AMERINDIA n° 11, 1986

TAXINOMIE, SYNTAXE ET ECONOMIE DÉS NUMÉRATIONS PARLÉES[1]

André CAUTY

UA 1026 (Paris) / NSF (L'Herbergement)

Tout vaudra mieux que la séparation essentielle
de syntaxe et sémantique, qui ramène, inéluctablement,
à une syntaxe avec un lexique muni de règles projectives.

A. CULIOLI (1968)

Deuxième partie : SYNTAXE

0. - Introduction

0.1. - L'objet de cette deuxième partie est de proposer une analyse syntaxique des numérations parlées. Ce travail est facilité par le fait que l'ensemble des paradigmes et celui des expressions numériques peuvent être déterminés de manière quasi exhaustive. Pourtant, cette analyse est rarement réalisée, même pour des numérations de langue de grande diffusion[2] ; et nous ne

---

[1] Je remercie mes maîtres A. Lentin et B. Pottier pour leur aide amicale et leurs remarques toujours riches d'enseignement. Je remercie également B.J. Hoff, M. Launey, C. Ntsadi, M-F Patte et A. Yaranga pour les communications qu'ils ont bien voulu me faire et M. Besada pour sa précieuse collaboration.
Cette recherche est financée depuis juillet 1984 par l'ATP 955318 du CNRS. La première partie est parue dans *Amerindia* n° 9.

[2] Un recueil de grammaire a toutefois été édité en 1968 par Brandt CORSTIUS sous le titre *grammars for number names,* Dordrecht: Reidel Publishing Company. Toutes les contributions formant ce recueil s'inscrivent dans le cadre des grammaires génératives.

connaissons, a priori, ni les catégories, ni les règles susceptibles de rendre compte des observables, par exemple des six expressions que l'on obtient en français par permutation des constituants mille, vingt et quatre :

mille vingt-quatre //1000/+/20/+/4// soit '1 024'

mille quatre-vingts //1000/+/4/x/20// soit '1 080'

vingt mille quatre //20/x/100/+/4// soit '20 004'

vingt-quatre mille //20/+/4/x/1000// soit '24 000'

quatre mille vingt //4/x/1000/+/20// soit '4 020'

quatre-vingt mille //4/x/20/x/1000// soit '80 000'

0.2. - Les mathématiciens versent au dossier les notions de *nombre,* d'*ordre,* d'*opération* et, plus spécifiquement, de *base* et de *numération de position.* Ces données rendent compte essentiellement des numérations écrites de position. Elles sont loin d'être négligeables. Cependant, elles ne permettent pas de répondre à plusieurs questions posées par l'étude des numérations, notamment des numérations parlées. Pourquoi, par exemple, **quatre-vingts** est-elle une expression bien formée (**quinze-vingts** ne l'est plus) bien que contredisant la propriété de commutativité de la multiplication (a x b = b x a) ? Pourquoi la lecture des nombres écrits dans le système décimal se fait-elle après un découpage préalable en tranches de deux, trois ou six chiffres ? Pourquoi a-t-on **soixante-dix** et non pas * *dix-soixante* ? ou *septante ?*

0.3. Les mathématiques, d'autre part, ne rendent pas compte des expressions numériques formées sur la base de principes non-arithmétiques. Par exemple, de l'usage d'un augmentatif (**mille**/**million**), de l'emploi d'oppositions grammaticales (comme en caribe : **oxto:ne** 'on one side' / **o:patoro** 'on both sides' cf. I.3.11.1.4.) ou d'oppositions pragmatiques (comme en panaré : **eña** 'main' / **pata** 'pied' -cf. I.3.10.5.), ou encore de l'usage des duels et des pluriels comme dans la numération hébraïque (cf. GUITEL, 1975, p.234 ou DHOMBRES, 1978, p. 114).

0.4. L'analyse linguistique montre rapidement que la formation des expressions numériques parlées repose sur un principe majeur lié à la contrainte de linéarisation : le principe de *concaténation.* Ainsi, par exemple, la mise en chaîne d'un préfixe (et/ou d'un suffixe) et d'une base lexicale fournit des dérivés (**mille**/**million**/**trillion**), celle de certains lexèmes (avec ou sans terme de relation) fournit des composés plus ou moins intégrés (**dix-sept**, **quatre-vingts**, **mille milliards** de **mille**).

0.5. Précisons cependant que la généralité et la simplicité du principe de concaténation recouvrent une grande diversité et une grande complexité de situations. Parfois on observe la chute d'un segment (**tri- + million -> trillion**) ; ou, au contraire, la présence d'un élément de redondance (//**vingt** / **ø** / **deux** // à côté de // **vingt** / **et** / **un** // ou encore // **ø** / **mille** // à côté de // **un** / **million** //). Il convient donc d'être attentif et de repérer ces faits qui reflètent à la fois le degré de liberté et le poids des contraintes qui règlent plus ou moins systématiquement la mise en signes. En un mot, de mettre suffisamment en évidence le fait le plus important, à savoir que les expressions numériques formées sont organisées, *structurées,* selon les règles d'une grammaire spécifique. Un constituant particulier (**dix**, par exemple) ne peut pas être suivi (ou précédé) de n'importe quel autre (***dix trois**, **dix-sept**) ; et, dans une chaîne attestée, un constituant particulier ne peut pas être remplacé par n'importe quel autre. Par exemple, dans l'expression **mille vingt-quatre**, on peut, sans devoir changer la structure additive de l'expression, substituer **trente**, **quarante**, **cinquante**, **soixante** ou **cent** au constituant **vingt**, tandis que la substitution de **million**, **billion**..., **sextillion** confère à l'expression une structure multiplicativo-additive, et que d'autres substitutions conduisent à des expressions non-attestées (***mille sept quatre**, ? **mille et quatre**). Ces contraintes n'existent pas dans une numération de position (c'est-à-dire en mathématiques), puisque tout chiffre y est concaténable et substituable à n'importe quel autre.

Les remarques précédentes suffisent à établir qu'une analyse linguistique doit nécessairement être mise en oeuvre dans l'étude des systèmes de numération.

0.6. L'analyse commence par le relevé du vocabulaire terminal de la numération envisagée. Nous supposons évidemment que des études préalables ont permis d'isoler les différents constituants effectivement engagés dans la formation des expressions numériques attestées. Nous avons vu (cf. I.1.6.) que ces analyses préalables peuvent être réalisées selon des critères plus ou moins exigeants et qu'il convient de distinguer différents types d'expressions composées selon que les procédés de formation relèvent de principes langagiers ou mathématiques. Par exemple, les expressions (manifestement composées) de cinq et dix en andoke sont inanalysables d'un point de vue strictement arithmétique **5** = **Ʌi-hako-domi-pãã** ('ce-côté-main-quantité') et **10** = **ka?-hako-domi-pãã** ('nous-côté-main-quantité') puisqu'elles font intervenir les oppositions *ce/autre, -ce/nous* et *main/pied* qui règlent la mise en signes des multiples de cinq.

Du point de vue de la numérosité, la signification de ces expressions andoke doit être saisie plus ou moins globalement (au même titre que celle de

deux, **treize**, **trente** ou **trillion** en français), mais certainement pas à partir des opérations arithmétiques élémentaires (qu'il convient, au contraire, de postuler pour des expressions comme **dix-sept** ou **quatre-vingts**).

Nous avons vu dans la première partie qu'il convient encore de distinguer les règles d'assemblage des numérations arithmétiques et celles des numérations ordinales. D'où les définitions suivantes :

0.7. Définition 1. Nous appelons *nombrant* tout signe à valeur numérique cardinale précise et qui n'est pas un composé arithmétique manifeste. Pour l'arithméticien, c'est un terme "constant" dont le signifiant est inanalysable ; ce terme est employé seul ou en tant qu'argument d'une opération arithmétique élémentaire (addition, multiplication, élévation à une puissance). Les nombrants constituent la partie principale de l'alphabet terminal d'un système de numération arithmétique. Du point de vue linguistique, les nombrants sont des éléments de première articulation ou des unités de niveau encore supérieur dont l'analyse relève alors de la morphologie.

0.8. Définition 2. De même, nous appellerons *repérant* tout signe à valeur ordinale précise et qui n'est pas analysable du point de vue arithmétique en termes de relation d'ordre ou de repérage. Comme le nombrant, un repérant ne peut être employé que seul ou en tant que terme d'une relation d'ordre. Toute comptine étant un instrument de comptage (cf. I.3-3 à 3-6), on peut admettre qu'un repérant est aussi un nombrant (et réciproquement) à condition de ne pas oublier que cette équivalence est abusive tant que l'on n'a pas établi que le nombre est effectivement saisi dans ses deux dimensions ordinale et cardinale.

0.9. Définition 3. Enfin, dans le cas d'une numération iconique, un élément du vocabulaire terminal sera dit *quantifiant ou qualifiant.* Par définition des numérations iconiques, il n'y a pas de composés syntaxiques dans ce type de numération. Il convient cependant de distinguer deux cas selon que les nombres sont mis en signes isolément et un par un, ou, au contraire, ensemble et globalement. Dans le premier cas, la comptine apparaît comme une liste hétéroclite de termes sans grand rapport les uns avec les autres ; dans le second, la liste des noms de nombres est transcendée par un procédé particulier susceptible de motiver une telle mise en signes globale ; il s'agit, par exemple, d'une technique gestuelle, ou du transfert au domaine numérique de l'organisation conceptuelle culturellement bien établie d'un autre domaine d'expérience.

0.10. Dans les définitions précédentes, la préoccupation essentielle est la détermination d'un niveau de référence pour l'observation et l'analyse des signes numériques; le but recherché étant que le niveau fondamental retenu permette la distinction et l'articulation des différents procédés de mise en signes attestés. Il serait incorrect, en effet, de traiter de la même manière les résultats de l'analyse de **million** (emprunt à l'italien où ce terme est un augmentatif), celle des dérivés non-transparents (**trente**, **quarante**, etc.), et celle des composés transparents (**dix-sept**, **quatre-vingts**, etc.) ou plus ou moins transparents (**billion**, **trillion**, etc.) selon les locuteurs ou les observateurs.

0.11. Le niveau fondamental de référence est déterminé à partir de l'évaluation d'une discontinuité dans le degré de liberté dont dispose le locuteur pour réaliser des chaînes d'unités. On sait, en effet, (grâce par exemple aux travaux de JAKOBSON, 1956:60), que le degré de liberté de combinaison des unités linguistiques augmente en même temps que leur niveau de complexité : quasiment nul pour la combinaison des traits distinctifs d'un phonème, et limité pour celle des constituants d'un morphème, il devient appréciable au niveau du fonctème et tout à fait important au-delà de ce niveau.

0.12. Dans le cas de l'étude de la mise en signes d'une notion comme celle de nombre, une telle discontinuité peut être déterminée sur des bases relativement objectives. Les nombres, en effet, à l'exception des tout premiers *(un, deux,* peut-être *trois),* doivent nécessairement être construits (cf. I.1.4.), en d'autres termes composés, et leur mise en signes doit évidemment permettre une communication non-ambiguë. Or, plus on s'élève dans l'échelle des nombres et plus augmentent les possibilités de (dé-) composition[3] et, corrélativement, les difficultés de choix et les risques d'ambiguïté en situation de reconnaissance.

Il est donc important de pouvoir maîtriser la multiplicité des formes que la non-unicité des décompositions engendre inexorablement[4]. L'analyse critique des moyens de réduire cette multiplicité fait apparaître une différence qui nous paraît devoir être soulignée. C'est celle qui oppose les procédés *arbitraires* et les procédés *conventionnels*[5] mis en oeuvre, d'une part, pour la sélection des procédés de mise en signes (décider par exemple d'une stratégie additive) et, d'autre part, pour le choix d'une forme canonique (décider par exemple d'utiliser l'appui **cinq** pour la formation des composés additifs compris entre cinq et dix).

0.13. Dans ce travail, le niveau des nombrants (respectivement, celui des repérants) est le niveau fondamental retenu pour l'analyse des systèmes de

[3] Par exemple, pour une simple décomposition en somme d'entiers on aura : 3 = 1 + 1 + 1 = 1 + 2 = 2 + 1 ; 4 = 1 + 1 + 1 + 1 = 1+ 1 + 2 = 1 + 3= 2 + 2 =...; 5 = 1+ 1 + 1 + 1 + 1 = 1 + 1 + 1 + 2 = 1 +1 + 3 = 1 + 4 =1 + 2 + 2 = 2 + 3 =... ; etc.

[4] On sait toute l'importance des théorèmes d'unicité en mathématiques.

[5] Le prototype de cette opposition pourrait être celle qui distingue le signe linguistique et le symbole mathématique.

numération arithmétique (resp. ordinale) En effet, dans le cas où il combine des nombrants, le locuteur peut raisonnablement être supposé conscient des mécanismes énonciatifs qu'il met en oeuvre pour produire des composés (**vingt et un**, **deux cents**) ; dans le cas de l'énonciation d'un nombrant, il est plus raisonnable, au contraire, de supposer qu'il n'a pas conscience que cet élément est en fait un composé plus ou moins laborieusement analysable (**trente**, **trillion**, par exemple).

Par abus de langage, nous appellerons *syntaxe* l'étude de la formation des expressions numériques composées dans une rationalité arithmétique ou ordinale par concaténation de nombrants ou de repérants; et nous appellerons *morphologie* l'analyse des nombrants, des repérants ou des quantifiants en constituants de niveau de complexité inférieur.

0.14. L'opposition *morphologie/syntaxe* ainsi introduite dépend du niveau fondamental retenu. Elle peut en conséquence être adaptée aux besoins de chaque étude particulière et évite de figer des oppositions dont l'usage n'est pas toujours absolument pertinent, comme les oppositions lexique/syntaxe ou syntaxe/ sémantique/pragmatique.

0.15. L'opposition *morphologie/syntaxe* permet de distinguer, dans certaines numérations, des composés qui dépendent d'une vision savante (consciente, volontaire, systématique) et des composés qui dépendent d'une vision naturelle ou langagière (moins consciente, moins volontaire, moins systématique, mais aussi souvent plus expressive). Du point de vue méthodologique, notons que les premiers relèvent d'une analyse du système et que l'interprétation des résultats n'est généralement possible que dans le cadre de la langue spécifique de l'arithmétique. Par exemple, il faut être mathématicien pour saisir la valeur numérique exacte d'un terme comme **trillion** et reconnaître dans *million* la base de l'opération d'élévation à la puissance trois. L'analyse des seconds relève par contre le plus souvent de techniques où interviennent l'étymologie, la comparaison avec des langues apparentées, la connaissance des valeurs culturelles et de l'état des techniques.

0.16. Dans cette deuxième partie, nous étudions la syntaxe des numérations parlées de type arithmétique, et nous posons les premiers jalons pour l'étude de la syntaxe des numérations ordinales. Un exemple des questions posées par l'analyse morphologique sera présenté en annexe.

1. - Syntaxe des numérations arithmétiques

1.0. Introduction

L'étude syntaxique que nous présentons dans cette deuxième partie développe assez largement l'exemple de la numération française. Le lecteur d'une revue consacrée aux langues amérindiennes est en droit de s'étonner d'une telle décision.

1.0.1. L'un des objectifs de nos recherches étant de dégager des concepts théoriques et de développer des outils d'analyse propres à l'étude des systèmes de numération, il nous a semblé prudent de commencer ce travail en choisissant la numération d'une langue bien connue. Le français est notre langue maternelle, mais c'est aussi une langue de grande diffusion qui permet d'exprimer de grands nombres et qui conserve des traces de plusieurs solutions à des problèmes de mise en signes que nous souhaitons traiter. Cette seule langue fournit donc autant d'exemples pour illustrer, de manière facilement diffusable, notre méthode et ses principaux résultats. S'agissant par ailleurs d'une numération enracinée dans des traditions culturelles bien connues, nous pouvions espérer, d'une part, que le système se soit spécialisé au point que certains de ses éléments possèdent au moins quelques caractéristiques du symbole mathématique[6], et, d'autre part, qu'il soit parfois possible d'étayer les interprétations proposées par des considérations non-linguistiques (historiques par exemple).

1.0.2. Le choix de la numération espagnole était tout aussi intéressant, voire davantage puisqu'une partie des communautés amérindiennes subissent encore l'influence de cette langue. Nous avons cependant opté pour la numération française : d'une part, en effet, les résultats obtenus sont pratiquement tous transposables immédiatement à la numération espagnole mais, d'autre part, la numération française présente une particularité intéressante, -qu'elle partage avec certaines numérations amérindiennes-, le fait qu'elle conserve des traces d'une numération à caractère vigésimal, c'est-à-dire des traces d'une situation de concurrence culturelle entre une tradition numérique vigésimale (probablement d'origine celte) et la tradition numérique décimale transmise par le conquérant romain.

