LN 71/4

Josette LECOMTE

TRAITEMENT AUTOMATIQUE DE LA DISCONTINUITE DANS LE SYNTAGME VERBAL HOMOGENE EN ANGLAIS SCIENTIFIQUE ECRIT

Thèse de Troisième Cycle

Linguistique Appliquée

UNIVERSITE NANCY BIBLIOTHEQUE

Nancy 1971

Josette LECOMTE

LN71/

TRAITEMENT AUTOMATIQUE DE LA DISCONTINUITE DANS LE SYNTAGME VERBAL HOMOGENE EN ANGLAIS SCIENTIFIQUE ECRIT

UNIVERSITE NANCY BIBLIOTHEQUE

A V A N T - P R O P O S

Cette recherche fait suite à un précédent travail sur "le syntagme verbal homogène en anglais écrit" (1). Il était apparu alors, que la structure du syntagme verbal se prêtait assez bien à un traitement automatique. Seuls les "syntagmes verbaux homogènes" correspondant à des "séquences contiguës réelles" avaient pu être traités; un organigramme de reconnaissance des syntagmes dans un texte quelconque avait été établi (2), et un programme d'analyse automatique des différentes catégories véhiculées avait pu être mis au point (3).

Restent à étudier les problèmes beaucoup plus complexes posés par les autres types de syntagmes verbaux et de séquences verbales a-temporelles (4), en particulier leur identification dans un texte quelconque, et la reconnaissance des catégories verbales.

La présente étude traite de la discontinuité des seuls syntagmes verbaux homogènes.

Notes :

(1) cf. notre publication : "Le syntagme verbal homogène en anglais écrit. Etude des problèmes en vue de l'analyse automatique". Publications Linguistiques de la Faculté des Lettres et des Sciences Humaines de Nancy, n° 6, 1967. (abrev. SVH).

(2) cf. SVH page 82

(3) cf. SVH pages 97 - 99, suivies d'un dépliant. Les résultats-machine obtenus depuis, à la suite d'une application expérimentale de ce programme sont donnés en Annexe 4 de c cette thèse.

(4) Sur les définitions et la distinction que nous établissons entre "syntagmes verbaux" et "séquences verbales", cf. infra "Rappel des définitions", pages 2 et suiv.

I N T R O D U C T I O N

1. RAPPEL DES DEFINITIONS

1.1. Terminologie

1.1.1. Nous appelons "séquence de formes verbales" une suite de formes verbales (mots graphiques), sans préjuger du mode de leur regroupement (1).
Exemples :

- The meridians continue uninterrupted across the map.
- This is pushing simplification too far.
- The glossary and grammar that we have developped cannot be transferred directly to any other language.

1.1.2. Nous appelons "syntagme verbal" toute séquence de formes verbales qui remplit la fonction dite "nodale" dans une proposition donnée, et porte à ce titre les marques des catégories primaires, notamment celle du temps grammatical(2).
Exemples :

- He became acquainted with Ostillio Ricci.
- They are being taught French

1.1.3. " Un syntagme verbal (...) est la résultante d'une structuration

a) homogène : les éléments qui entourent

---

(1) cf. SVH page 23
(2) cf. SVH pages 2-3, § II

le verbe, sans qu'interviennent les démarcateurs

b) hétérogène : la coordination " (1)

" Dans un complexe verbal, le dernier élément est dit "auxilié", les précédents "auxiliants". Tous les verbes de la langue peuvent être auxiliés, quelques uns seulement peuvent être auxiliants. D'où la division des verbes de la langue en "auxiliaires" (auxiliants et auxiliés) et verbes (auxiliés seulement)". (2)

Notre définition du "syntagme verbal homogène" est plus restrictive que celle de Monsieur Pottier :

- Nous n'y incluons pas l'objet nominal du verbe (3).

- Tout démarcateur (coordonnant ou préposition, sauf "to" dans les cas suivants : "ought to", "have to", "be to", "used to") est une limite à ce syntagme verbal homogène.

- Nous considérons le syntagme verbal homogène en anglais écrit comme étant composé uniquement d'un auxilié (4), précédé ou non d'auxiliaires (5).(6)

---

(1) cf. B.POTTIER "Introduction à l'étude des structures grammaticales fondamentales", 2° édition. Publications Linguistiques de la Faculté des Lettres et des Siences Humaines de Nancy, n°1, 1964 -- page 18, §11.3

(2) cf. B.POTTIER, o.c. page 19, § 12.2

(3) cf. B.POTTIER, o.c. page 18, § 11.1

(4) cf. SVH page 36, § 1.2

(5) cf. SVH page 35, § 1.1

(6) Sur le syntagme verbal homogène, voir aussi SVH page 33, § VII.

Exemples :

a) Syntagmes hétérogènes :

- He decided to come.
- He became acquainted with Ostillio Ricci.
- The colour map must select, identify and generalize in the same way as any other map of equivalent scale.

b) Syntagmes homogènes :

- They are being taught French.
- You will have been being surprized.

1.1.4. Nous appelons "Syntagme verbal homogène discontinu" toute suite verbale répondant aux critères de structuration du syntagme verbal homogène, mais dont le déroulement séquentiel est interrompu (à une ou plusieurs reprises) par l'insertion d'éléments externes.

Exemples :

- He is , therefore , specially fitted for the task.

  → Syntagme verbal homogène discontinu

- He will probably be pleasantly surprized at the wonderful colours ...

  → Syntagme verbal homogène discontinu

1.1.5. Une "Séquence verbale homogène a-temporelle" se différencie du syntagme verbal homogène par l'absence des marques du temps grammatical.

Exemples :

a) Séquence verbale a-temporelle continue :

- I happen recently to have been making some slight study of one of the great civilizations of the Ancient East.

b) Séquence verbale a-temporelle discontinue :

- The relief tends to be more determined by the tectonic structure of the subsoil.

1.1.6. L'expression "groupe verbal homogène" recouvre à la fois les syntagmes verbaux homogènes et les séquences verbales homogènes a-temporelles.

1.2. Les trois zones du syntagme verbal homogène (1)

Nous considérons que le syntagme verbal homogène comporte trois zones :

a) La zone A.T , qui est celle de l'auxiliaire temporel.

b) La zone A.A-T , qui est celle des auxiliaires a-temporels.

c) La zone N.A-T , qui est celle de l'auxilié a-temporel.

Exemple :

- No finer site could have been found.

zone A.T — could
zone A.A-T — have been
zone N.A-T — found

2. QUELQUES CONVENTIONS D'ECRITURE

Ving : forme verbale a-temporelle terminée par

(1) cf. SHV pages 49-50

les trois lettres -ing.
(exemples : "working", "hatching", ...)

Ved : forme verbale a-temporelle, dite "participe passé".
(exemples : "applied", "drawn", "made",..)

V : forme verbale a-temporelle, dite "infinitif".
(exemples : "make", "leave", ...)

⊕ : coordonnant ou élément de coordonnant discontinu.
(exemples : "and", "or", "either", ...

[xx] : "toutes les formes temporelles du verbe xx".
(exemple : be correspond à toutes les formes temporelles du verbe "to be", c'est à dire "am", "is", "are", "was", "were", mais pas "been", ni "be", ni "being").

3. CORPUS UTILISE.

3.1. Nos recherches prennent appui sur un Corpus de 1508 syntagmes verbaux homogènes discontinus, extraits de textes écrits en anglais scientifique.

Le nombre total de groupes verbaux homogènes recensés est d'environ 19000 à 20000; ils ont été relevés dans (1) :

a) deux manuels de géographie (niveau élémentaire), à l'usage d'écoliers anglais.

---

(1) Les références des ouvrages utilisés sont données en Annexe 6 de cette thèse, page 175.

b) des ouvrages ou revues scientifiques spécialisés : cartographie, mathématiques, géographie, biologie, botanique, archéologie, géologie, et physique.

De ces 20000 groupes verbaux homogènes, 1508 syntagmes verbaux homogènes discontinus ont été isolés :

- 418 dans les manuels à l'usage d'écoliers.
- 1090 dans les textes spécialisés.

3.2. Mode d'établissement du Corpus.

3.2.1. Au début, nous avions envisagé de faire traiter par ordinateur GAMMA 60 un CORPUS d'anglais scientifique de 132000 mots graphiques. Ce CORPUS étant codé, un programme EXSYN (1) permettait de reconnaître les groupes verbaux et d'en donner la liste. Pour des raisons indépendantes de notre volonté, un peu plus du tiers seulement de ces textes a pu être exploité. Le résultat a été une liste de 258 syntagmes discontinus, ce qui constituait une somme d'exemples beaucoup trop limitée.
Cependant, les conclusions obtenues à partir de ces 258 syntagmes verbaux homogènes discontinus sont données en Annexe 1 ("Statistiques Machine"), page 121.

3.2.2. Il a donc été nécessaire de continuer manuellement le relevé d'autres syntagmes. C'est ainsi que nous avons dépouillé deux manuels de géographie à l'usage d'écoliers.

a) Nous avons commencé par comptabiliser les différents syntagmes verbaux homogènes rencontrés dans les deux manuels. Il s'agissait de vérifier si la proportion

---

(1) Nous donnons en Annexe 5, page 171, un exemple des résultats que le programme EXSYN permettait d'obtenir.

entre les syntagmes verbaux continus et les syntagmes verbaux discontinus était la même que dans le précédent type de textes. Les statistiques-machine portant sur un total de 4388 syntagmes verbaux homogènes, nous avons compté un nombre à peu près équivalent de syntagmes verbaux homogènes dans le deuxième type de textes. La proportion s'est révélée sensiblement différente.

Une contre-épreuve portant sur 600 syntagmes homogènes supplémentaires, tirés de ces mêmes manuels, nous a montré que la proportion était cependant constante à l'intérieur des deux livres.

Le résultat de cette comptabilisation est donné en Annexe 1 ("Statistiques manuelles", § 1) page 127.

b) Ensuite, nous avons poursuivi le relevé des seuls syntagmes verbaux homogènes discontinus. Le nombre total obtenu a été de 418.

Les conclusions tirées à partir de ces 418 syntagmes sont indiquées en Annexe 1 ("Statistiques manuelles", § 2) page 127.

3.2.3. Le nombre total de syntagmes discontinus alors en notre possession était de 676. Ce Corpus de syntagmes verbaux homogènes discontinus nous est encore apparu insuffisant. C'est pourquoi nous avons décidé de l'augmenter une nouvelle fois, en relevant manuellement près de 400 syntagmes verbaux homogènes discontinus supplémentaires, dans la partie du CORPUS d'anglais scientifique qui n'avait pu être traitée par ordinateur. Ceci a porté notre total de syntagmes homogènes discontinus inventoriés à 1057.

3.2.4. Une dernière contre-épreuve a été effectuée sur d'autres textes scientifiques (1). Elle porte sur 451

---

(1) La liste des textes dépouillés est donnée en Annexe 6 page 175.

syntagmes discontinus supplémentaires. C'est à partir du nombre total de 1508 syntagmes homogènes discontinus qu'ont été établies les listes figurant en Annexe 2 (pages 141 et suivantes), et qu'a pu être réalisée la présente étude du syntagme verbal homogène discontinu.

4. CRITERES, METHODE, ET BUTS.

Les critères d'analyse sont de deux ordres. Il s'agit d'une part de rendre compte le plus fidèlement possible des faits de linéarité étudiés, et, d'autre part, de viser à l'efficacité.

Nous observons le comportement de l'unité syntaxique dans son contexte immédiat; aucune étude sémantique n'est envisagée : nous nous proposons seulement de recenser tous les rapports prévisibles entre syntagmes verbaux homogènes et chaines insérées, et de dégager les règles qui en permettront le traitement automatique.

Dans un premier temps, le problème de la démarcation automatique du syntagme verbal homogène est supposé résolu. L'étude qui suit porte donc sur des syntagmes verbaux isolés de leur contexte.

Exemples::

- ...⌈was locally superseded⌋...
- ...⌈Had the tilting envisaged by Shackleton occurred⌋...

début fin

Ceci nous a permis d'inventorier et de décrire les différents types d'insertion existant à l'intérieur du syntagme verbal discontinu, ainsi que les différents types de syntagmes verbaux admettant ces insertions.

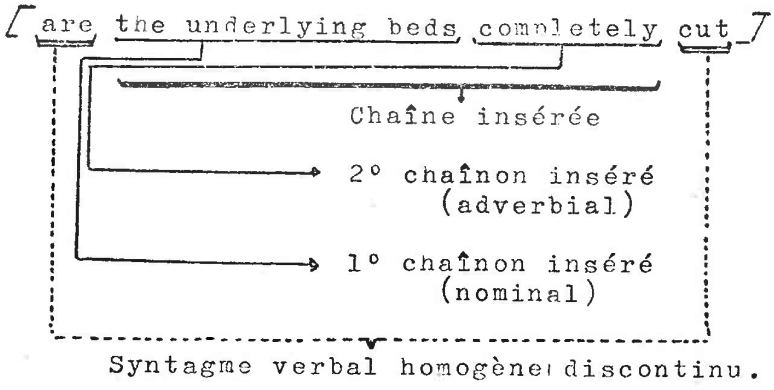
Dans un deuxième temps, à partir des conclusions obtenues, nous visons au traitement entièrement automatique du syntagme verbal homogène dans un texte quelconque (identification et regroupement des composantes du syntagme, et extraction automatique des catégories verbales).

LES CHAINES INSEREES

1. DEFINITIONS

1.1. Nous appelons "chaîne insérée" toute suite de mots graphiques provoquant la discontinuité du syntagme verbal homogène, et nous réservons le nom de "chaînon inséré" à toute partie constitutive d'une chaîne insérée (1).
Exemple :

- Since in neither case are the underlying beds completely cut out anywhere these breaks cannot be regarded as true unconformities.



1.2. Nous appelons "segments verbaux" les parties d'un syntagme verbal homogène disjointes par une ou plusieurs chaînes insérées. Le rapport du nombre des segments verbaux

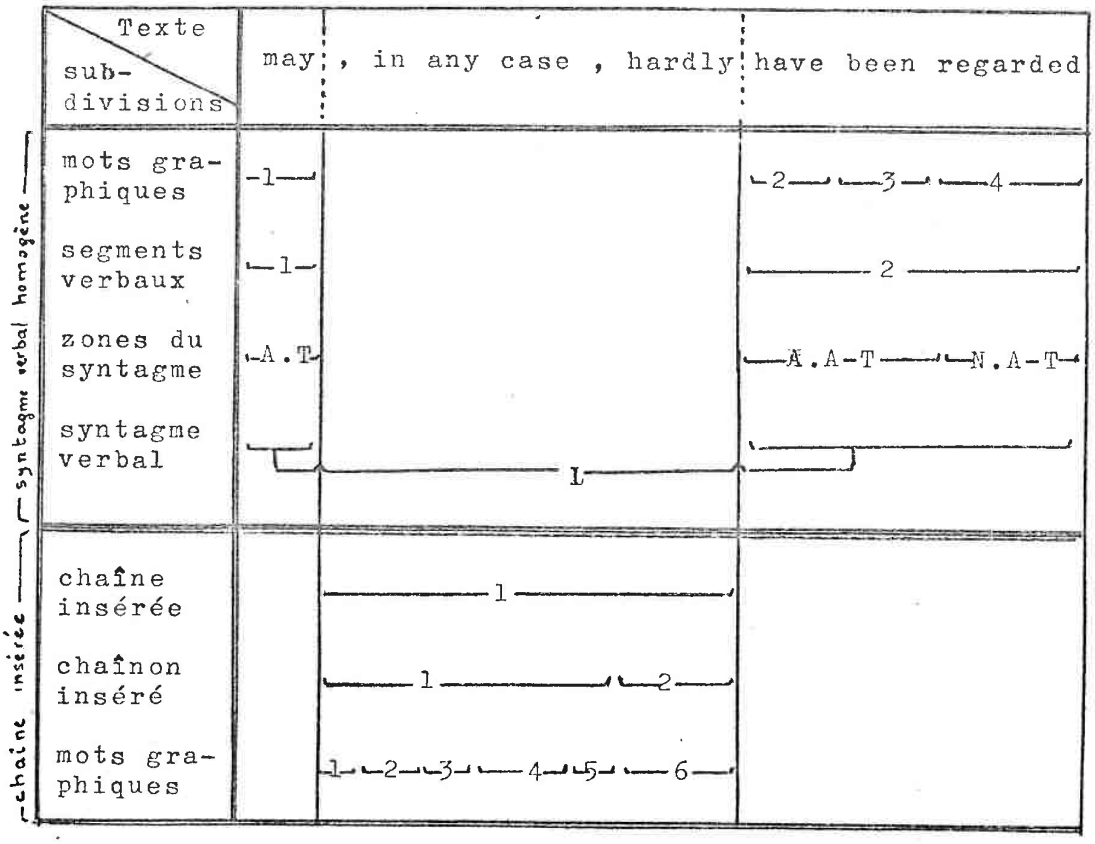
---

(1) Sur les critères de délimitation des chaînons insérés, cf. infra page 18.

au nombre des chaînes insérées est donné par la formule :

n chaines insérées → n+1 segments verbaux.

1.3. Tableau récapitulatif :

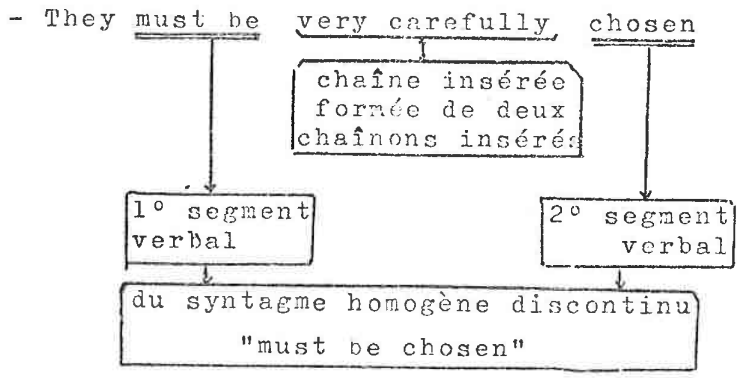


2. LES CHAINES INSEREES

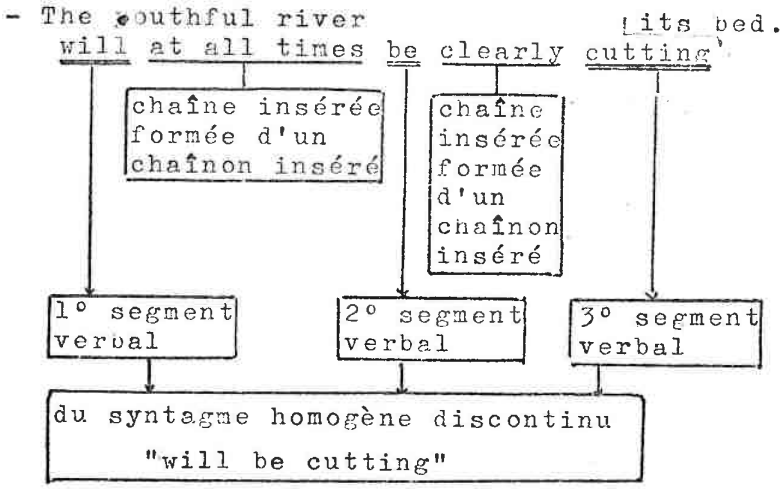
2.1. Nombre de chaînes insérées .

Un syntagme verbal homogène admet deux points d'insertion, donc peut accueillir au maximum deux chaînes insérées. Exemples :

A) Avec une seule chaîne insérée. Le syntagme verbal est alors subdivisé en deux segments verbaux.



B) Avec deux chaînes insérées. Le syntagme verbal est alors subdivisé en trois segments verbaux.



2.2. Lieux d'insertion

2.2.1. Dans un syntagme verbal homogène, il existe deux lieux d'insertion bien précis :

A) soit juste après l'auxiliaire temporel ou juste après la négation associée à cet auxiliaire.

Exemples :

a) - The same result [can still be obtained]

- The language [has in the meantime been outmoded] .

b) - Sheep [are not always found] on uplands.

- The thing [need not always have been named] directly.

B) soit juste avant l'auxilié a-temporel.

Exemples :

- The latter [might have been deliberately destroyed] at the time of his trials.

- It [may be all too readily assumed] that the sea has cut it.

2.2.2. Ces deux points d'insertion peuvent :

A) soit exister séparément, et être occupés simultanément ou non. C'est le cas lorsqu'il existe un ou plusieurs auxiliaires non temporels à l'intérieur du syntagme verbal homogène.

Exemple :

- The osmophilia [must now be partially distributed]

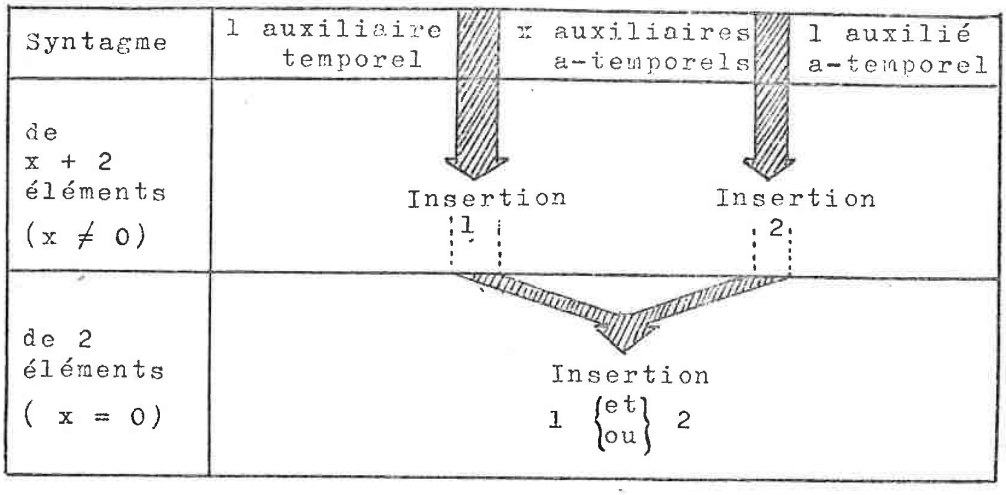
B) soit être confondus en un seul lorsque le syntagme verbal ne comporte aucun auxiliaire a-temporel.

Exemple :

- We [are more than fully occupied] with the elaboration of the project.

2.2.3. Schémas récapitulatifs

A)



B)

Syntagme	Insertion 1	Insertion 2	Discontinuité ?
plus de 2 éléments	+	-	oui. 2 segments verbaux
	-	+	oui. 2 segments verbaux
	+	+	oui. 3 segments verbaux
	-	-	non.
2 éléments	+		oui. 2 segments verbaux
	-		non.

2.3. Longueur des chaînes insérées.

2.3.1. La chaîne la plus courte comporte un seul mot graphique.

Exemples :

- /Let us see/ the limits of our region.
- Ornamental corners /should be strictly avoided/ .

2.3.2. La chaîne la plus longue comporte 16 mots graphiques, y compris les 2 virgules qui l'encadrent.

Exemple :

- In this way one /can, from the evidence presented by the spiral fabrics in the garnets of adjacent zones, designate/ the axial planes.

2.4. Composition des chaînes insérées.

2.4.1. La chaîne insérée peut être composée de :

A) Un seul chaînon inséré.

a) composé d'un seul mot graphique.

Exemple :

- /Does growth have to stop/ ?

b) composé de plusieurs mots graphiques.

Exemple :

- They /would one day be compared/.

B) Plusieurs chaînons insérés.

Exemples :

- Why /are volcanoes frequently found/ in such regions ?
- Not only /are the widely fluctuating trends of production accurately described/ , but the result is of manageable size

3. CONCLUSION

Il s'avère donc nécessaire de faire un inventaire des chaînons insérables et de la structure interne des chaînons.

Il faut ensuite étudier leurs combinaisons au sein d'une même chaîne insérée, et les influences que les deux chaînes exercent d'une part sûr leurs structures réciproques, et d'autre part, sur celle du syntagme verbal homogène.

INVENTAIRE DES CHAINONS INSERES

1. METHODE DE CLASSEMENT

1.1. Le classement que nous avons retenu est un classement en fonction de la structure linéaire du chaînon inséré.

1.1.1. C'est ainsi que nous avons été amenés à déterminer les bornes de chaque chaînon (bornes incluses dans chacun des chaînons).

A) Bornes de gauche.

Pour les déterminer, il suffit de recenser les chaînes insérées en fonction de leur premier mot graphique. Ce premier mot graphique peut être :

- un pronom
  - The boundaries /are all well defined/.
- une virgule
  - One /may , perhaps , be allowed/ to speculate.
- un substantif
  - /Does growth have to stop/ ?
- un adjectif
  - /Should large numbers be needed/ ...
- un introducteur de groupe nominal
  - /Do the mountains of the southern part of South America lie/ on the east or on the west of the Continent ?
- un adverbe

- Clinopyroxene /is typically associated/ with felspar fragments.

- une préposition

- It /should in any event be/ axiomatic that no chronological sequence can be regarded as established securely.

- un coordonnant

- They /have either directly or indirectly helped/ to identify ...

- une conjonction

- The compiler /should where possible secure/ the help of one or more critics.

- une locution adverbiale

- They /would no longer be compared/ to....

B) Bornes de droite

Pour les déterminer, nous avons recensé les derniers mots graphiques de chaque chaîne insérée. Nous avons retenu comme bornes :

- l'adverbe

- /Let us then simply assume/ that...

- la locution adverbiale

- The data /can at best indicate/ ...

- la virgule

- Only in one case /was a skull, minus the lower jaw ⊥ preserved/ intact.

- le pronom

- What rocks /do they flow/ over chiefly ?

- le substantif

- /Let the matter of impossibility be described/ first.

- le coordonnant

- These lines /are either drawn/ through sections with identical values, or are interpolated proportionnally between them.

1.2 Les chaînons insérés ont été classés comme suit :

G.A1 : Chaînon formé d'un adverbe simple. Borne de droite et borne de gauche sont superposées.
Exemple :

- /Have you ever tried/ to find it ?

G.A2 : Chaînon formé d'une lexie adverbiale reconnue comme telle dès la consultation de dictionnaire. Elle est considérée comme formant un seul "mot". Donc, borne de gauche et borne de droite sont superposées.
Exemple :

- This /would no longer be compared/

G.N2 : Chaînon formé d'un pronom ou d'une lexie pronominale. Borne de gauche et borne de droite sont superposées.
Exemple :

- The granite, /is itself cut/ by dolerite dykes.

G.N1 : Chaînon formé d'un groupe nominal simple ou tronqué.

La borne de gauche est un adjectif, un substantif, ou un introducteur de syntagme nominal. La borne de droite est le premier substantif rencontré. Par conséquent, bornes de droite et de gauche sont superposées dans le cas où le chaînon G.N1 débute par un substantif, et distinctes dans les autres cas.
Exemples :

- What easy transport facilities /do this factory area enjoy/ ?
  (G.N1) (G.N1)

G.V1 : Chaînon limité par deux virgules. La borne de gauche est une virgule, et la borne de droite est la virgule suivante. Jamais les deux bornes ne sont superposées.
Exemple :

- ..and /will, we feel, prove/ valuable.

G.V2 : Chainon limité par des parenthèses ou des crochets. La borne de gauche est donnée par le signe "ouvert" et la borne de droite par le signe "fermé". Jamais les deux bornes ne sont superposées.
Exemple :

- Nor /can a moraine topography (Fig. 17c) be shown/ on the map.

G.C1 : Chaînon débutant par une préposition autre que "of". La borne de gauche est cette préposition, et la borne de droite est le premier substantif ou le premier pronom rencontré. Jamais les deux bornes ne sont superposées.
Exemple :

- /Do the reserve-containing cells in the stem function/ in the same way ?

G.C2 : Chaînon débutant par la préposition "of".

La borne de gauche est "of", et la borne de droite est le premier substantif ou le premier pronom rencontré. Jamais les deux bornes ne sont superposées. (1)
Exemple :

- Only towards the Norwegian frontier /does the height <u>of the land</u> exceed/ 3,000 feet.

<u>G.I</u> : Chaînon formé d'un mot impersonnel, soit "there", soit "it", qui sont chacun leurs propres bornes.
Exemple :

-/Should <u>it</u> be/ possible to obtain a vanishinf first row ?

⊕ : Il s'agit du coordonnant, considéré comme un élément marginal, dont les deux bornes sont toujours superposées.

<u>1.3. Ce type d'analyse permet de résoudre</u>, sans tests de structure complexes, le problème des chaînes insérées dans un syntagme verbal homogène.
Il couvre la grande majorité des cas (1502 sur 1508). Cependant, il reste un certain nombre d'exemples (6 sur 1508) pour lesquels il faudrait recourir à une analyse autre que purement linéaire.

<u>1.3.1.</u> L'un de ces cas a été résolu en fixant arbitrairement la borne de droite au mot suivant la borne de gauche. Ce type de séquence, constitué d'une conjonction suivie d'un énoncé tronqué, a été appelé G.C3. Cet exemple a été rencontré une seule fois. Mais nous pensons qu'il est bon de prévoir la possibilité d'une telle chaîne insérée dans le syntagme verbal homogène discontinu.

---

(1) Sur G.C2, voir aussi <u>infra</u> page 26

Exemple :

- The compiler /should where possible secure/ the help of one or more critics.

1.3.2. Si notre analyse en bornes a permis de décomposer artificiellement la majorité des groupes nominaux hétérogènes en groupes simples ou tronqués, il reste 5 exemples qui demanderaient une approche et un type de classification faisant appel à une analyse syntaxique fonctionnelle (test du groupe nominal dans toute sa complexité).

Nous avons supposé déjà résolue l'identification de ces groupes nominaux hétérogènes (que nous appelons G.N3 dans nos organigrammes). Tous commencent par un chaînon G.N1 dont la borne de gauche est un introducteur de syntagme nominal. Mais la borne de droite du chaînon G.N3 ne peut être trouvée selon nos critères. Tout au plus aurions nous pu créer une série de "bornes de début de complication" :

- les relatifs
- les formes verbales Ved
- les formes verbales Ving
- les suites "to" + formes verbales V

Les exemples de G.N3 rencontrés sont au nombre de 5, et représentent 4 schémas verbaux différents qui sont :

defectif	G.N3	V
[do]	G.N3	V
[be]	G.N3	V
[have]	G.N3	V

En voici la liste complète :

a) - /Should the water which is intended to give depth to the colours dry/ too quickly, it ...

b) - /Do the quantum mechanics which have come early in this century exist/ independent of our measurement ?

c) - /Do the rivers which flow through the manufacturing areas cross/ very much limestone ?

d) - /Had the tilting envisaged by Shackleton occurred/ , the magnetic directions ...

e) - Why /is the capability of such a verb to govern other parts of speech ignored/ or suggested only in randomly selected examples ?

1.4. En résumé, 1503 syntagmes verbaux homogènes discontinus (soit 203 schémas sur 207) peuvent être traités de façon économique et rapide par une analyse de la structure linéaire, en testant les bornes des chaînons insérés. 5 syntagmes verbaux homogènes (soit 4 schémas) seulement demandent une approche différente.
Il nous semble que cette méthode, bien que naïve, corresponde assez bien au critère d'efficacité que nous nous étions imposé.

2. PROBLEMES POSES PAR LA DELIMITATION DES CHAINONS.

2.1. Le chaînon G.N1 (1)

2.1.1. Lorsque nous observons le Corpus de syntagmes verbaux homogènes discontinus, nous voyons que, chaque fois que 2 substantifs sont contigus linéairement (sans signe de ponctuation ni coordonnant entre les deux) dans une même chaîne insérée, ils appartiennent au même syntagme nominal. Il nous serait donc possible de dire que G.N1 peut recouvrir

(1) Pour une étude plus détaillée du cnaînon G.N1, cf.page 46.

des groupes nominaux comportant un ou plusieurs substantifs contigus linéairement. Par exemple :

- How /do the cotton tenant farmer and the chacra-farmer pay/ their rent ?

→ * G.N1

2.1.2. Si nous envisageons, au contraire, l'identification automatique du syntagme verbal homogène discontinu dans un texte quelconque, nous sommes obligés d'adopter pour G.N1 une définition beaucoup plus restrictive. Hors des chaînes insérées, deux substantifs contigus linéairement n'appartiennent pas nécessairement au même syntagme nominal. Tant que les différents éléments du syntagme verbal discontinu n'ont pas été identifiés, et que les substantifs n'ont pas été reconnus comme étant dans une chaîne insérée, il est impossible de dire s'ils entrent dans le même groupe nominal.

Dans ces conditions, il est donc méthodologiquement préférable de définir G.N1 comme une suite d'éléments nominaux ou pré-nominaux de syntagme nominal, s'arrêtant immédiatement après le premier substantif du groupe. Le chaînon obtenu correspondra, soit à un syntagme nominal simple, soit à un syntagme nominal tronqué. Par exemple :

- How /do the cotton tenant farmer and the chacra-farmer pay/ their rent

→ G.N1 (syntagme nominal simple)

→ G.N1 G.N1 G.N1 (tous syntagmes nominaux tronqués)

Ce n'est que dans un second temps, et en application des règles d'identification du syntagme nominal, que pourront être réunies en un seul syntagme nominal les parties artificiellement disjointes.

2.2. Le chaînon G.C2 fait l'objet d'un traitement particulier, différent de G.C1.
Dans 2 cas seulement sur 43, le chaînon G.C2 débute une chaîne insérée (exemple 1). Dans tous les autres cas (exemple 2), il est placé derrière un substantif et constitue un élargissement du syntagme nominal (l'ensemble formant alors un syntagme nominal hétérogène).
Exemples :

1) - .../do not of themselves suggest/ ...

2) - Nor /is the role of the government in the channeling of economic activity limited/ .

C'est en vue d'éventuels regroupements ultérieurs du syntagme nominal hétérogène que nous avons considéré G.C2 comme différent de G.C1.

2.3. Le chaînon G.V1.

Nous avons décidé de considérer les deux virgules comme bornes du chaînon, sans préjuger du contenu de la suite de mots graphiques limitée par ces virgules. Cette suite de mots est de nature souvent complexe, et nous aurait posé souvent des problèmes de regroupement. Par exemple :

- /will , I suppose, feel/ ...

regroupement correct

regroupement incorrect

Il devenait donc plus rentable de tenir compte uniquement des virgules.

2.4. Le coordonnant ⊕

Une importante difficulté est due au phénomène de coordination à l'intérieur des chaînes insérées.
Toujours par souci d'efficacité et de rentabilité, nous avons décidé de considerer tout coordonnant (≠ de la virgule) comme un démarcateur de groupe, comme un "séparateur" gardant sa valeur propre, étant à lui-même ses propres bornes. Exemple :

- What /does the colour of the men and the women , and their clothes, suggest/ about the weather ?

(schéma du syntagme verbal discontinu :
do G.N1 G.C2 ⊕ G.N1 G.V1 V)

What /does the colour — G.N1
of the men — G.C2
and — ⊕
the women — G.N1
, and their clothes, : G.V1
suggest/
about the weather ?

2.5. La particule "NOT"

Bien que "not" ait été considéré jusqu'à présent comme ne provoquant pas la discontinuité du syntagme verbal (c'est à dire qu'il ne devrait pas en être tenu compte dans la chaîne insérée), il nous a paru préférable de mentionner sa présence à la fois

a) dans le schéma du syntagme, puisqu'il ne provoque pas de discontinuité;

b) et dans le schéma de la chaîne insérée,

puisqu'il influe sur sa composition.

3. IMPORTANCE RELATIVE DES DIFFERENTS CHAINONS.

3.1. Longueur

Les plus courts sont G.I, G.A1, ou G.N2, qui ne comportent qu'un seul mot graphique.

Les plus longs sont les groupes G.V1 qui peuvent comporter jusqu'à 16 mots graphiques, et les groupes G.N3 qui atteignent 12 mots graphiques.

3.2. Fréquence

Les plus fréquents sont les chaînons G.A1 (1314 cas)

Les moins fréquents sont G.C3 (1 cas), et G.V2 (2 cas).

4. LES DIFFERENTS CHAINONS

(La liste des différents schémas de chaînons, ainsi qu'un relevé des fréquences d'apparition est donné en Annexe 2, page 142. De même, une liste des différents schémas de syntagmes discontinus, accompagnée des fréquences d'apparition est donnée aussi en Annexe 2, page 144)

4.1. Le chaînon G.I

4.1.1. Ses bornes ont été définies page 22.

4.1.2. Le chaînon G.I a été trouvé dans les schémas de syntagmes homogènes discontinus suivants :

defectif	G.I	V
[have]	G.I	Ved
[be]	G.I	Ved
[be]	G.I	Ving
defectif	G.I	be Ved

donc, dans des syntagmes verbaux homogènes discontinus de deux segments verbaux, ayant ou non un auxiliaire a-temporel.

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif [have] [be]	= 0 = 1. {be	V Ved Ving

POINT D'INSERTION
DE LA CHAINE
CONTENANT G.I

4.1.3. Inséré dans le syntagme verbal homogène, G.I est toujours isolé (c'est à dire qu'il forme à lui seul une chaîne insérée complète).

