system for mechanizing mathematics*



INRIA, rapport de recherches n° 401, 1985

- [28] J. Sammett *Programming languages*  
Prentice-Hall, 1969
- [29] M. Arbib *The Metaphorical Brain*  
John Wiley, 1972
- [30] P. Braffort *Le projet LAPAL*  
Action Poétique n° 95 (1984), p.51