---

[6] *"When a word has once been taken up to serve as a numeral, and is thenceforth wanted as a mere symbol, it becomes the interest of the language to allow it to break down into an apparent nonsense-word, from which all traces of original etymology have disappeared""* (TYLOR, 1891 cité par TAYLOR, D.).

1.1. Description de la numération parlée française.

1.1.0. Après la détermination du vocabulaire terminal, nous construisons un organigramme susceptible de modéliser la lecture des nombres supposés écrits dans le système décimal habituel (par ex. 1983). Par cet artifice, nous évitons la discussion des questions non-linguistiques de la conceptualisation du nombre, et ceci, afin de nous limiter à l'étude de la mise en signes du nombre déjà conceptualisé. Dans ces conditions, l'organigramme proposé apparaît comme un modèle métalinguistique de la mise en signes des conceptualisations numériques du sujet énonciateur. Implanté sur ordinateur, il transformerait tout nombre écrit en chiffres en une suite de mots constituant son expression en français standard, parlée ou écrite (nous négligeons ici les questions d'intonation et d'orthographe). Il s'agit donc d'un modèle de production et non d'un modèle de reconnaissance (cette question sera traitée plus loin).

1.1.1. Pour la clarté de l'exposé[7] :

a) nous n'envisageons que les expressions du système légal français (dit système de *l'échelle longue,* BOUVIER, 1979:200) ce système est caractérisé par la convention terminologique fixant le nom des puissances de dix de la manière suivante (noter particulièrement l'absence du terme **milliard**) 10^6 = **million**, 10^{12} = **billion**, 10^{18} = **trillion**, 10^{24} = **quadrillion**, 10^{30} = **quintillion** et 10^{36} = **sextillion**. Ce qui exclut aussi les formes d'énonciation qui reposent sur un découpage autre que celui en tranches de six chiffres (en usage par exemple pour la lecture des dates ou des numéros de téléphone : 1941 = **dix-neuf cent quarante et un**, 700 51 53 = **sept cent**, **cinquante et un**, **cinquante trois**).

b) Nous présentons d'abord un sous-algorithme limité à la lecture des nombres de trois chiffres (inférieurs à mille), en supposant ici que **quatre-vingts** est un nombre non-composé. Cette attitude est justifiée dans la mesure où les composés du même type (**onze-vingts**, **douze-vingts**, etc.) ne sont plus attestés en français.

1.1.2. Le vocabulaire terminal de la numération orale française (échelle longue) est formé des termes suivants : V_T = { un, deux, trois, quatre, cinq, six, sept, huit, neuf, dix, onze, douze, treize, quatorze, quinze, seize, vingt, trente, quarante, cinquante, soixante, quatre-vingts, cent, mille, million, billion, trillion, quadrillion, quintillion, sextillion, et } [8]. Il comporte trente termes : vingt-neuf nombrants et le coordonnant **et**. Notons que ce vocabulaire n'est en rien comparable à l'ensemble des

---

7 Cela n'enlève rien à la généralité du problème.

8 Ce vocabulaire est établi à partir de l'article *décimal* du *Dictionnaire des mathématiques* (BOUVIER, 1979:200). Pour rendre compte de certaines expressions non "légales", il conviendrait d'ajouter le terme *milliard,* la préposition *de* (mille milliards *de* mille...), les termes *septante, octante* ou *huitante, nonante,...*

dix chiffres de la numération décimale écrite. La numération parlée et la numération écrite ne sont pas des systèmes structurés de la même manière.

1.1.3. *Remarques*

a) L'exemple des numérations ordinales nous conduit à souligner que ce vocabulaire ne peut pas être assimilé à une *simple liste* d'items. Des procédés morphologiques de dérivation et de composition (repérages par tout locuteur compétent) le structurent[9] en classes plus ou moins systématiques d'éléments qui se correspondent et renvoient à une organisation profondément marquée par le nombre six, (la moitié des douze heures du jour, le nombre de faces d'un dé). Nous ignorons les raisons de la singularité du nombre six dans l'organisation sous-jacente à la numération française.

	dix	vingt	cent	mille
un	onze			million
deux	douze			billion
trois	treize	trente		trillion
quatre	quatorze	quarante		quadrillion
cinq	quinze	cinquante		quintillion
six	seize	soixante		sextillion[10]
sept		(septante)		septillion
huit		(huitante)		octillion
neuf		(nonante)		nonillion
				etc.

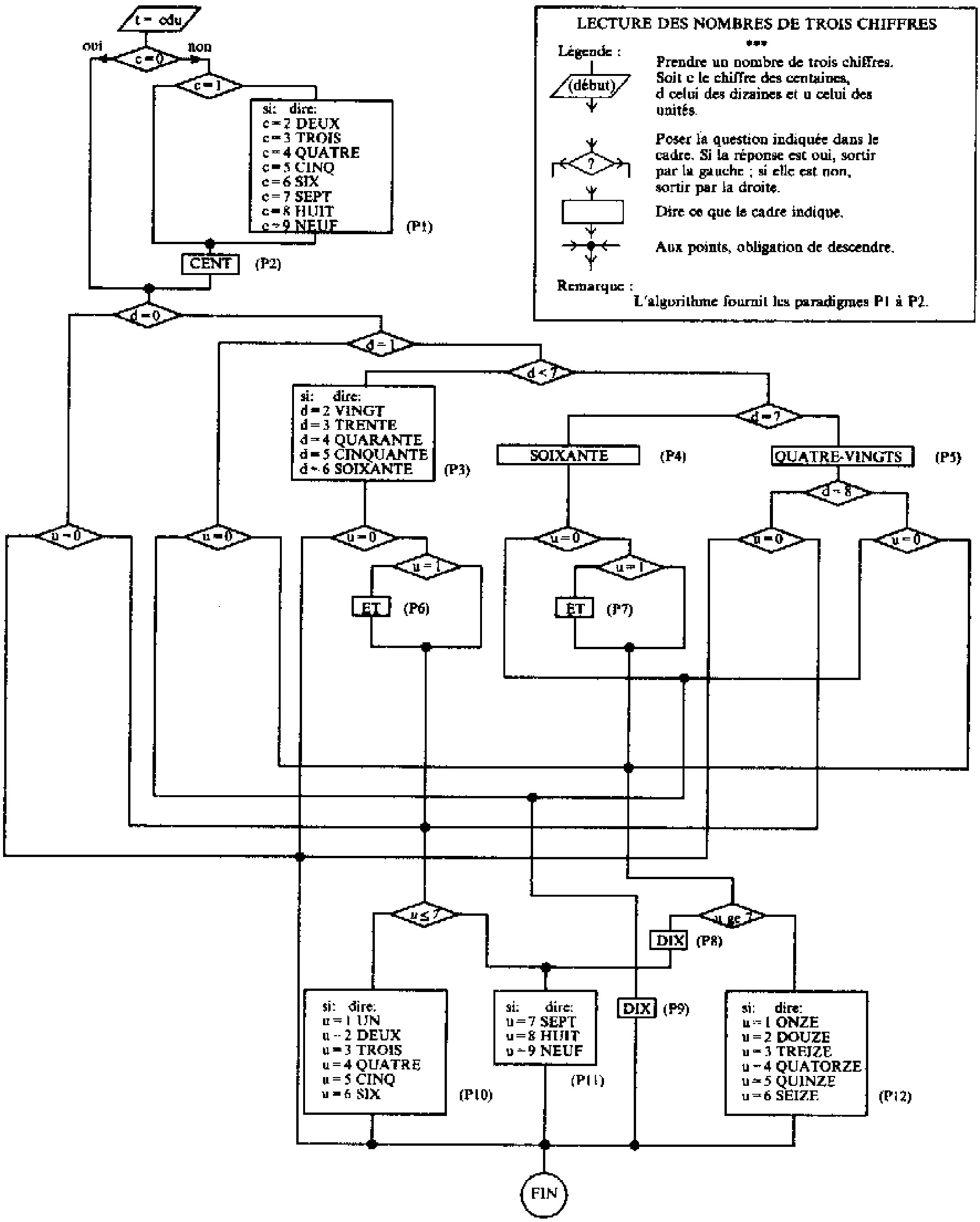
b) Le vocabulaire terminal permet de former l'expression de tous les entiers compris entre 1 et 999999 999999 999999 999999 999999 999999 999999. Ce nombre limite s'écrit avec 42 occurrences du chiffre 9 et s'énonce à l'aide de 97 occurrences de nombrants (28 occurrences de **neuf**, 14 occurrences de **cent**, **quatre**, **vingt**, **dix**, 7 occurrences de **mille** et 1 occurrence de **million**, **billion trillion quadrillion**, **quintillion**, **sextillion**). Notons que, comme tout système de numération non-positionnel, la numération française (échelle longue) admet une limite théorique qui ne peut être dépassée qu'à condition de créer de nouveaux termes ou d'introduire de nouveaux procédés de composition[11].

9 Le repérage et l'analyse de ces structures sont à la base de *ce qu'on appelle désormais les ethnosciences* (QUEIXALOS, 1982: 85), dont l'objet est de retrouver à partir de traces "morphologiques" quelque chose des visions du monde d'une communauté.

10 limite du système dit de l'échelle longue.

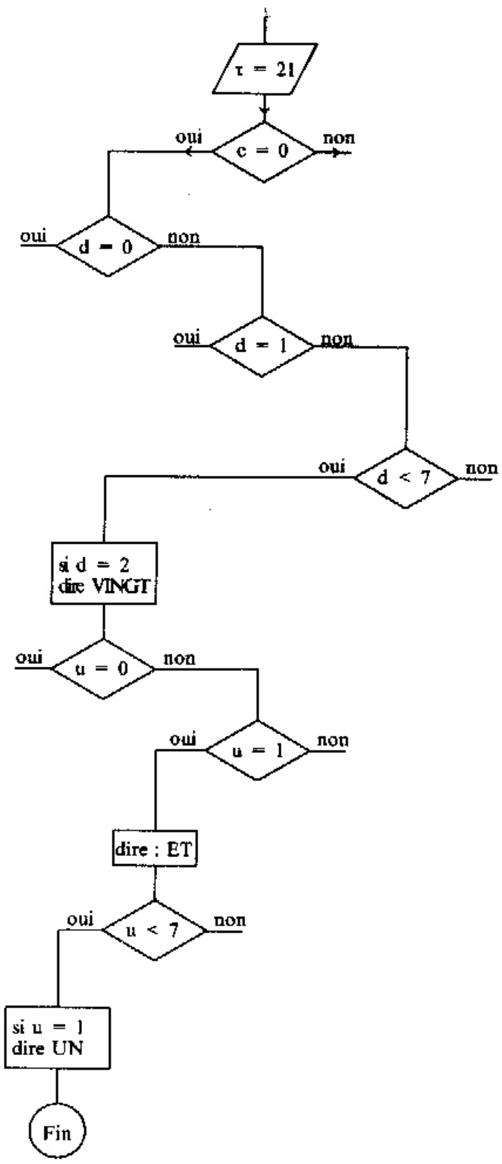
11 Les systèmes positionnels reposant sur des procédés récursifs n'ont pas de limite théorique, mais une limite *technique* liée à l'encombrement des expressions de plus en plus longues à mesure que l'on s'élève dans l'échelle des nombres.

1.1.4. Voici l'organigramme modélisant la mise en signes des nombres inférieurs à mille.



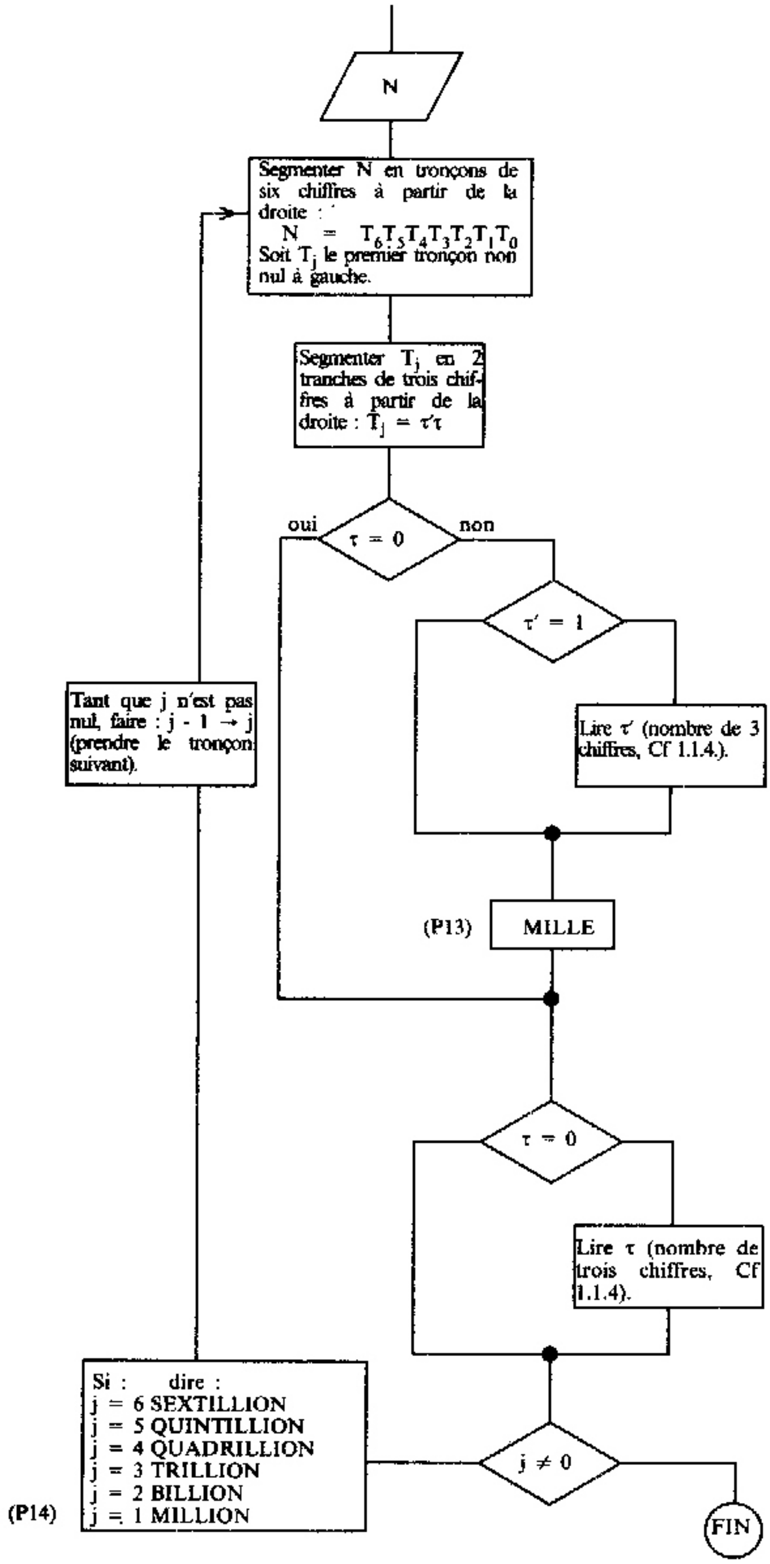
1.1.5. Pour utiliser cet organigramme, il faut partir de la donnée d'un nombre τ supposé écrit avec trois chiffres au plus dans le système décimal, soit $\tau = c\ d\ u$ (c représente le chiffre des centaines, d celui des dizaines et u celui des unités). Il suffit alors de suivre les flèches. Chaque fois que l'on rencontre une case-losange, il faut répondre à la question posée dans cette case et poursuivre le parcours en sortant par la gauche (resp. droite) lorsque la réponse est affirmative (resp. négative). Les cases rectangulaires contiennent en caractères majuscules les "mots" qu'il convient d'énoncer dans les conditions indiquées dans la case rectangulaire.

Exemple : lecture du nombre $\tau = 21$.



A la fin du parcours, on a énoncé : **vingt et un**.

1.1.6. Voici l'organigramme modélisant la mise en signes des nombres de l'échelle longue :



1.1.7. Les organigrammes précédents permettent de mettre en évidence les classes paradigmatiques et les contraintes d'enchaînement de la numération française. Voici par exemple quelques propriétés relevées pour le nombrant **cent** à partir du premier organigramme

a) son paradigme est un singleton (il ne contient que le nombrant **cent**).

b) **cent** peut être précédé de **ø**, **deux**, **trois**, **quatre**, **cinq**, **six**, **sept**, **huit**, **neuf**,

c) **cent** peut être suivi de **ø**, de tout nombrant autre que lui-même ou de toute expression désignant un nombre inférieur à cent.

1.1.8. Dans les composés // **x** / **cent** // et // **cent** / **x** //, les tactèmes d'ordre marquent respectivement la mise en signes d'une multiplication et d'une addition ; lorsque **x** est le morphème vide, celui-ci marque respectivement l'élément neutre de la multiplication (un) et l'élément neutre de l'addition (zéro).