4.1.4. Le schéma maximal de combinaison est par conséquent le suivant :

G.I

4.2. Le chaînon G.Vl

4.2.1. Ses bornes ont été définies page 21

4.2.2. Le chaînon G.Vl a été trouvé dans les schémas suivants :

defectif	G.Vl	V	
[be]	G.Vl	Ved	
[have]	G.Vl	Ved	
defectif not	G.Vl	V	
[do] not	G.Vl	V	
need not	G.Vl	V	
[have] not	G.Vl	Ved	
[be] not	G.Vl	Ving	
defectif not	G.Vl	G.Al	V

defectif	G.V1	G.A1		V
defectif	G.A1	G.V1		V
[be]	G.V1	G.A1		Ved
[be]	G.A1	G.V1		Ved
[be]	G.N1	G.V1		Ved
[be]	G.V1	G.A1	G.A1	Ved
[do]	G.N1	G.V1	⊕ G.N1	V
[do]	G.N1	G.C2	⊕ G.N1 G.V1	V
defectif	G.V1	be	Ved	
[have]	G.V1	been	Ved	
[have]	G.V1	not	been	Ved
defectif	G.V1	be	G.A1	Ved

c'est à dire dans des syntagmes discontinus de deux segments verbaux :

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
need {+ −} not		
[do] {+ −} not	= 0	V
[have] {+ −} not	= 1 {be / been}	Ved
def. {+ −} not		
[be] {+ −} not		

POINT D'INSERTION
DE LA CHAINE
CONTENANT G.V1

4.2.3. Le chaînon G.V1 se rencontre, soit isolé, soit combiné à d'autres chaînons dans une même chaîne insérée (juxtaposé ou coordonné). Etant donné la définition adoptée pour G.V1, ces autres chaînons ne seront jamais de même nature que lui.

Le chaînon G.V1 peut occuper n'importe quelle place dans la chaîne insérée (début, milieu, ou fin).

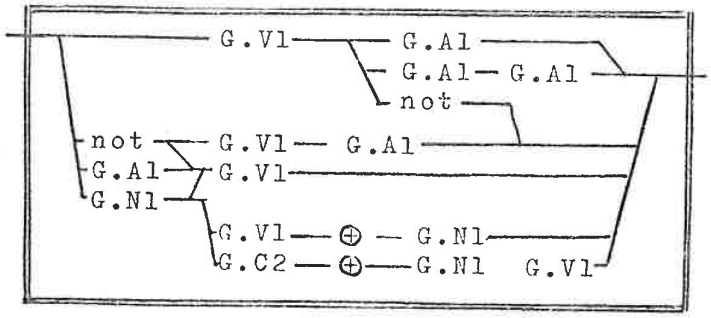
4.2.4. Schéma maximal de combinaison :

a) La zone A.A-T étant ≠ 0

Chaîne 1 ≠ 0, et Chaîne 2 ≠ 0

G.V1 (en Chaîne 1 uniquement)

b) La zone A.A-T étant = 0



4.3. Le chaînon G.V2

4.3.1. Les bornes du chaînon G.V2 ont été définies page 21. Ce groupe est très rare à l'intérieur du syntagme verbal homogène.

4.3.2. Il a été rencontré deux fois seulement, dans les deux schémas suivants :

defectif	G.N1	G.N1	G.V2	be Ved
[be]	G.N1	G.N1	G.V2	Ved

Dans les deux cas, il s'agit d'un syntagme verbal scindé en deux segments verbaux. Dans l'un des cas, le syntagme comporte un auxiliaire a-temporel.
Quand l'auxiliaire a-temporel est présent, le chaînon G.V2

se trouve dans la chaîne insérée au point 1 (juste après l'auxiliaire temporel):

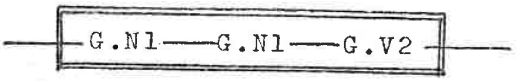
Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif [be]	= 0 = 1 {be	Ved

POINT D'INSERTION
DE LA CHAINE
CONTENANT G.V2

4.3.3. Le chaînon G.V2 n'a été trouvé que 2 fois, toujours accroché derrière G.N1 G.N1

4.3.4. Schéma maximal de combinaison :

(dans tous les cas, que la zone A.A-T soit égale à zéro ou différente de zéro)



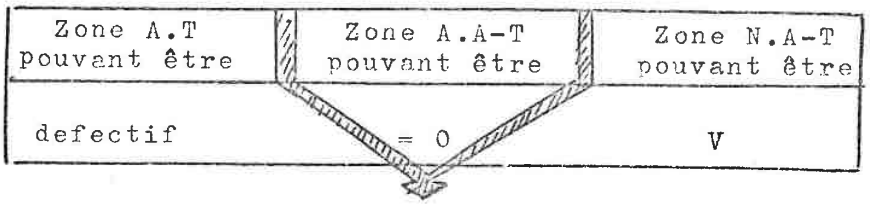
4.4. Le chaînon G.C3

4.4.1. Ses bornes ont été définies page 22, § 1.3.1.

4.4.2. Ce chaînon a été trouvé une seule fois dans le schéma suivant :

[defectif G.C3 V

donc, dans un syntagme verbal scindé en deux segments verbaux, ne comportant pas d'auxiliaire a-temporel:

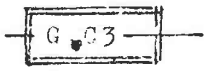


POINT D'INSERTION
DE LA CHAINE
CONTENANT G.C3

4.4.3. Inséré dans un syntagme verbal homogène discontinu, G.C3 est toujours isolé.

4.4.4. Schéma maximal de combinaison :

(la zone A.A-T étant = 0)



4.5. Le chaînon G.C2

4.5.1. Ses bornes ont été définies page 21

4.5.2. Le chaînon a été trouvé dans les schémas suivants :

defectif		G.C2	V			
[do]	not	G.C2	V			
[do]		G.N1	G.C2	V		
[do]		G.N2	G.C2	V		
defectif		G.N1	G.C2	not	V	
defectif		G.N1	G.C2	G.C2	V	
[do]		G.N1	G.C2	G.C2	V	
[do]		G.N1	G.C2	G.A1	V	
[do]		G.N1	G.C2	G.C1	V	
[be]		G.N1	G.C2	G.C1	Ved	
[be]		G.N2	G.C2	G.C1	Ved	
[be]		G.N1	G.C2	G.C1	G.C2	Ved
[do]		G.N1	G.N1	G.C1	G.C2	V
[do]		G.N2	G.C2	G.C1	G.C2	V
[do]		G.N1	G.C2 ⊕ G.N1	G.V1		V

let	G.N1	G.C2	be	Ved
[have]	G.C1	G.C2	been	Ved
defectif	G.N1	G.C2	have	Ved
[be]	G.N1	G.C2	to be	Ved
defectif not	G.N1	G.C2	be	Ved

donc, dans des schémas verbaux discontinus, scindés en 2 segments verbaux, comportant ou non un auxiliaire a-temporel ; lorsqu'il existe une zone A.A-T ≠ 0, la chaîne insérée contenant G.C2 est située au point 1 :

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif {+ / −} not	= 0	V
[do] {+ / −} not	= 1 { be, been, have, to be }	Ved
let		
[be]		
[have]		

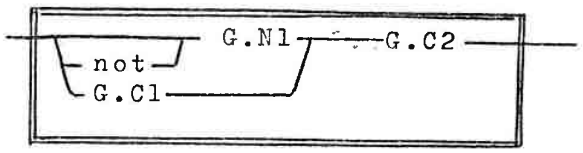
POINT D'INSERTION
DE LA CHAINE
CONTENANT G.C2

4.5.3. Le chaînon G.C2 existe rarement seul (2 exemples seulement; et dans l'un des cas, G.C2 est précédé de "not"). G.C2 suit presque toujours un substantif, (c'est à dire G.N1 ou G.C1, ou G.C2) ou un pronom (G.N2).

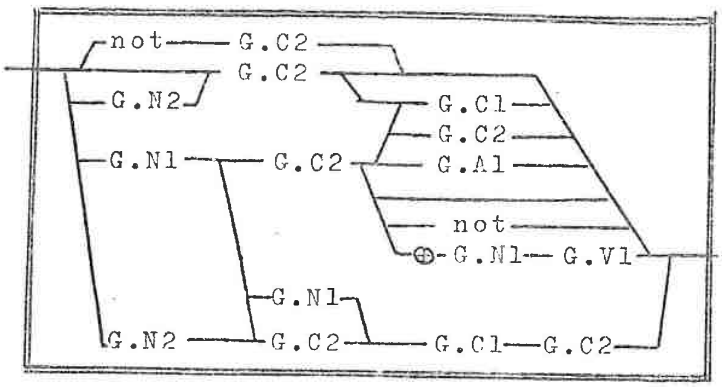
4.5.4. Schéma maximal de combinaison :

a) La zone A.A-T étant ≠ 0

(Chaîne 2 = 0)



b) La zone A.A-T étant = 0



4.6. Le chaînon G.C1

4.6.1. Les bornes du groupe G.C1 ont été définies page 21.

4.6.2. G.C1 a été rencontré dans les schémas verbaux suivants :

defectif		G.C1	V			
[have]		G.C1	Ved			
[be]		G.C1	Ved			
[do]	not	G.C1	V			
[be]		G.C1	not	Ved		
[do]		G.N1	G.C1	V		
[be]		G.N1	G.C1	Ving		
[be]		G.N1	G.C1	Ved		
[have]		G.C1	G.A1	Ved		
[be]		G.N1	G.C1	G.C2	Ved	
[be]		G.N1	G.C2	G.C1	Ved	
[be]		G.N2	G.C2	G.C1	Ved	
[do]		G.N1	G.C1	G.C1	V	
[do]		G.N1	G.C2	G.C1	V	
[do]		G.N1	G.N1	G.C1	G.C1	V
[do]		G.N1	G.N1	G.N1	G.C1	V
[do]		G.N1	G.N1	G.C1	G.C2	V
[do]		G.N2	G.C2	G.C1	G.C2	V
[be]		G.N1	G.C2	G.C1	G.C2	Ved
[do]		G.N1	G.C1	⊕ G.C1		V
[do]		G.N1	G.C1	⊕ G.N1		V

defectif	G.C1	be	Ved	
[have]	G.C1	been	Ved	
[have]	G.C1	G.C2	been	Ved
defectif not	G.N1	G.C1	be	Ved
defectif	G.N1	G.C1 ⊕ G.N1	be	Ved
defectif	G.C1	be	G.A1	Ving

donc, dans des schémas discontinus de deux ou trois segments verbaux, ayant ou non des auxiliaires a-temporels :

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif {+ / −} not	= 0	V
[do] {+ / −} not	= 1 {be / been	Ved
[be]		Ving
[have]		

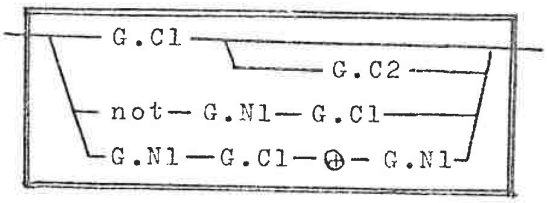
POINT D'INSERTION
DE LA CHAINE
CONTENANT G.C1

4.6.3. Le chaînon G.C1 existe seul ou combiné à d'autres chaînons dans une même chaîne insérée. Dans les cas où le syntagme verbal comporte deux chaînes insérées, seule la chaîne insérée au point 1 contient G.C1.
Nous remarquons également que plusieurs chaînons G.C1 peuvent se trouver dans une même chaîne (comme c'était déjà le cas pour G.C2).

4.6.4. Schéma maximal de combinaison :

a) La zone A.A-T étant ≠ 0

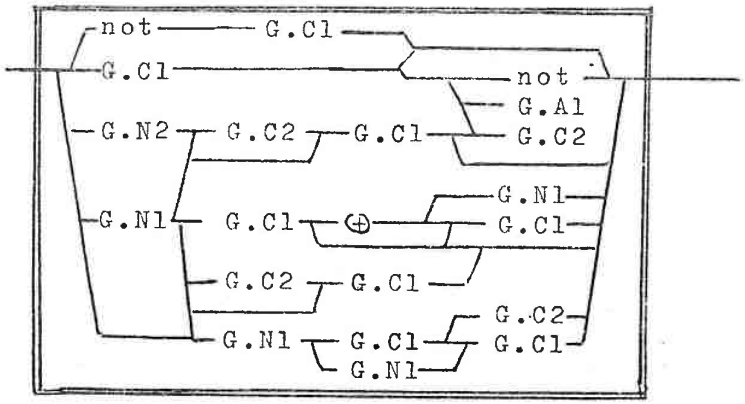
-- Chaîne 1 ≠ 0, et Chaîne 2 = 0



-- Chaîne 1 ≠ 0 et Chaîne 2 ≠ 0

G.C1 (en Chaîne 1 uniquement)

b) La zone A.A-T étant = 0



4.7. Le chaînon G.A2

4.7.1. Ses bornes ont été définies page 20.

4.7.2. Il a été trouvé à l'intérieur des schémas suivants :

defectif	G.A2	V
need	G.A2	V
[do]	G.A2	V
[be]	G.A2	Ved
[have]	G.A2	Ved
defectif	not G.A2	V
[have]	not G.A2	Ved
[be]	G.A2 not	Ved

defectif		G.A2	be	Ved
[have]		G.A2	been	Ved
[have]	not	G.A2	been	Ved
defectif		have	G.A2	Ved
defectif		be	G.A2 G.A1	Ved
defectif		G.A2	be G.A1	Ved

c'est à dire dans un syntagme verbal homogène discontinu scindé en deux ou trois segments verbaux, et contenant ou non un auxiliaire a-temporel :

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif {+ / -} not	= 0	V
need	= 1 {be / been / have}	Ved
[do]		
[have] {+ / -} not		
[be]		

2° POINT D'INSERTION DE LA CHAINE CONTENANT G.A2

1° POINT D'INSERTION D'UNE CHAINE CONTENANT G.A2

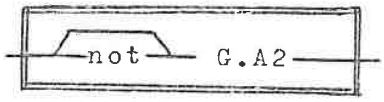
Il existe deux points d'insertion possibles pour le chaînon G.A2, mais ce chaînon n'apparaît pas en même temps dans les deux chaînes insérées.

4.7.3. Le chaînon G.A2 apparaît seul, ou associé seulement à "not" ou à G.A1.

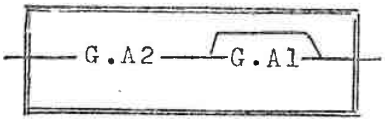
4.7.4. Schéma maximal de combinaison

a) la zone A.A-T étant ≠ 0

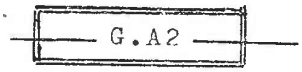
-- Chaîne 1 ≠ 0 et Chaîne 2 = 0



-- Chaîne 1 = 0 et Chaîne 2 ≠ 0

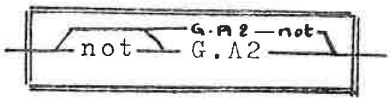


-- Chaîne 1 ≠ 0 et Chaîne 2 ≠ 0



(en Chaîne 1 uniquement)

b) la zone A.A-T étant = 0



4.8. Le chaînon G.A1

4.8.1. Les bornes du chaînon G.A1 ont été définies page 20.

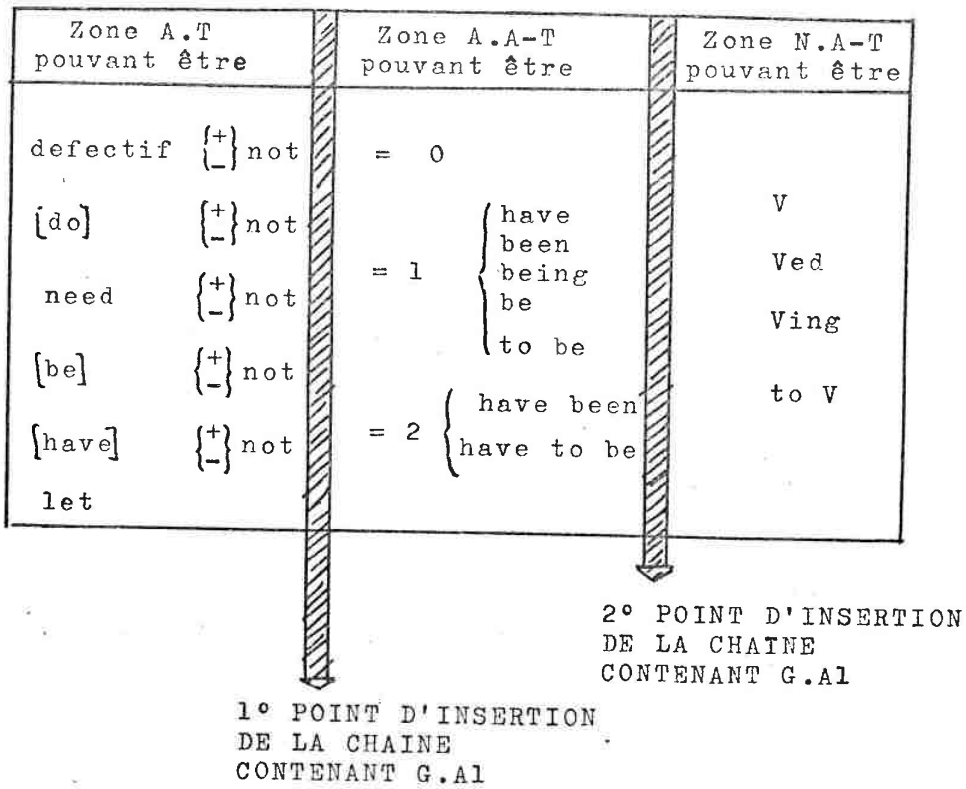
4.8.2. Il a été rencontré dans les très nombreux schémas suivants :

defectif		G.A1	V
[have]		G.A1	Ved
[be]		G.A1	Ved
[be]		G.A1	Ving
[be]		G.A1	to V
[have]		G.A1	to V
defectif	not	G.A1	V
[do]	not	G.A1	V

[have]	not	G.A1	Ved		
[be]	not	G.A1	Ved		
[be]	not	G.A1	Ving		
defectif		G.A1	not	V	
[be]		G.A1	not	Ved	
defectif		G.A1	G.A1	V	
defectif		G.V1	G.A1	V	
defectif		G.A1	G.V1	V	
let		G.N2	G.A1	V	
[do]		G.N1	G.A1	V	
[have]		G.A1	G.A1	Ved	
[have]		G.C1	G.A1	Ved	
[have]		G.A1	G.V1	Ved	
[have]		G.N2	G.A1	Ved	
[be]		G.V1	G.A1	Ved	
[be]		G.A1	G.V1	Ved	
[be]		G.N1	G.A1	Ved	
[be]		G.A1	G.A1	Ved	
[be]		G.A1	G.N1	Ved	
[be]		G.A1	G.A1	Ving	
[be]		G.N2	G.A1	Ved	
defectif	not	G.V1	G.A1	V	
[be]	not	G.A1	G.A1	Ved	
[be]	not	G.A1	G.A1	G.A1	Ving
[be]		G.A1	G.A1	G.A1	Ving
[be]		G.A1	G.A1	G.A1	Ved
[be]		G.V1	G.A1	G.A1	Ved
[be]		G.N1	G.C2	G.A1	Ved
[be]		G.N1	G.A1	G.A1	Ved
let		G.N2	G.A1	G.A1	V
[do]		G.N1	G.N1	G.A1	V
[do]		G.N1	G.C2	G.A1	V
[be]		⊕	G.A1		Ved
[have]		⊕	G.A1	⊕ G.A1	Ved
[be]		G.A1	⊕	G.A1	Ved
[be]		G.A1	G.A1	⊕ G.A1	Ved
[be]		G.A1	⊕	G.A1 G.A1	Ved
[have]		G.A1	been	Ved	
[be]		G.A1	being	Ved	
defectif		G.A1	be	Ved	
defectif		G.A1	be	Ving	
defectif		G.A1	have	Ved	
[have]		G.A1	had to	V	
[be]		G.A1	to be	Ved	

defectif	not	G.A1	be	Ved	
defectif	not	G.A1	have	Ved	
[have]	not	G.A1	been	Ved	
[have]		G.A1	G.A1	been	Ved
[have]		G.A1	G.N1	been	Ved
defectif		G.A1	G.A1	have	Ved
defectif		G.A1	G.A1	be	Ved
defectif	not	G.A1	G.A1	Be	Ved
[have]		G.A1	⊕ G.A1	been	Ved
defectif		G.A1	have	been	Ved
defectif		G.A1	have	been	Ving
defectif		G.A1	have	to be	Ved
need	not	G.A1	have	been	Ved
defectif	not	G.A1	have	been	Ved
defectif		have	been	G.A1	Ved
defectif		have	G.A1	Ved	
defectif		be	G.A1	Ved	
[have]		been	G.A1	Ved	
[be]		being	G.A1	Ved	
defectif	not	be	G.A1	Ved	
[have]	not	been	G.A1	Ved	
[have]		been	G.A1	G.A1	Ved
defectif		have	G.A1	G.A1	Ved
defectif		be	G.A1	G.A1	Ved
defectif		be	G.A2	G.A1	Ved
defectif		be	G.A1	G.A1	G.A1 Ved
defectif		be	G.A1 ⊕ G.A1	G.A1	Ved
defectif		be	G.A1	G.A1 ⊕ G.A1	Ved
[have]		been	G.A1 ⊕ G.A1		Ved
[be]		G.A1	being	G.A1	Ved
[have]		G.A1	been	G.A1	Ved
defectif		G.V1	be	G.A1	V
defectif		G.A1	be	G.A1	Ving
defectif		G.C1	be	G.A1	Ving
defectif		G.A1	be	G.A1	Ved
defectif		G.A2	be	G.A1	Ved
defectif		G.N2	be	G.A1	Ved
defectif		G.A1	be	G.A1	G.A1 Ved
[have]		G.N2	G.A1	be	G.A1 Ved

Le chaînon G.Al se rencontre donc dans des schémas beaucoup plus complexes que ceux des autres chaînons.
La zone d'auxiliaires a-temporels (zone A.A-T) peut être vide ou non, et le syntagme verbal homogène peut être scindé en deux ou trois segments verbaux.
De plus, le chaînon G.Al peut apparaître, soit à l'un ou l'autre des deux points d'insertion, soit aux deux, simultanément :

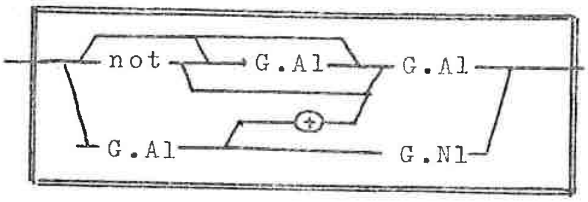


Le chaînon G.Al peut exister seul ou associé à d'autres chaînons dans une même chaîne insérée. Il leur est juxtaposé ou coordonné. De plus, les chaînons peuvent être de même classe que lui (d'autres G.Al) ou de classes différentes. En outre, G.Al peut occuper n'importe quelle position dans la chaîne insérée.

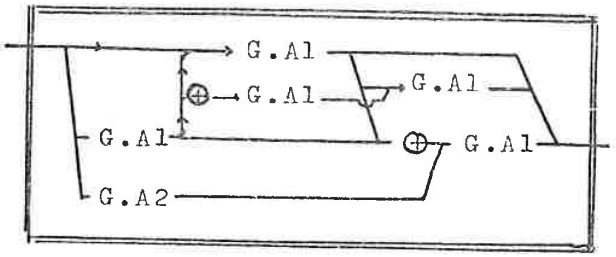
4.8.4. Schéma maximal de combinaison :

a) La zone A.A-T est ≠ 0

-- Chaîne 1 ≠ 0, et Chaîne 2 = 0

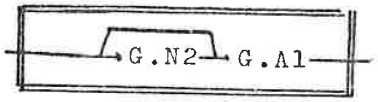


-- Chaîne 1 = 0, et Chaîne 2 ≠ 0

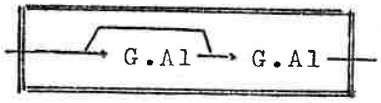


-- Chaîne 1 ≠ 0, et Chaîne 2 ≠ 0

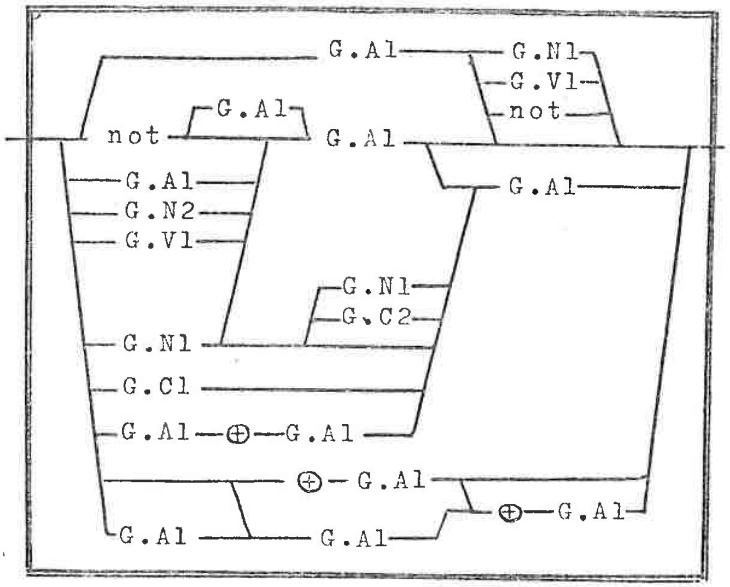
en Chaîne 1 :



en Chaîne 2 :



b) La zone A.A-T étant = 0



4.9. Le chaînon G.N2

4.9.1. Ses bornes ont été définies page 20.

4.9.2. Il a été rencontré dans les schémas verbaux suivants :

[be]		G.N2		Ving
defectif		G.N2		V
[have]		G.N2		Ved
[be]		G.N2		Ved
[do]		G.N2		V
let		G.N2		V
[be]		G.N2		to V
[have]		G.N2		to V
defectif	not	G.N2		V
[do]	not	G.N2		V
[be]	not	G.N2		Ved
defectif		G.N2	G.C2	V
[be]		G.N2	G.C2	Ved
[have]		G.N2	G.A1	Ved
let		G.N2	G.A1	V
[be]		G.N2	G.A1	Ved

let	G.N2	G.A1	G.A1	V	
[be]	G.N2	G.C2	G.C1	Ved	
[do]	G.N2	G.C2	G.C1	G.C2	V
defectif	G.N2		be	Ved	
let	G.N2		be	Ved	
[do]	G.N2		have to V		
[be]	G.N2		to be Ved		
let	G.N2	not	be	Ved	
defectif	G.N2	G.C2	have	Ved	
defectif	G.N2		be	G.A1	Ved
[have]	G.N2	G.A1	been	G.A1	Ved

donc, dans des syntagmes verbaux homogènes scindés en deux ou trois segments verbaux (la zone A.A-T pouvant être vide ou non).

De plus, le chaînon G.N2 n'apparaît qu'au premier point d'insertion :

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif {+ / -} not	= 0	
[do] {+ / -} not		V
[be] {+ / -} not	= 1 { been, be, have, to be, have to }	to V
let		
[have]		Ved

POINT D'INSERTION
DE LA CHAINE
CONTENANT G.N2

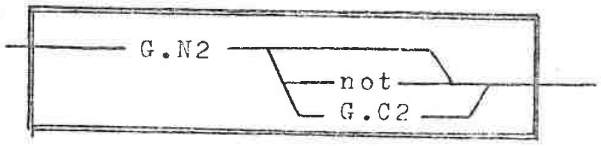
4.9.3. G.N2 apparaît seul ou juxtaposé à d'autres

groupes. Mais il occupe toujours la première place dans la chaîne insérée. Il ne se trouve jamais dans la chaîne insérée au 2° point.

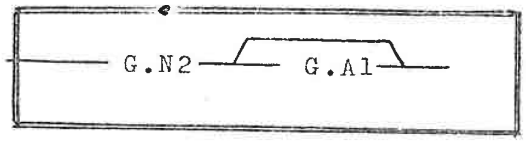
4.9.4. Schéma maximal de combinaison :

a) La zone A.A-T étant ≠ 0

-- Chaîne 1 ≠ 0, et Chaîne 2 =0

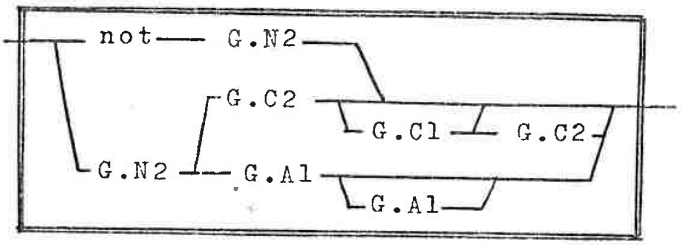


-- Chaîne 1 ≠ 0 et Chaîne 2 ≠ 0



(en Chaîne 1 uniquement)

b) La zone A.A-T étant = 0



4.10. Le chaînon G.N1

4.10.1 Ses bornes ont été définies page 20, et les problèmes posés par la délimitation de ce chaînon ont été évoqués page 24.

4.10.2. Il a été renconté dans les schémas verbaux suivants :

[do]	G.N1	V				
let	G.N1	V				
defectif	G.N1	V				
[have]	G.N1	Ved				
[be]	G.N1	Ved				
[be]	G.N1	Ving				
[do]	G.N1	G.A1	V			
[do]	G.N1	G.C1	V			
[do]	G.N1	G.N1	V			
[do]	G.N1	G.C2	V			
defectif	G.N1	G.N1	V			
[be]	G.N1	G.C2	Ving			
[be]	G.N1	G.C2	Ved			
[be]	G.A1	G.N1	Ved			
[be]	G.N1	G.A1	Ved			
[be]	G.N1	G.V1	Ved			
[be]	G.N1	G.N1	Ved			
[be]	G.N1	G.C1	Ved			
[do]	G.N1	G.N1	G.A1	V		
[do]	G.N1	G.C1	G.C1	V		
[do]	G.N1	G.C2	G.C1	V		
[do]	G.N1	G.C2	G.C2	V		
[do]	G.N1	G.C2	G.A1	V		
[be]	G.N1	G.N1	G.V2	Ved		
[be]	G.N1	G.C2	G.A1	Ved		
[be]	G.N1	G.C2	G.C1	Ved		
[be]	G.N1	G.C1	G.C2	Ved		
[be]	G.N1	G.A1	G.A1	Ved		
[be]	G.N1	G.C2	G.C1	G.C2	Ved	
[do]	G.N1	G.N1	G.C1	G.C2	V	
[do]	G.N1	G.N1	G.N1	G.C1	V	
[do]	G.N1	G.N1	G.C1	G.C1	V	
[do]	G.N1 ⊕	G.N1	V			
[be]	G.N1 ⊕	G.N1	Ved			
[do]	G.N1	G.C1 ⊕	G.C1	V		
[do]	G.N1	G.N1 ⊕	G.N1	V		
[do]	G.N1	G.C1 ⊕	G.N1	V		
[do]	G.N1	G.V1 ⊕	G.N1	V		
[do]	G.N1	G.C2 ⊕	G.N1	G.V1	V	
[do]	G.N1	G.N1	G.N1 ⊕	G.N1	V	
[do]	⊕ G.N1	⊕ G.N1	G.N1	⊕ G.N1	V	
defectif	G.N1	be	Ved			
[have]	G.N1	been	Ved			
[do]	G.N1	have to	V			
defectif	G.N1	G.C2	have	Ved		

defectif	G.N1	G.N1	be	Ved
let	G.N1	G.C2	be	Ved
[be]	G.N1	G.C2	to be	Ved
[have]	G.A1	G.N1	been	Ved
defectif not	G.N1	G.C2	be	Ved
defectif not	G.N1	G.C1	be	Ved
defectif	G.N1	G.N1	G.V2	be Ved
defectif	G.N1	G.C1 ⊕ G.N1		be Ved

c'est à dire qu'il n'apparaît que dans des syntagmes discontinus scindés en deux segments verbaux. La chaîne insérée contenant G.N1 ne se trouve jamais dans la chaîne insérée au 2° point (juste avant l'auxilié a-temporel):

Zone A.T pouvant être	Zone A.A-T pouvant être	Zone N.A-T pouvant être
defectif ± not do let have be	= 0 = 1 { have to be, been, have to be	V Ving Ved

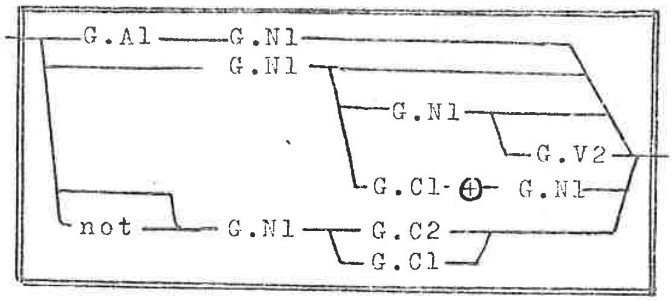
POINT D'INSERTION
DE LA CHAINE
CONTENANT G.N1

4.10.3. Le chaînon G.N1 apparaît seul ou associé à d'autres chaînons dans une même chaîne insérée, et il ne se trouve pas forcément en tête de cette chaîne insérée. Lorsqu'au premier point d'insertion du syntagme verbal, se trouve une chaîne comportant un ou plusieurs G.N1, il n'y a aucune insertion au point 2

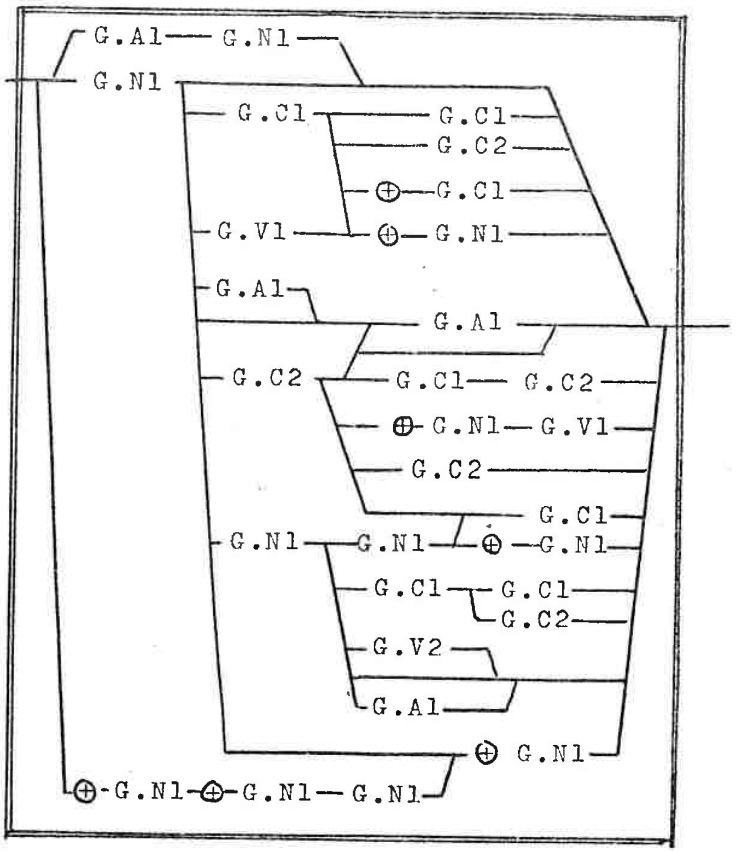
4.10.4. Schéma maximal de combinaison :

a) La zone A.A-T étant ≠ 0

-- Chaîne 1 ≠ 0, et Chaîne 2 = 0



b) La zone A.A-T étant = 0



4.11. Le chaînon G.N3

Nous l'avons déjà mentionné et décrit page 23, § 1.3.2.

Nous rappelons son schéma maximal de combinaison :

G.N3

5. CONCLUSION

Cet inventaire nous montre qu'il faut étudier :

a) les rapports entre les différents chaînons d'une même chaîne insérée (c'est à dire les rapports entre les différents chaînons de la Chaîne 1, et entre ceux de la Chaîne 2), afin de pouvoir éventuellement trouver des règles de compatibilité entre les différents chaînons.

b) Les rapports entre les deux chaînes insérées d'un même syntagme verbal, afin de déterminer les restrictions qu'elles s'imposent l'une à l'autre.

c) Les rapports entre les chaînes insérées et le syntagme verbal homogène, pour essayer d'établir des matrices de compatibilité / incompatibilité entre les différents types de syntagmes verbaux homogènes et les différents types de chaînes insérées.

d) Les rapports entre syntagmes verbaux, chaînes insérées et forme de la phrase.

COMBINATOIRE DES CHAINONS INSERES DANS UNE MEME CHAINE

1. NOMBRE, LONGUEUR, ET NATURE DES CHAINONS.

1.1. Nombre de chaînons dans une chaîne insérée :

1.1.1. Il existe des chaînes insérées formées d'un seul chaînon :

G.N3, G.I, G.C3 sont obligatoirement seuls dans la chaîne insérée,

G.V2 n'existe jamais seul

les autres chaînons peuvent être isolés ou non.

1.1.2. La chaîne insérée est souvent formée de plusieurs groupes. Dans le Corpus que nous avons analysé, le nombre maximum de ces chaînons est de 4.
Exemples :

- [be] G.N1 G.C2 G.C1 G.C2 Ved
- [do] ⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1 V

1.2. Longueur des chaînons :

1.2.1. Elle est fonction du chaînon lui-même :

Par exemple, un chaînon G.A1 n'est jamais composé que d'un seul mot graphique (un adverbe), alors qu'un chaînon G.N3 peut atteindre 12 mots graphiques (cf. exemple e) page 24).

1.2.2. Elle est fonction de la présence d'autres chaînons dans la chaîne insérée.

Par exemple, lorsque la chaîne insérée est réduite à un seul chaînon G.V1, ce dernier peut contenir jusqu'à 16 mots graphiques. Mais quand G.V1 est associé à

d'autres chaînons, sa longueur maximale n'est plus que de 7 mots graphiques.

Exemples :

- In this way one [can , from the evidence presented by the spiral fabrics in the garnets of adjacent zones , designate] the axial planes.

(16 mots graphiques)

- ..[will also , at least in some cases , be] subject to future modification.

(7 mots graphiques)

1.3. Nature des chaînons :

1.3.1. Une chaîne insérée peut être formée de plusieurs chaînons de même nature .

Exemples

- ...[have either directly or indirectly helped] ⊕ G.Al ⊕ G.Al

- Iceland [has only recently appeared] on the world tourist map. G.Al G.Al

1.3.2. Elle peut aussi être formée de chaînons de natures différentes.

Exemples :

- Why [are volcanoes frequently found] in such regions? G.N1 G.Al

- [Cannot some perfect system of translation be devised] ? G.N1 G.C2

2. POSITION DES CHAINONS

2.1. Position des chaînons dans la chaîne :

a) G.N3, G.C3, et G.I sont toujours en début de chaîne insérée, et non précédés de "not".

G.A2 est en début de chaîne insérée, précédé ou non de "not".

b) G.N3, G.C3, G.I, et G.V2 marquent toujours la fin d'une chaîne insérée.

c) Toutes les positions sont possibles pour les autres chaînons.

2.2. Position des chaînons les uns par rapport aux autres.

2.2.1. Chaîne composée de 2 chaînons :

Ils peuvent être :

a) Juxtaposés. Exemple :

- It was decided to use cultures in which the basidiocarps [would normally , with full illumination, require] about three days before the first discharge.

(schéma de la Chaîne : G.A1 G.V1)

b) Coordonnés. Exemples :

par un coordonnant simple :

- Some of the beds [are partly or completely dolomitized].

(schéma : G.A1 ⊕ G.A1)

par un coordonnant complexe :

- The upper more distinctive part [is conspicuously , though thickly bedded].

(schéma : G.A1 ⊕ G.A1)

par un coordonnant discontinu :

- They [have either directly or indirectly helped] to ...

(schéma : ⊕ G.A1 ⊕ G.A1)

2.2.2. Chaîne composée de plus de 2 chaînons :

Ils peuvent être :

a) Tous juxtaposés . Exemples :

- The valley [is ,however, particularly deeply cut].

(schéma : G.V1 G.A1 G.A1)

- [Do the mountains of the southern part of South America lie] on the east or on the west of the continent ?

(scnéma : G.N1 G.C2 G.C2)

b) Certains juxtaposés, d'autres coordonnés.

Exemples :

- Their functions [are sometimes inaccurately or incompletely described].

(schéma : G.A1 G.A1 ⊕ G.A1)

- The standard single-lens camera [is more ruggedly and accurately constructed] than the ordinary ground camera.

(schéma : G.A1 G.A1 ⊕ G.A1)

c) Tous coordonnés.

Nous n'avons pas trouvé d'exemple où tous les chaînons soient coordonnés.

3. ETUDE DE LA COMPOSITION D'UNE CHAINE INSEREE. (1)

3.1. Classement des chaînes insérées selon le nombre de chaînons qui y sont contenus. La présence de "not" et celle de coordonnants sont mentionnées séparément, devant les chaînons auxquels ces mots sont accolés.
(Il est à remarquer que, si "not" est présent, il n'y a pas de coordonnant dans la chaîne insérée)

3.1.1. Une chaîne insérée = un chaînon .

a) Chaînon seul :

- G.A1
- G.A2
- G.C1
- G.C2
- G.N1
- G.N2
- G.V1
- G.N3
- G.C3
- G.I

b) Présence de "not" :

avant :
- not G.A1
- not G.A2
- not G.C1
- not G.C2
- not G.N2
- not G.V1

après :
- G.A1 not
- G.C1 not
- G.N2 not
- G.V1 not

c) Présence de coordonnant :

- ⊕ G.A1
- ⊕

---

(1) La liste récapitulative des différentes chaînes insérées, ainsi que les indices de fréquence, sont donnés en Annexe 2 page 142.

(Nous avons trouvé plusieurs exemples où le premier élément d'un coordonnant discontinu ("either"), provoque une discontinuité du syntagme verbal homogène. Exemple :

- These lines are either drawn through sections with identical values, or are interpolated proportionally between them)

3.1.2. Une chaîne insérée = deux chaînons.

a) Chaînons seuls :

G.A1	G.A1
G.A1	G.N1
G.A1	G.V1
G.A2	G.A1
G.C1	G.C2
G.C1	G.A1
G.N1	G.A1
G.N1	G.C1
G.N1	G.C2
G.N1	G.N1
G.N1	G.V1
G.N2	G.A1
G.N2	G.C2
G.V1	G.A1

b) Présence de "not" :

avant :

not	G.A1	G.A1
not	G.N1	G.C2
not	G.N1	G.C1
not	G.V1	G.A1

après :

G.N1	G.C2	not

c) Présence de coordonnants :

G.N1	⊕	G.N1	
G.A1	⊕	G.A1	
⊕	G.A1	⊕	G.A1

3.1.3. Une chaîne insérée = trois chaînons .

a) Chaînons seuls :

G.A1	G.A1	G.A1
G.N1	G.A1	G.A1
G.N1	G.C1	G.C1

G.N1	G.C1	G.C2
G.N1	G.C2	G.A1
G.N1	G.C2	G.C1
G.N1	G.C2	G.C2
G.N1	G.N1	G.A1
G.N1	G.N1	G.V2
G.N2	G.A1	G.A1
G.V1	G.A1	G.A1
G.N2	G.C2	G.C1

b) Présence de "not" :

not	G.A1	G.A1	G.A1

c) Présence de coordonnants :

G.A1	G.A1	⊕	G.A1
G.N1	G.N1	⊕	G.N1
G.N1	G.C1	⊕	G.C1
G.N1	G.C1	⊕	G.N1
G.N1	G.V1	⊕	G.N1
G.A1	⊕	G.A1	G.A1

3.1.4. Une chaîne = quatre chaînons.

a) Chaînons seuls :

G.N1	G.C2	G.C1	G.C2
G.N1	G.N1	G.N1	G.C1
G.N1	G.N1	G.C1	G.C1
G.N1	G.N1	G.C1	G.C2
G.N2	G.C2	G.C1	G.C2

b) Présence de "not" :

Jamais "not" n'est apparu dans les chaînes insérées de plus de 3 chaînons.

c) Présence de coordonnants :

G.N1	G.N1	G.N1	⊕	G.N1
G.N1	G.C2	⊕	G.N1	G.V1
⊕	G.N1	⊕	G.N1	G.N1 ⊕ G.N1

3.2. Succession possible des chaînons :

3.2.1. Chaînes insérées à un seul chaînon.

	Chaînons seuls	Présence de "not"	
		avant	après
G.A1	+	+	+
G.A2	+	+	+
G.C1	+	+	+
G.C2	+	+	
G.C3	+		
G.I	+		
G.N1	+		
G.N2	+	+	+
G.N3	+		
G.V1	+	+	+
G.V2	-		
⊕	+		
⊕ G.A1	+		

3.2.2. Chaînes insérées à 2 chaînons

A) Chaînons juxtaposés :

2° chaînon / 1° chaînon	G.A1	G.C1	G.C2	G.N1	G.N2	G.V1	G.V2
G.A1	$+_a$			+		+	
G.C1	+		+				
G.C2							
G.N1	+	$+_a$	$+_c$	+		+	
G.N2	+		+				
G.V1	$+_a$						
G.A2	+						

a : cette chaîne peut être précédée de "not"

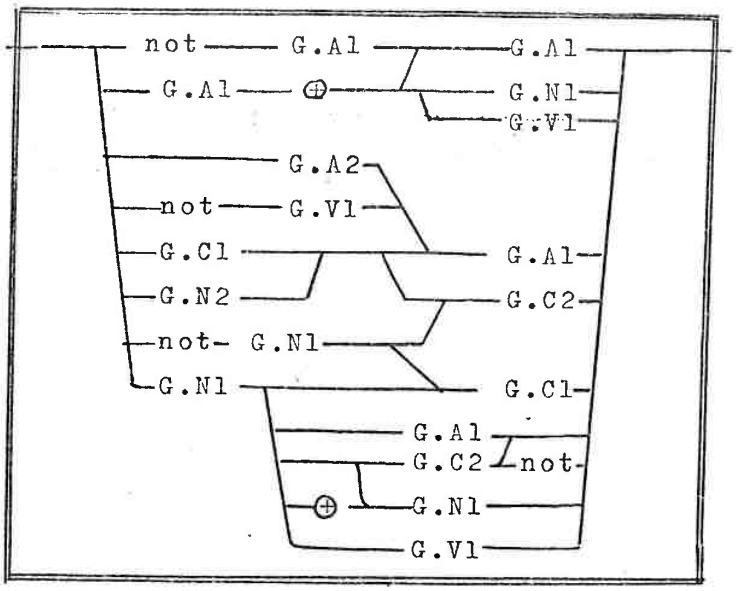
~~b : cette chaîne peut être suivie de "not"~~.

c : cette chaîne peut etre suivie ou précédée de "not".

B) Chaînons coordonnés :

2° chaînon / 1° chaînon	G.A1	G.C1	C.C2	G.N1	G.N2	G.V1	G.V2
G.A1	+						
G.C1							
G.C2							
G.N1				+			
G.N2							
G.V1							
G.A2							

C) Ce qui correspond à un schéma général des chaînes insérées de deux chaînons, qui est le suivant :



3.2.3. Chaînes insérées à trois chaînons

A) Les 2 premiers chaînons juxtaposés

2° chaînon / 1° chaînon	G.A1	G.C1	G.C2	G.N1	G.N2	G.V1	G.V2
G.A1	+						
G.C1							
G.C2							
G.N1	+	+	+	+		+	
G.N2	+		+				
G.V1	+						
G.A2							

a) Le 3° chaînon étant juxtaposé aux 2 premiers:

3° chaînon / 1° et 2°	G.A1	G.C1	G.C2	G.N1	G.N2	G.V1	G.V2
G.A1 G.A1	+						
G.N1 G.A1	+						
G.N1 G.C1		+	+				
G.N1 G.C2	+	+	+				
G.N1 G.N1	+						+
G.N1 G.V1							
G.N2 G.A1	+						
G.N2 G.C2		+					
G.V1 G.A1	+						

b) Le 3° chaînon étant coordonné aux 2 premiers:

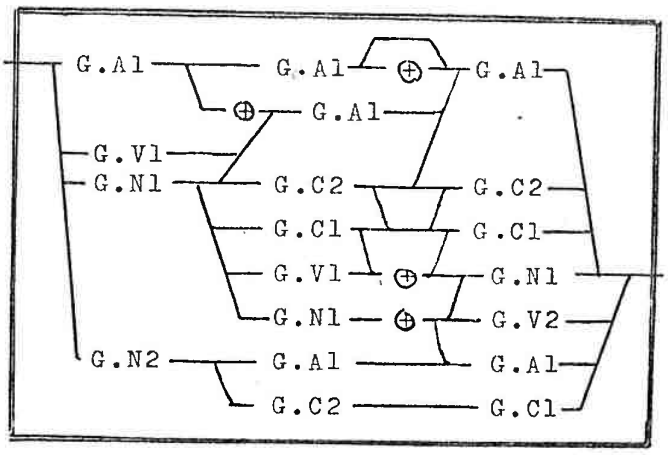
3° chaînon / 1° et 2°	G.A1	G.C1	G.C2	G.N1	G.N2	G.V1	G.V2
G.A1 G.A1	+						
G.N1 G.A1							
G.N1 G.C1		+		+			
G.N1 G.C2							
G.N1 G.N1				+			
G.N1 G.V1				+			
G.N2 G.A2							
G.N2 G.C2							
G.V1 G.A1							

B) Les 2 premiers chaînons coordonnés

Un seul schéma de ce type a été trouvé dans les chaînes insérées composées de 3 chaînons :

G.A1 ⊕ G.A1 G.A1

C) Schéma général de composition des chaînes insérées de 3 chaînons.



3.2.4. Chaînes insérées à quatre chaînons

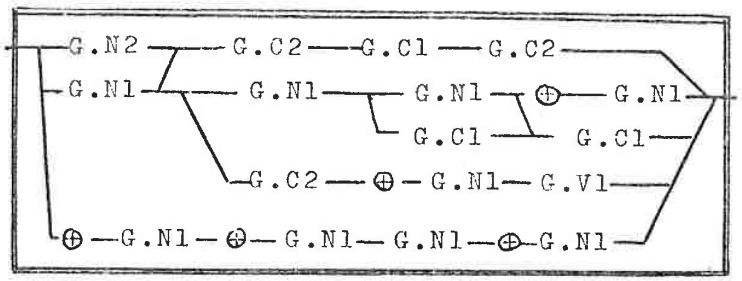
A) Sans coordination :

G.N1	G.N1	G.N1	G.C1
G.N1	G.N1	G.C1	G.C1
G.N1	G.N1	G.C1	G.C2
G.N1	G.C2	G.C1	G.C2
G.N2	G.C2	G.C1	G.C2

B) Avec coordination :

G.N1 G.N1 G.N1 ⊕ G.N1
G.N1 G.C2 ⊕ G.N1 G.V1
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1

C) Schéma général des chaînes composées de 4 chaînons insérés :

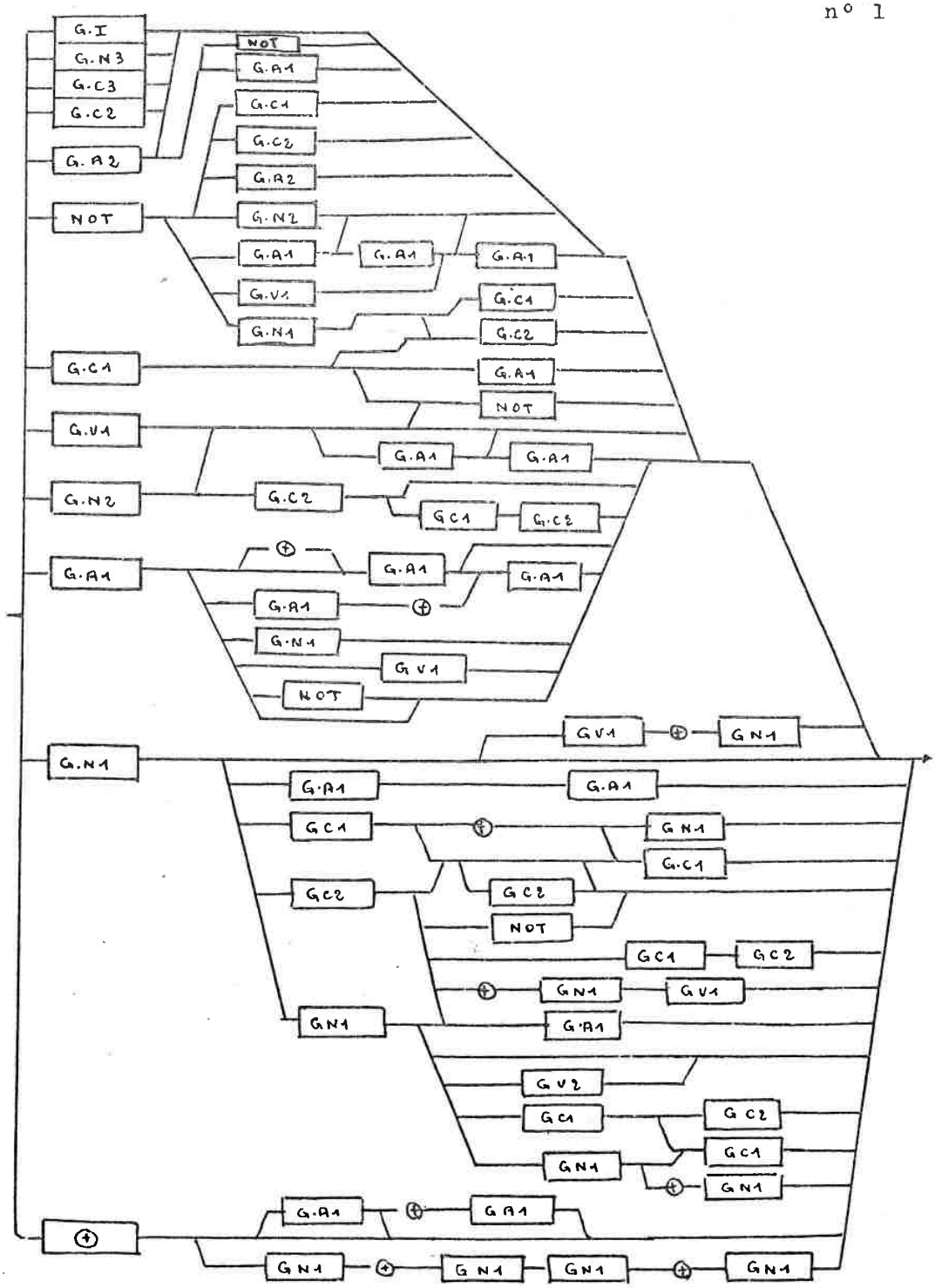


3.3. Organigramme général de prévision des chaînons dans une chaîne insérée . (Organigramme n° 1)

C'est une récapitulation des différents schémas donnés pages 59, 61, et 62.

Il donne un tableau de prévision à partir du premier chaînon de la chaîne insérée.

Organigramme
n° 1



RAPPORTS CHAINE 1 / CHAINE 2

1. RAPPEL

1.1. La Chaîne 1 est la chaîne insérée au 1° point, c'est à dire immédiatement derrière l'auxiliaire temporel.

La Chaîne 2 est la chaîne insérée au 2° point, c'est à dire immédiatement devant l'auxilié a-temporel.

1.2. Dans les syntagmes verbaux homogènes discontinus scindés en 2 segments verbaux, les 2 points d'insertion sont confondus en un seul. Nous appelons "Chaine unique" la chaîne insérée à ce point.

Le problème des rapports entre les deux chaînes insérées (Chaîne 1 et Chaîne 2) ne se pose que dans le cas où la zone A.A-T est ≠ 0

2. RELEVE DES CHAINES INSEREES QUAND LA ZONE A.A-T ≠ 0

Classement selon la structure interne des chaînes

2.1.1. Chaîne 1 ≠ 0 et Chaîne 2 = 0

G.A1

G.A1 G.A1

G.A1 G.N1

G.A1 ⊕ G.A1

G.A2

G.C1

G.C1 G.C2

G.N1

G.N1 G.N1

G.N1 G.C2

G.N1 G.C2 not

G.N1 G.N1 G.V2

G.N1 G.C1 ⊕ G.N1

G.V1
G.V1 not

G.I

G.N2
G.N2 not
G.N2 G.C2

not G.A1
not G.A2
not G.A1 G.A1
not G.N1 G.C2
not G.N1 G.C1

⊕

2.1.2. Chaîne 1 = 0, et Chaîne 2 ≠ 0

G.A2
G.A2 G.A1

G.A1
G.A1 G.A1
G.A1 G.A1 G.A1
G.A1 ⊕ G.A1
G.A1 ⊕ G.A1 G.A1
G.A1 G.A1 ⊕ G.A1

⊕

2.1.3. Chaîne 1 ≠ 0, et Chaîne 2 ≠ 0

Elles contiendront respectivement :

G.A1 et G.A1
G.A1 et G.A1 G.A1

G.A2 et G.A1

G.C1 et G.A1

G.N2 et G.A1
G.N2 G.A1 et G.A1

G.V1 et G.A1

3. LIMITATIONS RECIPROQUES DES 2 CHAINES INSEREES

3.1. Longueur et nombre des chaînons :

3.1.1. Au point 1 comme au point 2, les chaînes peuvent comporter 1 ou plusieurs chaînons insérés.

3.1.2. Il apparaît que la Chaîne 1 peut être beaucoup plus longue que la Chaîne 2. Ceci est dû à la nature différente des chaînons admis : un chaînon G.A1 comporte un seul mot graphique, alors qu'un chaînon G.V1 en comporte au moins trois.

a) Chaîne 2 maximale rencontrée :

[G.A1 G.A1 ⊕ G.A1
more faithfully and successfully

b) Chaîne 1 maximale rencontrée

1) [G.V1
, because of the reduction of the scale,

2) [G.N1 G.C1 ⊕ G.N1
reference to existing maps and published descriptions

3.1.3. Nous remarquons aussi que les longueur et nature respectives des deux chaînes insérées se conditionnent mutuellement.

Par exemple :

Quand Chaîne 1 = 0 et Chaîne 2 ≠ 0, la Chaîne 2 comporte de 1 à 3 chaînons insérés, soit de 1 à 4 mots graphiques.

Quand Chaîne 1 ≠ 0 et Chaîne 2 = 0, la Chaîne 1 comporte de 1 à 3 chaînons insérés, soit de 1 à 9 mots graphiques.

Quand Chaîne 1 ≠ 0 et Chaîne 2 ≠ 0, les 2 chaînes comportent respectivement 1 ou 2 chaînons (1 à 5

mots graphiques), et 1 ou 2 chaînons (1 à 2 mots graphiques). Jamais nous n'avons rencontré 2 chaînons en même temps dans chacune des 2 chaînes insérées (1) .

3.2. Composition des 2 chaînes insérées

3.2.1. L'examen des chaînes insérées relevées pages 64 et 65 montre que :

a) Les groupes G.C3 et G.N3 n'apparaissent jamais quand la zone A.A-T est $\neq 0$

b) Les autres groupes sont tous utilisés :

- G.V1, G.V2, G.N1, G.N2, G.C1 et G.C2 en Chaîne 1 uniquement

- G.A1, G.A2 en Chaîne 1 ou en Chaîne 2 (de même que le coordonnant isolé ⊕)

3.2.2. La Chaîne 1 a une composition variable, alors que la Chaîne 2 a une composition bien plus limitée, à base de G.A1 ou de G.A2.

La composition des deux chaines insérées est encore plus limitée quand les deux chaines existent simultanément dans le syntagme verbal homogène. Chacune des deux chaînes est alors formée d'un ou de deux chaînons simples, facilement identifiables.

Si la Chaîne 1 est seule présente, elle peut avoir un certain "poids" (de 1 à 3 chaînons); Si la Chaîne 2 est seule présente, elle peut avoir aussi une certaine importance (de 1 à 3 chaînons). Mais si les deux chaînes sont présentes en même temps, un certain équilibre s'établit entre elles; le "poids" est porté par l'ensemble des deux chaînes, et le nombre maximal de chaînons insérés dans le syntagme verbal homogène reste 3.

---

(1) cf. infra page 73 le tableau récapitulatif.

Exemples :

Chaîne 1 seule :

- Location cannot be given for any save a very few of the photos used, nor can reference to existing maps and published descriptions be made.

Chaîne 2 seule :

- ...must be more or less continuously readjusted ...

Chaîne 1 et Chaîne 2 présentes :

- Nor has it ever been adequately defined as such.

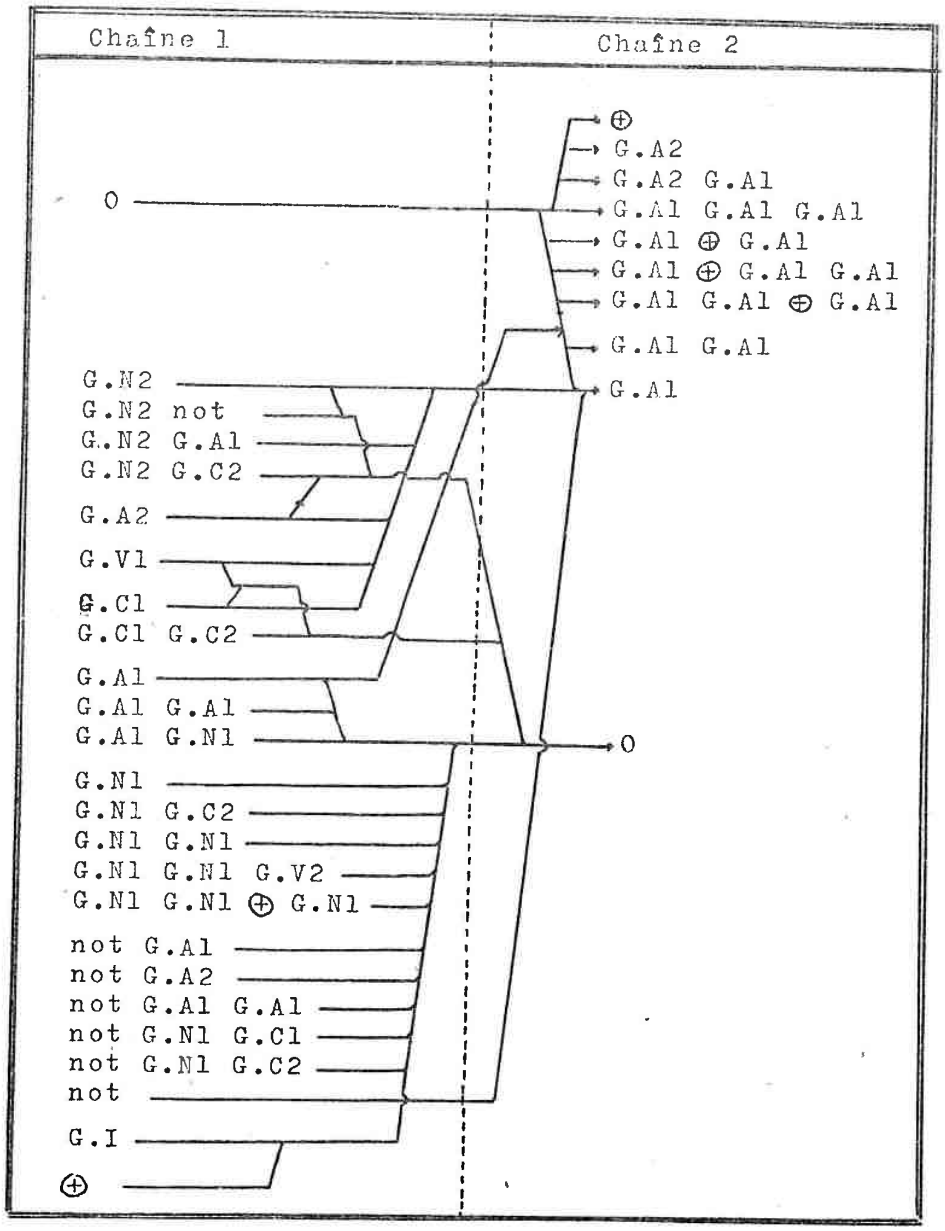
3.2.3. Chacun des chaînons inséré dans l'une ou l'autre des deux chaînes est relativement plus court et plus simple que ceux d'une Chaîne unique. Ainsi, G.V1, en Chaîne unique, peut atteindre 16 mots graphiques; mais G.V1, rencontré en Chaîne 1, a au maximum 9 mots graphiques.

3.2.4. Jamais la Chaîne 2 ne comporte de négation. "Not" n'apparaît qu'en Chaîne 1 ou en Chaîne unique, en début ou en fin de chaîne insérée (10 fois en fin, contre 64 en début, pour 1508 syntagmes verbaux homogènes discontinus) Des statistiques concernant les syntagmes discontinus à valeur négative sont données en Annexe 2 page 150)
Exemples :

- The thing need not always have been named directly.
- Let him not be called a liar.

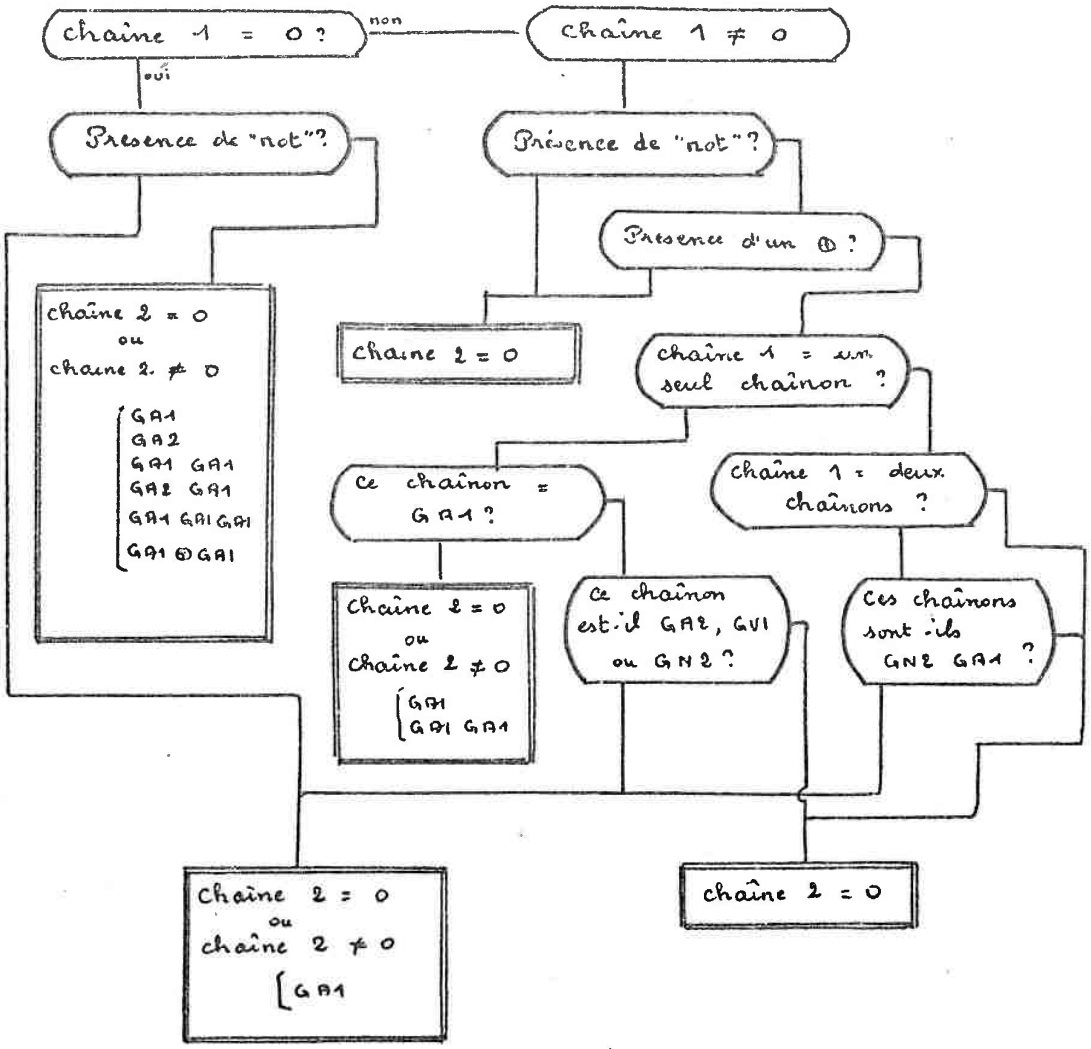
3.3. Prévision de la Chaîne 2 à partir d'une Chaîne 1 reconnue.

3.3.1. Tableau de prévision :



3.3.2. Organigramme de prévision de la Chaîne 2

Organigramme n° 2



4. COMPARAISON DES CHAINES 1 ET 2 AVEC LA CHAINE UNIQUE.

Les chaines insérées maximales que nous avons rencontrées sont des Chaînes uniques.

4.1. Peut on dire que la Chaîne unique est la somme des Chaînes 1 et 2 ?

4.1.1. Essayons de former une Chaîne unique avec une Chaîne 1 de longueur maximale, à laquelle nous ajoutons la Chaîne 2 de longueur maximale :

Chaîne 1	Chaîne 2
G.N1 G.C1 ⊕ G.N1	G.A1 G.A1 ⊕ G.A1

Il est impossible de trouver, dans la réalité, une Chaîne unique qui soit la somme de ces Chaînes : le nombre maximal de chaînons dans une chaîne insérée unique est 4.

4.1.2. Un autre test, par exemple, est d'essayer de regrouper 2 chaînes insérées telles que :

Chaîne 1	Chaîne 2
G.N2 not	G.A1

Ceci est encore impossible, "not" n'ayant jamais été rencontré en milieu de chaîne insérée.

4.1.3. En conséquence, il semble difficile de dire que la Chaîne unique soit la somme des Chaînes 1 et 2 quand la zone A.A-T = 0 .

4.2. Il semble plutôt que tout se passe comme si :

4.2.1. Un nombre maximal

a) de chaînes insérées soit permis à

l'intérieur du syntagme verbal homogène discontinu (ce nombre est 2).

b) de chaînons insérés soit permis à l'intérieur de ce même syntagme verbal homogène discontinu (ce nombre est 4).

La répartition en 1 ou 2 chaînes insérées dépend de la nature des chaînons insérés, de la présence ou de l'absence de "not", de la forme du syntagme verbal discontinu et de sa composition.

4.2.2. Toute suite de chaînons insérés peut devenir :

a) Une Chaîne unique quand la zone A.A-T = 0

b) Une Chaîne 1 quand la zone A.A-T $\neq$ 0. Dans certains cas, cette Chaîne 1 peut être déplacée du point 1 vers le point 2, et devenir une Chaîne 2, ou dissociée en Chaîne 1 et Chaîne 2.

4.2.3. La Chaîne unique n'est pas la somme des Chaînes 1 et 2; il est préférable de dire, au contraire, que ces dernières proviennent de l' "éclatement" de la chaîne insérée unique.