1.1.9. Les organigrammes précédents fournissent les paradigmes et les contraintes suivants[12] :

$P_1 = \{2, 3, 4, 5, 6, 7, 8, 9,\}$

Les éléments de P_1 doivent suivre le morphème /ø/ et être suivis immédiatement du nombrant **cent**. La concaténation de ces éléments a la valeur arithmétique P_1 x 100.

$P_2= \{100\}$.

L'élément de P_2 doit suivre le morphème /ø/ ou un élément de P_1 et précéder immédiatement le morphème /ø/ ou un élément de la réunion

$P_9 \cup P_{10} \cup P_{11} \cup P_{12} \cup P_3 \cup P_4 \cup P_5$.

La chaîne ø - 100 - x a pour valeur 1 x 100 + x ; et la chaîne P_1 - (100) - x a pour valeur P_1 x 100 + x. Lorsque x représente le morphème /ø/ il a la valeur numérique zéro.

$P_3 = \{20, 30, 40, 50, 60\}$

Les éléments de P_3 doivent précéder /ø/ ou un élément de la réunion $P_6 \cup P_{10} \cup P_{11}$, et suivre le morphème /ø/ ou l'élément **cent** de P_2. La concaténation de ces éléments a valeur additive.

$P_4 = \{60\}$

Le nombrant **soixante** suit le morphème /ø/ ou l'élément **cent** de P_2 il précède /ø/ ou un élément de $P_7 \cup P_8 \cup P_9 \cup P_{12}$. La concaténation avaleur additive.

[12] Pour des raisons de place, nous transcrivons dans ce paragraphe les nombrants en chiffres.

$P_5 = \{80\}$

L'élément de P_5 suit immédiatement le morphème /ø/ ou l'unique élément de P_2 ; il précède /ø/ ou un élément de $P_8 \cup P_9 \cup P_{10} \cup P_{11} \cup P_{12}$. La concaténation a valeur additive.

$P_6 = \{et\}$

Cet élément précède le nombrant **un** et suit un élément de P_3 ; c'est un marqueur d'addition.

$P_7 = \{et\}$

Cet élément précède le nombrant **onze** et suit l'unique élément **soixante** de P_4 ; il marque une addition.

$P_8 = \{10\}$

Cet élément précède un élément de P_{11} et suit /ø/, **cent**, l'unique élément **soixante** de P_4 ou l'unique élément **quatre-vingts** de P_5 ; la concaténation a valeur additive.

$P_9 = \{10\}$

Cet élément précède /ø/ et suit /ø/, **cent**, l'unique élément **soixante** de P_4 ou l'unique élément **quatre-vingts** de P_5 ; la concaténation a valeur additive.

$P_{10} = \{1, 2, 3, 4, 5, 6\}$

Les éléments de P_{10} précèdent /ø/ et suivent /ø/ ou un élément de la réunion $P_2 \cup P_3 \cup P_5 \cup P_6$; la concaténation a valeur d'addition.

$P_1 = \{7, 8, 9\})$

Les éléments de ce paradigme précèdent /ø/ et suivent un élément de $P_2 \cup P_3 \cup P_5 \cup P_8$ ou /ø/ ; la concaténation est additive.

$P_{12} = \{11, 12, 13, 14, 15, 16\}$

Les éléments de P_{12} précèdent le morphème /ø/ et suivent /ø/ ou un élément de $P_2 \cup P_4 \cup P_5 \cup P7$; la concaténation a valeur d'addition.

1.1.10. A ces douze ensembles, il faut ajouter $P_{13} = \{1000\}$ et $P_{14} = \{10^6, 10^{12}, 10^{18}, 10^{24}, 10^{30}, 10^{36}\}$ obtenus à partir de l'algorithme général (lecture des nombres de l'échelle longue).

1.2. Éléments pour une théorie des numérations arithmétiques

1.2.1. Nous proposons les définitions suivantes[13] :

Définition 4. *Nombre d'appui (additif)* : tout nombrant *a* du vocabulaire terminal (i.e. arithmétiquement non composé), concaténable à d'autres nombrants ou à des expressions composées exprimant des nombres plus petits que *a* pour former, au moins en partie, la suite des successeurs de *a* : *a + 1, a +2,...* jusqu'au double de *a* ou jusqu'au prédécesseur du double de *a.* En français, par exemple, **dix** est un nombre d'appui puisqu'il est arithmétiquement non-composé et qu'il est concaténable à **sept**, **huit**, **neuf** dans les composés additifs **dix-sept**, **dix-huit**, **dix-neuf**[14]. En andoke, *deux* est interprétable comme un appui additif intervenant dans la composition de *trois* et *quatre* (cf. I.3. II.2.3.).

Définition 5. *Nombre d'appui multiplicatif* : tout nombre d'appui *m* concaténable à quelques éléments au moins de la suite 1, 2, ..., *m* - 1, *m* pour former les expressions des multiples de *m*.

En français, par exemple, **cent** est un nombre d'appui multiplicatif puisqu'il est un appui additif (**cent un**, **cent deux**,..., **cent quatre-vingt-dix-neuf**) et qu'il entre dans les composés multiplicatifs (**deux cents**, **trois cents**, ..., **neuf cents**).

Définition 6. *Nombre d'appui systématique* : tout nombre d'appui multiplicatif *s* permettant la formation de la suite complète des multiples de *s* jusqu'au carré de *s*. C'est le cas de **mille** en français.

Définition 7. *Nombre d'appui parabasique* : tout nombrant *p* concaténable à *deux, trois, ...* pour former au moins quelques puissances de la suite *p2, p3,* etc...

On peut admettre que **million** est un nombre d'appui parabasique en français puisqu'il est concaténable à **deux**, **trois**,... pour former les puissances successives **billion**, **trillion**, etc.

Nous avons vu qu'en ngbaka (I.4.4.2.) le nombre *dix* est un appui parabasique 10 = **djomba**, 2 = **bissi** et 100 = **djomba-té-bissi**.

A notre connaissance, aucune numération amérindienne ne comporte d'appui parabasique.

---

[13] Dans la terminologie de J.P. Desclés, ces catégories seraient dites grammaticales (nous préférons le terme syntaxiques) car définies et construites indépendamment de la langue naturelle étudiée. Elles s'opposent aux catégories linguistiques obtenues, dans le cadre d'une langue naturelle particulière, par des procédés taxinomiques d'analyse distributionnelle (CULIOLI et DESCLES, 1979: 75-80).

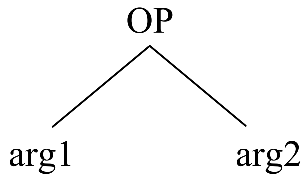
[14] Du point morphologique, dix est contaténable à un, deux,..., six pour donner onze, douze,..., seize.

Définition 8. *Base* : c'est un terme technique du vocabulaire arithmétique qui désigne le cardinal de la collection des chiffres d'une numération de position. (cf. note 16, partie I).

Il convient de réserver ce terme aux seules numérations de position, c'est-à-dire essentiellement à des numérations écrites.

Définition 9. La description de la mise en signes du nombre dans une numération arithmétique fait encore intervenir les notions *d'opérations* (addition, multiplication, puissance) et de *parenthésage* dont les définitions ne posent aucune difficulté.

1.2.2. Le schème conceptuel de base d'une composition arithmétique (addition, multiplication) est décrit par le schéma suivant qui représente une opération portant sur deux arguments : OP(arg1, arg2) :



Souvent[15], des contraintes de cooccurrence conduisent à distinguer l'un des arguments. Il s'agit alors d'un nombre d'appui et l'on pourrait noter : OP (arg, app).

1.2.3. La solution fondamentale du problème de la mise en signes de ce schème consiste, en français, à marquer l'opération par un tactème d'ordre[16] et à contraindre le choix de l'un des deux arguments dans une catégorie de nombres d'appui. C'est donc par la forme du signifié (Sy) que l'on pourra -en grammaire de reconnaissance- distinguer les structures additives **vingt-quatre** ou **cent deux** et les structures multiplicatives **quatre-vingts** ou **deux cents :** de même que *lit* est un substantif dans "il livre un lit" et un verbe dans "il lit un livre", **vingt** est un appui additif dans **vingt-quatre** et un appui multiplicatif dans **quatre-vingts**.

1.2.4. A notre connaissance, aucune numération arithmétique ne fait appel aux opérations réciproques (soustraction, division, extraction de racine), même dans les langues peu nombreuses où elles semblent attestées. Il serait maladroit d'introduire ces opérations du moins sans tenir compte des faits suivants a) l'extraction de racine n'est jamais attestée b) les rares "divisions" sont

[15] Il existe d'autres stratégies de mise en signes additives qui ne font intervenir aucun nombre.

[16] Il s'agit probablement d'une solution relativement moderne. En effet, l'addition -par exemple- semble avoir été marquée primitivement par le coordonnant *et* (que l'on trouve encore dans quelques expressions *vingt et un, trente et un,..., soixante et onze, mille et un)* puisque, en ancien français, *dis e set, dis e uit, dis e nuef* sont attestés. Quant à la multiplication, elle est encore marquée sporadiquement dans un pluriel audible grâce aux liaisons (quatre vingt-Z-ans). D'autres solutions sont possibles et attestées, en particulier dans d'autres langues.

toujours des "multiplications" dont le multiplicateur est une fraction simple (un demi, trois-quarts...) et le multiplicande un nombre d'appui multiplicatif systématique (la "base")[17] c'est d'ailleurs la généralisation 15 de Greenberg (GREENBERG, 1978: 261) ; c) quant à la "soustraction", la seule opération inverse moins exceptionnelle, elle est toujours soumise à de très fortes restrictions[18] (cf. les généralisations 11,12,13,14, et 15 de Greenberg) et renvoie ou peut au moins renvoyer soit à une composition ordinale soit à une technique de compte à rebours.

C'est le cas, par exemple, de 'dix-huit' **duo-dē-vīgintī** et de 'dix-neuf' **un-dē-vīgintī** en latin où ces nombres viennent rompre la série des composés additifs de onze à dix-sept : **un-decim**, **duo-decim**, etc. Nous pensons que *18 et 19 sont saisis dans une vision d'antériorité,* comme dans les phénomènes de "compte à rebours", par exemple le décompte des jours avant un événement jugé important (la "quille"; le "père cent", le jour "J moins 3", etc.). Le nombre formé soustractivement, en effet, est toujours très proche, voire le prédécesseur immédiat d'un nombre remarquable, prégnant dans l'esprit en raison d'une singularité quelconque (un nombre "rond", un nombre d'appui, la limite du système, etc.). Nous pensons *qu'il s'agit essentiellement de saisir un rapport de voisinage :* chrono-logiquement la vision est ordinale et non pas arithmétique comme tend à le prouver la présence du morphème **de** dans les exemples latins.

Voici un autre exemple rapporté par K. Zvelebil : *"Le système de numération dravidien était vraisemblablement de base 8* [...] *; on compte de la manière suivante : "un", "deux", "trois", " quatre", "cinq" , "six", "sept", "nombre", "beaucoup moins un", " beaucoup""* (in *Pour la science,* mai 1983, p. 22). Il ne fait aucun doute dans cet exemple que le successeur du composé apparemment soustractif **beaucoup moins un** soit remarquable : c'est la limite du système.

1.3. Application à la numération française

1.3.1. La structure fondamentale de l'expression du nombre en français fait intervenir trois éléments : un tactème d'ordre (pour marquer l'opération), un nombre d'appui (pour marquer l'argument distingué) et une place pouvant être saturée par **ø**, un nombrant ou une expression composée. Il est facile d'établir que :

---

[17] Pour un exemple non-amérindien, voir MAZAUDON, 1985 : 139 et sq. ou BRUNSCHVICG, 1981 : 21-22.

[18] Si le procédé soustractif était introduit, il faudrait souligner (et rendre compte de) son absence de productivité qui contraste avec la forte productivité du principe ordinal (attestée par exemple dans les numérations amérindiennes) et celle des opérations directes. La même remarque s'applique encore mieux à la division.

a) l'ordre croissant (resp. décroissant) est caractéristique de la multiplication (resp. addition).

b) la première (resp. seconde) occurrence dans une stucture additive (resp. multiplicative) doit être remplie par un nombre d'appui additif (resp. multiplicatif ou systématique), c'est l'occurrence de l'argument distingué.

c) la seconde (resp. première) occurrence doit être remplie par ø, un nombrant ou une expression composée désignant nécessairement un nombre inférieur (resp. supérieur) au nombre désigné par l'élément d'appui.

d) le morphème /ø/ s'interprète comme l'élément neutre de l'addition (resp. multiplication).

1.3.2. *Inventaire des structures de la numération française (échelle longue).*

Tout nombre *n* s'exprime par une chaîne d'au plus sept tronçons (un tronçon est l'équivalent parlé d'une suite de six chiffres en numération décimale écrite). La frontière de droite d'un tronçon est /ø/ ou l'un des termes en - *illion.* On peut écrire :

$$\mathbf{n} = \sum_{j=0}^{j=6} \mathbf{T}_j$$

Il existe trois types de tronçons que nous représentons par la formule :

$$T_j = \begin{pmatrix} \tau \\ \mu \\ \mu + \tau \end{pmatrix} \times \lambda_j$$

dans laquelle λ_0 = **ø**, λ_1, = **million**, λ_2 = **billion**, etc., λ_6 = **sextillion** ; et dans laquelle **μ** représente la séquence : **μ** = **τ'** x **mille** (ici **mille** est l'appui systématique **mille**, et **τ'** un nombre compris entre 2 et 999 ou le morphème /ø/), et **τ** un nombre compris entre 1 et 999 (**τ** représente une tranche de trois chiffres du système décimal).

Il existe trois types de tranches :

$$\tau = \begin{pmatrix} \alpha \\ \beta \\ \alpha + \beta \end{pmatrix}$$

formule dans laquelle **α** représente la séquence **u** x **cent**, **cent** étant un appui multiplicatif et **u** l'un des éléments du paradigme {**ø**, **un**, **deux**, **trois**, **quatre**, **cinq**, **six**, **sept**, **huit**, **neuf**}

L'élément ß prend de très nombreuses formes représentées par la formule suivante :

$$\beta = \left(\begin{array}{l} \mathbf{u} \\ \mathbf{v} \\ \mathbf{z} \\ \mathbf{dix} \pm \text{v} \\ \mathbf{d} \pm \begin{pmatrix} \text{u}' \\ \text{v} \end{pmatrix} \\ \mathbf{soixante} \pm \left(\begin{array}{l} \text{u}' \\ \text{v} \\ \text{z}' \\ \text{dix} \pm \text{v} \end{array} \right) \\ \text{quatre-vingt(s)} \pm \left(\begin{array}{l} \text{u} \\ \text{v} \\ \text{z} \\ \text{dix} \pm \text{v} \end{array} \right) \end{array} \right)$$

Dans cette formule, ± représente une concaténation facultative à signification additive ; **u** est un élément de {**un**, **deux**, **trois**, **quatre**, **cinq**, **six**} ; **v** est un élément de {**sept**, **huit**, **neuf**} ; **z** est un élément de {**onze**, **douze**, **treize**, **quatorze**, **quinze**, **seize**} **; u'** est un élément de {**et un**, **deux**, **trois**, **quatre**, **cinq**, **six**} **z'** est un élément de {**et onze**, **douze**, **treize**, **quatorze**, **quinze**, **seize**} ; **d** est un élément de {**vingt**, **trente**, **quarante**, **cinquante**} ; **dix**, **soixante**, **quatre-vingts** et les éléments **d** sont des appuis additifs. (Nous avons considéré ici quatre-vingts comme une forme arithmétiquement indécomposable).

Remarque. Dans l'hypothèse où l'on considère **quatre-vingts** comme une expression arithmétiquement composée, il convient de modifier la formule ß de la manière suivante :

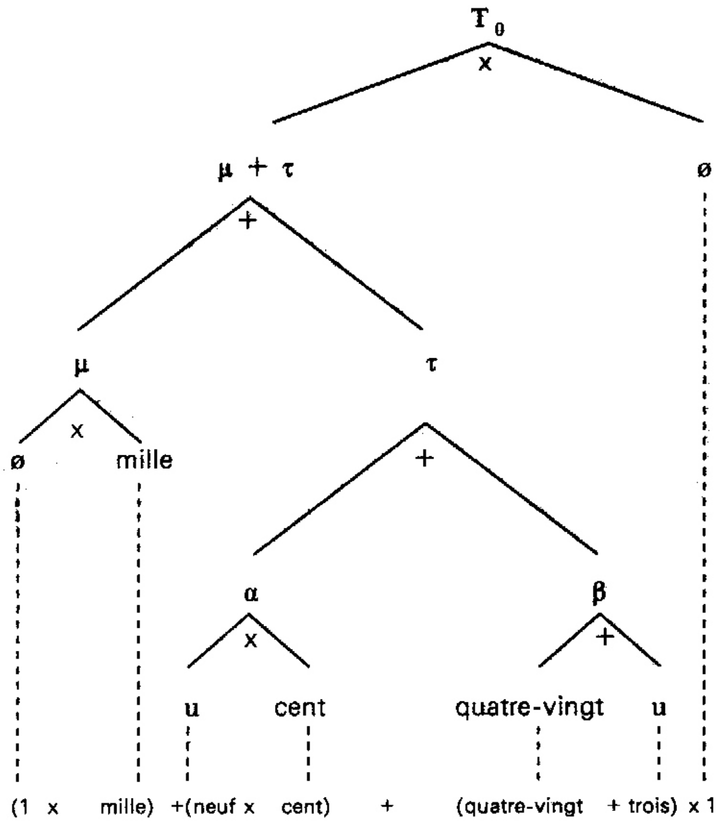
$$\beta = \left(\begin{array}{l} \text{u} \\ \text{v} \\ \text{z} \\ \text{dix} \pm \text{v} \\ \begin{pmatrix} \text{vingt} \\ \text{d} \end{pmatrix} \pm \begin{pmatrix} \text{u}' \\ \text{v} \end{pmatrix} \\ \text{soixante} \pm \left(\begin{array}{l} \text{u}' \\ \text{v} \\ \text{z}' \\ \text{dix} \pm \text{v} \end{array} \right) \\ \left(\begin{array}{l} \dots \\ \text{trois} \\ \text{quatre} \\ \dots \\ \text{six} \\ \text{sept} \\ \dots \\ \text{onze} \\ \dots \\ \text{quinze} \\ \dots \end{array} \right) \times \text{vingt(s)} \pm \left(\begin{array}{l} \text{u} \\ \text{v} \\ \text{z} \\ \text{dix} \pm \text{v} \end{array} \right) \end{array} \right)$$

Cette hypothèse fait évidemment apparaître le rôle singulier que joue **vingt** dans la numération orale française. Il appartient aux historiens de la langue de nous en donner les raisons ; cette singularité est peut-être une trace des origines celtiques (IFRAH, 1981: 51).

Le paradigme des coefficients de *vingt* contient des pointillés car nous n'y avons fait figurer que les multiplicateurs de vingt attestés dans l'ouvrage d'Ifrah (IFRAH, 1981:51).

1.3.3. *Exemple de dérivation.*

n = mille neuf cent quatre-vingt-trois.



1.3.4. *Grammaire de reconnaissance.*

Nous avons vu au cours de cette analyse de la numération française :

a) qu'un nombre d'appui additif ne peut être suivi que par une expression désignant un nombre inférieur à cet appui.

b) qu'un appui multiplicatif ou systématique ne peut être précédé que par une expression désignant un nombre inférieur à cet appui.

c) qu'il existe une hiérarchie des appuis multiplicatifs et systématiques permettant de rétablir le parenthésage (**sextillion**, **quintillion**, **quadrillion**, **trillion**, **billion**, **million**, **mille**, **cent**, **vingt**).

On peut, sur la base de ces données, énoncer les deux règles de reconnaissance suivantes :

R1) établir systématiquement la valeur additive ou multiplicative des concaténations selon les critères a) et b) ci-dessus ;

R2) a) reconnaître la hiérarchie des appuis, et b) parenthéser.

Il suffit d'ailleurs[19] de parenthéser les multiplications, et ceci seulement dans les chaînes qui comportent les deux opérations d'addition et de multiplication. Le problème du parenthésage se réduit donc à la détermination des frontières des facteurs (i.e. à celle de l'extension de l'expression des coefficients des appuis multiplicatifs et systématiques). La règle est simple ce coefficient s'étend vers la gauche jusqu'au premier élément d'appui supérieur rencontré ou jusqu'à la fin de la séquence s'il n'y a pas, vers la gauche, d'élément d'appui (multiplicatif ou systématique) supérieur.

Exemples

1) Soit la chaîne : **mille neuf cent quatre-vingt-trois**.

On a par application de R1) :

$$1000 + 9 \text{ x } 100 + 4 \text{ x } 20 + 3,$$

On a par R2 a) :

$$\underline{\underline{\underline{1000}}} + 9 \text{ x } \underline{\underline{100}} + 4 \text{ x } \underline{\underline{20}} + 3$$

On a par R2 b) :

$$(1 \text{ x } 1000) + (9 \text{ x } 100) + (4 \text{ x } 20) + 3.$$

2) Soit la chaîne : **quatre-vingt-deux millions trois cent quarante-huit mille sept cent quatre-vingt-dix-sept**.

On obtient :

$$\underline{\underline{4 \text{ x } 20} + 2 \text{ x } 10^6} + \underline{\underline{3 \text{ x } 100} + 40 + 8 \text{ x } 10^3} + \underline{7 \text{ x } 100} + \underline{4\text{x } 20} + 10 + 7$$

1.3.5. Remarque. Dans le premier exemple, nous avons restitué un morphème vide devant l'appui multiplicatif (systématique) *mille.* Cette position est justifiée par la comparaison avec des expressions comme, **deux mille**, **trois mille**, **neuf cent quatre-vingt-dix-neuf mille neuf cent quatre-vingt-trois**.

Cette remarque laisse pressentir que les numérations évoluent sous la pression du désir de nommer des nombres de plus en plus grands. Dans un état de développement donné, le système est en effet bloqué à une certaine limite[20].

---

[19] L'addition et la multiplication sont des opérations associatives : le parenthésage est donc inutile : 4 x (20 x 1000) = (4 x 20) x 1000.

[20] Le tout premier obstacle rencontré est généralement le nombre trois.

Supposons qu'il s'agisse de vingt. Cette limite doit d'abord être perçue comme telle et nommée de nombreuses numérations se terminent en tout cas par un terme qui renvoie sémantiquement à l'idée de beaucoup (chez les Aztèques, par exemple, 400 qui a pu être une de ces limites était représenté par le dessin schématique d'une natte de cheveux) ; dans beaucoup de numérations amérindiennes, c'est un terme signifiant "homme", atteint lorsque l'on a épuisé la liste des vingt doigts. Une fois la limite reconnue (**vingt**, par exemple), rien n'empêche de l'utiliser comme un terme d'appui additif, ce qui permet de former les expressions des nombres jusqu'au double de la limite initiale. On aurait par exemple **vingt-un**, **vingt-deux**, etc., le nouveau terme apparaissant en position d'argument distingué d'une opération d'addition. Dans une étape ultérieure, il pourrait être utilisé comme appui multiplicatif. Cette fois, **vingt** est utilisé en position d'argument distingué dans une opération de multiplication, comme dans la série :..., **quatre-vingts**, ..., **sept-vingts**, ..., **quinze-vingts**. Mais il convient de noter que l'habitude est prise de ne pas dire **un-vingt**, c'est-à-dire de ne pas mentionner le coefficient multiplicatif 1, et qu'il faudrait faire un retour sur l'acquis additif antérieur pour systématiser la composition multiplicative en introduisant le coefficient 1 ; on dirait alors : **un-vingt-trois**, **un-vingt-six**,...

Cette hypothèse paraît d'autant plus vraisemblable que ce retour systématisant est attesté, en français, à des stades de développement de la numération beaucoup plus avancés : le coefficient **1** est effectivement marqué devant les appuis multiplicatifs à partir de million. On dit en effet : **un million**, **un billion**, etc., mais pas ***un cent quarante**, ***un vingt et un**, ***un mille**.

1.4. *Syntaxe de la numération parlée française*[21]

1.4.1. *Le vocabulaire terminal* est comme en 1.1.2., on peut ajouter la forme une (féminin de un), et la préposition de.

1.4.2. *Le vocabulaire non-terminal*
$V_{NT} = \{\Sigma, \Sigma_0, A^0, A^1, A^2, A^3, A^{6N}, U_1, U_2, U_3, D_1, D_2, D_3, V, S, Q, fN,\}$
Note : Σ est un symbole initial.

[21] Les notations utilisées ici diffèrent de celles adoptées au paragraphe 1.3.2. pour l'inventaire des structures de la numération française. Les deux modèles, en effet, ne traduisent pas exactement les mêmes intentions méthodologiques.

1.4.3. *Les règles indépendantes du contexte*

(R1) $\Sigma \rightarrow \Sigma_0$ (fN) c'est-à-dire une expression numérique proprement dite suivie d'un fonctème nominal.

Note pour certains nombres, il convient de distinguer entre masculin et féminin (vingt-et-un hommes, vingt-et-une femmes).

(R 2) $$\Sigma_0 \rightarrow \begin{Bmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \\ A^{6N} \end{Bmatrix}$$

(R 3.1) $$A^0 \rightarrow \begin{Bmatrix} U_1 \\ U_2 \\ U_3 \end{Bmatrix}$$

(R 3.2) $$U_1 \rightarrow \begin{Bmatrix} \text{un} \\ U_2 \end{Bmatrix}$$

(R 3.3) $$U_2 \rightarrow \begin{Bmatrix} \text{deux} \\ \text{trois} \\ \text{quatre} \\ \text{cinq} \\ \text{six} \\ U_3 \end{Bmatrix}$$

(R 3.4) $$U_3 \rightarrow \begin{Bmatrix} \text{sept} \\ \text{huit} \\ \text{neuf} \end{Bmatrix}$$

(R 4.1) $$A^1 \rightarrow \begin{Bmatrix} D_1 \\ V \\ S \\ Q \end{Bmatrix}$$

(R 4.2) $$D_1 \rightarrow \begin{Bmatrix} \text{dix}(U_3) \\ D_2 \end{Bmatrix}$$

(R 4.3) $D_2 \rightarrow \begin{Bmatrix} \text{onze} \\ D_3 \end{Bmatrix}$ (R 4.4) $D_3 \rightarrow \begin{Bmatrix} \text{douze} \\ \text{treize} \\ \text{quatorze} \\ \text{quinze} \\ \text{seize} \end{Bmatrix}$

(R 5) $V \rightarrow \begin{Bmatrix} \text{vingt} \\ \text{trente} \\ \text{quarante} \\ \text{cinquante} \end{Bmatrix} \begin{pmatrix} \text{'et un'} \\ U_2 \end{pmatrix}$

(R 6) $S \rightarrow \{\text{sextillion}\} \begin{pmatrix} \text{'et un'} \\ U_2 \\ \text{dix } (U_3) \\ \text{'et onze'} \\ D_3 \end{pmatrix}$

(R 7) $Q \rightarrow \{\text{quatre} - \text{vingt}\} \begin{pmatrix} A^0 \\ D_1 \end{pmatrix}$

(R 8) $A^2 \rightarrow \begin{Bmatrix} \phi \\ U_2 \end{Bmatrix} \{\text{cent}\} \begin{pmatrix} A^0 \\ A^1 \end{pmatrix}$

(R 9) $A^3 \rightarrow \begin{Bmatrix} \phi \\ U_2 \\ A^1 \\ A^2 \end{Bmatrix} \{\text{mille}\} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \end{pmatrix}$

(R 10.1) $A^{6N} \rightarrow \begin{Bmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{Bmatrix} \{\text{N-illion}\} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix}$

avec $N = 1, 2, ..., 6$

(R 10.2) $\begin{bmatrix} \text{1 - illion} \\ \text{2 - illion} \\ \vdots \\ \text{6 - illion} \end{bmatrix} \rightarrow \begin{bmatrix} \text{million} \\ \text{billion} \\ \vdots \\ \text{sextillion} \end{bmatrix}$

Note : dans tout ce qui précède, les parenthèses ( ) indiquent un choix facultatif, les accolades { } un choix obligatoire, et les crochets [ ] le choix simultané des termes qui se correspondent.

1.4.4. *Règles dépendant du contexte*

En finale, un fN est réécrit un fN ou une fN selon que le fonctème est masculin ou féminin et N-illion fN se réécrit N-illion de fN. Il conviendrait encore de développer les règles morpho-phonologiques qui permettent par exemple de fixer la prononciation de **dix** (di, dis, diz), et, plus généralement, de résoudre la question des liaisons en particulier lorsqu'elles permettent de distinguer entre un singulier et un pluriel.

1.4.5. La numération française est une numération arithmétique (additivo-multiplicative parenthésée) à caractère décimal, mais aussi vicésimal. La dernière règle introduit un caractère récursif extrêmement original ; elle permet au locuteur d'étendre aussi loin qu'il le veut la capacité générative du système. Nous avons arbitrairement limité cette capacité en adoptant la valeur N = 6 comme maximale dans la règle (R_{10}), mais, Chuquet, -l'inventeur des termes en -illion-, poursuit jusqu'à 9 (**nonillion**) et dit explicitement que la suite peut être prolongée à volonté. Cette règle conduit à poser que la numération française est de type parabasique en raison de l'opération d'élévation à une puissance qu'elle implique.

Elle est aussi assez peu systématique.

1.5. *Exemples de numérations amérindiennes.*

1.5.0. Les numérations amérindiennes présentées ci-dessous sont toutes de type arithmétique (additivo-multiplicatives parenthésées) et possèdent une capacité générative importante permettant au moins d'atteindre une dizaine de milliers. Dans l'état actuel de nos recherches, il est souvent difficile de répondre catégoriquement à la question de savoir si ces numérations sont des creations précolombiennes ou si elles résultent d'une acculturation consécutive à la conquête espagnole. En tout état de cause, il convient de noter que les systèmes amérindiens se trouvent en pleine évolution et qu'ils sont, dans certaines communautés tout au moins, plus ou moins- abandonnés au profit de la numération espagnole. Beaucoup de linguistes soulignent le fait que très peu de locuteurs sont actuellement en mesure de maîtriser les numérations primitives en particulier lorsqu'il s'agit de la saisie de grands nombres. B.J. Hoff, par exemple, nous rapporte ce témoignage d'un informateur caribe, Mr Kiban :

> *"He then told me that he had once counted from 1 to 1000 in Carib, to see if he could do it, and he could. Yet he was fully aware of the fact that in carrying out this operation, it had not been possible, for him to turn off; as it were, his knowledge of modern arithmetics. Actually he had used -***mboto** *'times' in the sense of x, producing 1000 itself by translating 5 x 10 x 20 into Carib. Mr Kiban confirmed my belief, that in practice only the simple, manageable numbers were actually used. [....] where modern life demands higher numbers and abstract calculations, it provides at the same time its own numerical tools".* (HOFF, 1983).

1.5.1. *Numération quechua.*

1.5.1.0. Selon les dialectes et les communautés, les locuteurs quechua disposent actuellement d'une numération parlée dont la capacité générative théorique varie de 999 à 999 999 999 999. Cette limite est fonction de la valeur numérique du dernier nombre d'appui multiplicatif attesté dans la communauté : **syen** '100' (emprunt espagnol) dans le parler de Ferreñafe (source : Gerald Taylor), **uaranka** '1000' à Quito, Saquisiti et Otavalo (source : G. Tarlé) ; **unu** = **hunu** '1 000 000' en runasimi (parler quechua de Ayacucho, Apurimac, Huancavellica, source : A. Yaranga Valderrama).

1.5.1.1. *Exemples (Selon Yaranga).*

14 = **chunka tawa-yuq** soit // 10 / 4 - suff. poss. 3è pers. //
19 = **chunka isqun-niyuq** soit // 10 / 9 - suff.: poss. 3è pers. //
20 = **iskay chunka** soit // 2 / 10 //
85 = **pusaq chunka pichqa-yuq** soit // 8 / 10 / 5 - suff. poss. 3è pers. //
110 = **pachaq chunka** soit // 100 / 10 //
985 = **isqun pachaq pusaq chunka pichqa-yuq** soit // 9 / 100 / 8 / 10 / 5 - suff. poss. 3è pers. //
1985= **waranga isqun pachaq pusaq chunka pichqa-yuq** soit // 1000 / 9 / 100 / 8 / 10 / 5 - suff. poss. 3è pers. //.

1.5.1.2. *Autre exemple (selon Tarlé)*

3459 = **kimsa uaranka chusku patsak pichka chunka iskun** soit // 3 / 1000 / 4 / 100 / 5 / 10 / 9 /.

Note : Dans ce parler, le suffixe possessif de 3è personne (**niyuq**, **yuq**) n'est pas attesté.

1.5.1.3. *Autres exemples (selon G. Taylor)*

4 = **kwatru**, 5 = **singu**, 6 = **seys**, 7 = **syeti**, 8 = **uchu**, 9 = **nwevi**, 10 = **dyes**, 100 = **syen**.

Note : Dans ce parler, tous les noms de nombre sont des emprunts espagnols, sauf les trois premiers noms (**uk** 'l', **ishkay** '2', **kimsa** '3').

Ces faits suggèrent une hypothèse et posent une question. La première, que le modèle "primitif" des numérations quechua serait du type "1, 2, *beaucoup"*. La seconde est celle de la reconstruction du processus cognitif ou langagier qui a permis de dépasser cette toute première limite, c'est-à-dire de nommer trois.