4.3. Tableau récapitulatif

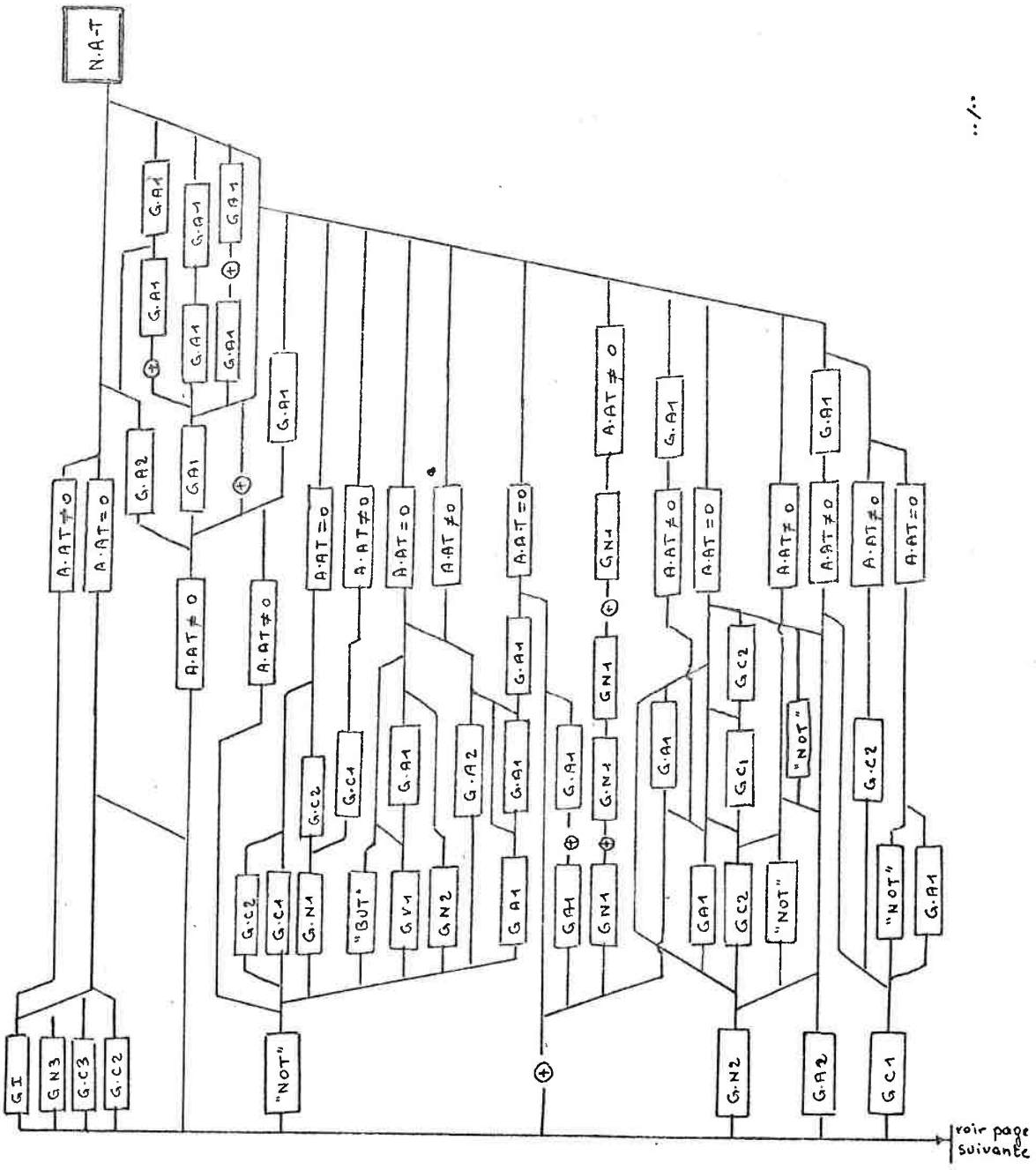
Zone	Chaînes insérées
A.A-T = 0	Chaîne unique {1 à 4 chaînons insérés {1 à 16 mots graphiques
A.A-T ≠ 0	Chaîne 1 = 0, Chaîne 2 ≠ 0 → {1 à 3 chaînons insérés {1 à 4 mots graphiques Chaîne 1 ≠ 0, Chaîne 2 = 0 → {1 à 3 chaînons insérés {1 à 9 mots graphiques Chaîne 1 ≠ 0, Chaîne 2 ≠ 0 • → {1 à 2 chaînons {1 à 2 mots graphiques → {1 à 2 chaînons insérés {1 à 5 mots graphiques

4.4. Organigramme d'apparition des chaînes insérées dans un syntagme verbal homogène discontinu.

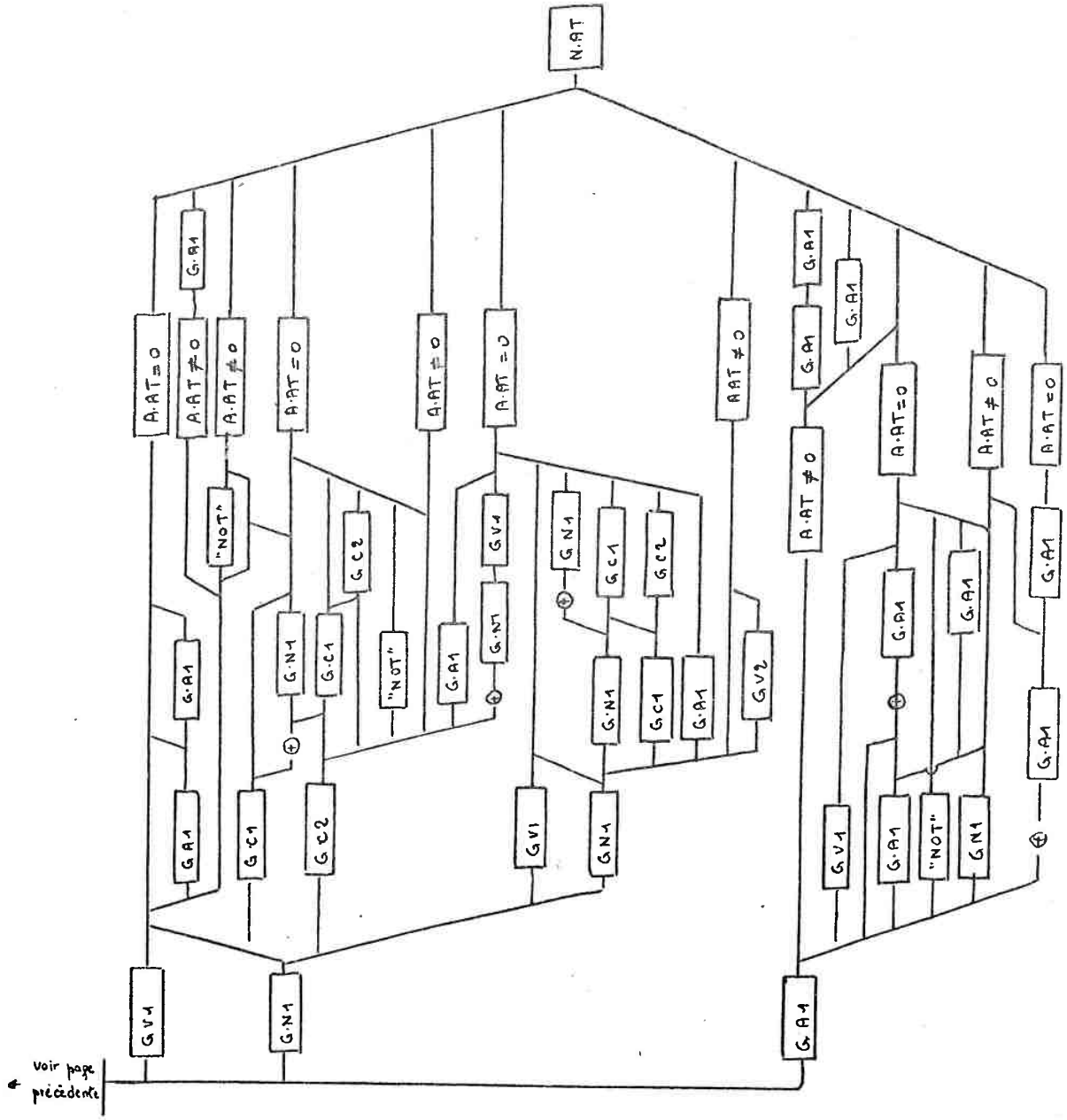
Il prévoit les différents types de chaînes insérées en fonction de la nature du premier chaînon inséré rencontré derrière l'auxiliaire temporel (nature du chaînon, ou "not", ou ⊕).

Voir pages 74 et 75

Organigramme
nº 3



Organigramme
n° 3



RAPPORTS CHAINES INSEREES / SYNTAGME VERBAL

Le problème est de voir si la forme de la chaîne insérée et celle du syntagme verbal homogène sont interdépendantes.

1. LES SYNTAGMES DISCONTINUS DONT LA ZONE A.A-T = 0

Ce sont les syntagmes qui admettent une seule chaîne insérée ("Chaîne unique"), entre l'auxiliaire temporel et l'auxilié a-temporel.

Ils sont au nombre de 1147 sur 1508.

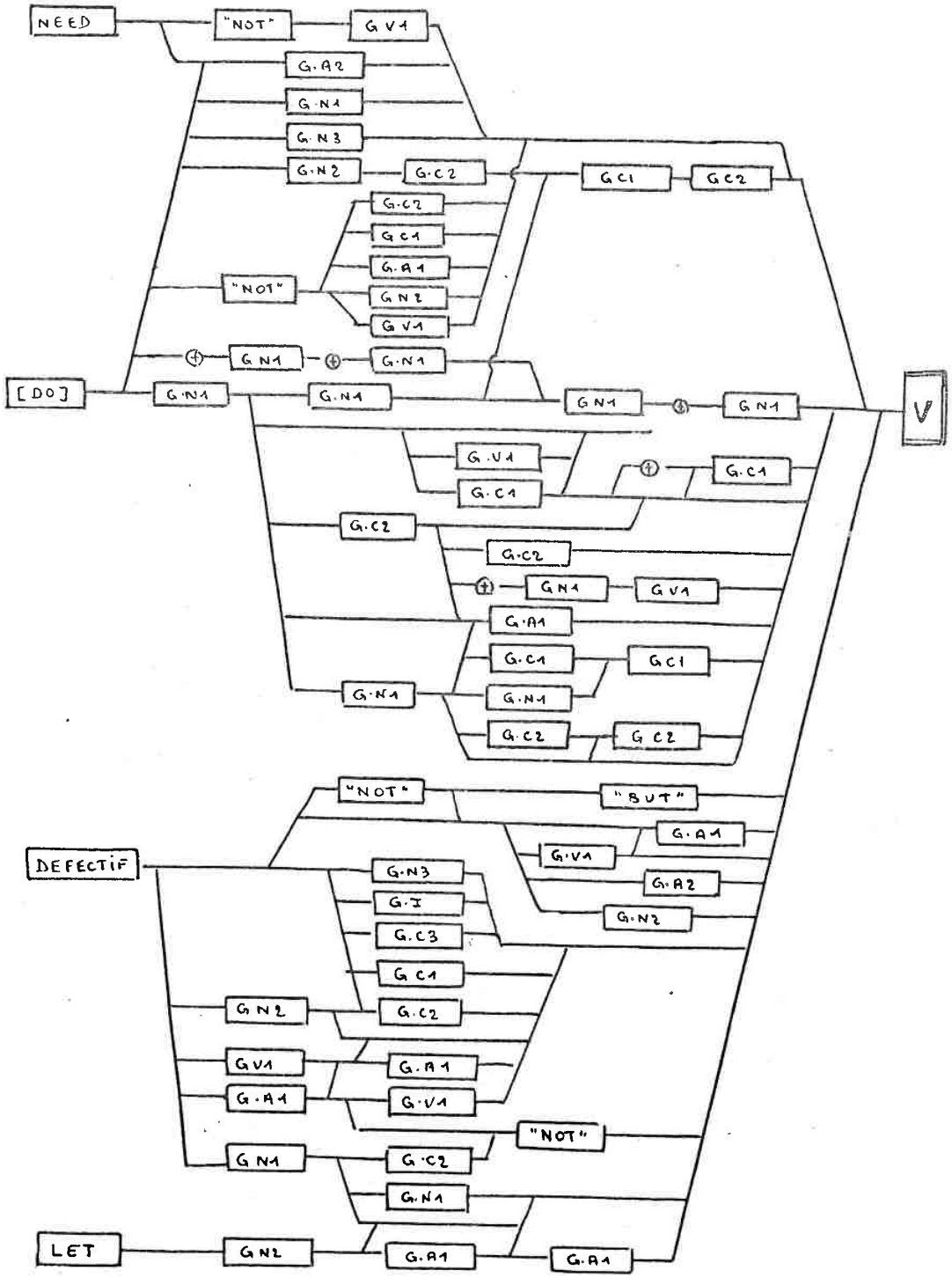
1.1. Les Chaînes uniques rencontrées sont celles données en Annexe 2 page 142, à l'exception des suivantes :

G.C1	G.C2	
G.A2	G.A1	
G.N2	not	
G.V1	not	
not	G.N1	G.C1
not	G.N1	G.C2

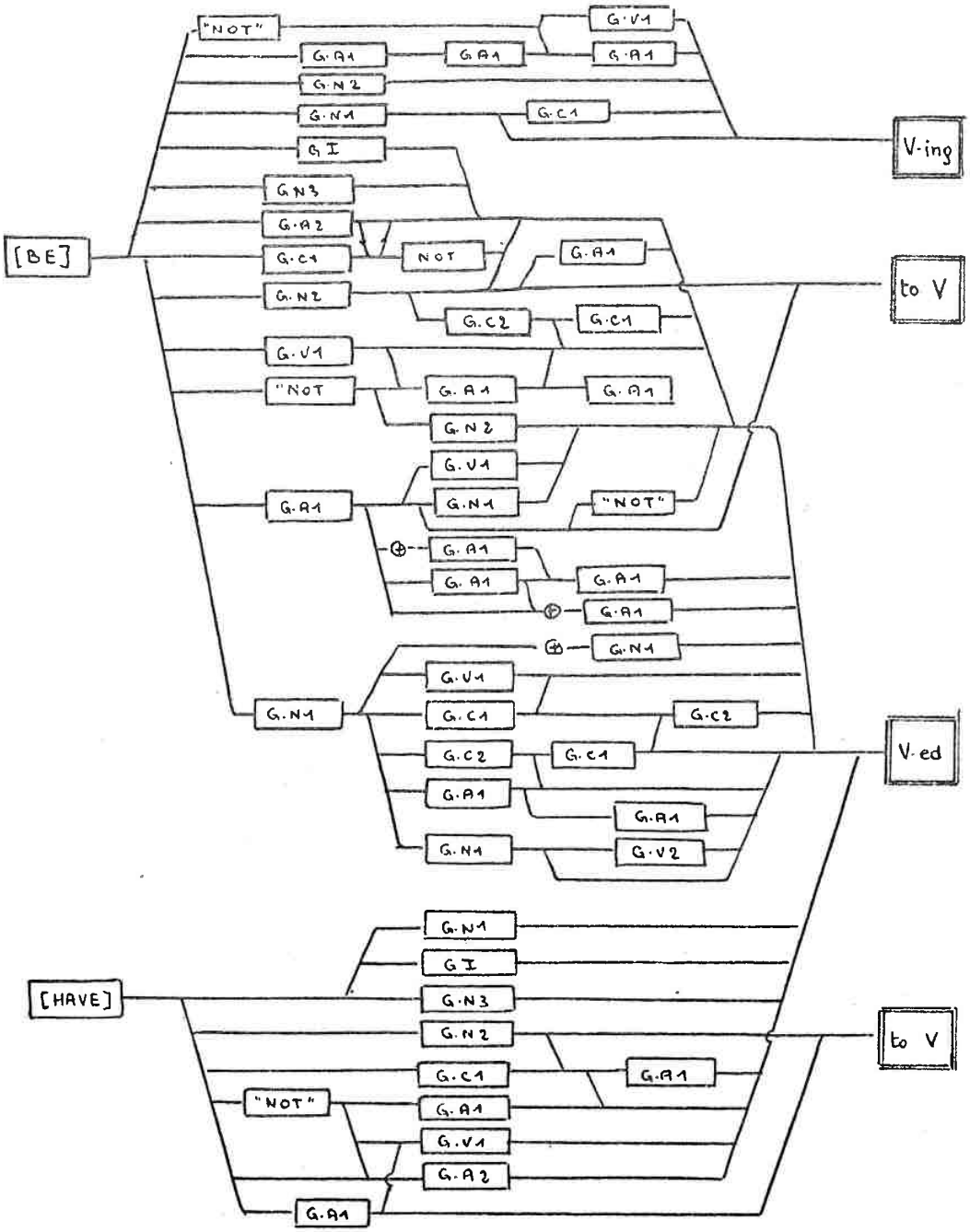
(c'est à dire que 6 schémas de chaînes insérées sur les 74 relevées ne figurent pas en Chaîne unique).

1.2. Les 1147 syntagmes verbaux homogènes discontinus dont la zone A.A-T = 0, sont décrits par 124 schémas différents.

1.3. Tableau de Chaînes uniques relevées dans ces 1147 syntagmes verbaux homogènes:


NEED
"NOT"
G V 1
G.A 2
G.N 1
G.N 3
G.N 2
G.C 2
G C1
G C2
G.C 2
G C 1
G.A 1
G N 2
G V 1
"NOT"
G N1
G.N 1
[DO]
G.N 1
G.N 1
G N1
G N 1
V
G.V 1
G.C 1
G.C 1
G.C2
G.C2
G N 1
G V 1
G.A1
G.C 1
G C1
G.N 1
G.N 1
G.C 2
G C2
"NOT"
"BUT"
G.A 1
G.V 1
G.A 2
G.N 2
DEFECTIF
G.N 3
G.I
G.C 3
G C 1
G N 2
G.C 2
G V 1
G.A 1
G.A 1
G.V 1
"NOT"
G N1
G.C 2
G.N 1
LET
G N2
G.A 1
G.A 1

../..


"NOT"
G.V1
G.A1
G.A1
G.A1
G.N2
G.N1
G.C1
G I
G.N3
G.A2
G.C1
NOT
G.A1
G.N2
G.C2
G.C1
G.V1
"NOT
G.A1
G.A1
G.N2
G.V1
G.A1
G.N1
"NOT"
G.A1
G.A1
G.A1
G.N1
G.V1
G.N1
G.C1
G.C2
G.C2
G.C1
G.A1
G.A1
G.N1
G.V2
[BE]
V-ing
to V
V-ed
[HAVE]
G.N1
G I
G.N3
G.N2
G.C1
G.A1
"NOT"
G.A1
G.V1
G.A2
G.A1
to V

1.4. Conclusion :

- Nous notons la très grande variété de chaînes insérées dans une très grande variété de syntagmes verbaux homogènes discontinus.

- Certaines chaînes insérées sont plus particulièrement associées à tel ou tel type de syntagmes. Par exemple :

. Entre LET et V, l'insertion ne peut que débuter par un groupe nominal (c'est à dire que la chaîne insérée débute toujours par un chaînon G.N2 ou G.N1)

. Entre [BE] et TO V, l'insertion n'est formée que d'un chaînon G.N2 ou G.A1(c'est à dire un groupe simple et très court).

- Nous avons rencontré peu d'auxiliés Ving; La plupart sont des formes Ved ou V.

2. LES SYNTAGMES DISCONTINUS DONT LA ZONE A.A-T ≠ 0

(Ils sont au nombre de 361 sur le total des 1508 syntagmes)

2.1. Syntagmes à 2 chaînes insérées.

2.1.1. Ils sont au nombre de 26 sur les 361 discontinus dont la zone A.A-T est ≠ 0.

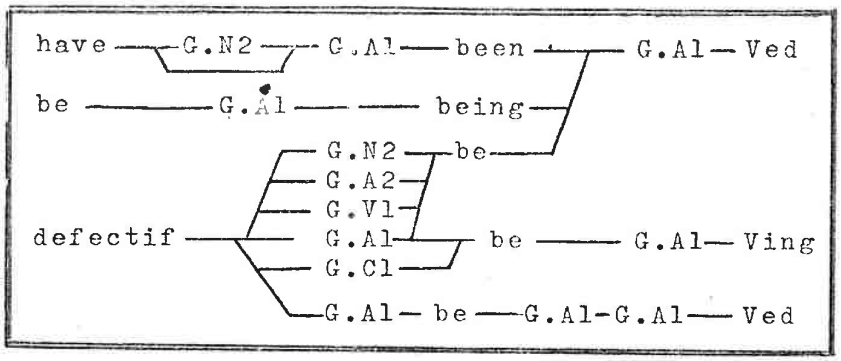
2.1.2. Table de compatibilité des 2 chaînes

Chaîne 1	Chaîne 2
G.A1	G.A1
G.A2	G.A1
G.C1	G.A1
G.N2	G.A1
G.V1	G.A1
G.A1	G.A1 G.A1
G.N2 G.A1	G.A1

2.1.3. Schémas de syntagmes à 2 chaînes insérées :

have	G.N2	G.Al	been	G.Al		Ved	(1)
have	G.Al		been	G.Al		Ved	(10)
be	G.Al		being	G.Al		Ved	(1)
defectif	G.Al		be	G.Al		Ving	(2)
defectif	G.Cl		be	G.Al		Ving	(1)
defectif	G.Al		be	G.Al		Ved	(7)
defectif	G.Vl		be	G.Al		Ved	(1)
defectif	G.A2		be	G.Al		Ved	(1)
defectif	G.Al		be	G.Al	G.Al	Ved	(1)
defectif	G.N2		be	G.Al		Ved	(1)

2.1.4. Le tableau d'ensemble suivant peut donc être établi :



2.1.5. En conclusion :

La "complication" apportée par la présence de deux chaînes insérées se traduit par une simplification à tous les niveaux (tant sur les chaînes insérées que sur le syntagme verbal homogène lui-même) :

- La zone A.A-T est alors composée d'un seul auxiliaire ("be", "been", ou "being")
- L'auxiliaire temporel ne peut être que [be], [have], ou un defectif.
- Chaque chaîne insérée est réduite à un ou quelquefois deux chaînons insérés.

- Jamais "not" n'est apparu.

- Ici encore, peu d'auxiliés sont des formes Ving (seulement 3 sur 26, les 23 autres étant des formes Ved).

- L'ensemble formé par le syntagme verbal homogène discontinu et les chaînes insérées n'a jamais comporté plus de 9 mots graphiques.

2.2. Syntagmes à 1 seule chaîne insérée.

(Ils sont au nombre de 335 sur les 361)

2.2.1. La chaîne insérée présente est une Chaîne 2.

A) C'est le cas pour 102 syntagmes sur les 335.

B) La chaîne insérée aura l'une des formes suivantes (cf. Page 65, § 2.1.2.)

G.A2
G.A2 G.A1

G.A1
G.A1 G.A1
G.A1 G.A1 G.A1

G.A1 ⊕ G.A1
G.A1 ⊕ G.A1 G.A1
G.A1 G.A1 ⊕ G.A1

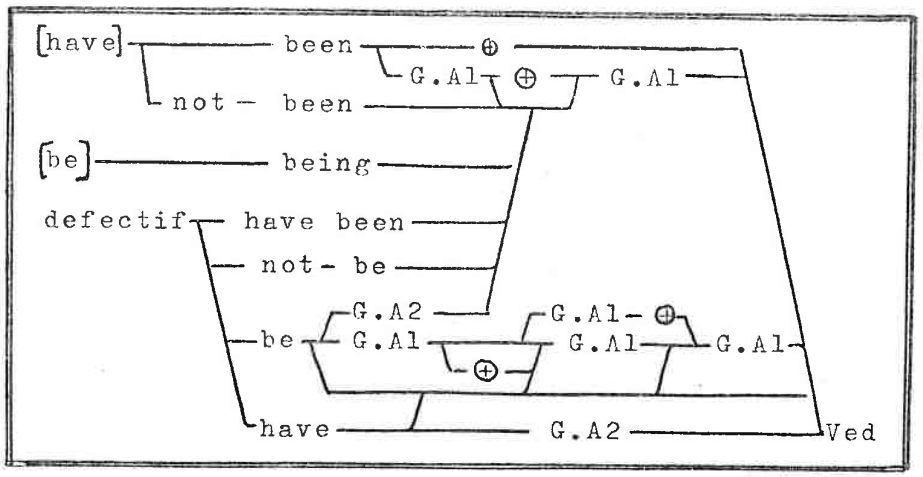
⊕

C) Les 17 schémas de syntagmes dans lesquels peut se trouver l'une de ces 9 chaînes insérées, sont les suivants :

defectif	have	been G.A1	Ved	(1)
defectif	have	G.A1 G.A1	Ved	(1)
defectif	have	G.A1	Ved	(6)
defectif	have	G.A2	Ved	(1)
defectif	be	G.A1 ⊕ G.A1 G.A1	Ved	(1)
defectif	be	G.A1 G.A1 ⊕ G.A1	Ved	(1)
defectif	be	G.A1 G.A1 G.A1	Ved	(1)
defectif	be	G.A1 G.A1	Ved	(6)
defectif	be	G.A2 G.A1	Ved	(1)

defectif	be	G.A1	Ved	(35)
[be]	being	G.A1	Ved	(6)
[have]	been	G.A1 ⊕ G.A1	Ved	(1)
[have]	been	G.A1 G.A1	Ved	(2)
[have]	been	G.A1	Ved	(29)
[have]	been	⊕	Ved	(1)
[have]	not been	G.A1	Ved	(2)
defectif	not be	G.A1	Ved	(7)

D) Les différentes possiblités d'apparition peuvent être prévues de la façon suivante :



E) En_conclusion :

- 86 des 102 syntagmes verbaux homogènes ont une chaîne insérée composée du seul chaînon G.A1.

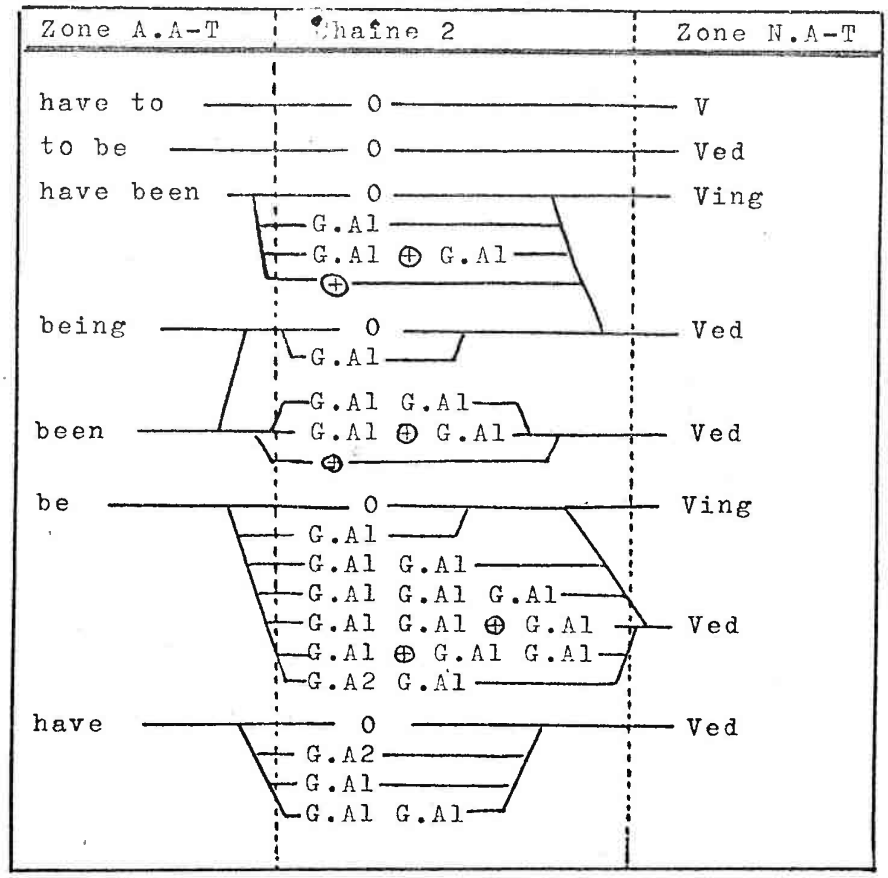
- Le nombre maximal de mots graphiques d'un tel syntagme verbal homogène discontinu et de sa chaîne insérée est 7.

- Nous n'avons trouvé aucun auxilié autre que Ved.

- Quand "not" est présent dans le syntagme verbal, la Chaîne 2 est réduite à G.A1. Il en est de même quand

la zone A.A-T comporte 2 auxiliaires. Par conséquent, il semblerait qu'un équilibre s'installe, entre la chaîne insérée et le syntagme verbal homogène (quand le nombre d'éléments composant le syntagme s'accroît, la chaîne insérée se simplifie).

- Il semble possible de trouver aussi des liens entre la zone A.A-T et la Chaîne 2. Par exemple, jamais nous n'avons rencontré de Chaîne 2 derrière une zone A.A-T comportant la particule "to" (donc quand l'un des auxiliaires du syntagme est "have to" ou "be to"). Le tableau qui suit est un tableau de prévision de la fin du syntagme verbal homogène discontinu, en fonction de la nature de la zone A.A-T :



2.2.2. La chaîne insérée est une Chaîne 1

A) C'est le cas pour 233 syntagmes sur les 361.

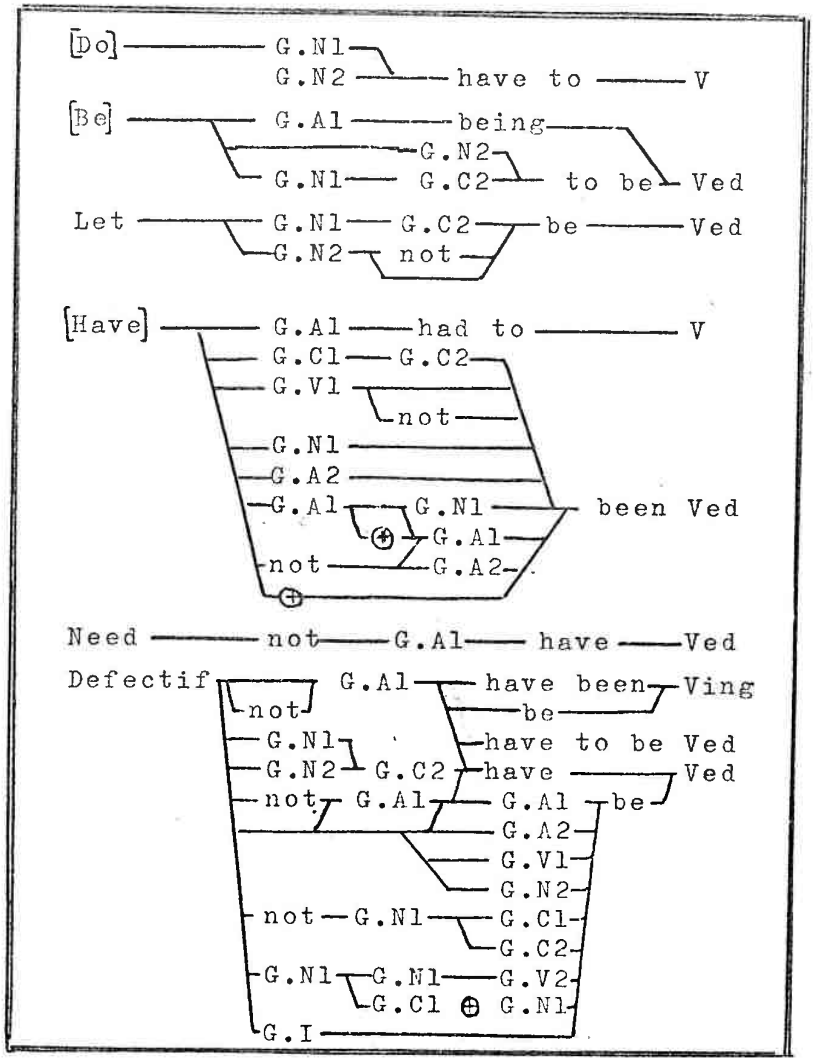
B) La chaîne insérée aura l'une des 25 formes mentionnées page 64 au § 2.1.1.

C) Les 49 schémas correspondant aux 233 syntagmes verbaux homogènes discontinus sont les suivants :

[Do]	G.N1		have to	V		(1)
[Do]	G.N2		have to	V		(3)
Need	not	G.A1	have	been	Ved	(1)
[Be]	G.A1		being	Ved		(18)
[Be]	G.N2		to be	Ved		(2)
[Be]	G.A1		to be	Ved		(3)
[Be]	G.N1	G.C2	to be	Ved		(1)
[Have]	G.A2		been	Ved		(1)
[Have]	G.A1		had to	V		(2)
[Have]	G.A1		been	Ved		(42)
[Have]	G.C1		been	Ved		(2)
[Have]	G.N1		been	Ved		(2)
[Have]	G.V1		been	Ved		(2)
[Have]	G.A1	G.A1	been	Ved		(1)
[Have]	G.A1	G.N1	been	Ved		(1)
[Have]	G.C1	G.C2	been	Ved		(2)
[Have]	G.V1	not	been	Ved		(1)
[Have]	not	G.A1	been	Ved		(4)
[Have]	not	G.A2	been	Ved		(1)
[Have]	⊕		been	Ved		(1)
[Have]	G.A1	⊕ G.A1	been	Ved		(1)
Let	G.N1	G.C2	be	Ved		(1)
Let	G.N2	not	be	Ved		(1)
Let	G.N2		be	Ved		(3)
Defectif	G.A1		have	Ved		(10)
Defectif	G.N1	G.C2	have	Ved		(1)
Defectif	G.N2	G.C2	have	Ved		(1)
Defectif	G.A1	G.A1	have	Ved		(1)
Defectif	not	G.A1	have	Ved		(1)
Defectif	G.A1		have been	Ved		(3)
Defectif	G.A1		have been	Ving		(1)
Defectif	not	G.A1	have been	Ved		(1)
Defectif	G.A2		be	Ved		(7)
Defectif	G.A1		be	Ved		(79)
Defectif	G.N1		be	Ved		(4)
Defectif	G.N2		be	Ved		(1)

Defectif	G.V1 be Ved	(9)
Defectif	G.C1 be Ved	(2)
Defectif	G.I be Ved	(1)
Defectif	G.N1 G.N1 be Ved	(1)
Defectif	G.A1 G.A1 be Ved	(3)
Defectif	G.N1 G.N1 G.V2 be Ved	(1)
Defectif	G.N1 G.C1 ⊕ G.N1 be Ved	(1)
Defectif	not G.A1 be Ved	(4)
Defectif	not G.A1 G.A1 be Ved	(1)
Defectif	not G.N1 G.C1 be Ved	(1)
Defectif	not G.N1 G.C2 be Ved	(1)
Defectif	G.A1 be Ving	(1)
Defectif	G.A1 have to be Ved	(1)

D) A partir de cette liste, nous pouvons établir le tableau suivant :



E) En conclusion :

- La Chaîne 1 est très libre quant à la nature de ses composantes; elle l'est toutefois moins que la Chaîne unique, puisqu'on dénombre seulement 23 schémas de chaînes insérées au lieu de 63.

- La Chaîne 1 peut se trouver dans un syntagme verbal homogène dont la zone A.A-T comporte 1 ou 2 auxiliaires. Si cette zone est égale à "have been", ou "haveto be", la Chaîne 1 est G.Al ou "not" G.Al. Ceci tend à confirmer qu'il existe une loi d'équilibre entre la forme du syntagme et la forme de la chaîne insérée.

- La négation "not" peut être présente, en fin ou en début de chaîne insérée, quand la zone A.A-T = 1 élément ou 2 éléments.

- La forme de l'auxilié est presque toujours Ved (seulement 2 Ving, et 5 V, contre 160 Ved)

- La nature du 1° mot graphique du syntagme verbal homogène peut parfois permettre de déterminer la chaîne insérée qui suivra; par exemple [Do] et "Let" sont suivis d'une chaîne insérée débutant par un chaînon nominal le plus souvent.

Dans certains cas, il est possible aussi de déterminer certaines correspondances entre la Chaîne 1 et la zone A.A-T; par exemple "being" est uniquement précédé de G.Al, "have been" est également précédé de G.Al.

3. RECAPITULATION

3.1. Rapports entre Chaînes insérées et Auxiliés

Zone A.A-T	Chaîne	Auxilié				
		Ving	V		Ved	
			to V	V		
= 0	unique	55	7	371	714	1147
≠ 0	Ch. 1	2	5	0	226	233
	Ch. 2	0	0	0	102	102
	Ch. 1 et Ch. 2	3	0	0	23	26
		60	12	371	1065	1508

3.2. Rapports entre Chaînes insérées et Zone A.A-T

3.2.1. La présence de Chaînes insérées influe sur la longueur et la nature de la zone A.A-T.

a) Dans un syntagme verbal homogène continu, il est théoriquement possible de trouver 4 auxiliaires différents dans cette zone. En pratique nous en avons trouvé un maximum de 3, qui se combinent de façons diverses. Exemples :

- You will have been being surprized.
- Houses may have had to be established.

b) Dans un syntagme discontinu, les mêmes possibilités théoriques existent. Pourtant, le nombre maximum d'auxiliaires a-temporels rencontrés n'a pas dépassé 2; Exemple :

- The latter might have been deliberately destroyed at the time of his trials.

3.2.2. Il semble exister une loi d'équilibre entre la longueur de la zone A.A-T et la longueur des chaînes insérées. La distribution se fait selon le tableau suivant, où l'on voit que :

Quand x décroit, le nombre de mots graphiques pou-

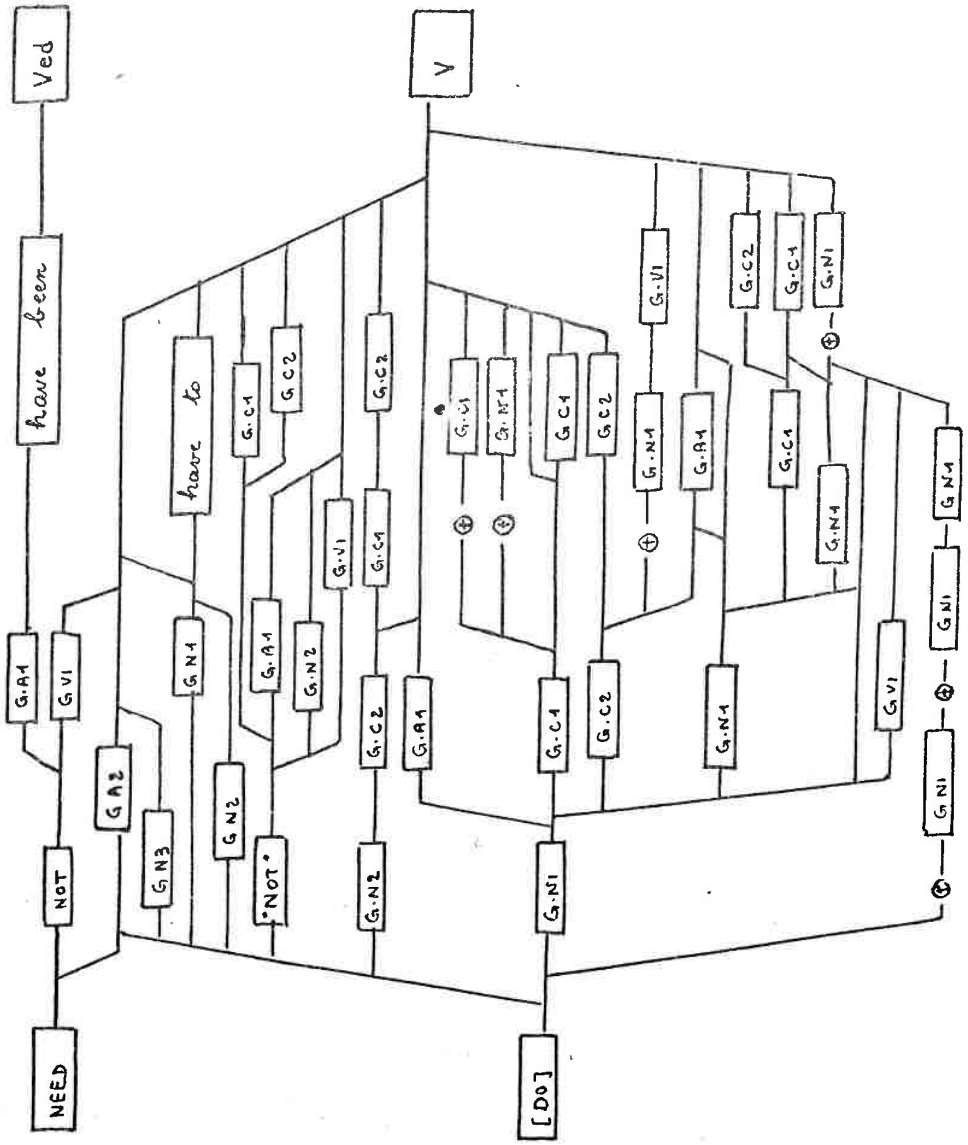
vant être insérés dans le syntagme verbal homogène tend à s'accroître :

x auxiliaires a-temporels	Chaînes insérées
x = 3	0
x = 2	1 Chaîne, de 1 Chaînon et 1 mot graphique
x = 1	1 Chaîne de 1 à 4 Chaînons et 1 à 12 mots graphiques ou 2 Chaînes, de chacune 1 à 2 Chaînons et 1 à 5 mots graphiques
x = 0	1 Chaîne, de 1 à 4 Chaînons et 1 à 16 mots graphiques.