1.5.1.4. *Le vocabulaire terminal.*

V_T = { **huk**, **iskay**, **kimsa**, **tawa**, **pichqa**, **suqta**, **qanchis**,
1 2 3 4 5 6 7
pusaq, **isqun**, **chunka**, **pachak**, **waranqa**, **wamani**,
8 9 10 100 1000 40 000
unu ~ hunu, -(ni)yuq
1 000 000 suff. poss. 3è pers.}

Note : ce vocabulaire est obtenu à partir des données de Yaranga. Cet auteur est le seul, à notre connaissance, à donner **wamani** = 40000 et à fixer précisément la valeur numérique de **unu**.

1.5.1.5. *Le vocabulaire non-terminal*

$V_{NT} = \{\Sigma, A^0, A^1, A^2, A^3, W, A^6, U\}$

Note : Σ est un symbole initial.

1.5.1.6. *règles syntaxiques (indépendantes du contexte).*

(R 1) $\Sigma \rightarrow \begin{Bmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \\ A^6 \end{Bmatrix}$ (R'1) $\Sigma \rightarrow W$

Note : La règle (R'1) est distinguée de la règle (R1) parce que les données attestées sur l'emploi de wamani sont insuffisantes.

(R 2) $A^0 \rightarrow \begin{Bmatrix} \text{huq '1'} \\ U \end{Bmatrix}$ $U \rightarrow \begin{Bmatrix} \text{iskay '2'} \\ \text{kimsa '3'} \\ \text{tawa '4'} \\ \text{pichqa '5'} \\ \text{suqta '6'} \\ \text{qanchis '7'} \\ \text{pusaq '8'} \\ \text{isqun '9'} \end{Bmatrix}$

(R 3) $A^1 \rightarrow \begin{Bmatrix} \phi \\ U \end{Bmatrix}$ {chunka '10'} (A^0 -niyuq)

Note : -niyuq ne se trouve qu'après une consonne ; après une voyelle, on a la forme -yuq

$$\text{(R 4)} \qquad A^2 \rightarrow \left\{\begin{matrix}\phi \\ U\end{matrix}\right\} \{\text{pachak '100'}\} \left(\left\{\begin{matrix}A^0 \\ A^1\end{matrix}\right\}\right)$$

$$\text{(R 5)} \qquad A^3 \rightarrow \left\{\begin{matrix}\phi \\ U \\ A^1\end{matrix}\right\} \{\text{waranka '1000'}\} \left(\left\{\begin{matrix}A^0 \\ A^1 \\ A^2 \\ A^3\end{matrix}\right\}\right)$$

$$\text{(R 6)} \qquad A^6 \rightarrow \left\{\begin{matrix}\phi \\ U \\ A^1 \\ A^2 \\ A^3\end{matrix}\right\} \{\text{unu '1000000'}\} \left(\left\{\begin{matrix}A^0 \\ A^1 \\ A^2 \\ A^3\end{matrix}\right\}\right)$$

$$\text{(R 7)} \qquad W \rightarrow \left\{\begin{matrix}\cdots \\ U \\ \cdots\end{matrix}\right\} \{\text{wamani '40000'}\} \text{ (?)}$$

Exemple : **soqta wamani** soit // 6 / 40 000 // = 240 000.

1.5.1.7. Les éléments du vocabulaire terminal se répartissent en trois sous-classes. La première ne comporte que les variantes phonétiques du suffixe possessif de 3ème personne (**-niyuq**, **-yuq**). La deuxième comprend les numéraux qui peuvent être utilisés seuls ou entrer dans une expression composée où ils remplissent une occurrence d'argument non-distingué dans une opération. Cette classe pourrait être appelée *classe des numéraux unités* puisqu'elle comprend les neuf premiers entiers. La troisième comprend les numéraux qui peuvent être utilisés seuls ou entrer dans une expression composée où ils remplissent une occurrence d'un argument distingué d'une opération. Cette classe pourrait être appelée *classe des numéraux unitaires* ou des classificateurs unitaires puisqu'elle ne comprend que des noms de nombre à caractère collectif.

1.5.1.8. Du point de vue fonctionnel, deux opérations doivent être distinguées. La première renvoie à la notion d'ajouter (en vision naturelle)[22] ou à la notion d'addition (en vision arithmétique). La seconde renvoie au dénombrement de groupements (en vision naturelle)[23] ou à la notion de multiplication (en vision arithmétique). Les deux opérations portent sur deux arguments, et nous avons noté plus haut que l'un des arguments est toujours distingué (en ce sens que certains nombres particuliers seulement peuvent remplir une occurrence d'argument distingué). Dans le cas de la première opération, le numéral distingué joue le rôle de ce que nous avons appelé un *nombre d'appui additif* par exemple **chunka** '10', dans la formation des composés de 11 à 19. Dans le cas de la seconde opération, le numéral distingué joue le rôle d'un *nombre d'appui multiplicatif.* Par exemple **pachak** '100', dans la formation des composés **isqay pachak** 'deux cents', **kimsa pachak** 'trois cents' etc.

1.5.1.9. Une contrainte générale limité la cooccurrence des arguments distingué et non distingué : l'élément non-distingué doit avoir une valeur numérique strictement inférieure à celle de l'argument distingué (c'est-à-dire à celle du nombre d'appui).

Une contrainte plus forte limite la cooccurrence des arguments distingué et non-distingué dans le cas du nombre d'appui multiplicatif **pachak** '100', à savoir que l'élément non-distingué doit être strictement inférieur à dix. Enfin, toujours en ce qui concerne le seconde opération, une contrainte bloque la cooccurrence de **huk** '1' et des nombres d'appui multiplicatif (**huk chunka**, **huk pachak** etc. ne sont pas attestés, mais on a **ø-chunka**, **ø pachak**, etc).

1.5.1.10. Si l'on néglige le rôle du relateur **(ni)yuq**, les deux opérations (addition, multiplication) ne peuvent être distinguées, qu'à partir d'un tactème d'ordre. Quand les constituants se trouvent dans la chaîne parlée dans l'ordre croissant (respectivement, décroissant), la concaténation a valeur multiplicative (respectivement, additive). On peut ainsi reconnaître la nature additive ou multiplicative d'un nombre d'appui. Par exemple, **pachak** '100' est un nombre

---

22 La formation des composés "additifs" peut renvoyer à une pratique arithmétique savante (hypothèse conceptuelle forte) ou à des opérations moins techniques et plus naturelles (hypothèse conceptuelle faible), celle d'ajouter ou de réunir des quantités et celle de dépasser un groupement ayant, dans la culture considérée, un caractère étalon ou standard. **chunka** '10' qui, selon Yaranga, signifie aussi 'beaucoup' pourrait avoir été un tel étalon.

23 Les composés multiplicatifs peuvent renvoyer à la multiplication arithmétique (hypothèse forte) ou à une pratique naturelle (hypothèse faible) de groupement standard et de dénombrement des groupes obtenus.

d'appui multiplicatif (respectivement additif) dans la chaîne croissante **iskay pachak** (respectivement dans la chaîne décroissante **pachak isquay**).

Note : les contraintes sur les arguments non-distingués étant différentes selon que l'opération envisagée est une addition ou une multiplication, il convient de partitionner la classe des numéraux unités en distinguant les nombres qui peuvent apparaître comme argument non-distingué d'une addition et ceux qui peuvent apparaître comme argument non-distingué d'une multiplication. C'est la raison de l'introduction du symbole U dans le vocabulaire non terminal.

1.5.1.11. Bilan. La description de la syntaxe de la numération quechua fait intervenir deux opérations (addition, multiplication) et les catégories suivantes :

– les *unités,* c'est-à-dire les éléments résultant de la réécriture du symbole A^0. Ces éléments peuvent tous jouer le rôle d'une expression numérique bien formée, ou celui de l'argument non-distingué dans une opération. Seuls les éléments de U et le morphème zéro peuvent être arguments non-distingués d'une multiplication.

– les *nombres d'appui,* c'est-à-dire les collectifs **chunka**, **pachak**, **waranka**, **unu**, **wamani**. Ces éléments peuvent tous jouer le rôle d'une expression numérique bien-formée, ou celui de l'argument distingué dans les opérations d'addition et de multiplication.

– les *complexes additivo-multiplicatifs d'ordre i,* c'est-à-dire les composés résultant de la réécriture des symboles A^i (Il serait évidemment possible de distinguer des complexes strictement additifs (des sommes) et des complexes strictement multiplicatifs (des multiples).

– *1'ordre* (au sens de celui des éléments d'une comptine) est indispensable à la saisie de la valeur numérique des unités.

– la *hiérarchie* des appuis est indispensable pour la segmentation et la formation des complexes additivo-multiplicatifs.

1.5.1.12. Les analyses précédentes conduisent à classer la numération quechua parmi les numérations arithmétiques (additivo-multiplicatives parenthésées) à caractère décimal (les nombres d'appui **chunka** '10', **pachak** '100', **waranqa** '1 000', **unu** '1 000 000' sont tous extraits d'une progression géométrique, celle des puissances de 10).

Cette numération parlée ne peut pas être considérée comme une numération décimale parlée de position pour différentes raisons (pas de terme pour zéro, pas de terme pour 10^4, ni pour 10^5, présence de **wamani** = 2^6 x 5^4 dans la liste des nombres d'appui).

Les caractéristiques précédentes étant presque exactement celles de la numération parlée espagnole, la question se pose de savoir si la numération quechua est un emprunt ou une construction amérindienne originale. Un élément de réponse à cette question se trouve peut-être dans une meilleure connaissance de la morphologie du terme **wamani**, c'est-à-dire du seul nombre d'appui dont l'origine ne peut être espagnole. Selon les données de Yaranga ce nombre d'appui a pour valeur numérique 40 000, c'est-à-dire 26 x 54 ce qui permet, par exemple, les décompositions suivantes 4 x 10 000, 100 x 400, 5 x 8 000. L'étymologie et les études comparatives permettraient sans doute de tester ces hypothétiques analyses. Si l'on établissait, par exemple, la reconstruction **tawa** + **mani** → **wamani**, on pourrait admettre l'existence d'un terme quechua (**mani**) pour désigner 104, ce qui serait un argument en faveur d'une origine amérindienne du système de numération et de son caractère décimal. Si l'analyse révélait au contraire l'existence d'un constituant renvoyant à une puissance de vingt (400 ou 8 000) on établirait du même coup le caractère originellement vigésimal de la numération quechua, ce qui serait encore un argument en faveur d'une origine amérindienne (mais aussi de la profonde évolution subie par cette numération).

1.5.2. *Numération tarahumara.*

1.5.2.1. *Exemples (selon Merrifield)* .

1 = **biré**, 2 = **okuá**, 3 = **bikiyá**, 4 = **nawó**, 5 = **marí**, 6 = **usáni**, 7 = **kičáo**, 8 = **o-sá nawó**, 9 = **kimakói**, 10 = **makói**, 11 = **makói waminá biré**, 19 = **makói waminá kimakói**, 20 = **o-sá makói**, 21 = **o-sá makói waminá biré** 30 = **bai-sá makói**, 45 = **nawó-sá makói waminá marí**, 100 = **biré siénto**, 227 = **okuá siénto waminá o-sá makói waminá kičáo**, 1000 = **míli**, 9999 = **kimakói míli waminá kimakói siénto waminá kimakói-sá makói waminá kimakói.**

1.5.2.2. *Le vocabulaire terminal.*

V_T = { **biré**, **okuá**, **bikiyá**, **nawó**, **marí**, **usáni**, **kičáo**, **o-sá nawó**,

1	2	3	4	5	6	7	8

kimakói, **makói**, **siénto**, **míli**, **-sá**, **waminá** }

9	10	100	1000	suf-ordinal	déictique

Note : Dans le cadre d'une analyse *morphologique* fine, il conviendrait de considérer **o-sá nawó** '8' comme un composé, et donc d'éliminer ce terme du vocabulaire terminal. De même, il conviendrait d'estimer la productivité de la composition éventuelle d'un hypothétique **ki-** (**makói** '10', **ki-makói** '9', **ki-čáo**

'7') qui pourrait renvoyer à un procédé de repérage (vision ordinale d'anticipation) ou de compte à rebours (vision soustractive). Dans les deux hypothèses, il faudrait restituer l'argument **biré** 'un' : 9 = **(biré)-ki-makói** 'un-de-dix'. Vingt = **o-sá makói** 'deuxième dix' est, du point de vue *syntaxique,* un composé et ne figure donc pas dans le vocabulaire terminal.

1.5.2.3. *Le vocabulaire non-terminal.*

$V_{NT} = \{ \Sigma, A^0, A^1, A^2, A^3, U_1, O, R \}$

Note : Σ est un symbole initial.

1.5.2.4. *Les règles de réécriture.*

(R 1) $\Sigma \rightarrow \begin{Bmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{Bmatrix}$

(R 2) $A^0 \rightarrow \begin{Bmatrix} \text{biré} & \text{'1'} \\ U_1 & \end{Bmatrix}$ $\qquad U_1 \rightarrow \begin{Bmatrix} \text{okuá} & \text{'2'} \\ \text{bikiyá} & \text{'3'} \\ \text{nawó} & \text{'4'} \\ \text{marí} & \text{'5'} \\ \text{usáni} & \text{'6'} \\ \text{kičáo} & \text{'7'} \\ \text{o-sá nawó} & \text{'8'} \\ \text{kimakói} & \text{'9'} \end{Bmatrix}$

(R 3) $O_1 \rightarrow U_1 O$ $\qquad O \rightarrow \{\text{-sá}\}$

Note : {okuá-sá} → {o-sá} 'deuxième' ; {bikiyá-sá} → {bai-sá} 'troisième' {nawó-sá} 'quatrième'.

(R 4) $A^1 \rightarrow \begin{Bmatrix} \phi \\ O_1 \end{Bmatrix} \{\text{makói} \ \ \text{'10'}\} \ (R \ A^0)$ $\qquad R \rightarrow \{\text{waminá}\}$

(R 5) $A^2 \rightarrow \begin{Bmatrix} \phi \\ A^0 \end{Bmatrix} \{\text{siénto} \ \ \text{'100'}\} \ (R \begin{Bmatrix} A^0 \\ A^1 \end{Bmatrix})$

(R 6) $A^3 \rightarrow \begin{Bmatrix} \phi \\ A^0 \\ ?\ A^1 \\ ?\ A^2 \end{Bmatrix} \{\text{míli} \ \ \text{'1000'}\} \ (R - \begin{Bmatrix} A^0 \\ A^1 \\ A^2 \end{Bmatrix})$

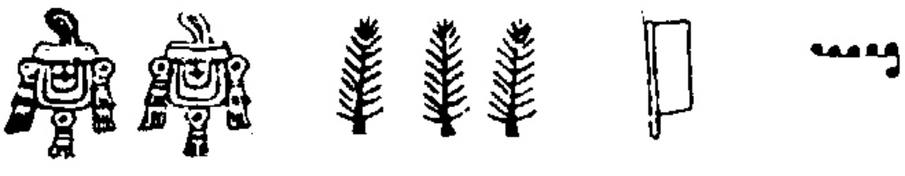
1.5.2.5. Cette numération moderne doit être classée, du point de vue syntaxique, dans les numérations arithmétiques (additivo-multiplicatives parenthésées) à caractère décimal.

L'emprunt à l'espagnol des nombres d'appui **siénto** '100' et **míli** '1000' suffit à prouver que cette numération a subi très fortement l'influence de cette culture. Elle offre cependant des traces importantes de procédés de formation ordinale dans lesquelles **nawó** 'quatre' et **makói** 'dix' ont pu jouer le rôle de repère privilégié. Noter aussi le déictique **waminá** 'farther' interprétable comme une marque d'ajout.

1.5.3. *Numération nahuatl*

1.5.3.0. *"On ne saurait oublier que les aztèques n'ont jamais écrit de nombres égaux ou supérieurs à 20^4 = 160 000 [....]. Pour un mathématicien, le contraste entre l'usage du calendrier par les mayas et sa simple utilisation pratique par les aztèques, fait penser que ceux-ci n'ont été que des imitateurs de la science maya ou d'une science antérieure"* (GUITEL, 1975 :139). Ce jugement sévère ne s'adresse, il est vrai, qu'à la numération *écrite* et ne saurait s'appliquer à la numération parlée mais il permet aussi, et peut-être surtout, de souligner que les Aztèques savaient manipuler de grands nombres[24] et que l'origine de leur système de numération est manifestement précolombienne.

Rappelons que 1a numération écrite utilise quatre symboles : un petit rond (ou un point), parfois un doigt, pour représenter le nombre un , un dessin ressemblant à un petit drapeau pour vingt ; une natte de cheveu pour quatre-cents ; et un sac pour huit mille. Dans ce système, un nombre (décomposé selon les "puissances" de vingt) est représenté de manière fort "primitive" simplement par juxtaposition d'autant de symboles de chaque type qu'il est nécessaire. Par exemple, 17 226 = $2 \times 20^3 + 3 \times 20^2 + 1 \times 20 + 6$ sera représenté par le dessin suivant :



Note : Dans l'écriture d'un nombre, les symboles de même nature sont généralement regroupés cinq par cinq, mais cette habitude de regroupement n'est pas rigide.

24 Et ceci, pas seulement dans le cade des computs calendaires car les aztèques tenaient à jour le cadastre des terres, les comptes de l'impôt, etc. (cf. Herbert R. HARVEY et Barbara WILLIAMS 'l'arithmétique aztèque', *La Recherche,* n° 126, octobre 1981).

1.5.3.1. *Exemples d'expressions numériques composées.*

Les données et les transcriptions sont, pour l'essentiel, celles de Raoul de la Grasserie (1903).[25]

1 = **ce**, 2 = **ome**, **on**, 3 = **ye**, **ei**, 4 = **naùi**, 5 = **ma-cuilli** (< **maitl** 'main' + **cui** 'prendre'), 6 = **chiqua-cen**, 10 = **ma-tlactli** (< **maitl** 'main' + **tlactli** 'buste, tronc'), 11 = **matlactli o-ce**, 12 = **matlactli om-ome**, 15 = **caxtolli**, 16 = **caxtolli o-ce**, 17 = **caxtolli om-ome**, 20 = **cem-poalli** (**cem** 'un' **poa** 'compter'), 23 = **cem-poalli om-ei**, 40 = **om-poalli**, 60 = **e-poalli**, 400 = **cen-tzontli** (**tzontli** 'barbe, poil, cheveu'), 800 = **ome-tzontli**, 16000 = **ome-xikipilli** (d'après J.Luna Cárdenas 'bolsa').

1.5.3.2. Raoul de la Grasserie signale un certain nombre de classificateurs numériques : **tetl** pour les objets ronds ; **olotl** pour le maïs, les troncs, les piliers, **tlamantli** pour les paires, **pantli** pour les rangées, les sillons, les murs, les objets ou personnes disposés en rang... Nous ne traiterons pas de cette question dans le cadre de cette étude et nous n'analyserons que les expressions numériques proprement dites.

1.5.3.3. *Le vocabulaire terminal.*

V_T= {	**ce**,	**ome**,	**ye**,	**naùi**,	**macuilli**,	**chiqua**,	**matlactli**,	**caxtolli**,
	1,	2,	3,	4,	5,	'fraction',	10	15
	poalli,	**tzontli**,	**xikipilli**,	**om**. }				
	20	400	8 000	'relateur'				

1.5.3.4. *Le vocabulaire non-terminal.*

$V_{NT} = \{ \Sigma, B^0, B^1, B^2, B^3, U, F, D, Q, R \}$

---

25 Selon Michel Launey, on trouve en nahuatl classique

1 :	**sē**	14 :	**maʔλāk-λi on-nāwi**
2 :	**ōme**	15 :	**kaštol-li**
3 :	**ēyi**	16 :	**kaštol-li on-sē**
4 :	**nāwi**	17 :	**kaštol-li om-ōme**
5 :	**māk**^{w}**il-li**	20 :	**sem-pōwal-li** 'un compte'
6 :	**cik**^{w}**asē**	21 :	**sem-pōwal-li on sē**
7 :	**cikōme**	40 :	**ōm- pōwal-li**
8 :	**cik**^{w}**ēyi**	60 :	**ē-pōwal-li**
9 :	**cik**^{w}**nāwi**	80 :	**nāw-pōwal-li**
10 :	**maʔλāk-λi**	100 :	**māk**^{w}**il-pōwal-li**
11 :	**maʔλāk-λi on-sē**	200 :	**maʔλāk-pōwal-li**
12 :	**maʔλāk-λi om-ōme**	400 :	**sen-con-λi** 'une touffe de cheveux'
13 :	**maʔλāk-λi om-ēyi**	8000 :	**sen-šikipil-li** 'un sac de grains'

1.5.3.5. *Les règles syntaxiques.*

(R 1) $$\Sigma \rightarrow \begin{Bmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{Bmatrix}$$

(R 2) $$B^0 \rightarrow \begin{Bmatrix} U \\ \text{macuili} \\ F \\ D \\ Q \end{Bmatrix} \qquad U \rightarrow \begin{Bmatrix} \text{ce} & '1' \\ \text{ome} & '2' \\ \text{ye} & '3' \\ \text{naùi} & '4' \end{Bmatrix}$$

$$F \rightarrow \{\text{chiqua}\}\ \{U\} \qquad D \rightarrow \{\text{matlalli}\}\ (RU)$$

$$Q \rightarrow \{\text{caxtolli}\}\ (RU) \qquad R \rightarrow \{\text{om-}\}$$

Remarque : devant les consonnes autres que s ou m on trouve -on

(R 3) $$B^1 \rightarrow \{B^0\}\ \{\text{poalli '20'}\}\ (R\,B^0)$$

(R 4) $$B^2 \rightarrow \{B^0\}\ \{\text{tzontli '400'}\}\ \left(R \begin{Bmatrix} B^0 \\ B^1 \end{Bmatrix}\right)$$

(R 5) $$B^3 \rightarrow \{B^0\}\ \{\text{xikipilli '8000'}\}\ \left(R \begin{Bmatrix} B^0 \\ B^1 \\ B^2 \end{Bmatrix}\right)$$

Note : Il convient de distinguer 5, 10, 15 (qui ne peuvent jouer le rôle d'appui multiplicatif) et 20, 400, 8 000 (qui peuvent jouer ce rôle). D'où la partition suivante du vocabulaire terminal { 1, 2, 3, 4 }, { 5, 10, 15 }, { 20, 400, 8 000 } { **om** }

Il conviendrait d'établir les règles phonétiques.

1.5.3.6. La numération parlée nahuatl doit être classée parmi les numérations arithmétiques (additivo-multiplicatives parenthésées) vigésimales (le paradigme des coefficients des "puissances de vingt" est systématiquement limité aux dix-neuf premiers nombres). Il s'agit donc d'une numération polynomiale[26] qui ne diffère d'une numération de position que par l'absence du zéro et par l'existence d'un sous-système additif permettant la formation des "chiffres", c'est à dire des vingts premiers nombres. Ce sous-système utilise trois appuis additifs (5, 10, 15) qui renvoient peut-être à une gestuelle des doigts.

1.5.4. *Numération otomi*

1.5.4.0. Cette numération présente la particularité d'être de caractère vigésimal jusqu'à cent, et de caractère décimal à partir de cent.

1.5.4.1. *Exemples (selon Merrifield):*

1 = **ʔna**, 2 = **yoho**, 3 = **hñu**, 4 = **goho**, 5 = **kïʔa**, 6 = **ʔrato**, 7 = **yoto**, 8 = **hñato**, 9 = **gïto**, 10 = **ʔroetʔa**, 11 = **ʔroetʔa ma-ʔra**, 12 = **ʔroetʔa ma-yoho**, 19 = **ʔroetʔa ma-gïto**, 20 = **ʔnate**, 21 = **ʔnate ma-ʔra**, 40 = **yo-ʔrate**, 41 = **yo-ʔrate ma-ʔra**, 99 = **goho ʔrate ma-ʔroetʔa ma-gïto**, 100 = **ʔna nthebe**, 247 = **yoho nthebe ne-yo-ʔrate ma-yoto**, 1 000 = **ʔna ʔmo**, 2488 = **yoho ʔmo ne-goho nthebe ne-goho ʔrate ma-hñato**, 9999 = **gïto ʔmo ne-gïto nthebe ne-goho ʔate ma-ʔroeta ma gïto**.

1.5.4.2. *Le vocabulaire terminal.*

V_T = {	**ʔna**,	**yoho**,	**hñu**,	**goho**,	**kïtʔa:**,	**ʔroetʔa**,	**ʔnate**,	**nthebe**,
	1	2	3	4	5	10	20	100
	ʔmo,	**-to**,	**ma**,	**ne**				
	1 000	suffixe	adjectif 'autre'	conj. coord. }				

1.5.4.3. *Le vocabulaire non terminal.*

V_{NT} = { Σ, U, U_1, U_2, u, B^0, B^1, V, Q, A^2, A^3, D, R, R' }

[26] C'est à dire une numération dans laquelle le paradigme des multiplicateurs (coefficients) des diverses puissances de la "base" est l'ensemble des premiers nombres limité au prédécesseur immédiat de la "base" (les chiffres).

1.5.4.4. *Les règles de composition syntaxique*

(R 1)

$$\Sigma \rightarrow \left\{\begin{matrix} \mathrm{U} \\ \mathrm{B}^0 \\ \mathrm{B}^1 \\ \mathrm{A}^2 \\ \mathrm{A}^3 \end{matrix}\right\}$$

(R 2)

$$\mathrm{U} \rightarrow \left\{\begin{matrix} \text{?nate} & \text{'1'} \\ \mathrm{U}_1 & \\ \text{kït?a} & \text{'5'} \\ \mathrm{U}_2 & \end{matrix}\right\} \qquad \mathrm{U}_1 \rightarrow \left\{\begin{matrix} \text{yoho} & \text{'2'} \\ \text{hnu} & \text{'3'} \\ \text{goho} & \text{'4'} \end{matrix}\right\}$$

$$\mathrm{U}_2 \rightarrow \left\{\begin{matrix} \text{?ra} & \text{'1b'} \\ \text{yo} & \text{'2b'} \\ \text{hna} & \text{'3b'} \\ \text{gï} & \text{'4b'} \end{matrix}\right\} \{\mathrm{D}\}$$

$$\mathrm{D} \rightarrow \{\text{to}\}$$

(R 3)

$$\mathrm{B}^0 \rightarrow \{\text{?rœt?a} \ \text{'10'}\} \ (\mathrm{R} \ u)$$

$$u \rightarrow \left\{\begin{matrix} \text{?ra} & \text{'1b'} \\ \mathrm{U}_1 & \\ \text{kït?a} & \text{'5'} \\ \mathrm{U}_2 & \end{matrix}\right\}$$

$$\mathrm{R} \rightarrow \{\text{ma} \ \text{'other'}\}$$

(R 4)

$$\mathrm{B}^1 \rightarrow \left\{\begin{matrix} \mathrm{V} \\ \mathrm{Q} \end{matrix}\right\} \qquad \mathrm{V} \rightarrow \{\text{?nate} \quad \text{'20'}\} \ \left(\mathrm{R}\left\{\begin{matrix} u \\ \mathrm{B}^0 \end{matrix}\right\}\right)$$

$$\mathrm{Q} \rightarrow \left\{\begin{matrix} \text{yo} & \text{'2b'} \\ \text{hnu} & \text{'3'} \\ \text{goho} & \text{'4'} \end{matrix}\right\} \{\text{?rate} \quad \text{'20b'}\} \ \left(\mathrm{R}\left\{\begin{matrix} u \\ \mathrm{B}^0 \end{matrix}\right\}\right)$$

(R 5)

$$\mathrm{A}^2 \rightarrow \{\mathrm{U}\} \ \{\text{nthebe} \ \text{'100'}\} \ \left(\mathrm{R'}\left\{\begin{matrix} u^{27} \\ \mathrm{B}^0 \\ \mathrm{B}^1 \end{matrix}\right\}\right)$$

$$\mathrm{R'} \rightarrow \{\text{ne} \quad \text{'and'}\}$$

(R 6)

$$\mathrm{A}^3 \rightarrow \left\{\begin{matrix} \mathrm{U} \\ ? \end{matrix}\right\} \{\text{?mo} \ \text{'1000'}\}\left(\mathrm{R'}\left\{\begin{matrix} u^{27} \\ \mathrm{B}^0 \\ \mathrm{B}^1 \\ \mathrm{A}^2 \end{matrix}\right\}\right)$$

1.5.4.5. Selon Merrifield, les "relateurs" **ne** 'and' et **mo** 'other' seraient respectivement une conjonction de coordination et un adjectif. Cet auteur cite E. Wallis qui suggère les analyses suivantes : "**-to** *'double',* **tʔa** *'only',* ʔ**roe** *an alternant of* ʔ**ye** *'hands' (thus:* ʔ**roetoea** = *'hands only'* = 10 *fingers) and* **te** *'life, person' (thus :* ʔ**nate** *'one person'* = 20 *digits)."*

1.5.4.6. La numération otomi doit être classée comme les précédentes dans les numérations arithmétiques (additivo-multiplicatives parenthésées) à caractère mixte vigésimal (jusqu'à cent) et décimal (au-delà de cent). La productivité du caractère vigésimal a été (volontairement ?) limitée à 99, comme le montre le fait que dans la règle (R4) le paradigme U1 des multiplicateurs de vingt n'est pas maximal (on s'attendrait à trouver B^0). Ce qui plaide en faveur d'une origine vigésimale et donc à interpréter en termes d'emprunt (à la culture espagnole) le caractère décimal.[27]

La deuxième originalité est d'ordre morphologique plutôt que syntaxique : la dérivation des noms des nombres 6, 7, 8, 9, à partir de celui des nombres 1, 2, 3, 4. Une hypothèse de dérivation pour les nombres 5 et 10 peut également être formulée si l'on convient de considérer **-tʔa** comme un suffixe dérivationnel. Mais l'analyse morphologique reste à faire.

2. - Éléments pour une syntaxe des numérations ordinales.

2.0. Dans l'état actuel de nos recherches, les données factuelles sont trop récentes pour qu'il soit possible de brosser un panorama consistant de la syntaxe des numérations ordinales[28] Nous y reviendrons dans la partie III, après avoir établi que la numération maya est ordinale.[29] Cependant, un certain nombre d'éléments peuvent être proposés et servir de jalons pour les recherches ultérieures.

2.1. A l'origine des procédés de saisie ordinale, il convient probablement de postuler un principe de *discrétisation* d'un continuum (indifféremment, sans doute spatial ou temporel), ce qui revient à la capacité d'isoler un point sur une droite ou un instant dans une durée.

---

27 *"...Otomi... presumably had vigesimal numbering system which was able to name the positive integers from 1 to at least 159 999. The partially decimal system of today for numbers above 99 very likely dates from the introduction of the spanish* **sientu** *'hundred' into the system... the old vigesimal system persisted for counting* **tortillas**..." (MERRIFIELD 1968 : 95)

28 Une première synthèse sera présentée au 45ème Congrès des américanistes (Bogota, Colombie)

29 Voir CAUTY 1986 pour ce point.

2.2. La capacité de discrétisation peut être postulée dans toutes les civilisations et semble pouvoir être mise en relation avec le fait rappelé par J. Dhombres que dans toutes les langues on sait exprimer au moins quelques fractions : *"Il n'existe pas de civilisation sans arithmétique rudimentaire (...) Généralement quelques nombres fractionnaires font partie du bagage commun (1/3, 1/2, etc)"* (DHOMBRES, 1978 : 58).

2.3. Il convient probablement encore de distinguer les systèmes selon que le continuum discrétisé est saisi comme une droite indéfinie ou comme un cycle qui se referme sur lui-même. Le temps par exemple possède dans les cultures amérindiennes un caractère cyclique qui le distingue de notre notion de temps linéaire ; l'étonnante maîtrise du calcul des congruences par les prêtres maya n'est sans doute pas un hasard.

2.4. La mise en place d'une numération ordinale repose sur la saisie d'une relation d'ordre (suivre, être le successeur de...) ou de sa réciproque (précéder, être le prédécesseur de...) permettant de définir une *structure* dans laquelle pourront être définis un premier élément, le successeur et le prédécesseur d'un élément donné, et éventuellement une borne (une limite).. En d'autres termes, tout repose sur la saisie d'une relation de coïncidence et sur la maîtrise d'antonymies comme haut/bas, avant/après, si/alors dans les domaines spatial, temporel ou notionnel, antonymies au moins aussi primitives que les oppositions beaucoup/peu, très/peu des domaines quantitatif et intensif.

2.5. L'exemple des numérations amérindiennes conduit à souligner l'importance de ce que nous pourrions appeler les changements d'échelle de graduation. Les *comptines* apparaissent, en effet, comme des graduations unitaires, puis quinaires, allant ensuite de 20 en 20, puis de 400 en 400 etc.., (dans une perspective cardinale, il s'agit du passage des groupements aux groupements de groupements).

2.6. Il convient encore de postuler un principe pour rendre compte de l'usage des comptines en tant qu'instrument de comptage, c'est-à-dire du passage du point de vue ordinal au point de vue cardinal. Tous les auteurs soulignent que ces changements de point de vue sont primitifs en ce sens, par exemple, qu'ils sont très tôt acquis par le jeune enfant (PIAGET, 1941 : 163).