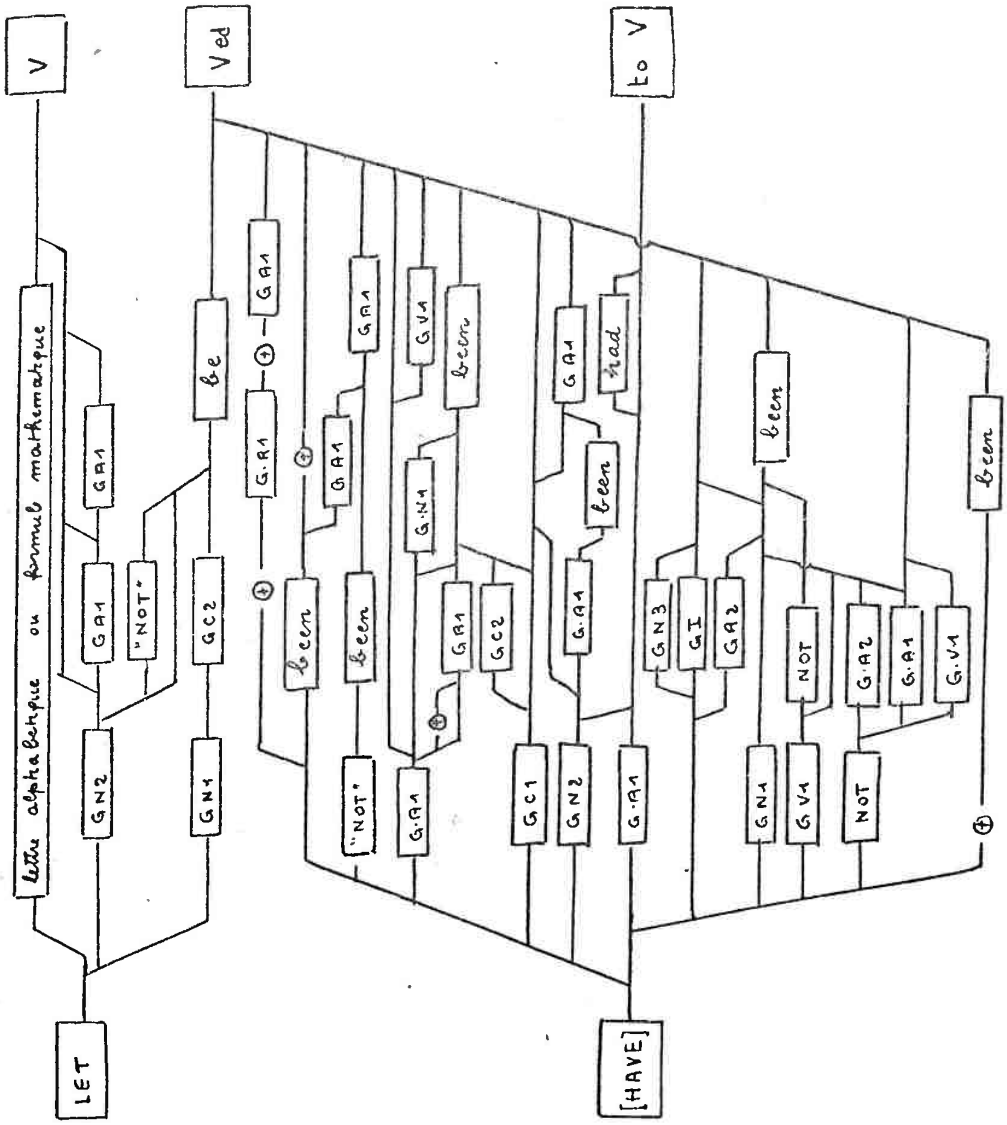
3.3. Organigramme de prévision des chaînes insérées, à partir de l'auxiliaire temporel (1° mot du syntagme verbal homogène).

voir pages 89, 90, 91, et 92.

Organigramme
n° 4

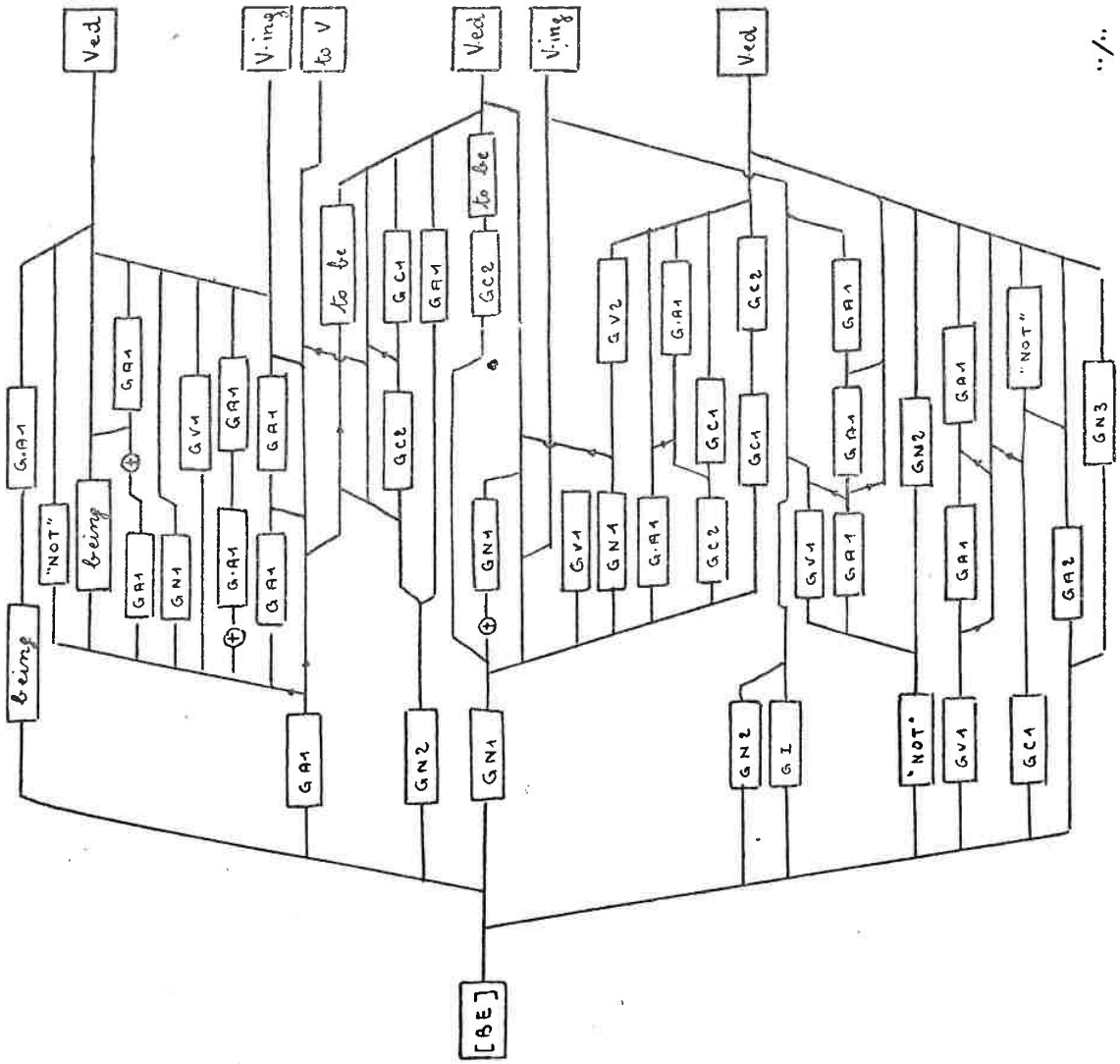


Organigramme
n° 4



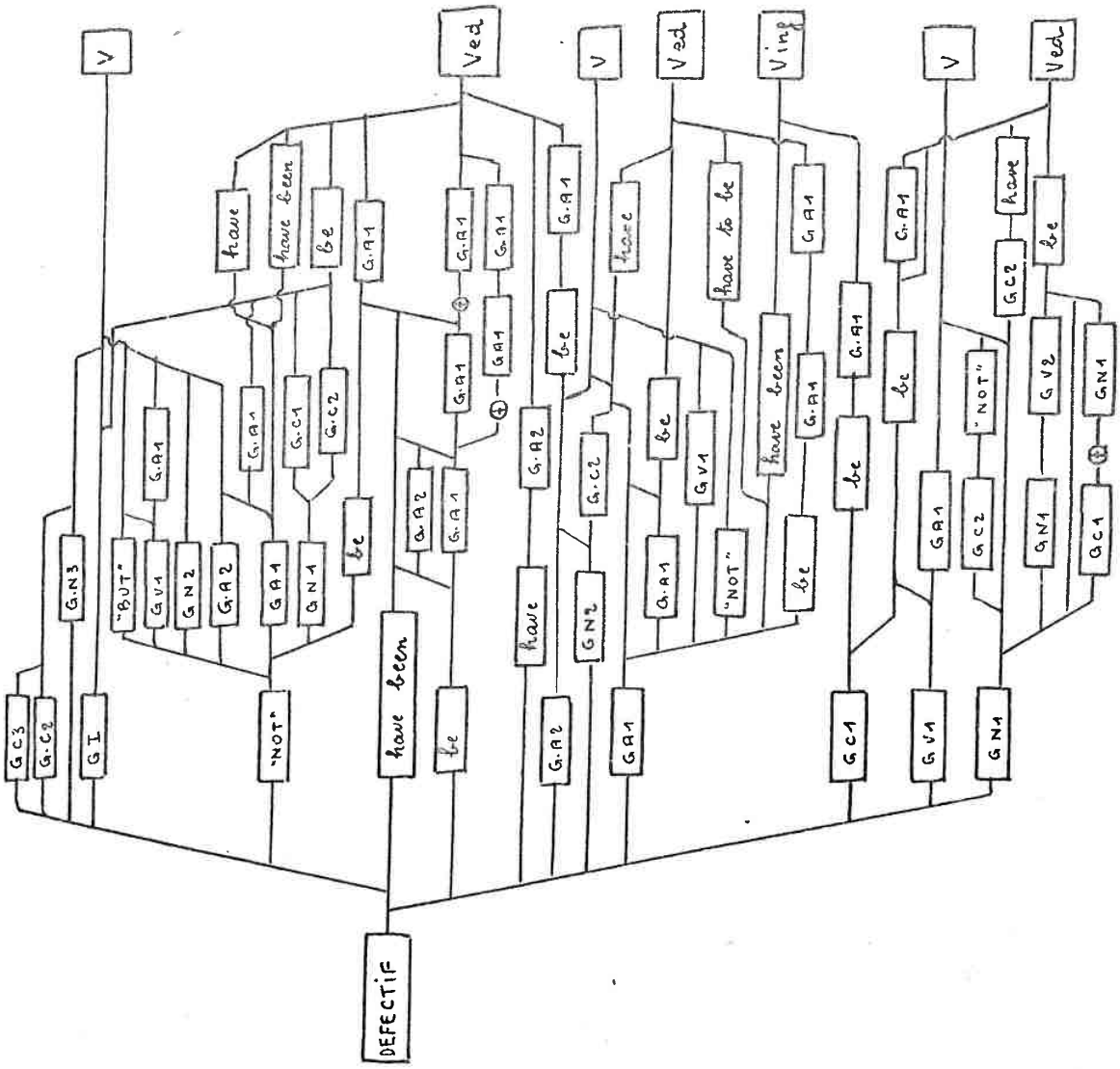
../..

Organigramme
n° 4



../..

Organigramme
n° 4



RAPPORTS CHAINES INSEREES / SYNTAGME VERBAL / PHRASE

Nous avons constaté l'interdépendance formelle des chaînes insérées et des syntagmes verbaux homogènes.

Il semble aussi que, lorsque la présence d'un point d'interrogation peut être testée en fin de phrase, certains schémas de syntagmes verbaux homogènes discontinus doivent être particulièrement attendus.

La liste récapitulative (Annexe 2 page 144) nous montre que certains schémas de syntagmes ne sont jamais employés que dans un seul type de phrases (soit interrogatives, soit non-interrogatives), et que d'autres peuvent être "bivalents".

1. LES SCHEMAS VERBAUX "MONOVALENTS".

1.1. Schémas utilisés en phrases non-interrogatives.

1.1.1. Ils sont au nombre de 149 (sur 207 schémas au total). Ces 149 schémas correspondent à 1276 syntagmes sur les 1508 recensés.

1.1.2. Tout auxiliaire temporel peut débuter un syntagme verbal homogène discontinu en phrase non-interrogative. Mais seuls "need" et "let" débutent obligatoirement une phrase non-interrogative.

1.1.3. Certains schémas de chaînes insérées se trouvent obligatoirement dans un syntagme verbal homogène discontinu non-interrogatif. Ce sont toutes les chaînes débutant par G.A1, G.A2, G.V1, G.C1, G.C2, et G.C3, ainsi que le montre le tableau suivant :

Chaîne insérée débutant par :	Fréquence dans phrase	
	avec "?"	sans "?"
G.A1	0	1080
G.A2	0	35
G.N1	90	37
G.N2	84	54
G.N3	3	2
G.V1	0	37
G.C1	0	26
G.C2	0	1
G.C3	0	1
G.I	2	4
not	2	63
⊕	1	6
	182	1346

1.1.4. Les différents schémas de chaînes insérées et les différents schémas de syntagmes verbaux homogènes discontinus, ainsi que leurs fréquences en phrases interrogatives et phrases non-interrogatives sont indiqués en Annexe 2, pages 142,et 144 .

1.2. Schémas utilisés uniquement en phrases interrogatives.

1.2.1. Ils sont au nombre de 43 (sur 207 schémas au total). Ces 43 schémas correspondent à 64 syntagmes sur les 1508 recensés. Ces schémas peuvent être retrouvés dans la liste donnée en Annexe 2, page 153.

1.2.2. Les auxiliaires "need" et "let" n'introduisent jamais de syntagme verbal homogène discontinu interrogatif. Tous les autres auxiliaires temporels le peuvent.

1.2.3. Les 43 schémas de syntagmes verbaux homogènes discontinus interrogatifs ne comportent jamais plus d'un auxiliaire a-temporel, et encore ce cas est-il relativement peu fréquent : 8 schémas sur 43, correspondant à 12 syntagmes sur 64 :

[Do]	G.N1	have to	V			(2)	
[Do]	G.N2	have to	V			(3)	
[Be]	G.N2	to be	Ved			(2)	
[Be]	G.N2	G.C2	to be	Ved		(1)	
Defectif	G.N2		be	Ved			(1)
Defectif	not	G.N1	G.C2	be	Ved		(1)
Defectif	not	G.N1	G.C1	be	Ved		(1)
Defectif	G.N1	G.C2		have	Ved		(1)

1.2.4. Nous n'avons jamais trouvé 2 chaînes insérées dans un syntagme verbal discontinu interrogatif. Et la chaîne insérée est à chercher immédiatement derrière l'auxiliaire temporel.

Exemples :

- /Was he to be admonished/ only if he failed to comply with the request ?
- Why /do the cattle have to move/ from Maiduguri ?
- /Am I to stand/ here for ever ?

1.2.5. Les chaînes insérées entrant dans des schémas interrogatifs ne débutent jamais par G.A1, G.A2, G.V1, G.C1, G.C2, ni G.C3. Elles commencent toujours par un Chaînon nominal ou impersonnel :

G.I	:	1 cas sur 64
G.N1	.. :	37 cas sur 64
G.N2	.. :	21 cas sur 64
G.N3	:	2 cas sur 64
not G.N1	:	2 cas sur 64
Θ G.N1	:	1 cas sur 64

Cette forme de la chaîne insérée en construction interrogative était prévisible, puisque cette construction entraîne une inversion du sujet. Si d'autres chaînons sont insérés entre auxiliaire temporel et auxilié, ils sont toujours situés à la suite du groupe nominal.

1.2.6. La particule "not" n'est presque jamais rencontrée dans les syntagmes homogènes discontinus en phrases interrogatives. Elle a été trouvée 3 fois en tout :

- 2 fois en début de chaîne insérée, et
- 1 fois en fin de chaîne insérée, dans les exemples suivants :

- Cannot rules for automatic translation be elaborated ?
- Cannot some perfect system of translation be devised ?
- Should the initial set of constraints not be compatible ?

1.2.7. Composition de la chaîne insérée interrogative en fonction de l'auxiliaire temporel :

Chaîne insérée	Auxiliaire temporel				
	Do	Be	Have	Defectif	
G.I	0	1	0	0	1
G.N3	2	0	0	0	2
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1	1	0	0	0	1
not G.N1 + 1 chaînon	0	0	0	2	2
G.N1 seul	2	2	0	0	4
+ 1 chaînon	10	4	0	2	16
+ 1 ch. + not	0	0	0	1	1
+ 2 chaînons	8	3	0	0	11
+ 3 chaînons	5	0	0	0	5
G.N2 seul	3	4	1	1	9
+ 1 chaînon	2	4	3	1	10
+ 2 chaînons	0	1	0	0	1
+ 3 chaînons	1	0	0	0	1

2. LES SCHEMAS VERBAUX "BIVALENTS"

2.1. Inventaire des schémas bivalents

2.1.1. Schémas des chaînes insérées

Chaînes insérées	Fréquence dans phrases	
	avec "?"	sans "?"
G.N1	39	16
G.N2	63	27
G.N3	1	1
G.I	1	1
G.N1 G.C2	8	3
G.N1 G.A1	2	1
G.N1 G.N1	4	1

2.1.2. Ces 7 schémas de chaînes insérées entrent dans 15 schémas de syntagmes verbaux homogènes discontinus bivalents qui sont :

Syntagmes discontinus	Fréquence dans phrases	
	avec "?"	sans "?"
[Do] G.N1 V	19	4
[Do] G.N2 V	38	7
[Do] G.N1 G.C2 V	6	1
[Be] G.N1 Ved	10	5
[Be] G.N2 Ved	4	9
[Be] G.N3 Ved	1	1
[Be] G.N1 G.A1 Ved	2	1
[Be] G.N1 G.N1 Ved	4	1
[Be] G.N1 G.C2 Ved	2	2
[Have] G.N1 Ved	1	2
[Have] G.N2 Ved	2	5
[Have] G.N1 been Ved	1	1
Defectif G.N1 V	8	4
Defectif G.N2 V	19	6
Defectif G.I V	1	1

2.1.3. Les schémas verbaux bivalents sont au nombre de 15, sur les 59 schémas verbaux utilisés en phrases interrogatives. Mais ces 15 schémas sont très productifs et correspondent à 118 syntagmes verbaux discontinus interrogatifs, et à 168 syntagmes verbaux au total.

2.1.4. Les chaînes insérées bivalentes sont toutes des Chaînes 1 ou des Chaînes uniques.

Des 15 schémas verbaux bivalents, un seul possède une zone A.A-T $\neq$ 0 (occupée par "been")

Dans un schéma bivalent, jamais la particule "not" n'a été relevée.

2.2. Les schémas ne sont bivalents que sous certaines conditions.

2.2.1. Dans toute phrase interrogative, ils sont prévisibles. En effet, ils possèdent toutes les caractéristiques des schémas monovalents interrogatifs :

- Ils ne sont pas introduits par "need" ni par "let".
- Leur zone A.A-T n'est jamais supérieure à 1.
- Il n'existe jamais 2 chaînes insérées dans le même syntagme verbal homogène.
- Les chaînes insérées commencent toutes par un groupe nominal ou impersonnel.
- La particule "not" n'a pas été trouvée.

2.2.2. En phrases non-interrogatives, ils n'apparaîtront que dans certains cas bien précis, qui sont :

a) Les phrases où est employé "let".
b) Les phrases qui débutent
   - par un auxiliaire conditionnel ("Had" ou "Should")

- par une suite de mots graphiques qui commence par "very" ou "only".
- par un coordonnant négatif ("neither", "nor") ou, plus rarement, par une locution adverbiale ou un adverbe négatif ("no longer", "in neither case").

c) Dans tous les cas mentionnés en a) et b), le syntagme verbal en cause doit être le premier de la phrase.

En conséquence, lors de l'apparition d'un schéma bivalent dans une phrase non-interrogative, il estnécessaire d'effectuer un certain nombre de tests.

2.3. Tests à effectuer en phrase non-interrogative lors de la rencontre d'un schéma bivalent :

2.3.1. Tests sur la nature de l'insertion.

A) G.N2 :

Nous avons pu voir, par exemple, que derrière "let" auxiliaire temporel, le groupe G.N2 ne pouvait être composé que d'un pronom personnel objet ou ambigu (sujet/objet), mais jamais d'un pronom sujet. Si, par hasard, "let" était suivi d'un pronom sujet, il fonctionnerait lui-même comme auxilié et non comme auxiliaire.

De même, dans une phrase non-interrogative à schéma normal (différent de ceux indiqués au § 2.3.2., page suivante), un pronom sujet ou objet ne peut se trouver dans une chaîne insérée; seuls ont été rencontrés des indéfinis ou des refléchis. Par exemple :

- The granite /is itself cut/ by dolerite dykes.
- Boulder beds/are both interbedded/with and appear to be ....

C'est ainsi que nous sommes amenés à créer des

sous-classes à l'intérieur du chaînon G.N2 :

- G.N2S : pronom personnel sujet ("I", "he", "they" ...)
- G.N2O : pronom personnel objet ("me", "them", "us" ...)
- G.N2A : pronom personnel ambigu S/O ("it", "you")
- G.N2R : pronom réfléchi ("itself" ..)
- G.N2I : autre pronom ("much", "each", "one" ...)

Cette division, et la prévision d'apparition des différents sous-groupes G.N2 dans les différents contextes, permettra d'éviter des regroupements erronés tels que le suivant :

- He suggests that bishops who have not forgotten how to do so should compose their prayers in Latin, and then have them translated.

B) G.N1

Nous avons constaté qu'il existait certaines relations entre le chaînon G.N1 et la forme de l'auxiliaire temporel qui le précède.
Ainsi, derrière l'auxiliaire temporel [be] , jamais (dans le Corpus que nous avons analysé) G.N1 ne débute par un adjectif. Par conséquent, [be] suivi d'un adjectif peut être automatiquement reconnu comme auxilié, et non comme auxiliaire possible.

Il nous semble donc utile de créer aussi des sous-classes à l'intérieur du chaînon G.N1 :

- G.N1A : chaînon débutant par un adjectif.
- G.N1I : chaînon débutant par un introducteur de syntagme nominal.
- G.N1S : chaînon débutant par un substantif.

(Cette subdivision pourra en outre aider plus tard à réunir plusieurs G.N1 situés en séquence).

2.3.2. Tests sur les mots qui précèdent l'auxiliaire.

Certaines constructions appellent un schéma discontinu bivalent semblable à celui des phrases interrogatives :

A) Phrases qui débutent par un auxiliaire conditionnel "had" ou "should".

Ces deux auxiliaires , placés en tête de phrase, commencent obligatoirement un schéma bivalent.
Exemples :

- /Had we had/ time for further journeys, we could have included the great forest belts.
- /Should a driver lose/ his way, he can find his direction by using instruments.

B) Phrases qui débutent par une locution ou un mot négatif, le schéma discontinu en cause étant le premier de la phrase.

a) Ce mot négatif peut être un coordonnant tel que "neither", ou "nor" . Exemple :

- Nor /could any of the textiles have survived/ the temperature.

b) Ce mot peut être un adverbe. Exemple :

- Nowhere else /is the stair made/ such an important feature.

C) Phrases qui débutent par une suite de mots graphiques dont le premier est "only" ou "very". Ici encore, le schéma discontinu bivalent ne sera reconnu comme valable, que si le syntagme verbal homogène est le premier de la phrase. Exemple :

- Only towards the Norwegian frontier /does the height of the land exceed/ 3,000 feet.

2.3.3. Test sur la présence de "there"

Ce test se fait lors de la rencontre d'un schéma de chaîne insérée bivalent, derrière le verbe "to be". (forme temporelle ou non).

Les seuls cas où la construction impersonnelle a été prise en considération sont ceux où "there" et "it" provoquaient eux-mêmes une discontinuité du syntagme verbal homogène.

Exemples :

- /Should it be/ possible that ...
- Only in Iraq /is there found/ any extensive area of lowland.

Mais si "there" précède une suite verbale temporelle dans laquelle "to be" peut fonctionner comme auxilié, le schéma impersonnel normal est retrouvé.

Exemple :

There [are / have been / should be / ⋮] hotels catering for tourists.

Dans ce cas , le syntagme n'a pas été considéré comme discontinu. "Are" (ou "have been", ou "should be" ...) sont considérés comme syntagmes verbaux homogènes continus, et "catering" comme une adjectivation postposée à un substantif.

De même, les schémas impersonnels interrogatifs n'ont pas été inventoriés, donc ne sont pas prévus.

Exemple :

- Are there many farms named on the map?

* G.I G.Nl

→ est une chaîne insérée non prévue.

3. CONCLUSION

Cette division des schémas verbaux en schémas interrogatifs et schémas non-interrogatifs entraîne une modification de l'organigramme n°4 (pages 89 à 92).

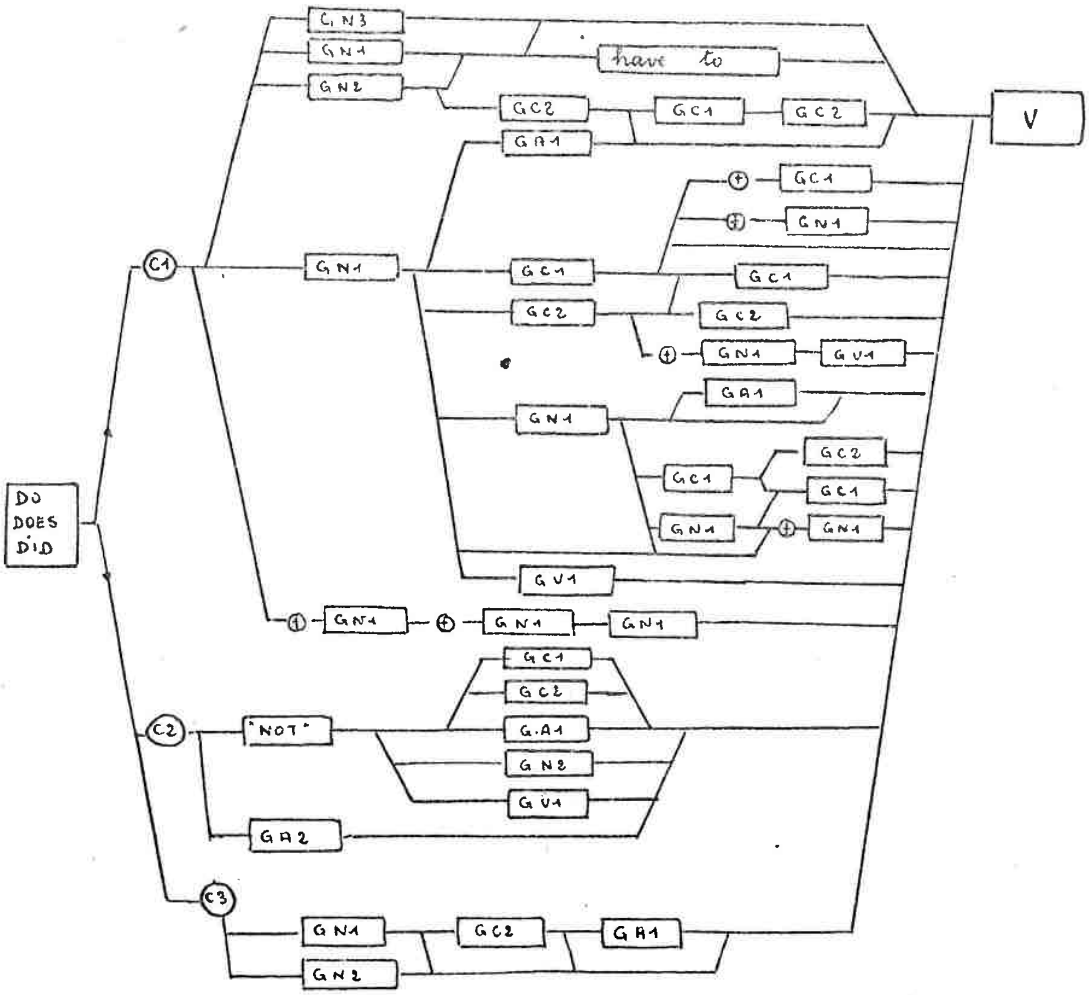
Sur l'organigramme n°5 (pages 104 à 107) nous indiquons aussi les regroupements de syntagmes obligatoires par un trait double, et les regroupements probables ou possibles par un trait simple.

Organigramme n°5 :

(C1) , (C2) , (C3) , (C4) , sont des conditions.

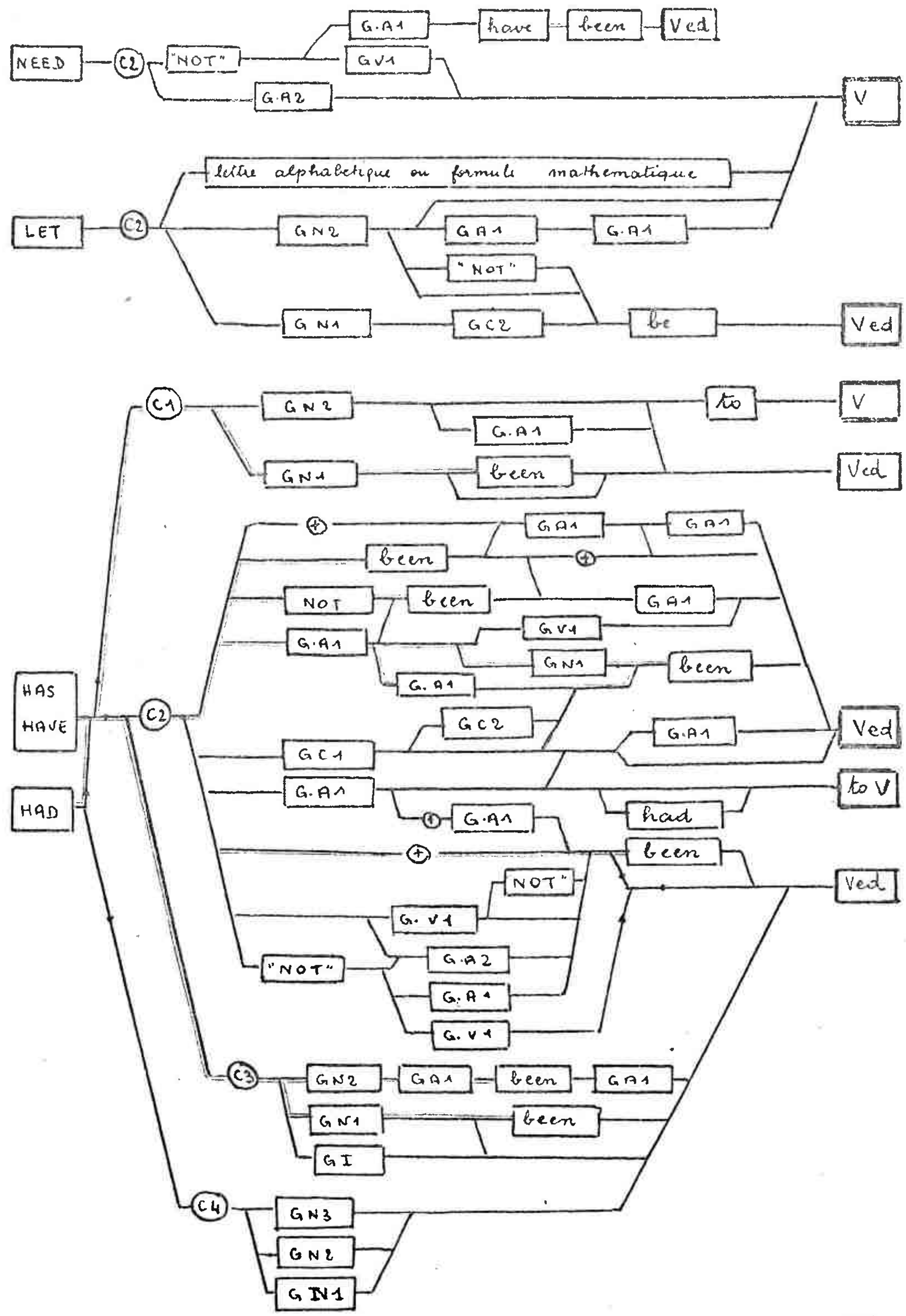
(C1) : présence du point d'interrogation
(C2) : absence du point d'interrogation
(C3) : condition C2 + présence immédiatement avant l'auxiliaire temporel, d'une locution ou d'un mot négatif;
ou condition C2 + présence, en début de phrase, de "very" ou de "only".
(C4) : condition C2 + présence d'un point immédiatement avant l'auxiliaire "had" ou "should" (ou "be")

Organigramme
n° 5



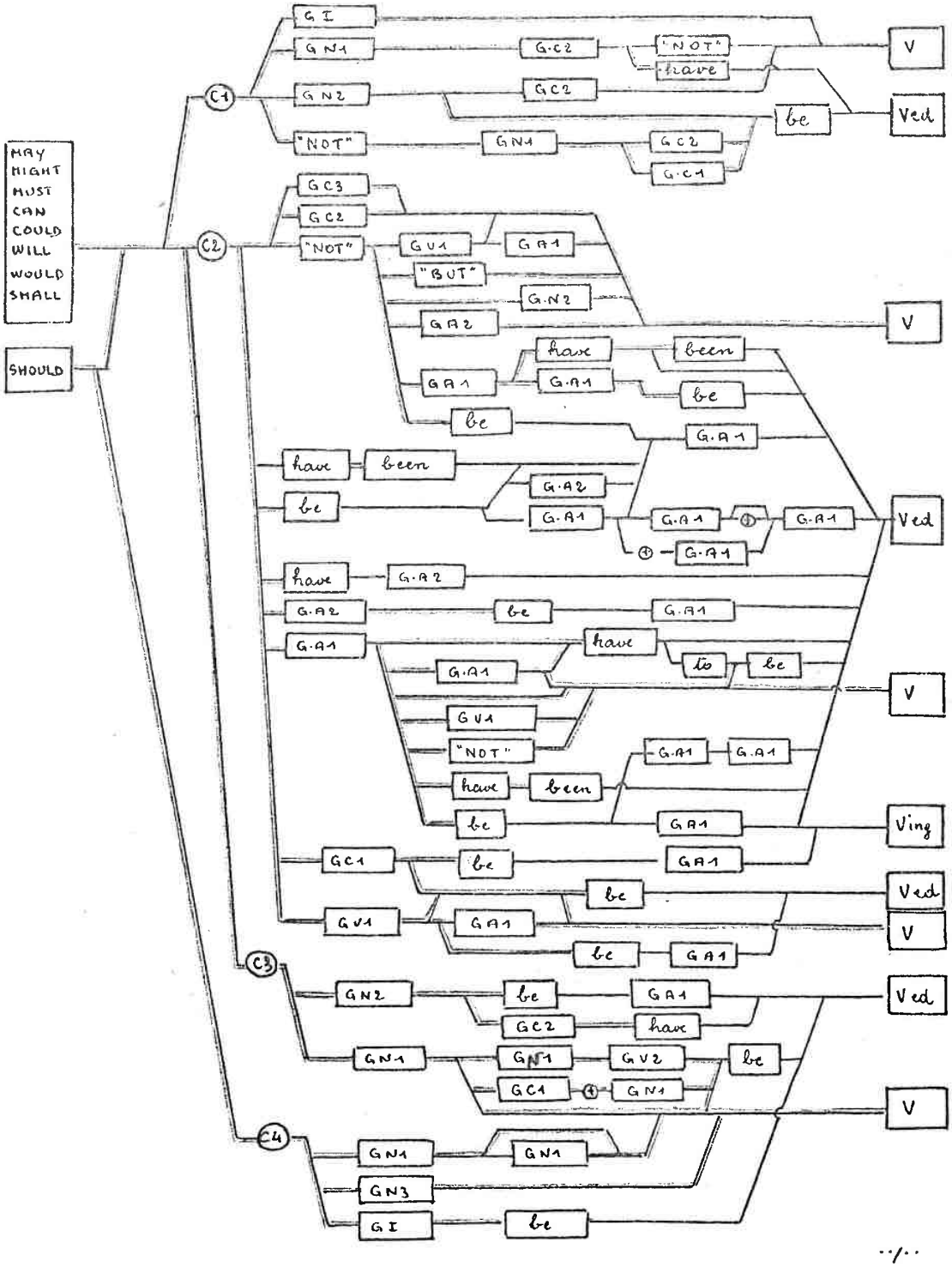
../..