2.7. En première approximation, la syntaxe d'une numération ordinale fait intervenir les catégories[30] suivantes

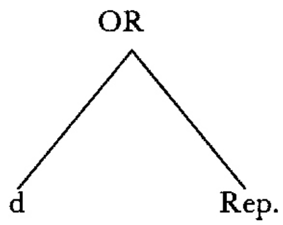
2.7.1. une relation d'ordre (directe et réciproque)

2.7.2. des divisions. Il s'agit des éléments d'une graduation unitaire, c'est-à-dire de la suite ordonnée des premiers nombres (au moins jusqu'à deux, par exemple jusqu'à cinq).

2.7.3. des repérants, c'est-à-dire des éléments d'une graduation non-unitaire. En andoke, par exemple, 5, 10, 15, 20 seront dits des repérants. En maya, on trouvera les puissances de 20.

2.7.4. le schème conceptuel de base d'une composition ordinale est l'évaluation, en termes de divisions, de la distance à un repérant, distance qui peut être saisie en vision d'antériorité ou en vision de postériorité

Contrairement au schème conceptuel de base d'une composition arithmétique (cf. 1.2.1. et 1.2.4.), il semble difficile ici de ne pas postuler que l'un des éléments de la relation soit distingué, (d'où son caractère de repérant). Ce schème peut être représenté par la figure suivante qui représente une relation entre une division d et un repérant, OR(d,rep) :



2.7.5. La mise en signes de ce schème conceptuel conduit parfois à des expressions qui s'interprètent sans difficulté dans le cadre ordinal, et qu'il est impossible, au contraire, d'interpréter en termes d'opérations arithmétiques

[30] Dans la terminologie de J.P. Desclés, ces catégories seraient dites grammaticales (nous préférons le terme syntaxiques) car définies et construites indépendamment de la langue naturelle étudiée. Elles s'opposent aux catégories linguistiques obtenues dans le cadre d'une langue naturelle particulière, par des procédés taxinomiques d'analyse distributionnelle (CULIOLI et DESCLES, 1979 : 75-80).

élémentaires (addition, multiplication, etc.). C'est le cas par exemple de 35 en maya (yucatèque) ou de 11 en andoke dont les gloses sont respectivement :

35 = **holhu ca kal** 'quinze deux(ième) vingtaine(s)' soit 35 = (15,<u>40</u>)

11 **= ka dnka-ka-nisidé** 'un pour aller à nos pieds' soit 11 = (1,<u>20</u>)
et, plus généralement, des numérations qui utilisent les procédés connus dans la littérature sous les appellations de *back-counting* et *under-counting.*

2.7.6. Tous les moyens grammaticaux habituels pour marquer une relation dans les domaines spatial, temporel et notionnel sont susceptibles d'être utilisés, ainsi que les ressources de la pragmatique ou de la rhétorique. Soulignons cependant que parfois la relation d'ordre n'est pas marquée autrement que par un tactème d'ordre. Quand le même phénomène affecte l'opération dans un composé arithmétique, il devient particulièrement difficile de décider si un composé est arithmétique ou ordinal.

2.7.7. Des exemples de composés ordinaux sont attestés évidemment aussi en dehors des langues amérindiennes nous avons fait allusion au latin à propos des formes apparemment soustractives. Il en serait de même en étrusque et dans de nombreuses langues africaines. Des composés de ce type ont été attestés récemment au Tibet par M. Mazaudon (cf note 17).

3. - Conclusions.

3.0. Les recherches présentées dans cette dernière partie se situent dans la tradition de la linguistique européenne qui a toujours réservé une place importante à la sémantique et maintenu l'exigence de fonder ses réflexions *"sur plusieurs langues naturelles, dans un va-et-vient constant entre le particulier et le général"* (POTTIER, 1974: Préface).

3.1. L'originalité éventuelle des méthodes que nous avons mis en oeuvre provient de la décision d'articuler *dès le moment de l'observation,* le savoir arithmétique et les théories linguistiques[31]. Cette attitude conduit à distinguer deux types de catégories. Les premières sont obtenues langue par langue par l'analyse distributionnelle d'un corpus d'expressions numériques. Il n'est pas indifférent de noter que les corpus peuvent ici être rassemblés de manière quasi-exhaustive. Les secondes ne sont pas obtenues à partir d'une langue particulière,

---

[31] Cette articulation est possible dans les recherches sur le nombre parce que le domaine d'expérience du quantitatif est conceptuellement organisable et par le locuteur et par le mathématicien : le domaine numérique fait partie de l'expérience commune avant même de faire partie de l'expérience propre du mathématicien.

c'est l'arithmétique qui permet de les construire et de les fonder dans une rationalité qui transcende la diversité des langues.

Les unités de niveau fondamental, c'est-à-dire les sombrants et les repérants, sont ainsi analysables soit comme des signes linguistiques de première articulation, soit comme des symboles arithmétiques.

En tant que signe, une unité du niveau fondamental est constituée d'un signifiant et d'un signifié lui-même analysable selon une composante syntaxique et une composante sémantique (POTTIER, 1974:26-31)

$$(1)\ \mathrm{Si} = \frac{\mathrm{Se'/Sy}}{\mathrm{Sa}}$$

Ce signe est souvent encore susceptible d'une analyse morphologique ou étymologique.

En tant que symbole, et sauf si l'on posait que les symboles sont toujours totalement vides de sens, une unité du niveau fondamental possède une charge sémantique. Cette charge est inscrite dans le jeu des combinaisons arithmétiques permises ; en d'autres termes, elle est portée par la seule composante syntaxique du signifié, mais portée comme en creux car le mathématicien la maintient méthodologiquement séparée dans un système interprétant

$$(2)\ \mathrm{SI} = \frac{\mathrm{Sy}}{\mathrm{Sa}}\ //\mathrm{Se'}$$

Le symbole n'est pas susceptible d'une analyse en constituants de première articulation. Par exemple, les termes en *-illion* ne livrent leur valeur numérique exacte qu'au mathématicien ou au locuteur qui (re)connaît en eux l'opération d'élévation à une puissance et ses deux arguments ; la valeur arithmétique de billion ('million à la puissance deux') n'est pas révélée par l'analyse des expressions de la langue formées sur le même modèle de composition (bicéphale, biplan, etc.), mais est saisie *après coup,* c'est-à-dire après une "initiation" mathématique.

Une caractéristique importante du langage mathématique peut en être déduite : la charge sémantique du symbole est conventionnelle : *elle est portée par la seule composante syntaxique.* Cette charge sémantique peut être dite le reflet des possibilités opératoires permises par une axiomatique[32], elle-même construite dans le but de modéliser le domaine d'expérience propre, ou devenu propre, au mathématicien.

3.2. Il nous semble encore important de souligner que les catégories syntaxiques (construites dans la rationalité arithmétique) rendent possible, non pas la comparaison directe des catégories linguistiques d'une langue avec celles d'une autre langue, mais l'évaluation de la distance où elles se trouvent par rapport aux premières (cf. par exemple, la conclusion des analyses de la numération nahuatl). L'évaluation de cette distance permettrait de contourner trois obstacles traditionnels : a) l'extrême difficulté de comparer les catégories linguistiques de langues éloignées b) la difficulté de fonder la construction de catégories susceptibles de transcender la diversité des langues ; c) la quasi-impossibilité d'interpréter en termes de processus cognitif (et en évitant le danger des projections ethnocentriques) ce que reflètent cependant la diversité attestée des mises en signes. Notons enfin que les distinctions introduites (par exemple, l'opposition morphologique/syntaxe) ouvrent une perspective prometteuse pour l'étude de l'enracinement dans la langue naturelle des procédés de formation lexicale que privilégie la langue mathématique spécialisée. Ces procédés pourraient être décrits comme des dé-re-contextualisations continuelles.

3.3. Le lecteur pourrait s'interroger à propos de la relative complexité de la syntaxe des numérations parlées, en particulier en comparaison de la grande simplicité de la grammaire des numérations de position qui ne comporte qu'une seule règle (sans aucune exception) et un simple alphabet terminal (cf. I.4.5.1.).

Une hypothèse relativement simple permet d'expliquer cette différence. Elle consiste à poser qu'il existe deux modes de fonctionnement pour les numérations : un mode "savant" et un mode "populaire".

Dans le mode savant, la numération est utilisée de manière critique[33], comme un outil que la pensée se forge pour résoudre un défi lancé par le

---

32 Dans la querelle historique pour laisser entrer les nombres *irrationnels* dans le domaine mathématique, la pomme de discorde n'était pas la définition de ces nombres (les mathématiciens grecs connaissaient déjà des algorithmes pour en calculer des valeurs approchées), mais bien la question d'oser leur appliquer les règles de calcul habituelles dans le domaine des nombres rationnels. La vraie difficulté était d'écrire $\sqrt{a}+\sqrt{b}$ et non pas de concevoir $\sqrt{a}$ ou $\sqrt{b}$ (cf. DHOMBRES, 1978).

33 Le célèbre *"Il faut que j'y pense encore"* prononcé par Lagrange à propos d'un mémoire sur les parallèles devant l'Académie des Sciences est révélateur de cette attitude.

développement d'une science toujours en construction (cf. I.4.5.5.). Le système de mise en signes est totalement subordonné aux besoins de l'arithmétique : une solution particulière est immédiatement sacrifiée au profit d'une solution plus générale, même quand celle-ci s'avère plus lourde dans quelques circonstances particulières. Par exemple, on marquera systématiquement les deux arguments d'une opération même lorsque l'un d'eux est redondant (cf. ø-vingt, quatre-vingts ; un million, quatre millions ; l'invention du zéro de position).

Le fonctionnement savant conduit à des solutions de plus en plus systématiques, en particulier en raison du principe de *réfutabilité* qui rend possible l'abandon des solutions devenues caduques dans la communauté. L'évolution peut aller jusqu'à la création de systèmes d'écriture *ad hoc* pratiquement totalement autonomes par rapport à la langue naturelle[34] mais sous le contrôle de la rationalité et de la communauté qui les ont faites naître. On passe du régime de la morphologie à celui de la syntaxe, de l'arbitraire au conventionnel, du singulier expressif au général performant. Le mathématicien cependant ne joue pas dans des espaces axiomatiques vides de sens, car la *composante syntaxique* du symbole a subi le long travail du temps et de la communauté qui l'ont patiemment et systématiquement modelée pour que le sens puisse y être, à volonté, rabattu.

Dans le mode populaire, toutes les solutions empiriques au problème de la mise en signes du nombre ont le même statut et sont acceptées sans critique possible (parfois pour des questions de mode, de culture, de tabou, pour des arguments d'autorité, etc.). Lorsqu'une solution est entrée dans le système de la langue, on ne peut plus l'abandonner volontairement, même si elle entre en conflit avec une solution plus générale ou plus simple. Le fonctionnement populaire fait ainsi surgir un ensemble hétéroclite de procédés d'origines les plus diverses. On observe par exemple l'usage de nombreux nombres d'appui renvoyant à des opérations plus ou moins bien dégagées, des exceptions assez nombreuses, d'où des règles peu systématiques. Les compositions (dont l'origine et la signification se perdent) glissent au niveau morphologique des mécanismes inconscients, dont le contrôle échappe au locuteur. En un mot, les catégories linguistiques prolifèrent, la morphologie se complique en laissant au locuteur

[34] Notons que l'écrit (ou tout autre support souple et non fugace de l'information : abaque, boulier, quipu, etc.) est particulièrement adapté à la prise de distance qu'implique toute activité de contrôle ou de critique, que celle-ci soit individuelle ou collective.

imaginatif la liberté de jouer des procédés sélectifs d'où surgissent de pittoresques nouvelles expressions (mille milliards de mille sabords !).

Les deux modes de fonctionnement ne semblent pas pouvoir se réguler spontanément l'un l'autre en raison du fossé qui sépare les préoccupations savantes et les besoins populaires. Lorsque Chuquet, par exemple, invente en 1484 la série des termes en *-illion,* ses travaux sur les progressions arithmétiques et géométriques font de lui un précurseur de la théorie des logarithmes. Pendant ce temps, *vingt* est encore en France un nombre d'appui (peut-être systématique) et les mathématiciens commencent à peine à accepter l'emprunt de million à l'italien. (GUITEL, 1975 : 567).

Apparemment, il faut une puissante volonté politique pour diffuser le savoir savant, comme celle des révolutionnaires de 1789 qui entreprirent, malgré les émeutes que cela provoquait, d'imposer le système métrique des poids et mesures et le système décimal des nombres (cf. MAREC, 1982 : 175-200). Une simple transposition didactique du savoir, bien que nécessaire, ne suffit pas à en assurer la diffusion et encore moins à garantir la prise de cette greffe chez les sujets.

La partie III, *Taxinomie des numérations parlées*[35], est consacrée à l'étude des aspects dynamiques susceptibles d'éclairer l'évolution des systèmes de numération et les principaux mécanismes de leur acquisition.

L'Herbergement, novembre 1984 et juillet 1985.

ANNEXE : Les enjeux de l'analyse morphologique.

> *"- Mon fils n'a jamais eu trois ans.*
> *- ?*
> *- Eh bien oui, à deux ans il savait dire 'deux' et réciter 'un, deux'. A quatre ans, il savait dire 'un, deux, trois, quatre' et 'j'ai quatre ans'. Mais pendant sa troisième année, il n'a jamais été capable de dire 'trois' ni de compter jusqu'à trois".*

0. - Introduction

A partir d'exemples d'analyse morphologique, cette annexe propose quelques réflexions sur l'origine de la notion de nombre et des processus susceptibles de conduire à sa mise en signes systématique.

---

35 A paraître dans *Chantiers Amerindia*.

0.1. *De l'origine de la notion de nombre.*

La mise en signes du nombre suppose que cette notion ait été conceptualisée (sans poser évidemment qu'il s'agit d'une conceptualisation nécessairement claire et distincte). En d'autres termes, les noms de nombre peuvent être considérés comme des traces d'un acte énonciatif de mise en signes, et, en tant que telles, elles renvoient à l'origine et à la construction de la notion de nombre.

Selon B. Pottier, le déclencheur (stimulus) est le monde de référence *"réel ou imaginaire, non-fini et non-discret",* c'est-à-dire un domaine d'expérience (celui du qualitatif) conceptuellement organisé en représentations (qui peuvent être déduites, construites, imaginées, transmises, empruntées,...) dont une caractéristique essentielle est leur constant renouvellement. *"L'émetteur doit en faire une saisie mentale pour sélectionner un certain nombre d'éléments de la perception :tout ce qui est imaginé ou perçu n'est pas dit. C'est le phénomène fondamental de la conceptualisation, ou réduction sélective de la référence"* (POTTIER, 1974 : 21). *"de toutes façons, le message produit n'exprime qu'une petite partie de l'intention de communication...*

$$\textit{communication} = \frac{\textit{CONTEXTE}}{\textit{SITUATION}} + \textit{message"} \text{ (id. : 25).}$$

La *conceptualisation* (ou réduction sélective de le référence) n'est pas seulement en amont de l'acte d'énonciation car le sujet est tour à tour émetteur et récepteur *"le locuteur, qui va de la conceptualisation au message, suit un processus onomasiologique, tandis que l'auditeur, qui part du message vers une conceptualisation, suit un processus sémasiologique"* (ibid. : 36).

Il est très difficile de cerner ces transformations du linguistique en conceptuel (et réciproquement). Dans une perspective sémasiologique, par exemple, *"il est difficile de savoir ce qu'est comprendre un texte. On conceptualise des tranches de discours, constamment remodelées par la conceptualisation des tranches suivantes. L'oubli d'une partie sensible du texte lu ou entendu est la condition même de la rétention mémorielle"* (ibid. : 36)

Cette difficulté est d'autant plus grande que le niveau conceptuel n'est guère accessible à l'observation directe, et qu'il est fermé, pour des raisons déontologiques évidentes, à l'investigation expérimentale systématique.

On est conduit 1) à maintenir fermement l'hypothèse conceptuelle selon laquelle *"la structure d'entendement, très profonde, lieu de la connaissance"* est par nature *"déliée des langues naturelles",* même si le *"stimulus conceptualisé"* peut et *"doit être encodé dans la langue naturelle pour pouvoir être exprimé"* (ibid. : 21) ; 2) à poser qu'il n'est pas possible de remonter des traces attestées aux processus conceptuels réels car ceux-ci ont été filtrés et sélectionnés, et, de plus pondérés dans la situation.

Ces rappels permettent de poser le problème de l'origine de la notion de nombre : étant donné un nom de nombre, que peut-on dire de la notion de nombre que présuppose cette mise en signes particulière ? Notre propos n'est pas d'engager un débat épistémologique sur l'origine de la notion de nombre, mais de préciser le statut que l'on peut accorder aux résultats de l'analyse morphologique des expressions numériques ; celles-ci étant considérées comme des traces de processus cognitifs qu'il convient de maintenir, méthodologiquement et sans paradoxe, séparées de leurs représentations linguistiques. Nous verrons en particulier que ce problème se pose dès la mise en signes du nombre un et du nombre deux.