(n° 5)


NEED
C2
"NOT"
G.A1
have
been
Ved
GV1
G.A2
V
LET
lettre alphabetique ou formule mathematique
GN2
GA1
G.A1
"NOT"
GN1
GC2
be
Ved
HAS
HAVE
HAD
C1
GN2
to
V
G.A1
GN1
been
Ved
C2
GA1
GA1
been
NOT
been
GA1
G.A1
GV1
GN1
been
G.A1
GC2
G.A1
GC1
Ved
G.A1
to V
G.A1
had
been
NOT
Ved
G.V1
"NOT"
G.A2
G.A1
G.V1
C3
GN2
GA1
been
GA1
GN1
been
GI
C4
GN3
GN2
GN1

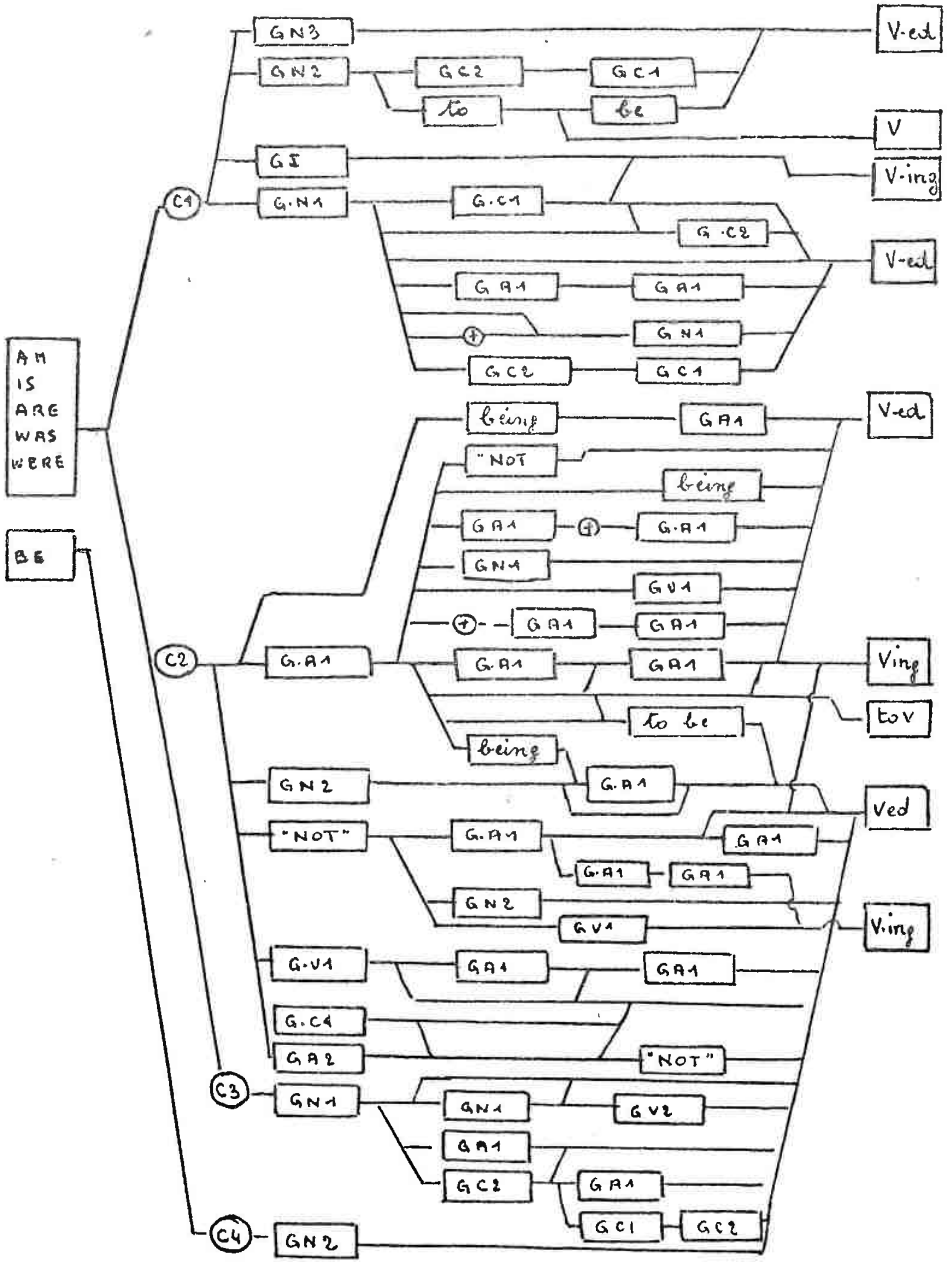
../..

(n° 5)


MAY
MIGHT
MUST
CAN
COULD
WILL
WOULD
SHALL
SHOULD
C1
C2
C3
C4
GI
GN1
GN2
"NOT"
G.C2
GC2
GN1
GC2
G.C1
have
be
V
Ved
GC3
GC2
"NOT"
GV1
GA1
"BUT"
G.N2
GA2
have
been
GA1
G.A1
be
have
been
be
G.A2
G.A1
G.A1
G.A1
G.A1
have
G.A2
G.A2
be
G.A1
G.A1
have
to
be
G.A1
GV1
"NOT"
have
been
G.A1
G.A1
be
GA1
Ving
GC1
be
GA1
be
GV1
GA1
be
GA1
GN2
be
GA1
GC2
have
GN1
GN1
GV2
be
GC1
GN1
GN1
GN1
GN3
GI
be

../..

Organigramme
nº 5



CONSEQUENCES MECANISABLES

Cette étude des différents syntagmes verbaux homogènes discontinus, ainsi que des éléments provoquant la discontinuité, nous a déjà permis de prévoir :

- La chaîne insérée maximale.
  (Organigramme n° 1, page 63)
- La possibilité d'apparition d'une Chaîne 2 en fonction de la Chaîne 1.
  (Organigramme n° 2 page 70)
- La possibilité d'apparition et l'organisation des chaînes insérées à partir du premier chaînon rencontré.
  (Organigramme n° 3, pages 74 et 75)
- La possibilité d'apparition et l'organisation des chaînes insérées, en fonction de l'auxiliaire temporel du syntagme verbal homogène.
  (Organigramme n° 4, pages 89 à 92)
- L'ordre des constituants prévisibles entre l'auxiliaire temporel et l'auxilié, selon la forme de la phrase.
  (Organigramme n° 5, pages 103 à 107)

Il reste à établir l'organigramme définitif de réunion des différents éléments d'un syntagme verbal homogène, et à prévoir l'extraction deses catégories verbales.

1. ORGANIGRAMME DE REGROUPEMENT DES ELEMENTS DU SYNTAGME VERBAL.

1.1. Délimitation des chaînons insérés

Les bornes des différents chaînons ont été définies pages 20 à 22, et les sous-classes établies pour les chaînons G.N1 et G.N2 sont indiquées page 100.

1.2. Commentaire de l'organigramme

1.2.1. Il tient compte des résultats acquis dans la publication précédente (1), et des résultats acquis dans la présente, concernant le syntagme verbal homogène discontinu.

1.2.2. Il doit permettre de réunir dans une même "chaîne résultat" les différents éléments d'un syntagme verbal homogène (continu ou discontinu), afin de passer, dans un deuxième temps, à l'extraction des catégories verbales portées par ce syntagme verbal homogène.

1.2.3. Il n'indique qu'une démarche générale; la plupart des tests indiqués correspondent chacun à un sous-programme particulier (par exemple la reconnaissance du "début de G.N1").

1.2.4. Le premier test à effectuer porte sur l'auxiliaire présumé, chacun de ces auxiliaires appelant des schémas de syntagmes et de chaînes insérées différents.
Le test sur la présence du point d'interrogation se fait à des moments différents selon la nature de l'auxiliaire et les schémas rencontrés.

1.2.5. Le programme établi à partir de cet organigramme devra être passé après un programme de consultation de dictionnaire.
La consultation de dictionnaire aura permis de regrouper les différentes lexies (adverbiales, pronominales, ou nominales), et d'affecter automatiquement aux mots du texte des codes correspondant à leur nature grammaticale. Le programme testera ce code, ou tantôt le code et tantôt le mot du texte.

1.2.6. La liste des codes utilisés dans l'organigramme est la suivante :

---

(1) cf. SVH

G.N1	subdivisé en	G.N1A G.N1I G.N1S
G.N2	subdivisé en	G.N2S G.N2O G.N2A G.N2R G.N2I
G.A1		
G.A2		
G.C1		
G.C2		
G.C3		
G.V1		
G.V2		
G.N3		

1.3. Organigramme

1.3.1. L'organigramme de traitement que nous avons établi est "long". Nous n'en donnons ici qu'un fragment, à titre d'exemple : le cas du traitement des syntagmes commençant par [do] et [be] .

1.3.2. Voir Annexe 3, page 157.

2. EXTRACTION DES CATEGORIES VERBALES

2.1. Un programme d'extraction des catégories verbales du syntagme verbal homogène existe (1).

2.1.1. Il a été établi pour les syntagmes verbaux homogènes continus, et fonctionne pour tout syntagme débutant par les auxiliaires :

(1) cf. SVH, pages 96 et suivantes, ainsi que dépliant, et cf infra Annexe 4 pages 166 et suivantes.

[need, dare, shall, will, should, would, may, might, can, could, must, ought, used, have, be, get, do.

2.1.2. Pour le syntagme verbal homogène discontinu, deux cas se sont présentés :

A.) - Après regroupement des divers éléments du syntagme verbal homogène discontinu, nous obtenions un schéma semblable à celui d'un syntagme verbal continu. Dans ce cas, l'organigramme était directement applicable.

- Après regroupement des divers éléments, nous obtenions un schéma verbal non prévu. Dans ce cas, l'organigramme n'était pas directement applicable, et a nécessité quelques aménagements.

B.) Selon que la forme de la phrase est interrogative ou non, les indices de la catégorie "valeur" sont différents. Ceci nous a obligés à prévoir de nouveaux tests.

2.2. Modifications apportées à l'organigramme existant, pour son utilisation avec les syntagmes verbaux homogènes discontinus.

2.2.1. Phrases non-interrogatives.

A.) Tout schéma discontinu regroupé commençant par un des auxiliaires cités plus haut (§ 2.1.1.), peut être traité par l'organigramme non modifié.

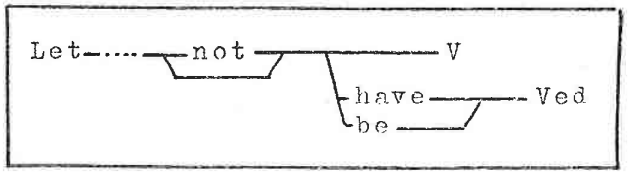
B.) Reste le cas des schémas verbaux commençant par "let", pour lesquels il faut prévoir des branches supplémentaires à l'organigramme.

Cas des schémas commençant par "let" :

En prenant modèle sur les correspondances établies entre syntagmes verbaux et catégories verbales dans l'étude précé-

dente (1), nous établissons les tableaux suivants :

A.) Ordre des constituants :



B.) Correspondance Syntagme/Catégories

Nous utilisons les mêmes symboles que dans la publication précédente (1), mais nous sommes amenés à en ajouter de nouveaux.
En effet, nous n'avons pas encore de symbole NP (indice de nombre et personne) correspondant à l'ambiguité suivante : 1° et 3° personne du singulier et 1° et 3° personne du pluriel. Nous décidons de coder cette ambiguité NP 09.

Nous obtenons les correspondances suivantes :

Let	V	V1	A1	F1	P1	M3	D0	T1	NP09
Let	be Ved	V2	A1	F1	P1	M3	D0	T1	NP09
Let	have Ved	V1	A1	F1	P2	M3	D0	T1	NP09

(Dans tous les cas, la présence de "not" transforme l'indice F1 en F2)

2.2.2. Phrases interrogatives.

A.) Nous avons dit dans la publication précédente (2) qu'il existait 4 "valeurs" différentes : affirmative, négative, interrogative, et interro-négative. La présence du point d'interrogation influe donc sur l'indice de valeur.

---

(1) cf. SVH pages 83 et suivantes
(2) cf. SVH page 64, § II.1.

Mais il n'a aucune influence sur les autres catégories verbales.

B.) Le test sur la présence du point d'interrogation peut être fait en début ou en fin de traitement. Si le schéma verbal n'est pas interrogatif, les indices de valeur restent F1 ou F2 suivant les cas; si le schéma verbal est interrogatif, nous ajoutons 2 au chiffre accolé à F, de façon à obtenir F3 ou F4 suivant les cas.

2.3. En conclusion :

Les modifications à apporter sont peu nombreuses :

a) ajouter quelques tests permettant de traiter les syntagmes verbaux homogènes commençant par "let".

b) prévoir, par exemple en fin de traitement, un test sur la présence du point d'interrogation.

C O N C L U S I O N

1. CONTRE-EPREUVE ET RESULTATS

1.1. L'établissement de notre Corpus de syntagmes verbaux discontínus s'est fait par une série de contre-épreuves :

a) la première série de 258 syntagmes homogènes discontinus nous a donné 49 schémas différents (1).

b) la deuxième série (2) nous a donné, pour 418 syntagmes supplémentaires, 60 nouveaux schémas; soit un total de 119 schémas pour 676 syntagmes.

c) une troisième série de 381 syntagmes nous a apporté 64 nouveaux schémas; soit un total de 183 schémas différents pour 1057 syntagmes verbaux homogènes.

d) une dernière contre-épreuve nous a permis de relever 451 syntagmes verbaux homogènes discontinus, apportant 24 nouveaux schémas; soit un total de 207 schémas différents pour 1508 syntagmes verbaux.

1.2. Nous avons tiré les conclusions suivantes :

1.2.1. Au fur et à mesure que l'échantillonage s'accroît, la proportion de schémas non prévus tend à décroître :

a) 1° série : 49 schémas différents pour 258 syntagmes discontinus; soit 1 schéma pour 5,2 syntagmes.

b) 1° et 2° séries : 119 schémas différents

---

(1) cf. infra "Statistiques machine", Annexe 1, page 121
(2) cf. infra "Statistiques manuelles", Annexe 1, page 127

pour 676 syntagmes; soit 1 schéma pour 5,6 syntagmes.

c) 1°, 2° et 3° séries : 183 schémas verbaux pour 1057 syntagmes; soit 1 schéma pour 5,7 syntagmes.

d) 1°, 2°, 3° et 4° séries : 207 schémas verbaux pour 1508 syntagmes; soit 1 schéma pour 7,2 syntagmes.

1.2.2. Les schémas verbaux nouveaux qui apparaissent, ne sont pas très productifs, alors que le chaînon G.A1 confirme sa très nette avance : 940 chaînes insérées composées du seul chaînon G.A1, pour 1508 syntagmes verbaux homogènes discontinus.

1.2.3. Les risques de regroupements erronés, ou de regroupements non effectués sont relativement rares puisque

a) les schémas les plus productifs ont été prévus.

b) nous avons pu déterminer certains moyens de les éviter (subdivision des groupes G.N1 et G.N2, tests sur les mots qui précèdent l'auxiliaire temporel, etc..

1.2.4. Nous avons eu la confirmation que nos critères de délimitation des chaînons insérés étaient rentables.

2. TEST DE LA METHODE de reconnaissance du syntagme verbal homogène discontinu.

A défaut de pouvoir présenter un programme accompagné de ses résultats, nous avons tenu à extraire quelques phrases d'un certain nombre de textes, et à montrer l'aboutissement théorique du programme de reconnaissance automatique des composantes du syntagme verbal homogène.

1) Where are the mulberry trees grown ?

Le schéma /[be] G.N1 G.N1 Ved/ interrogatif existe, et est prévu. Donc, /are grown/ formera

un syntagme verbal homogène possible.

2) Where is the shortest sea crossing ?

Le schéma [be G.N1 Ving] interrogatif existe. Donc [is crossing] sera regroupé en un syntagme verbal homogène possible.

MAIS le regroupement est mauvais. Un moyen d'éviter ce genre de regroupements, est d'indiquer au dictionnaire le plus grand nombre possible de lexies (adverbiales et nominales), et surtout d'indiquer que ces regroupements sont possibles et non obligatoires.

3) Are most of the coal mining areas north or south of 52° N ?

Le schéma [be G.N2 G.C2 Ving] interrogatif n'a pas été rencontré, donc n'est pas prévu.

Le regroupement [are mining] ne sera pas fait.

[are] reste seul et constitue à lui seul un syntagme verbal homogène; ce quiest la solution correcte dans cette phrase.

4) Why is it in a good position to defend the City ?

Le schéma [be G.N2 G.C1 to V] interrogatif n'est pas prévu. Donc, [is to defend] ne sera, heureusement, pas considéré comme un syntagme verbal homogène.

[is] constitue à lui seul un syntagme verbal homogène.

5) There could not have been much dirty mud coming into the sea at that time.

Il y a ici deux raisons de ne pas effectuer le regroupement [could not have been coming] :

a) le schéma [defectif not have been G.N1 Ving] n'existe pas (aucune Chaîne 2 ne peut être composée de G.N1).

b) l'auxiliaire temporel est précédé de "there".

Donc [could not have been] sera reconnu comme seul syntagme verbal homogène temporel possible.

6) There are two types of vegetation shown in the area ..

Ici encore, nous avons deux raisons de ne pas effectuer le regroupement [are shown] :

a) le schéma [be G.N1 G.C2 Ved] n'existe qu'en phrases interrogatives.

b) "are" est précédé de "there".

Donc "are" reste le seul syntagme verbal homogène possible, ce qui est la solution correcte.

7) What is the chief use of the crops grown on the farms ?

Le schéma [be G.N1 G.C2 Ved] est un schéma interrogatif prévu. Le regroupement erroné [is grown] sera donc considéré comme possible.

Il faudra attendre que des études sur les schémas de phrases, sur l'ordre des constituants des différents types de phrases, soient faites, pour empêcher ce genre de regroupements.

8) The east bay of the choir has its roof raised above the remainder.

Le schéma [have G.N1 Ved] est un schéma non-

interrogatif prévu sous certaines conditions seulement. Les conditions requises n'étant pas remplies, le regroupement /has raised/ nesera pas possible, et /has/ constitue à lui seul un syntagme verbal homogène.

3. CONCLUSION.

Notre étude, fondée sur l'observation des données d'un Corpus, a permis de prévoir l'ordre des constituants d'un syntagme verbal homogène, et la nature des insertions qui peuvent s'y trouver. Elle montre qu'il est possible à un ordinateur de détecter dans un texte non pré-édité les composantes d'un syntagme verbal, de les regrouper, et de les analyser.

Nous avons tenté de montrer qu'il est fréquemment possible de résoudre une part importante des problèmes de décryptage du groupe verbal dès le premier niveau d'analyse, c'est à dire à partir du test de toute forme graphique pouvant constituer son premier élément.

L'avantage de la méthode est de fournir des résultats immédiatement ré-utilisables pour le décryptage syntaxique du reste de la phrase. En revanche, pour résoudre les cas (statistiquement minoritaires) laissés en suspens, il faudra recourir aux résultats d'une analyse syntaxique du reste de la phrase.

Les règles de combinatoire linéaire dégagées par nous, permettent de faire l'économie de tests syntaxiques couteux.

A N N E X E S

Annexe 1 . . . . . page 120

Annexe 2 141

Annexe 3 157

Annexe 4 166

Annexe 5 171

Annexe 6 174

A N N E X E 1

"Statistiques Machine" et "Statistiques manuelles" ont été faites sur destextes différents, et dans une optique quelque peu différente. Une récapitulation, sous forme de tableaux comparatifs, a cependant pu être faite pour certains aspects particuliers.

Statistiques Machine . . . . . page 121

Statistiques manuelles 127

Récapitulation 137

STATISTIQUES MACHINE

Elles portent sur 35 textes différents, extraits du CORPUS d'anglais scientifique qui a servi de base à notre étude. Ces textes traitent principalement d'archéologie, de physique, de biologie et de cartographie.

Un nombre total de 48 322 cartes-mots a été traité par l'ordinateur GAMMA 60, à l'aide d'un programme EXSYN qui extrayait automatiquement tous les groupes verbaux mineurs, temporels ou non. (voir en Annexe 5).

Ceci a permis d'obtenir les résultats suivants :

1. RESULTATS GLOBAUX :

Groupes verbaux	A-TEMPORELS	TEMPORELS	Nombre total
CONTINUS	1166	2886	4052
DISCONTINUS	4	258	262
COMPLEXES	13	61	74
Nombre total	1183	3205	4388

2. FREQUENCE D'APPARITION DES DIFFERENTS SYNTAGMES DISCONTINUS :

(Il s'agit des syntagmes discontinus temporels, au nombre de 258).

[do]		G.A2	V				1	
[do]		G.N2	V				2	
[do]		G.N1	V				1	8
[do]	not	G.A1	V				4	
"let"		G.N2	V				3	3
[be]		G.A1	Ving				1	
[be]	not	G.A1	Ving				1	
[be]		G.A1	Ved				101	
[be]		G.A2	Ved				1	
[be]		G.V1	Ved				1	
[be]		G.N2	Ved				2	
[be]		G.A1	G.A1	Ved			9	126
[be]		G.N1	G.V1	Ved			1	
[be]		G.A1	G.A1	G.A1	Ved		1	
[be]		G.V1	G.A1	G.A1	Ved		1	
[be]	not	G.A1	Ved				5	
[be]	not	G.A1	G.A1	Ved			1	
[be]		G.A1	being	Ved			1	
[have]		G.A1	Ved				10	
[have]		G.I	Ved				1	
[have]		G.A1	been	Ved			11	
[have]	not	G.A1	been	Ved			2	
[have]	not		been	G.A1	Ved		1	35
[have]			been	G.A1	Ved		8	
[have]			been	G.A1 ⊕ G.A1		Ved	1	

defectif		G.A1	V				21	
defectif		G.A2	V				3	
defectif		G.V1	V				5	
defectif		G.N2	V				2	
defectif		G.N3	V				1	
defectif		G.A1	G.A1	V			3	
defectif	not	G.A2	V				1	
defectif	not	but	V				1	
defectif		G.A1	be	Ved			19	
defectif		G.A2	be	Ved			2	
defectif		G.N1	be	Ved			1	
defectif		G.N2	be	Ved			1	
defectif		G.V1	be	Ved			3	
defectif		G.N1	G.N1	be	Ved		1	86
defectif	not	G.A1	be	Ved			2	
defectif	not		be	G.A1	Ved		2	
defectif			be	G.A1	Ved		11	
defectif		G.A1	be	G.A1	Ved		2	
defectif		G.A2	be	G.A1	Ved		1	
defectif		G.N1	G.C2	have	Ved		1	
defectif			have	G.A1	Ved		1	
defectif		G.A1	have	been	Ved		1	
defectif			have	been	G.A1	Ved	1	

soit,

a) 49 schémas différents pour 258 syntagmes verbaux discontinus.

b) 4 syntagmes seulement, sur ces 258, possèdent 2 séquences insérées.

3. FREQUENCE D'APPARITION DES CHAINES INSEREES.

(18 schémas différents pour 262 chaînes insérées)

3.1. Classement par fréquence décroissante :

G.A1	196
not G.A1	14
G.A1 G.A1	12
G.V1	9
G.N2	9
G.A2	8
G.N1	3
G.N3	1
G.I	1
G.N1 G.V1	1
G.N1 G.N1	1
G.N1 G.C2	1
not G.A2	1
not but	1
not G.A1 G.A1	1
G.A1 G.A1 G.A1	1
G.V1 G.A1 G.A1	1
G.A1 ⊕ G.A1	1

3.2. Classement selon la forme de la chaîne insérée :

G.A1	196
G.A2	8
G.I	1
G.N1	3
G.N2	9
G.N3	1
G.V1	9
G.A1 G.A1	12
G.N1 G.V1	1
G.N1 G.N1	1
G.N1 G.C2	1
G.A1 G.A1 G.A1	1
G.V1 G.A1 G.A1	1
not G.A1	14
not G.A2	1
not but	1
not G.A1 G.A1	1
G.A1 ⊕ G.A1	1

4. FREQUENCE DE L'INSERTION, EN RAPPORT AVEC LA ZONE A.A-T.

Chaîne insérée	Zone A.A-T = 0	Zone A.A-T ≠ 0				
		insertion au point 1		insertion au point 2		
		2° point = 0	2° point ≠ 0	1° point = 0	1° point = not	1° point ≠ not 0
G.A1	133	32	3	21	3	4
G.A2	5	2	1			
G.I	1					
G.N1	2	1				
G.N2	8	1				
G.N3	1					
G.V1	6	3				
G.A1 G.A1	12					
G.N1 G.V1	1					
G.N1 G.N1		1				
G.N1 G.C2		1				
G.A1 G.A1 G.A1	1					
G.V1 G.A1 G.A1	1					
not G.A1	10	4				
not G.A2	1					
not but	1					
not G.A1 G.A1	1					
G.A1 ⊕ G.A1				1		
	184	45	4	22	3	4

Le tableau précédent permet en outre de déduire :

	1 Chaîne insérée		2 Chaînes insérées	
	Point 1	Point 2	Points 1 et 2	
Zone A.A-T = 0	184		-	184
Zone A.A-T = 1	45	25	2	72
Zone A.A-T = 2	0	2	0	2
	256		2	258

STATISTIQUES MANUELLES

1. RESULTAT NUMERIQUE.

Ces statistiques portent sur 2 livres de géographie à usage scolaire.

Seuls les syntagmes verbaux homogènes (donc temporels), ont été relevés.

Le tableau ci-dessous donne un résultat comparatif partiel. En effet, les 2 livres n'ontpas été dépouillés totalement pour cette statistique (cf. supra page 7, § 3.2.2.).

SYNTAGMES	CONTINUS	DISCONTINUS	Nombre total
dans Bk. 1	2637	177	2814
dans Bk. 2	2168	112	2280
Nombre total	4805	289	5094

2. STATISTIQUES PORTANT SUR LA TOTALITE DES SYNTAGMES VERBAUX DISCONTINUS RELEVES.

Nous avons inventorié 418 syntagmes verbaux homogènes discontinus dans ces deux livres.

Ils sont distribués comme suit :

Bk. 1 : 168 syntagmes verbaux homogènes discontinus, dont 7 dans des phrases interrogatives.

Bk. 2 : 250 syntagmes verbaux homogènes discontinus, dont 141 dans des phrases interrogatives.

2.1. FREQUENCE D'APPARITION DES 418 SYNTAGMES

2.1.1. Le tableau ci-dessous (§ 2.1.2.) est divisé en deux parties :

a) à gauche une colonne "schémas de syntagmes"

b) à droite une colonne "fréquence d'apparition", qui se subdivise en 6 sous-colonnes :

1) Fréquence, dans Bk. 2, phrases non interrogatives.

2) Fréquencedans Bk. 2, phrases terminées par un point d'interrogation.

3) Fréquence dans Bk. 1, phrases non interrogatives.

4) Fréquence dans Bk. 1, phrases terminées par un point d'interrogation.

5) Nombre total d'occurrences pour chaque type de syntagme (addition, dans le sens horizontal, des contenus des colonnes 1, 2, 3, et 4)

6) Nombre total d'occurrences pour chaque catégorie de syntagme, catégorie qui est fondée sur le premier mot du syntagme (le nombre inscrit en colonne 6 correspond à l'addition, dans le sens vertical, des nombres contenus dans la colonne 5 pour la catégorie de syntagmes considérée).

2.1.2. Tableau de fréquence :

SCHEMAS DE SYNTAGMES		FREQUENCE					
		1	2	3	4	5	6
DEFECTIF	G.N1 V	1			2	3	
	G.N2 V				14	14	
	G.A1 V	8		17		25	
	G.A2 V	2				2	
	G.V1 V			1		1	
	not G.A1 V	1				1	
	not G.N2 V			1		1	
	G.N2 G.C2				1	1	
	G.A1 be Ved	5		1		6	
	G.A1 be G.A1 Ved	1				1	67
	be G.A1 Ved	3				3	
	be G.A1 G.A1 Ved			1		1	
	G.A1 have Ved	1		1		2	
	have G.A2 Ved	1				1	
	have G.A1 G.A1 Ved	1				1	
	G.A1 be G.A1 Ved			2		2	
	G.C1 be G.A1 Ved	1				1	
	G.A1 have been Ving	1				1	
[HAVE]	G.A1 Ved	10		12		22	
	G.N2 Ved		2			2	
	G.C1 Ved	1				1	
	not G.A1 Ved			1		1	
	not G.V1 Ved			1		1	
	G.A1 G.A1 Ved	1				1	
	G.N2 G.A1 Ved		3			3	
	G.C1 G.A1 Ved	1				1	54
	G.A1 been Ved	7		4		11	
	G.N1 been Ved				1	1	
	G.C1 been Ved	2				2	
	G.A1 been G.A1 Ved	1		2		3	
	been G.A1 Ved	2		1		3	
	⊕ G.A1 ⊖ G.A1 Ved			1		1	
	been ⊕ Ved	1				1	

	1	2	3	4	5	6
LET G.N2 V	1				1	1
[DO] G.N1 V				19	19	
G.N2 V		1		37	38	
G.N3 V				1	1	
G.N1 G.C2 V		1		5	6	
G.N1 G.A1 V				1	1	
G.N2 G.C2 V				2	2	
G.N1 G.N1 V				4	4	
G.N1 G.C1 V				2	2	
G.N1 G.C2 G.C2 V				1	1	
G.N1 G.C1 G.C1 V				1	1	
G.N1 G.C2 G.C1 V				1	1	
G.N1 G.N1 G.A1 V				1	1	
G.N1 G.N1 G.N1 G.C1 V				1	1	
G.N1 G.N1 G.C1 G.C1 V				1	1	
G.N2 G.C2 G.C1 G.C2 V				1	1	91
G.N1 ⊕ G.N1 V				1	1	
G.N1 G.N1 ⊕ G.N1 V				1	1	
G.N1 G.C1 ⊕ G.C1 V				1	1	
G.N1 G.V1 ⊕ G.N1 V				1	1	
G.N1 G.C2 ⊕ G.N1 G.V1 V				1	1	
G.N1 G.N1 G.N1 ⊕ G.N1 V				1	1	
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1 V				1	1	
G.N1 have to V				1	1	
G.N2 have to V				3	3	

	1	2	3	4	5	6
[BE] G.A1 Ved	77		44		121	
G.A2 Ved	3		1		4	
G.N1 Ved				10	10	
G.N2 Ved				4	4	
G.I Ved				1	1	
not G.A1 Ved			1		1	
G.A1 G.A1 Ved	7		7		14	
G.N1 G.N1 Ved				4	4	
G.N1 G.C1 Ved				1	1	
G.N1 G.A1 Ved				1	1	
G.N1 G.C2 Ved				2	2	
G.N2 G.C2 Ved				4	4	
G.N1 G.C2 G.A1 Ved				1	1	
G.N1 G.C1 G.C2 Ved				1	1	205
G.N1 G.A1 G.A1 Ved				1	1	
G.A1 G.A1 G.A1 Ved	1		1		2	
not G.A1 G.A1 Ved	2				2	
G.N1 ⊕ G.N1 Ved				1	1	
G.N1 Ving				2	2	
G.A1 Ving	10	4			14	
G.N1 G.C1 Ving				1	1	
G.A1 G.A1 Ving	1				1	
G.A1 G.A1 G.A1 Ving	1				1	
G.A1 being Ved	5		2		7	
being G.A1 Ved	1		3		4	
	161	7	109	141	418	418

2.1.3. Conclusions amenées par cette série de tableaux :

a) 83 schémas verbaux différents, pour 418 syntagmes.

b) Sur ces 83 schémas, 42 sont utilisés uniquement dans des phrases interrogatives directes.

c) Sur ces 418 syntagmes verbaux discontinus,

363 ont une zone A.A-T = 0
54 ont une zone A.A-T = 1 élément
1 a une zone A.A-T = 2 éléments.

d) Ces 418 syntagmes se répartissent commesuit :

Syntagmes à	Zone A.A-T			
	= 0	= 1	= 2	
Schéma toujours interrogatif	141	5	0	146
Schéma jamais interrogatif	219	49	1	269
Schéma bivalent	3	0	0	3
	363	54	1	418

e) De plus, sur les 55 syntagmes ayant une zone A.A-T différente de 0, 7 possèdent 2 Chaînes insérées, contre 48 qui n'en possèdent qu'une seule :

Syntagmes à	une seule Chaîne		deux Chaînes
	Point 1	Point 2	
Schéma toujours interrogatif	5	0	0
Schéma jamais interrogatif	29	14	7
Schéma bivalent	0	0	0

f) Les schémas interrogatifs sont bien distincts des autres. Seul un schéma est bivalent :

[defectif G.N1 V

2.2. FREQUENCE D'APPARITION DES 425 CHAINES INSEREES :

2.2.1. Résultats ordonnés selon la forme de la Chaîne :

Schéma de la Chaîne insérée	dans phrase avec "?"	dans phrase sans "?"	
G.A1		232	232
G.A2		7	7
G.N1	35	1	36
G.N2	61	1	62
G.N3	1		1
G.C1		4	4
G.V1		1	1
G.I	1		1
G.A1 G.A1		18	18
G.C1 G.A1		1	1
G.N1 G.A1	2		2
G.N1 G.C1	4		4
G.N1 G.C2	8		8
G.N1 G.N1	8		8
G.N2 G.A1	3		3
G.N2 G.C2	7		7
not G.A1		3	3
not G.N2		1	1
not G.V1		1	1
not G.A1 G.A1		2	2
G.A1 G.A1 G.A1		3	3
G.N1 G.A1 G.A1	1		1
G.N1 G.C1 G.C1	1		1
G.N1 G.C1 G.C2	1		1
G.N1 G.C2 G.A1	1		1
G.N1 G.C2 G.C1	1		1
G.N1 G.C2 G.C2	1		1
G.N1 G.N1 G.A1	1		1
G.N1 G.N1 G.C1 G.C1	1		1
G.N1 G.N1 G.N1 G.C1	1		1
G.N2 G.C2 G.C1 G.C2	1		1
G.N1 ⊕ G.N1	2		2
⊕ G.A1 ⊕ G.A1		1	1
G.N1 G.V1 ⊕ G.N1	1		1
G.N1 G.N1 ⊖ G.N1	1		1
G.N1 G.C1 ⊕ G.C1	1		1
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1	1		1
G.N1 G.N1 G.N1 ⊕ G.N1	1		1
G.N1 G.C2 ⊕ G.N1 G.V1	1		1
⊕	1		1
	148	277	425

2.2.2. Résultats ordonnés suivant la fréquence décroissante d'apparition :

G.A1	232
G.N2	62
G.N1	36
G.A1 G.A1	18
G.N1 G.C2	8
G.N1 G.N1	8
G.A2	7
G.N2 G.C2	7
G.C1	4
G.N1 G.C1	4
not G.A1	3
G.N2 G.A1	3
G.A1 G.A1 G.A1	3
not G.A1 G.A1	2
G.N1 G.A1	2
G.N1 ⊕ G.N1	2
G.I	1
G.N3	1
G.V1	1
⊕	1
G.C1 G.A1	1
not G.N2	1
not G.V1	1
G.N1 G.A1 G.A1	1
G.N1 G.C1 G.C2	1
G.N1 G.C2 G.A1	1
G.N1 G.C1 G.C1	1
G.N1 G.C2 G.C1	1
G.N1 G.C2 G.C2	1
G.N1 G.N1 G.A1	1
G.N1 G.N1 G.C1 G.C1	1
G.N1 G.N1 G.N1 G.C1	1
G.N2 G.C2 G.C1 G.C2	1
G.N1 G.C1 ⊕ G.C1	1
G.N1 G.N1 ⊕ G.N1	1
G.N1 G.V1 ⊕ G.N1	1
G.N1 G.C2 ⊕ G.N1 G.V1	1
G.N1 G.N1 G.N1 ⊕ G.N1	1
⊕ G.A1 ⊕ G.A1	1
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1	1

2.2.3. Résultats ordonnés suivant le type de phrase (interrogative ou non), et la fréquence :

a) Phrases sans "?" :

G.A1	232
G.A1 G.A1	18
G.A2	7
C.C1	4
not G.A1	3
G.A1 G.A1 G.A1	3
not G.A1 G.A1	2
not G.N2	1
not G.V1	1
G.V1	1
G.N1	1
G.N2	1
G.C1 G.A1	1
⊕	1
⊕ G.A1 ⊕ G.A1	1

b) Phrases sans "?" :

G.N2	61
G.N1	36
G.N1 G.C2	8
G.N1 G.N1	8
G.N2 G.C2	7
G.N1 G.C1	4
G.N2 G.A1	3
G.N1 G.A1	2
G.N1 ⊕ G.N1	2
G.N3	1
G.I	1
G.N1 G.C2 G.A1	1
G.N1 G.C2 G.C2	1
G.N1 G.C2 G.C1	1
G.N1 G.C1 G.C2	1
G.N1 G.C1 G.C1	1
G.N1 G.A1 G.A1	1
G.N1 G.N1 G.A1	1
G.N1 G.N1 G.C1 G.C2	1
G.N1 G.N1 G.N1 G.C1	1
G.N2 G.C2 G.C1 G.C2	1
G.N1 G.C1 ⊕ G.C1	1
G.N1 G.N1 ⊕ G.N1	1
G.N1 G.V1 ⊕ G.N1	1
G.N1 G.C2 ⊕ G.N1 G.V1	1
G.N1 G.N1 G.N1 ⊕ G.N1	1
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1	1

2.2.4. Conclusions amenées par cette deuxième série de tableaux :

a) 40 schémas différents de Chaînes insérées, pour 425 Chaînes .

b) Ces schémas de Chaînes insérées sont assez bien différenciés, selon que la phrase est interrogative ou non.

c) Dans les phrases quine sont pas terminées par un point d'interrogation, G.A1 apparaît 232 fois sur 277.
Dans les phrases terminées par un point d'interrogation, c'est G.N2 qui apparaît le plus souvent : 61 fois sur 148.

QUELQUES TABLEAUX COMPARATIFS

1. SYNTAGMES VERBAUX COMPTABILISES :

	Groupes verbaux	Syntagmes temporels	
		continus	discontinus
Statistiques machine	4052	2886	258
Statistiques manuelles	5094	4805	289
	9146	7691	547

Proportion :

a) dans les textes scientifiques pour adultes :
258 discontinus pour 2886 continus, soit
1 discontinu pour 10,11 continus.

b) dans les livres de géographie pour enfants :
289 discontinus pour 4805 continus, soit
1 discontinu pour 16,62 continus.

La proportion moyenne de syntagmes verbaux discontinus est donc de 1 pour 14 syntagmes continus environ.

2. SYNTAGMES VERBAUX DISCONTINUS RELEVES :

(pour ces deux séries de statistiques)

RELEVE	SCHEMAS DIFFERENTS	SYNTAGMES DIFFERENTS
Machine : 258	49 dans phrases sans "?"	258 { A.A-T = 0 : 184 ; A.A-T = 1 : 72 ; A.A-T = 2 : 2 }
Main : 418	83 { 41 dans phrases sans "?" ; 43 dans phrases avec "?" ; (1 schéma bivalent) }	269 { A.A-T = 0 : 219 ; A.A-T = 1 : 49 ; A.A-T = 2 : 1 } ; 149 { A.A-T = 0 : 144 ; A.A-T = 1 : 5 ; A.A-T = 2 : 0 }

c'est à dire :

a) dans les textes scientifiques :

74 discontinus dont la zone A.A-T est $\neq 0$, pour
258 syntagmes discontinus au total;
soit 1 pour 3,4

b) dans les livres de géographie :

55 discontinus dont la zone A.A-T est $\neq 0$, pour
418 syntagmes discontinus au total;
soit 1 pour 7,6

STATISTIQUES MACHINE	STATISTIQUES MANUELLES
Textes scientifiques pour adultes	textes géographiques pour enfants
258 syntagmes relevés	418 syntagmes relevés
49 schémas de syntagmes différents tous schémas non-interrogatifs	83 schémas de syntagmes différents c'est à dire 42 schémas interrogatifs uniquement 40 schémas non-interrogatifs 1 schéma bivalent
1 discontinu pour 10,11 continus	1 discontinu pour 16,62 continus
74 discontinus dont la zone A.A-T est ≠ 0 , sur un total de 258 soit, 1 pour 3,4 4 avec 2 Chaines insérées 70 avec 1 Chaîne insérée 45 insérées au 1° point 25 insérées au 2° point (suite page suivante)	55 discontinus dont la zone A.A-T est ≠ 0 , sur un total de 418 soit, 1 pour 7,6 7 avec 2 Chaînes insérées 48 avec 1 Chaîne insérée 34 insérées au 1° point 14 insérées au 2° point

<table>
<tr><th>STATISTIQUES MACHINE</th><th>STATISTIQUES MANUELLES</th></tr>
<tr><td>262 Chaînes insérées relevées</td><td>425 Chaînes insérées relevées</td></tr>
<tr><td>18 schémas de Chaînes différents

tous dans des schémas de syntagmes verbaux discontinus interrogatifs.</td><td>40 schémas de Chaînes différents
c'est à dire
12 dans des syntagmes à schéma interrogatif uniquement
26 dans des syntagmes à schéma non-interrogatif uniquement
2 schémas de chaînes insérées bivalents.</td></tr>
<tr><td>Chaînes les plus fréquentes
<table><tr><th>phrases sans "?"</th></tr><tr><td>G.A1 196</td></tr><tr><td>not G.A1 14</td></tr><tr><td>G.A1 G.A1 12</td></tr><tr><td>G.V1 9</td></tr><tr><td>G.N2 9</td></tr></table></td><td>Chaînes les plus fréquentes
<table><tr><th>phrases sans "?"</th><th>phrases avec "?"</th></tr><tr><td>G.A1 232</td><td>G.N2 61</td></tr><tr><td>G.A1 G.A1 18</td><td>G.N1 36</td></tr><tr><td>G.A2 7</td><td>G.N1 G.C2 8</td></tr><tr><td>G.C1 4</td><td>G.N1 G.N1 8</td></tr></table></td></tr>
<tr><td colspan="2">au total, pour les 2 types de textes :
45 schémas différents de chaînes insérées
110 schémas différents de syntagmes discontinus.</td></tr>
</table>

A N N E X E 2

Listes de Fréquence

Schémas de chaînes insérées . . . . . page 142

Schémas de syntagmes verbaux homogènes discontinus 144

Schémas de syntagmes verbaux homogènes discontinus comportant la particule "not" 150

Schémas de syntagmes verbaux homogènes discontinus trouvés en phrases interrogatives 153

LISTE DE TOUS LES SCHEMAS DE CHAINES INSEREES

phrases avec "?"	phrases sans "?"	Schémas des chaînes insérées
0	1	G.C2
0	1	G.C3
2	4	G.I
0	32	G.A2
0	2	G.A2 not
0	1	G.A2 G.A1
3	3	G.N3
0	22	G.C1
0	1	G.C1 G.A1
0	2	G.C1 G.C2
0	1	G.C1 not
0	31	G.V1
0	4	G.V1 G.A1
0	1	G.V1 G.A1 G.A1
0	1	G.V1 not
72	45	G.N2
0	1	G.N2 not
3	6	G.N2 G.A1
7	1	G.N2 G.C2
0	1	G.N2 G.A1 G.A1
1	0	G.N2 G.C2 G.C1
1	0	G.N2 G.C2 G.C1 G.C2
0	940	G.A1
0	3	G.A1 not
0	102	G.A1 G.A1
0	2	G.A1 G.N1
0	4	G.A1 G.V1
0	15	G.A1 G.A1 G.A1
42	20	G.N1
11	5	G.N1 G.C2
3	1	G.N1 G.A1
5	0	G.N1 G.C1
9	3	G.N1 G.N1
0	1	G.N1 G.V1
1	0	G.N1 G.C2 not
0	2	G.N1 G.N1 G.V2
1	0	G.N1 G.A1 G.A1

1	0	G.N1 G.C1 G.C2
1	0	G.N1 G.C2 G.C2
1	0	G.N1 G.N1 G.A1
1	0	G.N1 G.C1 G.C1
2	0	G.N1 G.C2 G.C1
0	2	G.N1 G.C2 G.A1
1	0	G.N1 G.N1 G.N1 G.C1
1	0	G.N1 G.N1 G.C1 G.C1
1	0	G.N1 G.N1 G.C1 G.C2
0	1	G.N1 G.C2 G.C1 G.C2
0	4	not G.A2
0	6	not G.V1
0	1	not G.C2
0	3	not G.N2
0	1	not G.C1
0	39	not G.A1
0	5	not G.A1 G.A1
1	0	not G.N1 G.C2
1	0	not G.N1 G.C1
0	1	not G.V1 G.A1
0	1	not G.A1 G.A1 G.A1
0	2	not but
0	8	G.A1 ⊕ G.A1
0	3	G.A1 G.A1 ⊕ G.A1
0	3	G.A1 ⊕ G.A1 G.A1
2	0	G.N1 ⊕ G.N1
1	0	G.N1 G.N1 ⊕ G.N1
1	0	G.N1 G.C1 ⊕ G.C1
1	1	G.N1 G.C1 ⊕ G.N1
1	0	G.N1 G.V1 ⊕ G.N1
1	0	G.N1 G.N1 G.N1 ⊕ G.N1
1	0	G.N1 G.C2 ⊕ G.N1 G.V1
0	4	⊕
0	1	⊕ G.A1
0	1	⊕ G.A1 ⊕ G.A1
1	0	⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1
0	2	lettre alphabétique
0	4	formule mathématique
182	1352	

soit 74 schémas de chaînes insérées pour un total de 1534 chaînes insérées dans 1508 syntagmes verbaux homogènes.

LISTE DE TOUS LES SCHEMAS DE SYNTAGMES DISCONTINUS

Phrases avec "?"	Phrases sans "?"	Schémas de syntagmes
0	1	[DO] G.A2 V
19	4	G.N1 V
38	7	G.N2 V
2	0	G.N3 V
0	14	not G.A1 V
0	1	not G.N2 V
0	1	not G.C1 V
0	1	not G.C2 V
0	5	not G.V1 V
2	0	G.N2 G.C2 V
1	0	G.N1 G.A1 V
3	0	G.N1 G.C1 V
6	2	G.N1 G.C2 V
5	0	G.N1 G.N1 V
1	0	G.N1 G.C2 G.C2 V
1	0	G.N1 G.C2 G.C1 V
0	1	G.N1 G.C2 G.A1 V
1	0	G.N1 G.C1 G.C1 V
1	0	G.N1 G.N1 G.A1 V
1	0	G.N1 G.N1 G.C1 G.C1 V
1	0	G.N1 G.N1 G.N1 G.C1 V
1	0	G.N1 G.N1 G.C1 G.C2 V
1	0	G.N1 G.C2 G.C1 G.C2 V
2	0	G.N1 have to V
3	0	G.N2 have to V
1	0	G.N1 ⊕ G.N1 V
1	0	G.N1 G.C1 ⊕ G.N1 V
1	0	G.N1 G.C1 ⊕ G.C1 V
1	0	G.N1 G.N1 ⊕ G.N1 V
1	0	G.N1 G.V1 ⊕ G.N1 V
1	0	G.N1 G.C2 ⊕ G.N1 G.V1 V
1	0	G.N1 G.N1 G.N1 ⊕ G.N1 V
1	0	⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1 V
0	1	NEED G.A2 V
0	3	not G.V1 V
0	1	not G.A1 have been Ved

0	41	[BE] G.A1 Ving
0	1	G.N2 Ving
2	0	G.N1 Ving
1	0	G.I Ving
0	5	G.A1 G.A1 Ving
1	0	G.N1 G.C1 Ving
0	1	G.A1 G.A1 G.A1 Ving
0	1	not G.A1 Ving
0	1	not G.V1 Ving
0	1	not G.A1 G.A1 G.A1 Ving
0	18	G.A1 being Ved
0	6	being G.A1 Ved
0	1	G.A1 being G.A1 Ved
0	373	G.A1 Ved
0	9	G.A2 Ved
0	3	G.V1 Ved
0	5	G.C1 Ved
10	5	G.N1 Ved
4	9	G.N2 Ved
1	1	G.N3 Ved
0	1	G.I Ved
0	2	G.A1 not Ved
0	2	G.A2 not Ved
0	11	not G.A1 Ved
0	1	not G.N2 Ved
0	3	not G.A1 G.A1 Ved
0	70	G.A1 G.A1 Ved
0	1	G.A1 G.V1 Ved
0	1	G.A1 G.N1 Ved
0	3	G.V1 G.A1 Ved
2	1	G.N1 G.A1 Ved
0	1	G.N1 G.V1 Ved
4	1	G.N1 G.N1 Ved
1	0	G.N1 G.C1 Ved
2	2	G.N1 G.C2 Ved
0	2	G.N2 G.A1 Ved
4	0	G.N2 G.C2 Ved
0	1	G.N1 G.C2 G.A1 Ved
1	0	G.N1 G.C2 G.C1 Ved
1	0	G.N1 G.C1 G.C2 Ved
1	0	G.N1 G.A1 G.A1 Ved
0	1	G.N1 G.N1 G.V2 Ved
0	13	G.A1 G.A1 G.A1 Ved
0	1	G.V1 G.A1 G.A1 Ved
1	0	G.N2 G.C2 G.C1 Ved
0	1	G.N1 G.C2 G.C1 G.C2 Ved

1	0	G.N1 ⊕ G.N1 Ved
0	6	G.A1 ⊕ G.A1 Ved
0	2	G.A1 ⊕ G.A1 G.A1 Ved
0	2	G.A1 G.A1 ⊕ G.A1 Ved
0	1	⊕ G.A1 Ved
0	2	⊕ Ved
0	6	G.A1 to V
2	0	G.N2 to V
2	0	G.N2 to be Ved
0	3	G.A1 to be Ved
1	0	G.N1 G.C2 to be Ved
0	104	[HAVE] G.A1 Ved
0	1	G.A2 Ved
0	4	G.C1 Ved
0	1	G.I Ved
1	2	G.N1 Ved
2	5	G.N2 Ved
0	1	G.N3 Ved
0	3	G.V1 Ved
0	3	not G.A1 Ved
0	2	not G.A2 Ved
0	1	not G.V1 Ved
0	5	G.A1 G.A1 Ved
0	1	G.A1 G.V1 Ved
3	0	G.N2 G.A1 Ved
0	1	G.C1 G.A1 Ved
0	1	⊕ G.A1 ⊕ G.A1 Ved
0	1	G.A2 been Ved
0	42	G.A1 been Ved
0	2	G.C1 been Ved
1	1	G.N1 been Ved
0	2	G.V1 been Ved
0	1	G.A1 G.A1 been Ved
0	1	G.A1 G.N1 been Ved
0	2	G.C1 G.C2 been Ved
0	4	not G.A1 been Ved
0	1	not G.A2 been Ved
0	1	G.V1 not been Ved
0	1	G.A1 ⊕ G.A1 been Ved
0	1	⊕ been Ved
0	2	not been G.A1 Ved
0	10	G.A1 been G.A1 Ved
0	1	G.N2 G.A1 been G.A1 Ved

0	29	been G.Al Ved
0	1	been ⊕ Ved
0	1	been G.Al ⊕ G.Al Ved
0	2	G.Al to V
1	0	G.N2 to V
0	1	G.Al had to V
0	13	LET G.N2 V
0	3	G.N2 G.Al V
0	1	G.N2 G.Al G.Al V
0	1	G.N1 G.C2 be Ved
0	1	G.N2 not be Ved
0	3	G.N2 be Ved
0	4	formule mathématique V
0	1	lettre alphabétique V
0	1	lettre alphabétique "="
0	120	DEFECTIF G.Al V
0	10	G.A2 V
8	4	G.N1 V
19	6	G.N2 V
0	1	G.N3 V
0	8	G.C1 V
0	1	G.C2 V
0	1	G.C3 V
1	1	G.I V
0	26	G.V1 V
0	1	G.Al not V
0	1	not G.V1 V
0	2	not but V
0	3	not G.Al V
0	1	not G.N2 V
0	1	not G.A2 V
0	1	not G.V1 G.Al V
1	0	G.N1 G.C2 not V
1	0	G.N2 G.C2 V
0	1	G.N1 G.N1 V
0	6	G.Al G.Al V
0	2	G.Al G.V1 V
0	1	G.V1 G.Al V
0	79	G.Al be Ved
0	6	G.A2 be Ved
0	1	G.I be Ved

0	4	G.N1 be Ved
1	0	G.N2 be Ved
0	9	G.V1 be Ved
0	2	G.C1 be Ved
0	3	G.A1 G.A1 be Ved
0	1	G.N1 G.N1 be Ved
0	4	not G.A1 be Ved
0	1	not G.A1 G.A1 be Ved
1	0	not G.N1 G.C2 be Ved
1	0	not G.N1 G.C1 be Ved
0	1	G.N1 G.N1 G.V2 be Ved
0	7	not be G.A1 Ved
0	35	be G.A1 Ved
0	6	be G.A1 G.A1 Ved
0	1	be G.A2 G.A1 Ved
0	1	be G.A1 G.A1 G.A1 Ved
0	7	G.A1 be G.A1 Ved
0	1	G.A2 be G.A1 Ved
0	1	G.N2 be G.A1 Ved
0	1	G.V1 be G.A1 Ved
0	1	G.A1 be G.A1 G.A1 Ved
0	1	G.N1 G.C1 ⊕ G.N1 be Ved
0	1	be G.A1 ⊕ G.A1 G.A1 Ved
0	1	be G.A1 G.A1 ⊕ G.A1 Ved
0	1	G.A1 be Ving
0	2	G.A1 be G.A1 Ving
0	1	G.C1 be G.A1 Ving
0	10	G.A1 have Ved
0	1	G.N2 G.C2 have Ved
0	1	G.A1 G.A1 have Ved
2	0	G.N1 G.C2 have Ved
0	1	not G.A1 have Ved
0	6	have G.A1 Ved
0	1	have G.A2 Ved
0	1	have G.A1 G.A1 Ved
0	1	have been G.A1 Ved
0	3	G.A1 have been Ved
0	1	not G.A1 have been Ved
0	1	G.A1 have been Ving
0	1	G.A1 have to be Ved
182	1326	

Conclusion :

a) 207 schémas verbaux différents pour 1508 syntagmes verbaux homogènes discontinus.

b) Sur ces 207 schémas,

- 43 sont utilisés uniquement dans des phrases interrogatives.
- 149 ne se trouvent jamais en phrases interrogatives.
- 15 sont des schémas bivalents.

c) Sur les 1508 syntagmes discontinus :

1°)
- 1147 possèdent une zone A.A-T = 0
- 353 possèdent une zone A.A-T = 1
- 8 possèdent une zone A.A-T = 2

2°)
- 1482 possèdent une seule chaîne insérée :
  - 1147 en Chaîne unique
  - 233 en Chaîne 1
  - 102 en Chaîne 2
- 26 possèdent 2 chaînes insérées.

LISTE DE TOUS LES SCHEMAS DE SYNTAGMES DISCONTINUS

COMPORTANT LA PARTICULE "NOT"

Phrases avec "?"	Phrases sans "?"	Schémas de syntagmes
0	9	[Do] not G.A1 V
0	1	not G.N2 V
0	4	not G.V1 V
0	1	not G.C1 V
0	1	not G.C2 V
0	3	Need not G.N1 V
0	1	not G.A1 have been Ved
0	1	[Be] not G.A1 Ving
0	1	not G.V1 Ving
0	1	not G.A1 G.A1 G.A1 Ving
0	9	not G.A1 Ved
0	1	not G.N2 Ved
0	3	not G.A1 G.A1 Ved
0	2	G.A1 not Ved
0	2	G.A2 not Ved
0	3	[Have] not G.A1 Ved
0	2	not G.A2 Ved
0	1	not G.V1 Ved
0	4	not G.A1 been Ved
0	1	not G.A2 been Ved
0	2	not been G.A1 Ved
0	1	G.V1 not been Ved
0	1	Let G.N2 not be Ved
0	2	Defectif not but V
0	3	not G.A1 V
0	1	not G.A2 V
0	1	not G.N2 V
0	1	not G.V1 V
0	1	not G.V1 G.A1 V
0	4	not G.A1 be Ved
0	1	not G.A1 G.A1 be Ved
1	0	not G.N1 G.C2 be Ved
1	0	not G.N1 G.C1 be Ved
0	1	not G.A1 have Ved
0	1	not G.A1 have been Ved
0	6	not be G.A1 Ved
0	1	G.A1 not V
1	0	G.N1 G.C2 not V
3	78	

A) De ce tableau, nous déduisons que :

a) 81 syntagmes discontinus comportent la particule "not". Ils sont répartis en 38 schémas verbaux différents qui correspondent à 18 chaînes insérées différentes.

b) Sur ces 81 syntagmes verbaux discontinus :

- 7 ont la particule "not" séparée de la chaîne insérée; "not" se trouve immédiatement après l'auxiliaire temporel, et la chaîne insérée est une Chaîne 2 de la forme "G.Al".
- 74 ont la particule "not" liée à la chaîne insérée :
  - 10 ont "not" en fin de chaîne insérée.
    - 6 dans des syntagmes dont la zone A.A-T = 0; "not" précède alors l'auxilié.
    - 4 dans des syntagmes dont la zone A.A-T = 1; "not" précède l'auxiliaire a-temporel.
  - 64 ont "not" en début de chaîne insérée.
    - 2 dans des syntagmes dont la zone A.A-T = 2.
    - 13 dans des syntagmes dont la zone A.A-T = 1.
    - 49 dans des syntagmes dont la zone A.A-T = 0

B) Le nombre de mots graphiques d'une chaîne insérée comportant la particule "not" peut atteindre 10, y compris "not".

Exemple :

- Minerals do not , so far as is at present known ,, exist in any appreciable quantities for commercial exploitation.

LISTE DE TOUS LES SCHEMAS DE SYNTAGMES DISCONTINUS

TROUVES EN PHRASES INTERROGATIVES

phrases avec "?"	phrases sans "?"	Schémas de syntagmes
19	4	[Do] G.N1 V
38	7	G.N2 V
2	0	G.N3 V
2	0	G.N2 G.C2 V
1	0	G.N1 G.A1 V
3	0	G.N1 G.C1 V
6	2	G.N1 G.C2 V
5	0	G.N1 G.N1 V
1	0	G.N1 G.C2 G.C2 V
1	0	G.N1 G.C2 G.C1 V
1	0	G.N1 G.N1 G.A1 V
1	0	G.N1 G.C1 G.C1 V
1	0	G.N1 G.N1 G.C1 G.C1 V
1	0	G.N1 G.N1 G.N1 G.C1 V
1	0	G.N1 G.N1 G.N1 G.C2 V
1	0	G.N2 G.C2 G.C1 G.C2 V
2	0	G.N1 have to V
3	0	G.N2 have to V
1	0	G.N1 ⊕ G.N1 V
1	0	G.N1 G.C1 ⊕ G.C1 V
1	0	G.N1 G.C1 ⊕ G.N1 V
1	0	G.N1 G.N1 ⊕ G.N1 V
1	0	G.N1 G.V1 ⊕ G.N1 V
1	0	G.N1 G.C2 ⊕ G.N1 G.V1 V
1	0	G.N1 G.N1 G.N1 ⊕ G.N1 V
1	0	⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1 V
1	0	Defectif G.N2 be Ved
1	0	not G.N1 G.C2 be Ved
1	0	not G.N1 G.C1 be Ved
2	0	G.N1 G.C2 have Ved
8	4	G.N1 V
19	5	G.N2 V
1	1	G.I V
1	0	G.N2 G.C2 V
1	0	G.N2 G.C2 not V

2	5	[Have] G.N2 Ved
1	2	G.N1 Ved
3	0	G.N2 G.A1 Ved
1	1	G.N1 been Ved
1	0	G.N2 to V
2	0	[Be] G.N1 Ving
1	0	G.I Ving
1	0	G.N1 G.C1 Ving
10	5	G.N1 Ved
4	9	G.N2 Ved
1	1	G.N3 Ved
2	1	G.N1 G.A1 Ved
4	1	G.N1 G.N1 Ved
1	0	G.N1 G.C1 Ved
2	2	G.N1 G.C2 Ved
4	0	G.N2 G.C2 Ved
1	0	G.N1 G.C2 G.C1 Ved
1	0	G.N1 G.C1 G.C2 Ved
1	0	G.N1 G.A1 G.A1 Ved
1	0	G.N2 G.C2 G.C1 Ved
1	0	G.N1 ⊕ G.N1 Ved
2	0	G.N2 to V
2	0	G.N2 to be Ved
1	0	G.N1 G.C2 to be Ved
182	50	

De ce tableau, nous déduisons :

A) 59 schémas différents peuvent apparaître en phrases interrogatives :

Syntagmes débutant par	[DO]	[BE]	[HAVE]	DEFECTIF	
Nombre de syntagmes discontinus	96	42	8	36	182
Nombre de schémas de synt. discontinus	26	19	5	9	59

B) Ces 59 schémas se divisent en

- 44 schémas uniquement interrogatifs apparaissant dans 64 phrases

  et

- 15 schémas bivalents apparaissant dans 168 phrases
  - 118 interrogatives
  - 50 non-interrogatives

C) Place de la chaîne insérée dans les 182 syntagmes homogènes discontinus interrogatifs :

Auxiliaire temporel / Chaîne insérée	DO	BE	HAVE	DEFECTIF	
Chaîne 1	5	3	1	5	14
Chaîne 2	0	0	0	0	0
Chaîne unique	91	39	7	31	168
	96	42	8	36	182

D) Composition des chaînes insérées dans les syntagmes verbaux homogènes interrogatifs, établie en fonction de l'auxiliaire temporel du syntagme :

Chaîne insérée	Auxiliaire temporel				
	DO	BE	HAVE	DEFECTIF	
G.I	0	1	0	1	2
G.N3	2	1	0	0	3
⊕ G.N1 ⊕ G.N1 G.N1 ⊕ G.N1	1	0	0	0	1
not G.N1 + 1 chaînon	0	0	0	2	2
G.N1 seul	21	12	2	8	43
+ 1 chaînon	16	12	0	3	31
+ 1 chaînon + "not"	0	0	0	0	1
+ 2 chaînons	8	3	0	0	11
+ 3 chaînons	5	0	0	0	5
G.N2 seul	41	8	3	19	71
+ 1 chaînon	2	4	3	1	10
+ 2 chaînons	0	1	0	0	1
+ 3 chaînons	1	0	0	0	1
	97	42	8	35	182

A N N E X E 3

Organigramme partiel de reconnaissance automatique des différentes composantes du syntagme verbal homogène.

Explications complémentaires . . . . . page 158

Dépliant 161

ORGANIGRAMME DE RECONNAISSANCE DES SYNTAGMES VERBAUX HOMOGENES

L'organigramme que nous présentons ici est un organigramme partiel, montrant le traitement à effectuer lors de la rencontre d'une forme [do] ou [be] .

(cf. supra page 108, § 1)

1. CONVENTIONS D'ECRITURE :

RESU : il s'agit d'une "chaîne résultat" devant contenir, après reconnaissance des diverses composantes du syntagme, le syntagme verbal homogène regroupé.

MOT : il s'agit d'une "chaîne de travail" devant contenir successivement chaque mot du texte.

:= : symbole signifiant "affecter à".

donc ; par exemple :

RESU := [do] signifie :

ranger dans la chaîne RESU la forme temporelle de l'auxiliaire [do] que l'on vient de tester.

RESU := RESU suivi de "BEEN" signifie:

ajouter "been" à la suite du contenu de RESU.

MOT := MOT + 1 signifie :

effacer le contenu de MOT, et y faire entrer le mot suivant du texte.

2. "LISTES" MENTIONNEES DANS L'ORGANIGRAMME

Elles sont de deux types :

2.1. Listes qui sont des "chaînes de travail" à consulter par l'ordinateur :

Liste AA : elle contient les mots ou locutions négatives (précédés d'un point) pouvant permettre l'existence de schémas bivalents :

- Nor
- Neither
- Not only
- No longer
- In neither case

⋮

2.2. Listes que le programmeur devra prendre en compte au moment de la rédaction du programme, car ce sont des listes de chaînons insérés à tester.

Liste BA : Liste des différents schémas de chaînes insérées à faire tester après un chaînon G.N1 identifié :

G.C1
G.C1 G.C1
G.C1 ⊕ G.N1
G.C1 ⊕ G.C1
G.C2
G.C2 G.C2
G.C2 G.C1
G.C2 ⊕ G.N1 G.V1
G.V1 ⊕ G.N1
G.N1
G.N1 G.A1
G.N1 G.C1 G.C1
G.N1 G.N1 G.C1
G.N1 G.C1 G.C2
G.N1 ⊕ G.N1
G.N1 G.N1 ⊕ G.N1
⊕ G.N1

Liste BB :

G.A1
G.C1
G.C1 G.C2

Liste BC :

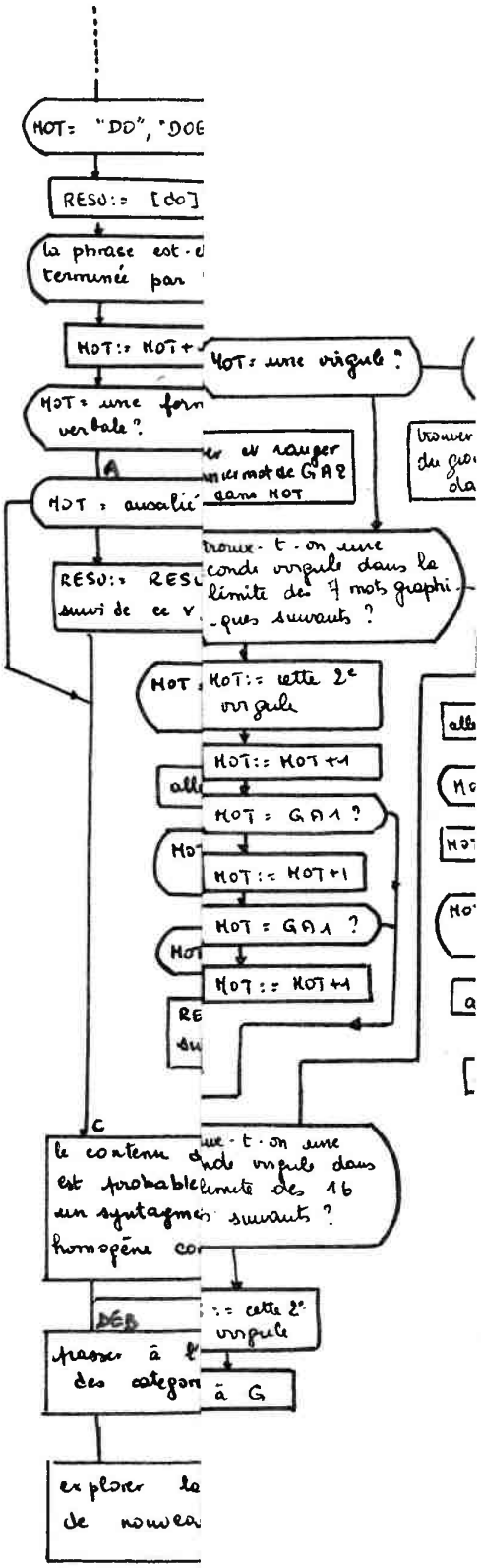
G.C1
G.N1
G.A1
G.V1
G.C1 G.C2
G.A1 G.A1
G.N1 G.V2
⊕ G.N1

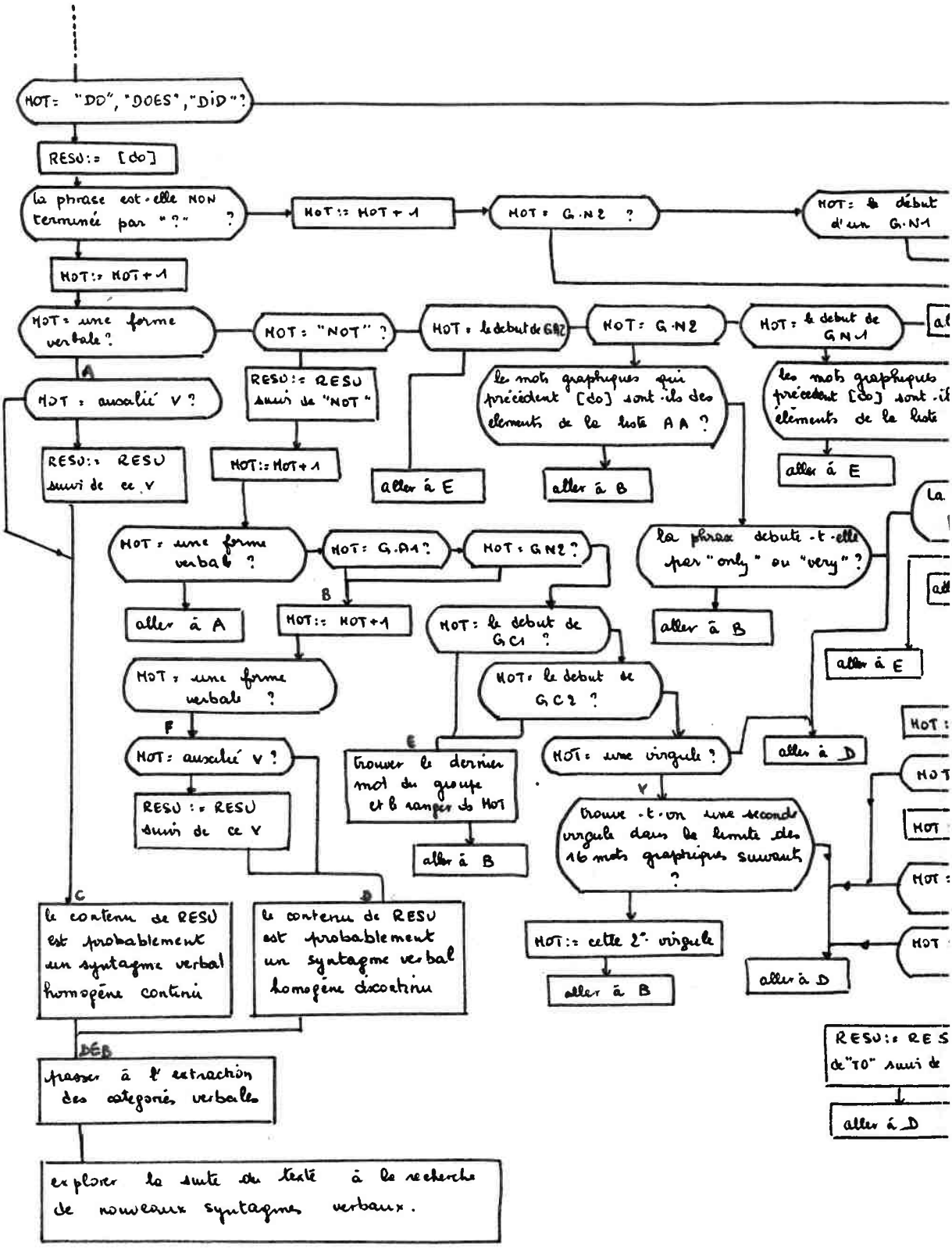
3. DEPLIANT

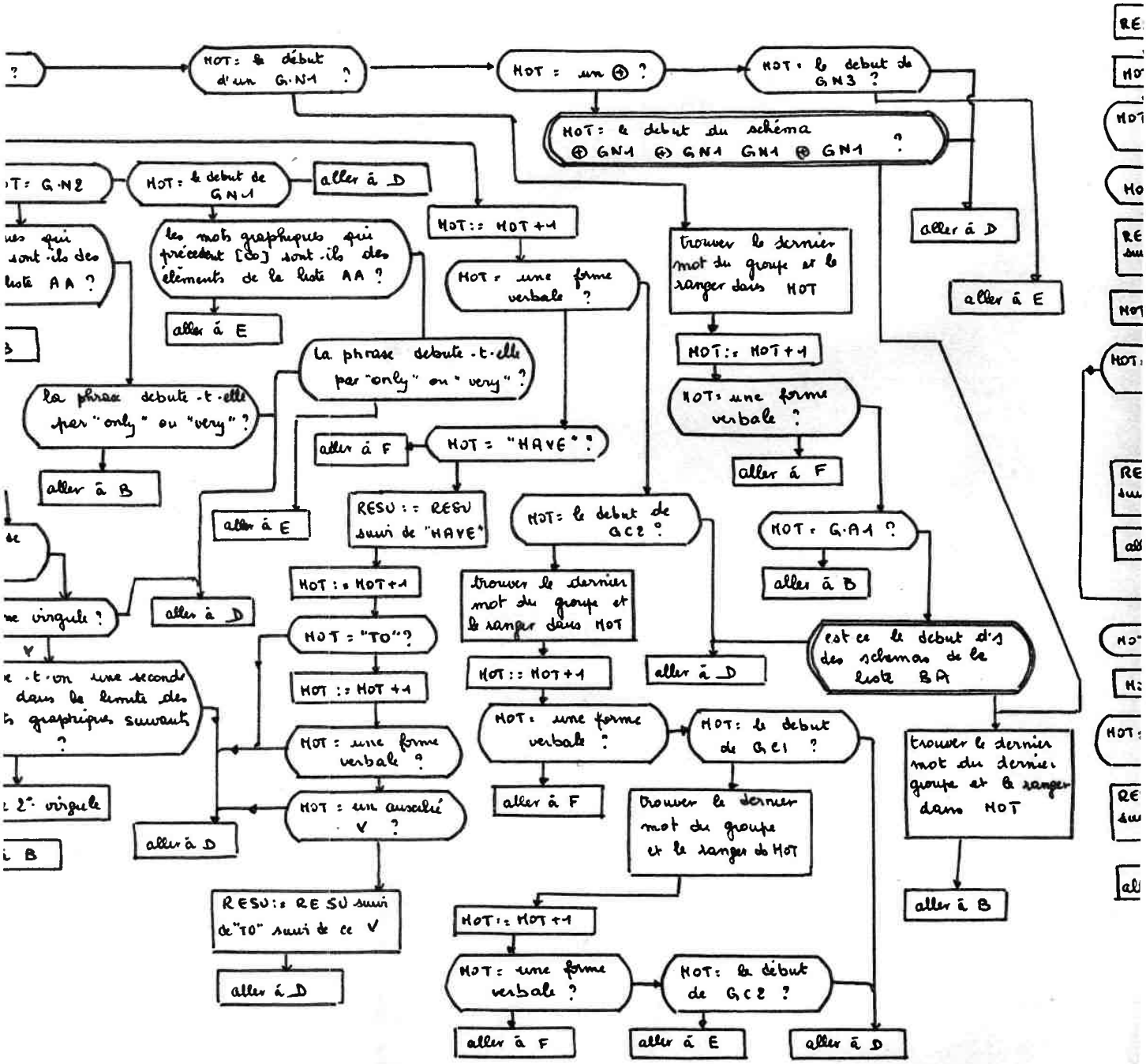
pages 161 à 165.