0.2. *Du problème de la numération.*

Supposons maintenant que la notion de nombre ait été dégagée. Nous entendons par là que le sujet est capable de distinguer et de nommer un et deux. A priori, rien ne l'oblige à la conceptualisation d'autres nombres, et son système de numération est du type *'un, deux, beaucoup'.* C'est le cas par exemple des indiens Yanomami du Venezuela qui n'éprouvent pas le besoin d'exprimer d'autres conceptualisations numériques. Dans d'autres communautés, le sujet éprouve ce besoin, et se trouve confronte au problème de conceptualiser et de nommer un au-delà de deux, trois par exemple.

Notre propos est de présenter quelques solutions sémantiques attestées et de préciser leur statut d'un point de vue conceptuel. Nous verrons en particulier que les solutions sémantiques attestées conduisent à distinguer les mises en signes iconique et systématique. Dans le premier cas, il s'agit de nommer isolément un nouveau nombre ; dans le second, de construire une règle (voire un système) pour nommer simultanément plusieurs nouveaux nombres.

1. - L'analyse morphologique des noms de nombre en añún.

1.1. *Les données.*[36]

Les données sur la numération añún sont peu nombreuses (il s'agit d'une langue en voie de disparition) et se limitent à l'inventaire suivant :

1 **manaï**			
2 **piyomu**	4 **piinchi**	6 **piyaami**	8 **pínyour**
3 **apani**			
5 **haata**	7 **?**		

1.2. *Les analyses de M.F. Patte.*

"1. Seul **manai** *est sensible au genre et a une forme de masculin :* **manei**. *En outre, ce numéral peut s'apocoper sous la forme* **mana** *et s'utilise parfois comme article indéfini [....] de plus,* **mané** *'autre' montre une parenté évidente avec 'un' ; il est parfois précédé du déictique* **shi mané** *'cet autre' [....].*

2. On remarque que les signifiants des nombres pairs présentent entre eux une analogie que nous aimerions rapprocher du pronom de seconde personne : **pia**. *Le signifiant du nombre 'huit'* **pínyour** *attesté par A. Jahn commence également par* **pi.** *A tout le moins, cette analogie régulière de la syllabe initiale des référents numériques de deux en deux [....] permet d'inférer qu'ils contiennent très certainement un élément commun, ce que semblent corroborer d'autres données provenant de différentes langues de la même famille. [....] Il semble très probable que les nombres pairs aient été construits à partir du référent de deuxième personne [....].*

3. Quant au signifiant de 'trois', **apani**, *l'étymologie proposée par D. Taylor, qui rapproche le terme correspondant dans plusieurs langues arawak du mot référant à la main, pourrait s'appliquer à la langue añún (cf. 'main'* **aapï**) *[....]. Pour notre part, nous aimerions rapprocher* **apani** *de* **apïnincha** *'peu', où nous pouvons clairement reconnaître en terminaison, le suffixe diminutif'* (PATTE, 1985).

1.3. *Le nombre añún à l'état naissant.*

D'après les données précédentes, les noms de nombre sont inanalysables du point de vue arithmétique, et ne peuvent être soumis à une analyse syntaxique car ils n'entrent dans aucune expression numérique composée. En ce sens, il n'y a pas (ou plus) de système de numération añún (sauf à la reconstruire sur la base de données comparatives, par exemple avec le système guajiro de type arithmétique et à caractère décimal).

Du point de vue linguistique, les analyses morphologiques de Patte apportent un éclairage intéressant sur le nombre à l'état naissant. Elles suggèrent

[36] Les données présentées ont été recueillies sur le terrain par M.F. Patte (PATTE, 1985).

en effet, au moins au titre d'hypothèse, que la notion de nombre (au sens de capacité à distinguer un et deux) pourrait bien résulter, chez les Añún, de la conceptualisation de distinctions très primitives comme l'opposition un/autre, ou l'opposition je/tu. De plus, le fait que **manaïi** 'un' soit sensible au genre renforce, selon nous, l'idée que la conceptualisation des oppositions d'où surgit le nombre s'enracine profondément dans la conceptualisation d'oppositions plus primitives et véhiculées par la langue, les oppositions de genre, de nombre (grammatical) et de personne (masculin/non-masculin/pluriel, je/tu). Soulignons le fait que la mise en signes de la notion de nombre en añún engage la déixis et la personne énonciative.

1.4. *Reconnaissance et dépassement de la limite trois.*

1.4.1. En añún, il n'est pas absurde de rapprocher **apani** 'trois' de **aapï** 'main' et de **apïnincha** 'peu'. Sur la base de ces données, deux types de reconstruction peuvent être envisagés. Dans le premier, on rapproche **apani** de **apïnincha**, et on conclut que le modèle primitif de la numération añún serait le modèle *"un, deux, un peu"* dans lequel une limite est simplement reconnue. Dans le second, on rapproche **apani** de **aapi** 'main' (en tenant compte ou non du rapprochement avec **apïnincha** 'peu'). Dans ce cas, un problème sémantique se pose immédiatement : pourquoi la main renvoie-t-elle au nombre trois et non pas au nombre cinq ?

Cette question a été traitée par Taylor et la solution qu'il propose peut s'appliquer, selon Patte, à la langue añun. Elle consiste à reconstruire une forme ancienne composée de l'élément 'main' et d'un élément renvoyant à l'idée de partie principale (comme 'noyau', 'coeur', 'semence', etc.). Toute la force du raisonnement de D. Taylor provient du fait que le rapprochement n'est pas fortuit et limité à une seule langue, mais attesté dans un ensemble de langues et qu'il est compatible avec les contraintes phonétiques. *"Taken by itself, the resemblance between Lokono* **-kabo** *'hand' and* **kabuhin** > **kabuin** > **kabun** *'three' might well be passed over as a coincidence ; but it becomes significant on comparison with equivalent forms of related languages, such as Goajiro* **ahapu** *'hand' and* **apuni** *'three', Araua [...], Paumarí [...]"* (TAYLOR D. : 185).

1.4.2. D'autres procédés de mise en signes de la limite trois sont attestés. Ifrah signale des solutions iconiques (trois est synonyme d'un terme comme trèfle). On trouve plus souvent des solutions systématiques dans lesquelles trois est obtenu par concaténation de deux et de un (par exemple en andoke), ou encore des solutions que nous qualifions de techniques ou de culturelles parce que la valeur numérique des noms de nombre découle d'une technique énumérative par exemple, une gestuelle en papou (**anusi** 'un' 'petit doigt', **doro** 'deux' 'doigt (annulaire)', **doro** 'trois' 'doigt (majeur)') ou, comme en añún, une technique (de duplication ?) permettant de dériver tout un ensemble de nombres pairs dont les signifiants *"présentent entre eux une analogie que nous aimerions rapprocher du pronom de seconde personne".*

1.4.3. Malgré ses limites (on ne sait rien des nombres impairs, ni du second constituant des nombres pairs **pi-yomu**, **pi-inchi**, **pi-yaami**, **pi-nyour**) l'analyse morphologique de la numération añún apporte des éléments substantiels : a) à propos de l'origine de la notion de nombre, b) sur la reconnaissance et le dépassement de la première limite trois, c) sur la formation des premiers nombres pairs et d) sur le surgissement inattendu du nombre en relation avec la déixis et l'appréhension du corps au sens de la personne énonciative.

1.5. Dans les exemples précédents, l'analyse morphologique permettait d'établir des conjectures sur les stratégies hypothétiquement mises en oeuvre pour appréhender conceptuellement certains nombres et assurer l'unicité (cf. 0.12.) de leur représentation.

En retour, la connaissance des stratégies possibles a priori dans un contexte culturel donné permet parfois d'orienter les recherches morphologiques. On sait par exemple qu'un nombre peut être saisi comme la "somme" d'autres nombres (cf. 0.12. et 1.2.2.), et qu'il existe une stratégie désignée dans la littérature anglo-saxonne par l'expression *"representation by two equal or quasi-equal summands" (cf.* ZASLAVSKY, 1970: 350 et SEIDENBERG, 1960: 227).

Cette stratégie est attestée par la mise en signes de nombres jusqu'à une dizaine généralement, et se trouve souvent associée à une gestuelle des doigts particulière (comme par exemple celle des Deni-Dindjé du Canada que rapporte Lévy-Bruhl dans LEVY-BRUHL, 1921: 233). Elle se caractérise par le fait que les "additions" retenues portent sur deux nombres égaux (quand le nombre visé est pair) ou sur deux entiers successifs (quand le nombre visé est impair)

"A good example is the Ekoi language of the Cameroons :

3	**esa**	7	**eniresa**
4	**eni**	8	**enireni**
5	**elon**	9	**eloneni**
6	**esaresa**		

This variation is widespread east of Lake Tanganyika, through the Congo River area, and in the western Senegal [....] Evidence of this principle is found in the word for eight in Swahili" (ZASLAVSKY, 1970: 350).

Cette stratégie est différente de celle où les sommes sont formées après qu'un nombre d'appui particulier ait été choisi, bien qu'elle conduise pour certains nombres particuliers aux mêmes décompositions. Par exemple, en prenant cinq comme nombre d'appui, on aura respectivement :

"equal summands"	*"avec 5 comme nombre d'appui"*
6 = 3+3	6 = 5+ 1
7 = 4+ 3	7 = 5+ 2
8 = 4+ 4	8 = 5+ 3
9 = 5+ 4	9 = 5+ 4
10 = 5 + 5	10 = 5 + 5

Dans ce cas, ce serait par l'analyse de sis, sept et de huit que l'on pourrait distinguer les deux stratégies.

Les données culturelles précédentes conduisent à formuler l'hypothèse que le nombre huit **naana** en Kikongo (langue d'Afrique centrale) pourrait bien avoir été construit selon le principe de représentation en somme de deux termes égaux. Cette hypothèse implique que le signifiant de quatre soit **na**, ce qui n'est pas le cas en Kikongo où quatre se dit **ya**. L'hypothèse proposée ouvre ainsi le champ des recherches : la décomposition de **naana** est-elle possible ? L'alternance **na/ya** peut-elle être envisagée ? Pour C. Ntsadi l'hypothèse est crédible car dans d'autres dialectes kikongo quatre se dit **na** et *"on pourrait dire que* **naana** *'huit' dérive de //* **na + ø-a + na** *// dérivant lui-même de //* **na** 'quatre' + **na** *'et'* + **na** *'quatre'* //" (N'TSADI, 1984).

2. - Conclusion

Pour conclure cette annexe, nous dirons que l'analyse morphologique permet d'établir des conjectures sur les relations mutuelles les plus probables susceptibles d'exister (au sens de l'observation) entre le linguistique (les signes) et le conceptuel (les représentations cognitives internes). Il conviendrait d'accorder à ses résultats le statut général des procédures expérimentales une capacité à suggérer des hypothèses et un pouvoir de réfutation des théories.

Le résultat le plus notable, selon nous, consiste en ceci que les quelques exemples rapportés obligent à enraciner la notion de nombre très profondément dans la langue et la culture, en tenant compte des distinctions de genre, de nombre grammatical, de personne, et surtout, de l'appréhension du corps au sens énonciatif des personnes.

Les données présentées sont évidemment insuffisantes pour tirer des conclusions définitives. Elles suggèrent néanmoins que le référent primitif du nombre ne serait pas la multiplicité des choses, mais plus probablement l'appréhension de soi en tant que locuteur charnel. En d'autres termes, l'accès à la notion de nombre ferait sans doute intervenir une sorte de rupture dans la construction du référent imaginaire du je et du schéma corporel, une sorte de moment de symbolisation où *l'opposition je/tu,* par exemple, *en viendrait à représenter l'opposition un/deux.*

Un deuxième résultat de l'analyse morphologique concerne les modes d'instauration de la numération. Nous retiendrons d'abord qu'elle suppose posé le problème de la numération. On sait que ce problème se pose dès le nombre trois[37], limité qu'il s'agit de reconnaître et de nommer puis de dépasser. Ce dépassement n'a rien d'automatique comme l'attestent les communautés amérindiennes dans lesquelles le calcul (au sens arithmétique) n'est pas une valeur et encore moins un besoin : comme les Yanomami, on compte alors selon le modèle "un, deux, beaucoup". Quand il se fait, le dépassement est réalisé selon différentes modalités (iconique, arithmétique, ordinale, etc.) propres à chaque langue et à chaque culture ; mais chaque fois, le référent imaginaire le plus important semble bien être le je et le corps (au sens énonciatif des personnes). Ce n'est peut-être pas un "accident de la nature" qui conduit aussi invariablement un grand nombre de peuples à privilégier certains nombres d'appui (3, 5, 10, 20) ou certaines stratégies d'appréhension des nombres en relation avec telle ou telle gestuelle.

L'Herbergement, septembre 1985

[37] Ce point est obtenu indépendamment par Fisher, dans le cadre de ses recherches en didactique expérimentale des mathématiques (FISHER, 1985).

REFERENCES[38]

6. Documents

CAUTY, A. (1985) 'Sémantique de la mise en signes du nombre : une vision ordinale', *45è Congrès International des Américanistes*, Bogota (Colombie).

(1986) 'Contribution ethnoarithmétique à l'histoire des sciences, à propos de la numération maya' in *Sciences et Techniques en perspective,* n° 12, ed. par J. Dhombres, Université de Nantes.

GREENBERG, J. (1978) 'Generalizations about numeral systems' in *Universals of human language*, ed. by J.H. Greenberg, C. Ferguson and E. Moravcsik, Standford : University Press, p. 249-295.

GRASSERIE, R (de la) (1903) *Le nahuatl langue des aztèques conquérants du Mexique précolombien,* bibliothèque linguistique américaine, XXV, Nendeln (Liechtenstein) : Krauss reprint.

HARVEY, H.R. et WILLIAMS, B. (1981) 'L'arithmétique aztèque', *La Recherche* 126, p. 1058-1081.

MAREC, Y. et alt. (1982) 'L'introduction du calcul décimal et du système métrique à Rouen pendant la Révolution', *La Rigueur et le Calcul,* Paris : Cedic, p. 175-200.

MERRIFIELD, RW. (1968) 'Number Names in four Languagues of Mexico', *Grammars for Number Names,* edited by H. Brandt Corstius, Dordrecht (Hollande) : Reidel PuBlishing Company, p. 91-102.

NTSADI, C. (1985) *La numération parlée en Kikongo, étude des noms de nombre à la recherche de leur identité linguistique et lexicale*, Meaux : dactylographié, 62 p.

SEIDENBERG, A. (1960) 'The diffusion of counting practices', *University of California Publications in Mathematics,* Berkeley : University of California press.

TARLÉ, G. (1980) 'Método de matemáticas para la alfabetización en lengua quechua', *Revista de la Universidad Católica,* 25° año VIII, Ecuador : Pontificia Universidad Católica, p. 87-97.

[38] Pour compléter les références de la partie I (*Amerindia* n° 9).

TAYLOR, D. (19..) 'On the Etymology of some Arawakan words for three', *International Journal of American Linguistics, vol. XXI, p.* 185-187.

TAYLOR, G. (1982) *Aspectos de la dialectología quechua, introducción al quechua de Ferreñafe*, Chantiers Amerindia, Paris : Association d'Ethnolinguistique Amérindienne, 40 pp.

ZASLAVSKY, C. (1970) 'Black African traditional Mathematics', *The Mathematics Teacher,* edited by Howard Eves, University of Maine (USA), p. 345-356.

7. Auteurs cités

BRUNSCHVICG, L. (1912) *Les étapes de la philosophie mathématique,* Paris: Blanchard, édition 1981, 592 pp.

DHOMBRES, J. (1978) *Nombre, mesure et continu,* Paris : Cedic/Nathan, 338 pp.

JAKOBSON, R. (1956) 'Two Aspects of Language and two types of aphasic Disturbances', *Fundamentals of Language,* La Hague (Hollande) Mouton, p. 55-82.

LEVY BRUHL, L. (1921) *Les fonctions mentales dans les sociétés primitives,* Paris : Alcan, 9è édition, pp. 204-257.

MAZAUDON, M. (1985) 'Dzongkha number systems' in *Southeast Asian linguistic studies presented to André-G. Haudricourt*, ed. by Suriya Ratanakül, David Thomas, Suwilài Premsriat, Mahidol University, Bangkok (Thaïlande), 1985.

QUEIXALOS, F. (1982) 'Le regard et le réel', *Amerindia, 7,* Paris: A.E.A., p.85-105.

8. Communications personnelles

HOFF, B. J. (1983)

LAUNEY, M. (1985)

NTSADI, C. (1984)

PATTE, M.F. (1985)

YARANGA A. (1984)