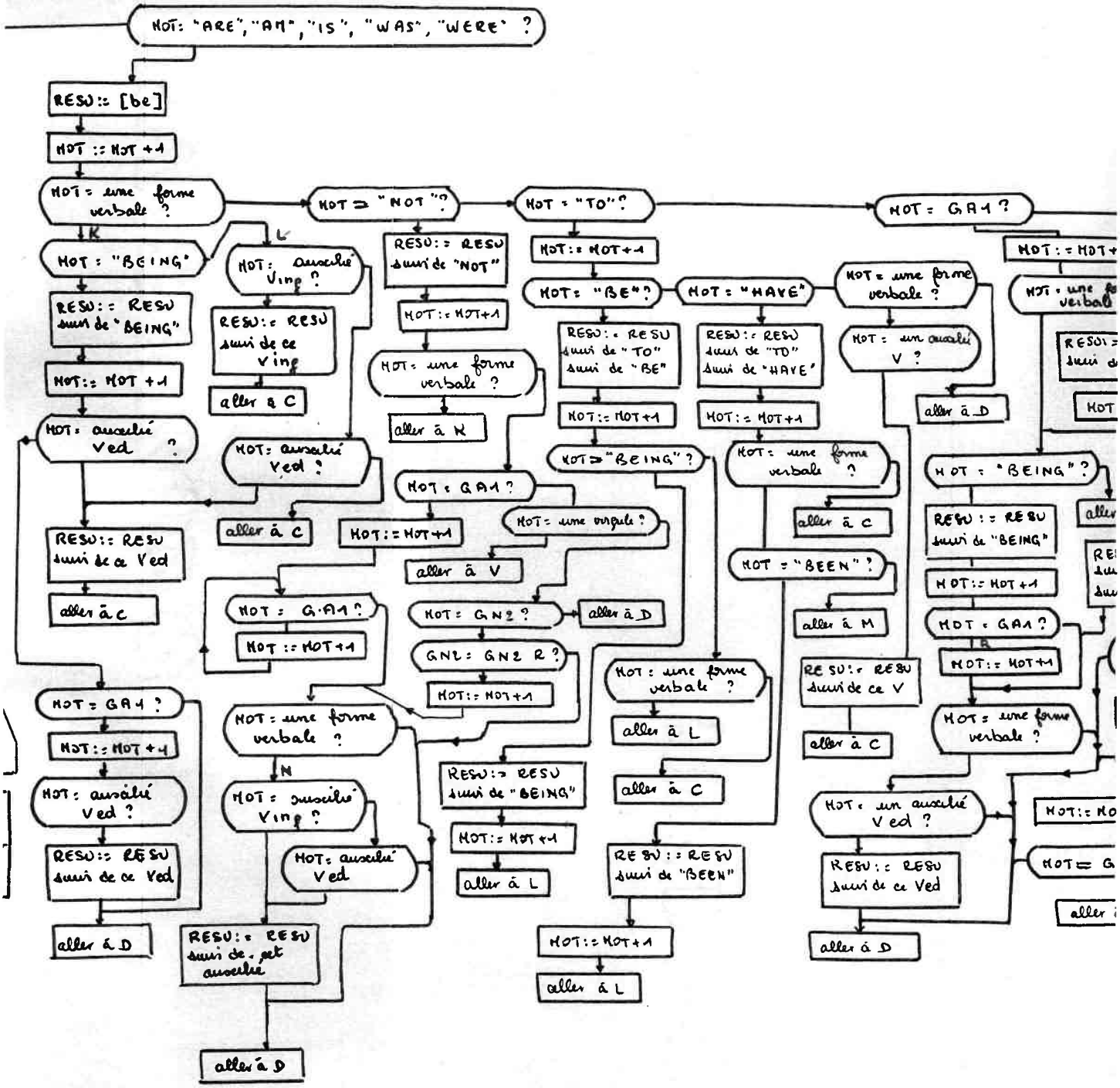
RECONNAISSANCE
DES
SYNTAGMES VERBAUX
HOMOGENES

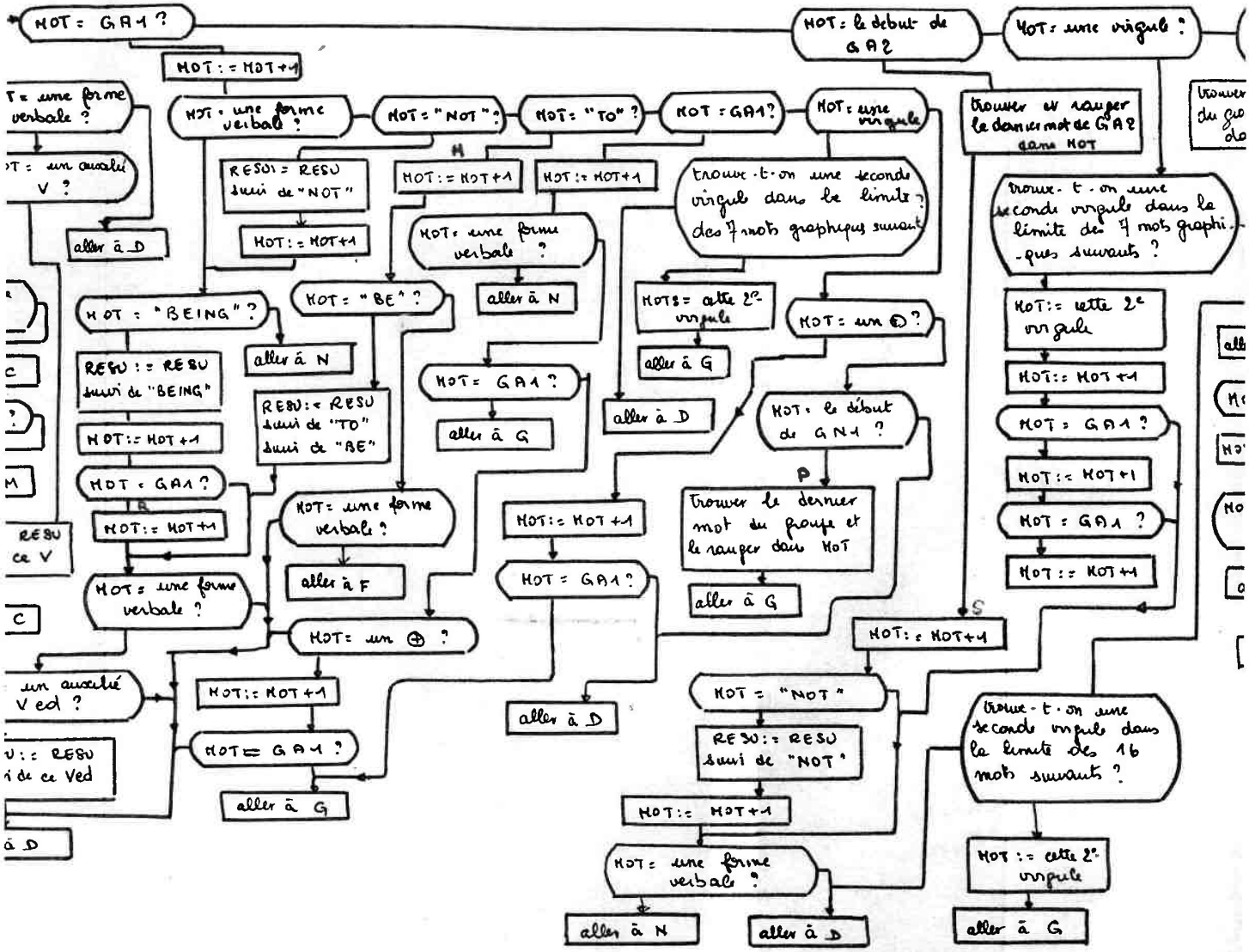
(ORGANIGRAMME PARTIEL)

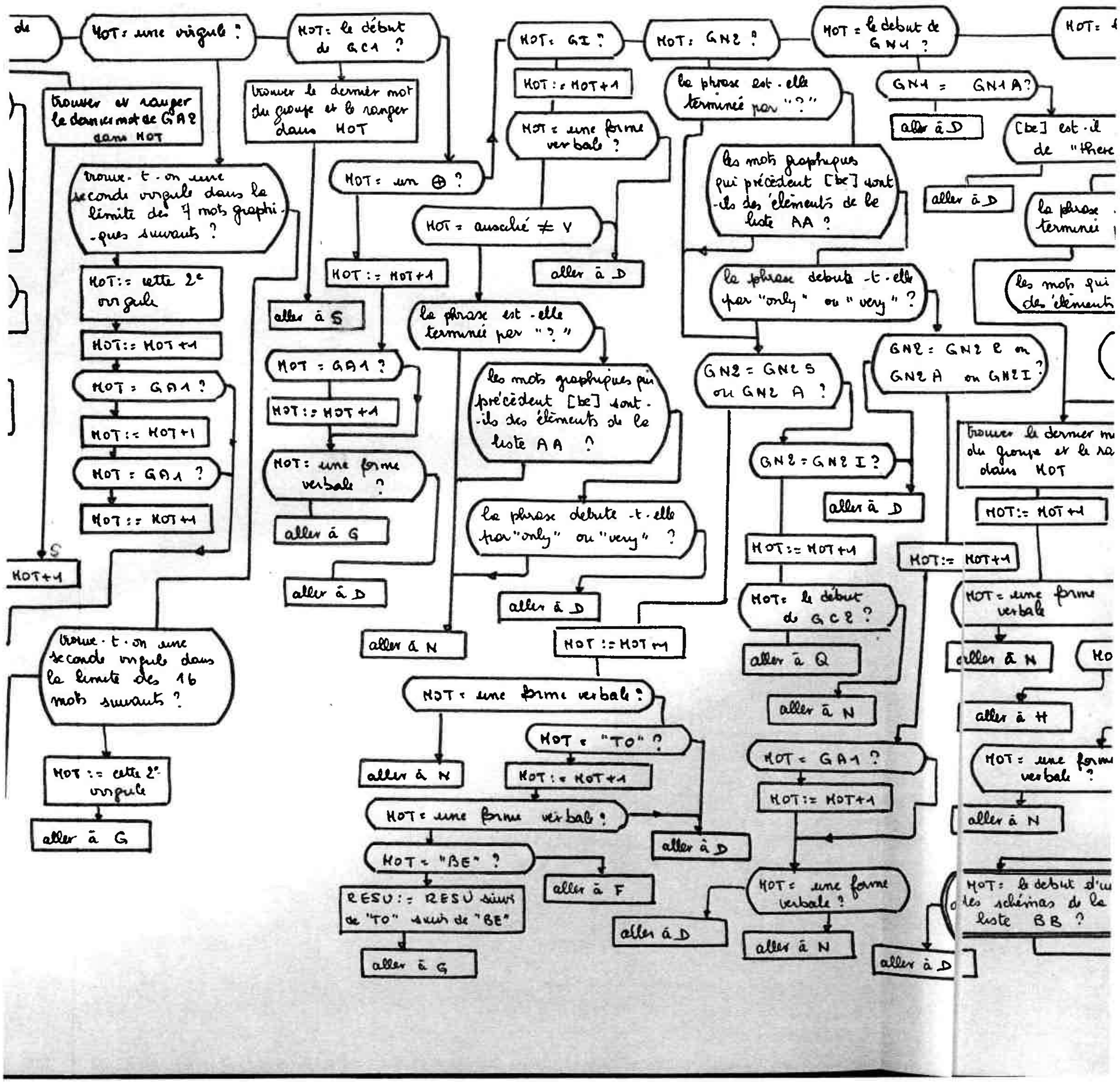


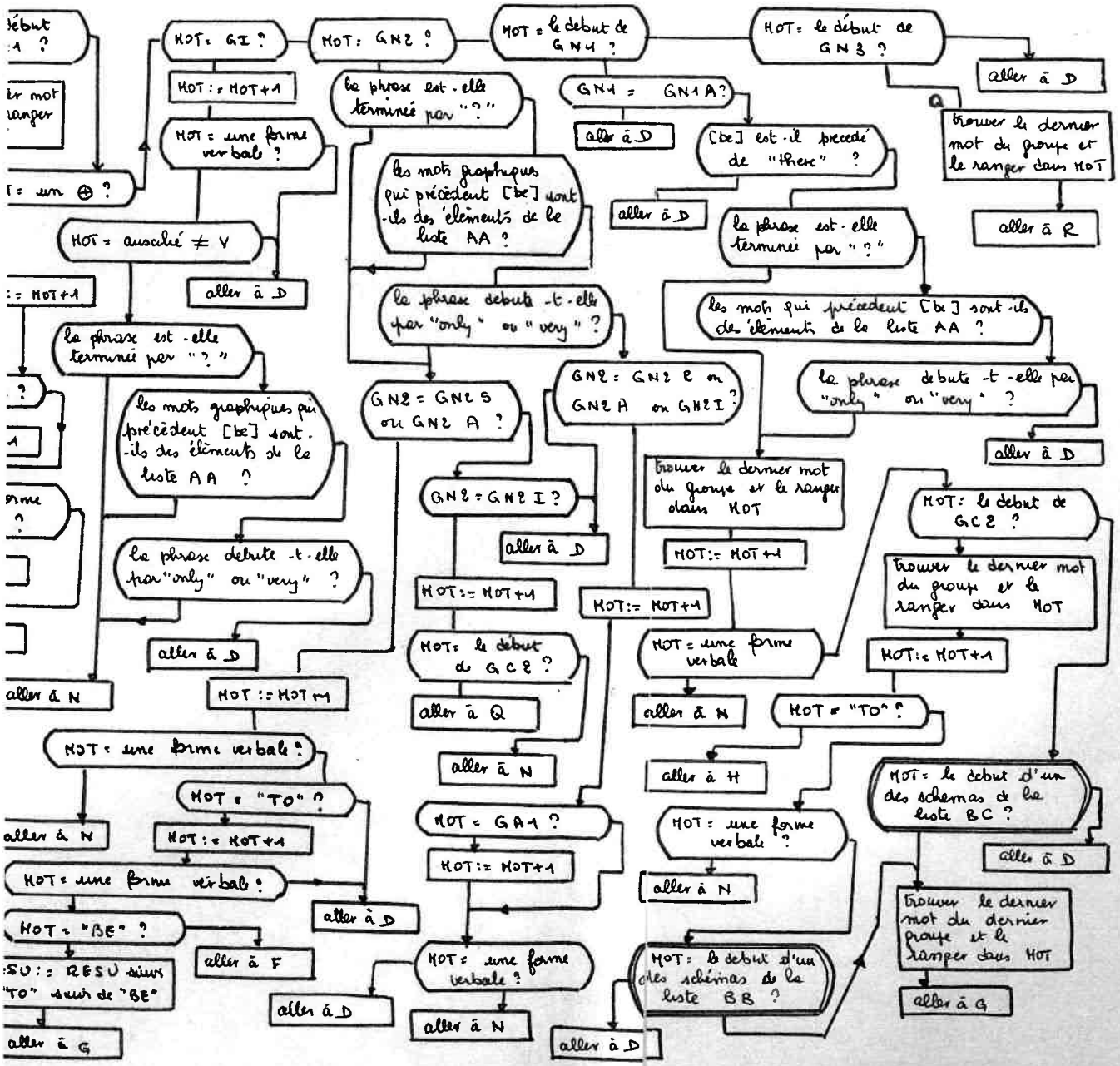

MOT= "DO", "DOES", "DID" ?
RESU:= [do]
la phrase est-elle NON terminée par "?" ?
MOT:= MOT+1
MOT = G.N2 ?
MOT: le début d'un G.N1
MOT:= MOT+1
MOT= une forme verbale ?
MOT = "NOT" ?
MOT = le début de G.N2
MOT= G.N2
MOT= le début de G.N1
A
MOT = auxilié V ?
RESU:= RESU suivi de ce V
RESU:= RESU suivi de "NOT"
MOT:= MOT+1
aller à E
les mots graphiques qui précèdent [do] sont-ils des elements de la liste AA ?
aller à B
les mots graphiques précèdent [do] sont-il elements de la liste
aller à E
MOT = une forme verbale ?
MOT= G.A1 ?
MOT = G.N2 ?
la phrase debuti-t-elle par "only" ou "very" ?
aller à A
B
MOT:= MOT+1
MOT = le début de GC1 ?
aller à B
aller à E
MOT = une forme verbale ?
MOT = le début de GC2 ?
F
MOT= auxilié V ?
E
trouver le dernier mot du groupe et le ranger ds MOT
MOT= une virgule ?
aller à D
RESU := RESU suivi de ce V
aller à B
V
trouve-t-on une seconde virgule dans la limite des 16 mots graphiques suivants ?
C
le contenu de RESU est probablement un syntagme verbal homogène continu
D
le contenu de RESU est probablement un syntagme verbal homogène discontinu
MOT:= cette 2e virgule
aller à B
aller à D
RESU:= RES de "TO" suivi de
aller à D
DÉB
passer à l'extraction des catégories verbales
explorer la suite du texte à la recherche de nouveaux syntagmes verbaux.


MOT : le début d'un G.N1 ?
MOT : un ⊕ ?
MOT : le début de GN3 ?
MOT : le début du schéma ⊕ GN1 ⊕ GN1 GN1 ⊕ GN1 ?
aller à D
aller à E
MOT : le début de GN1
aller à D
les mots graphiques qui précèdent [∞] sont-ils des éléments de la liste AA ?
aller à E
MOT := MOT + 1
MOT : une forme verbale ?
trouver le dernier mot du groupe et le ranger dans MOT
MOT := MOT + 1
MOT : une forme verbale ?
aller à F
la phrase débute-t-elle par "only" ou "very" ?
la phrase débute-t-elle par "only" ou "very" ?
aller à B
aller à F
MOT = "HAVE" ?
aller à E
RESU := RESU suivi de "HAVE"
MOT : le début de GC2 ?
MOT : G.A1 ?
aller à B
MOT := MOT + 1
trouver le dernier mot du groupe et le ranger dans MOT
aller à D
MOT = "TO" ?
MOT := MOT + 1
MOT := MOT + 1
MOT : une forme verbale ?
MOT : le début de GC1 ?
est ce le début d's des schémas de la liste BA
trouver le dernier mot du dernier groupe et le ranger dans MOT
aller à B
MOT : une forme verbale ?
MOT : un auxiliaire V ?
aller à D
aller à F
trouver le dernier mot du groupe et le ranger ds MOT
RESU := RESU suivi de "TO" suivi de ce V
aller à D
MOT := MOT + 1
MOT : une forme verbale ?
MOT : le début de GC2 ?
aller à F
aller à E
aller à D


MOT: "ARE", "AM", "IS", "WAS", "WERE" ?
RESU:= [be]
MOT := MOT +1
MOT = une forme verbale ?
MOT = "BEING"
RESU:= RESU suivi de "BEING"
MOT:= MOT +1
MOT: auxilié Ved ?
RESU:= RESU suivi de ce Ved
aller à C
MOT: auxilié Ving ?
RESU:= RESU suivi de ce Ving
aller à C
MOT = "NOT" ?
RESU:= RESU suivi de "NOT"
MOT:= MOT+1
MOT: auxilié Ved ?
aller à C
MOT = une forme verbale ?
aller à K
MOT = GA1 ?
MOT:= MOT+1
MOT = GA1 ?
MOT:= MOT+1
MOT: auxilié Ved ?
RESU:= RESU suivi de ce Ved
aller à C
MOT = GA1 ?
MOT:= MOT+1
MOT: auxilié Ving ?
MOT: auxilié Ved
RESU:= RESU suivi de "get auxilié
aller à D
MOT = "TO" ?
MOT:= MOT+1
MOT = "BE" ?
RESU:= RESU suivi de "TO" suivi de "BE"
MOT:= MOT+1
MOT = "BEING" ?
MOT = une forme verbale ?
aller à K
MOT = une virgule ?
aller à V
MOT = GN2 ?
aller à D
GN2 = GN2 R ?
MOT:= MOT+1
MOT = une forme verbale ?
N
RESU:= RESU suivi de "BEING"
MOT:= MOT+1
aller à L
MOT = "HAVE"
RESU:= RESU suivi de "TO" suivi de "HAVE"
MOT:= MOT+1
MOT = une forme verbale ?
aller à C
MOT = "BEEN" ?
aller à M
MOT = une forme verbale ?
aller à L
aller à C
RESU:= RESU suivi de "BEEN"
MOT:= MOT+1
aller à L
MOT = une forme verbale ?
MOT = un auxilié V ?
aller à D
RESU:= RESU suivi de ce V
aller à C
MOT = un auxilié Ved ?
RESU:= RESU suivi de ce Ved
aller à D
MOT = GA1 ?
MOT:= MOT+1
MOT = "BEING" ?
RESU:= RESU suivi de "BEING"
MOT:= MOT+1
MOT = GA1 ?
MOT:= MOT+1
MOT = une forme verbale ?


MOT = GA1 ?
MOT := MOT+1
MOT = une forme verbale ?
MOT = "NOT" ?
MOT = "TO" ?
MOT = GA1 ?
MOT = une virgule
MOT = le début de GA2
MOT = une virgule ?
trouver et ranger le dernier mot de GA2 dans MOT
RESU := RESU suivi de "NOT"
MOT := MOT+1
MOT := MOT+1
MOT := MOT+1
trouve-t-on une seconde virgule dans la limite des 7 mots graphiques suivants ?
trouve-t-on une seconde virgule dans la limite des 7 mots graphiques suivants ?
aller à D
MOT = une forme verbale ?
aller à N
MOT = "BEING" ?
MOT = "BE" ?
MOT := cette 2e virgule
MOT := un ⊕ ?
MOT := cette 2e virgule
aller à N
aller à G
RESU := RESU suivi de "BEING"
RESU := RESU suivi de "TO" suivi de "BE"
MOT = GA1 ?
aller à D
MOT := MOT+1
MOT = GA1 ?
MOT := MOT+1
MOT = GA1 ?
MOT := MOT+1
MOT = GA1 ?
MOT := MOT+1
MOT = GA1 ?
MOT := MOT+1
MOT = une forme verbale ?
MOT = une forme verbale ?
aller à F
MOT := MOT+1
MOT = GA1 ?
MOT = un ⊕ ?
MOT := MOT+1
MOT = "NOT"
RESU := RESU suivi de "NOT"
MOT := MOT+1
MOT = une forme verbale ?
aller à N
aller à D
aller à G
trouve-t-on une seconde virgule dans la limite des 16 mots suivants ?
MOT := cette 2e virgule
aller à G
aller à D
MOT = GA1 ?
aller à G
aller à D


MOT: une virgule ?
MOT: le début de GC1 ?
MOT: GI ?
MOT: GN2 ?
MOT = le debut de GN1 ?
trouver et ranger le dernier mot de GA2 dans MOT
trouver le dernier mot du groupe et le ranger dans MOT
MOT := MOT+1
MOT = une forme verbale ?
la phrase est-elle terminée par "?"
GN1 = GN1A?
aller à D
[be] est-il de "there
trouve-t-on une seconde virgule dans la limite des 7 mots graphiques suivants ?
MOT = un ⊕ ?
les mots graphiques qui précèdent [be] sont-ils des éléments de la liste AA ?
MOT = associé ≠ V
la phrase
aller à D
MOT := cette 2e virgule
aller à S
MOT := MOT+1
la phrase est-elle terminée par "?"
la phrase debute-t-elle par "only" ou "very" ?
les mots qui des éléments
MOT := MOT+1
MOT = GA1 ?
GN2 = GN2 S ou GN2 A ?
GN2 = GN2 R ou GN2 A ou GN2 I ?
MOT := MOT+1
MOT := MOT+1
les mots graphiques qui précèdent [be] sont-ils des éléments de la liste AA ?
MOT = GA1 ?
GN2 = GN2 I ?
trouver le dernier m du groupe et le ra dans MOT
MOT := MOT+1
MOT: une forme verbale ?
aller à D
MOT := MOT+1
la phrase debute-t-elle par "only" ou "very" ?
MOT := MOT+1
MOT := MOT+1
aller à G
MOT = une forme verbale
aller à D
aller à D
MOT: le début de GC2 ?
trouve-t-on une seconde virgule dans la limite des 16 mots suivants ?
aller à N
MOT := MOT+1
aller à N
aller à Q
MOT = une forme verbale ?
aller à N
aller à H
MOT = "TO" ?
MOT = GA1 ?
MOT := cette 2e virgule
aller à N
MOT := MOT+1
MOT := MOT+1
MOT = une forme verbale ?
aller à G
MOT = une forme verbale ?
aller à D
aller à N
MOT = "BE" ?
aller à F
MOT = une forme verbale ?
MOT: le debut d'un des schémas de la liste BB ?
RESU := RESU suivi de "TO" suivi de "BE"
aller à D
aller à N
aller à D
aller à G


MOT = le début de GN1 ?
MOT = GI ?
MOT = GN2 ?
MOT = le début de GN3 ?
MOT := MOT+1
MOT = une forme verbale ?
la phrase est-elle terminée par "?" ?
GN1 = GN1A ?
aller à D
[be] est-il précédé de "there" ?
aller à D
trouver le dernier mot du groupe et le ranger dans MOT
aller à R
les mots graphiques qui précèdent [be] sont-ils des éléments de la liste AA ?
la phrase est-elle terminée par "?" ?
MOT = auxilié ≠ V
aller à D
la phrase debute-t-elle par "only" ou "very" ?
les mots qui précèdent [be] sont-ils des éléments de la liste AA ?
la phrase est-elle terminée par "?" ?
GN2 = GN2S ou GN2A ?
GN2 = GN2P ou GN2A ou GN2I ?
la phrase debute-t-elle par "only" ou "very" ?
aller à D
les mots graphiques qui précèdent [be] sont-ils des éléments de la liste AA ?
GN2 = GN2I ?
aller à D
trouver le dernier mot du groupe et le ranger dans MOT
MOT := MOT+1
MOT: le début de GC2 ?
trouver le dernier mot du groupe et le ranger dans MOT
MOT := MOT+1
la phrase debute-t-elle par "only" ou "very" ?
aller à D
MOT := MOT+1
MOT := MOT+1
MOT = le début de GC2 ?
aller à Q
aller à N
MOT = une forme verbale
aller à N
MOT = "TO" ?
aller à H
MOT = une forme verbale ?
MOT := MOT+1
MOT = GA1 ?
MOT := MOT+1
MOT = une forme verbale ?
aller à N
MOT = le début d'un des schémas de la liste BC ?
aller à D
MOT = "BE" ?
aller à F
MOT = une forme verbale ?
aller à N
trouver le dernier mot du dernier groupe et le ranger dans MOT
aller à G
MOT = le début d'un des schémas de la liste BB ?
aller à D
aller à D
aller à N
aller à G

A N N E X E 4

Extraction des catégories verbales

EXTRACTION DES CATEGORIES VERBALES

Nous présentons ici les résultats d'une application expérimentale du programme d'extraction des catégories verbales d'un syntagme verbal homogène donné.

Ce programme CATEVERB a été écrit à la suite de l'organigramme présenté en fin de notre précédente publication. Il a été mis au point sur ordinateur GAMMA 60.

Il est actuellement applicable à tout syntagme verbal homogène continu ou regroupé, non-interrogatif, et dont l'auxiliaire temporel est différent de "let".

Il sera applicable à tout syntagme verbal homogène après deux modifications :

a) addition des branches de traitement des syntagmes discontinus dont le 1° mot graphique est "let". (28 cas sur 1508, soit 2 à 3% d'erreurs si nous ne prévoyons pas ce cas).

b) addition du test sur la présence d'un point d'interrogation. (182 phrases interrogatives sur 1508, soit environ 13 % d'erreurs si ce test n'est pas prévu).

1. RAPPEL DES DIFFERENTS SYMBOLES apparaissant dans les résultats.

V2	:	voix passive
V1	:	voix active
A1	:	aspect non-progressif
A2	:	aspect progressif
F1	:	valeur affirmative
F2	:	valeur négative
F3	:	valeur interrogative
F4	:	valeur interro-négative

D0 : pas d'auxiliaire de modalité
D1 : un auxiliaire de modalité
D2 : deux auxiliaires de modalité
D3 : ambiguïté : pas ou une modalité
D4 : ambiguïté : une ou deux modalités

M1 : mode indicatif
M2 : mode subjonctif
M3 : mode imperatif
M4 : mode ambigu : indicatif ou subjonctif
M5 : mode ambigu : indicatif, subjonctif, ou imperatif
M6 : mode ambigu : subjonctif ou imperatif.

T1 : temps non passé
T2 : temps passé

P1 : phase non perfective
P2 : phase perfective

NP00 : Nbre et Pers. indéterminés
NP01 : 1° pers. singulier
NP02 : 1° pers. sing. et 1° pers. pluriel
NP03 : 3° pers. sing.
NP04 : 2° pers. sing. et 1°, 2°, 3° pers. plur.
NP05 : 1°, 2° pers. sing. et 1°, 2°,,3° pers. plur.
NP06 : 2° pers. sing. et 2°, 3° pers. plur.
NP07 : 1° et 3° pers. sing.
NP08 : 2°, 3° pers. sing. et 2°, 3° pers. plur.
NP09 : 1°, 3° pers. sing. et 1°, 3° pers. plur.

NP91 : si mode ambigu M = M1 : NP05
M = M2 : NP00
M = M3 : NP00

NP92 : si mode ambigu M = M1 : NP02
M = M2 : NP00

NP93 : si modalité ambiguë D3 = D0 : NP02
D3 = D1 : NP00

NP94 : si modalité ambiguë D4 = D1 : NP02
D4 = D2 : NP00

NP95 : si modalité ambiguë D3 = D0 : NP08
D3 = D1 : NP00

NP96 : si modalité ambiguë D4 = D1 : NP08
D4 = D2 : NP00

2. EXEMPLES DE RESULTATS

HAS BEEN DEVELOPED A I D O F I M I N P 0 3 P 2 T I V 2

CAN LEAVE A I D I F I M I N P 0 0 P I T I V I

OUGHT TO BE LEFT A I D I F I M I N P 0 0 P 2 T 2 V 2

ATE A I D O F I M 4 N P 0 0 P I T 2 V I

HAS SUNG A I D O F I M I N P 0 3 P 2 T I V I

DID SAY A I D I F I M 4 N P 0 0 P I T 2 V I

IS COMING A 2 D O F I M I N P 0 3 P I T I V I

ANSWERED A I D O F I M 4 N P 0 0 P I T 2 V I

NEED NOT COME A I D I F 2 M I N P 0 0 P I T I V I

SLEEPS A I D O F I M I N P 0 3 P I T I V I

SHOULD HELP A I D I F I M 4 N P 0 0 P I T 2 V I

IS A I D O F I M I N P 0 3 P I T I V I

MAY NOT HAVE HAD TO JUMP A I D 2 F 2 M I N P 0 0 P 2 T I V I

HAS◇NOT◇HAD◇TO◇BE◇FLYING◇A2◇D1◇F2◇M1◇NP03◇P2◇T1◇V1◇

MUST◇NOT◇HAVE◇BEEN◇BEING◇ESTABLISHED◇A2◇D1◇F2◇M1◇NP00◇P2◇T1◇V2◇

MIGHT◇HAVE◇TO◇BE◇BEING◇HELPED◇A2◇D2◇F1◇M4◇NP00◇P1◇T2◇V2◇

SHOULD◇NOT◇HAVE◇HAD◇TO◇BE◇BEING◇LEFT◇A2◇D2◇F2◇M4◇NP00◇P2◇T2◇V2◇

ERREUR
IS◇BEEN◇LEFT
F1◇A1◇D0◇P2◇V2◇

HAD◇BEEN◇GETTING◇A2◇D0◇F1◇M4◇NP00◇P2◇T2◇V1◇

RECEIVE◇A1◇D0◇F1◇M5◇NP91◇P1◇T1◇V1◇

HAVE◇NOT◇TO◇GO◇A1◇D1◇F2◇M4◇NP91◇P1◇T1◇V1◇

MIGHT◇BE◇JUMPING◇A2◇D1◇F1◇M4◇NP00◇P1◇T2◇V1◇

WERE◇BEING◇ESTABLISHED◇A2◇D0◇F1◇M4◇NP92◇P1◇T2◇V2◇

USED◇TO◇SAY◇A1◇D1◇F1◇M1◇NP00◇P1◇T2◇V1◇

CAN◇A1◇D1◇F1◇M1◇NP00◇P1◇T1◇V1◇

A N N E X E 5

Identification des syntagmes verbaux homogènes dans un texte pré-codé.

IDENTIFICATION DES SYNTAGMES VERBAUX HOMOGENES DANS UN TEXTE PRE-CODE

Nous avons mis au point un programme EXSYN qui permet d'extraire tous les groupes verbaux homogènes, continus ou discontínus, temporels ou a-temporels, d'un texte préalablement indexé.

Le système d'indexation utilisé a été décrit dans une publication du groupe de Traduction automatique de Nancy : "Exploitation d'un Corpus d'Anglais scientifique écrit; Système de notation".

Grâce au programme EXSYN, nous avons pu traiter une partie de ce Corpus d'Anglais scientifique écrit (cf. supra page 7 § 3.2.1. et page 121).

Il s'est révélé très utile dans notre approche du problème; mais sa valeur reste limitée, puisqu'il teste un codé qui a été manuellement affecté à chaque mot graphique d'un texte.

Voici un exemple des résultats obtenus :

(voir pages 172 et 173)

MOST OF THE BONES IN THE UPPER PORTION WERE BLACKISH AS IF HEAVILY ROASTED ., OTHERS WERE OF THE LIGHT BROWN COLOR FOUN D IN INHUMATIONS IN OTHER CATAL HOUSES .

GROUPE VERBAL CONTINU
WERE BLACKISH ◊

SYNTAGME HOMOGENE TEMPOREL
WERE◊

GROUPE VERBAL CONTINU
ROASTED ◊

SEQUENCE A-TEMPORELLE

GROUPE VERBAL CONTINU
WERE ◊

SYNTAGME HOMOGENE TEMPOREL
WERE◊

GROUPE VERBAL CONTINU
FOUND ◊

SEQUENCE A-TEMPORELLE

SOME OF THE BONES ON THE BOTTOM WERE STILL EMBEDDED IN THEIR
CHARRED MUSCLES WITH THE SKIN PRESERVED , WHILE SEVERAL WH
OLF LIMBS WERE FOUND FOLDED UP AND WRAPPED .

GROUPE VERBAL CONTINU
WERE FOUND FOLDED UP AND WRAPPED ◊

SYNTAGME HOMOGENE TEMPOREL
WERE FOUND◊

GROUPE VERBAL DISCONTINU
WERE EMBEDDED◊

GROUPE INSERE :
I: STILL ◊

SYNTAGME HOMOGENE TEMPOREL
WERE EMBEDDED◊

A N N E X E 6

Liste des ouvrages d'où sont tirés nos exemples et notre Corpus de syntagmes verbaux homogènes.

Manuels de géographie à usage scolaire :

Honeybone R.C. & M.C. Goss
BRITAIN AND OVERSEAS, Bk. 1
Heinemann, 1959

Warden H.R.
CONCENTRIC WORLD GEOGRAPHY, Bk. 2
Cassel, 1960

Ouvrages scientifiques spécialisés, d'où sont extraits les textes composant le CORPUS d'Anglais scientifique:

a) ouvrages :

Bradford J.
ANCIENT LANDSCAPES
London, G.Bell, 1957

Cookson M.B.
PHOTOGRAPHY FOR ARCHAEOLOGISTS
London, Parrisk, 1954

Monkhouse F.J. & H.R. Wilkinson
MAPS AND DIAGRAMS
Methuen, 1961

Smith H.T.U.
AERIAL PHOTOGRAPHS AND THEIR APPLICATIONS
New York, Appleton, 1943

Wheeler M. (Sir)
ARCHAEOLOGY FROM THE EARTH
Oxford, atthe Clarendon Press, 1955

b) revues :

THE ANNALS OF BOTANY
Oxford, at the Clarendon Press, 1963, 27

JOURNAL OF CELL BIOLOGY
Rockfeller Institute, 1963, 18

ANNALS OF MATHEMATICS
Princeton University Press, 1963, 78

THE CANADIAN GEOGRAPHICAL JOURNAL
Montreal, 1963

GEOLOGICAL MAGAZINE
Hertford, G.B., 1961

GEOGRAPHICAL JOURNAL
London, Royal Geographical Society, 1961

AMERICAN JOURNAL OF PHYSICS
New York, vol. 32, n° 3

ANNUAIRE INTERNATIONAL DE CARTOGRAPHIE
Armand Colin, 1962, II

Ouvrages scientifiques spécialisés, utilisés pour notre dernière contre-épreuve :

a) ouvrage :

Herdan G.
TYPE-TOKEN MATHEMATICS
Mouton, 1960

b) revue :

SCIENTIFIC AMERICAN
New York, sept. 1963

TABLE DES MATIERES

AVANT-PROPOS 1

INTRODUCTION 2

1. LES CHAINES INSEREES 11

2. INVENTAIRE DES CHAINONS INSERES 18

3. COMBINATOIRE DES CHAINONS INSERES 51

4. RAPPORTS CHAINE 1/ CHAINE 2 64

5. RAPPORTS CHAINES INSEREES / SYNTAGME VERBAL . 76

6. RAPPORTS CHAINES / SYNTAGMES / PHRASE 93

7. CONSEQUENCES MECANISABLES 108

CONCLUSION 114

ANNEXE 1 120

ANNEXE 2 141

ANNEXE 3 157

ANNEXE 4 166

ANNEXE 5 171

ANNEXE 6 174