Morris Salkoff

ON USING SEMANTIC DATA IN AUTOMATIC SYNTACTIC ANALYSIS

0. Introduction.

A program for the syntactic analysis of French text has been developed which is used in conjunction with a string grammar of French\(^1\) and a dictionary of the grammatical properties of the words in the sentences analyzed. The program has been written in FORTRAN so that it can be used on any computer having a FORTRAN compiler. It has been tested on the IBM 360-91 at the Centre d'Etudes Nucléaires in Saclay, and on the IBM 370-165 at the computing center of the C.N.R.S. in Orsay and yields satisfactory analyses of sentences in a reasonable computing time.

The analyses obtained for the first eight sentences of a text by Jacob and Monod in molecular biology are presented in the figures of the appendix. My object in this paper is to discuss what semantic data have been incorporated in the string grammar in order to produce these analyses and to prevent other false analyses from being obtained. A discussion of certain features of these analyses, as well as of those analyses not obtained, will bring out the nature of the semantic data that I have in mind. The text analyzed is reproduced in Fig. 1, and the analyses in the figures that follow.

1. The analyses.

Sentence JM-1 Remarks on the analyses.

a) The sequence *quelques années sur la structure...* is analyzed as a noun phrase in which *sur la structure ...* is a right adjunct (modifier)

\(\text{\textcopyright Ce travail a pu être réalisé en partie grâce au concours apporté par un contrat du Comité de Recherche en Informatique.}\)

\(\text{\textcopyright Z. S. Harris, \textit{String Analysis of Sentence Structure}, Mouton, La Haye, 1961.}\)

\(\text{\textcopyright M. Salkoff, \textit{Une Grammaire en Chaine du Français}, Dunod, Paris, 1973.}\)
of années. This analysis is a possible one (although not correct in this sentence), for a sentence such as *Quelques années sur ce problème m'ont convaincu que ...* is correct.

b) The sequence *Les connaissances ... acides nucléiques et protéines* is the subject in the third and fourth analyses of JM-1 (figs. 4 and 5).

c) The pair of parentheses (), e.g. in line 7 of Fig. 2, represent a zeroed indefinite subject of the verb *comprendre*: *(permettent)* à quelqu'un de comprendre cela → *(permettent)* () de comprendre cela. The double parentheses in line 13 of Fig. 3 represent an element that has been reduced to zero in the domain of a conjunction.

The following analyses were rejected:

d) The analysis in which the sequence

(1) *depuis quelques années ..., acides nucléiques et protéines*

would have been analyzed as a prepositional phrase consisting of *depuis* followed by a conjunction of three noun phrases. In effect, a concrete noun, a time noun, or a nominalization is possible after *depuis*:

(2) 
   a) *Depuis le coin de la rue, (je vous dis cela)*
   b) *Depuis quelques années, (je vous dis cela)*

but two different types cannot be conjoined after *depuis*:

(3) *Depuis le coin de la rue et quelques années, (je vous dis cela)*

Since (1) is similar to (3) in this respect, this analysis was rejected.

Sentence JM-2 Remarks

a) *bien* is analyzed as a sentence adjunct only when it occurs to the right of a form of *être*, as in this sentence, or in such sentences as

(4) *c'est bien lui; l'idée était bien à moi.*

b) The sequence à la suite de should be analyzed as a complex preposition, but the appropriate mechanism to accomplish this has not yet been added to the analyzer.

Rejected analyses.

c) The prepositional phrase à N cannot be the right adjunct of a proper noun:
(5) *Pierre à Paris (est mon ami); mais: Pierre de Paris (est mon ami)

This restriction prevents the program from obtaining the incorrect analysis in which à la suite ... is a right adjunct of Crick.\(^5\)

d) the analyser rejects the analysis in which the sequence

(6) qu'avaient proposé Watson et Crick à la suite

is taken as similar to the sequence

(7) qu'avaient proposé Watson et Crick au gouvernement

The verb proposer belongs to a subclass \(V_{15}\) that requires a human-like noun in the position \(N_s\) of its object \(N_i \rightarrow N_s\):

(8) proposer quelque chose [au jury / à cette femme / à la foule / ...]

*proposer quelque chose à la table

The verb apporter also belongs to \(V_{15}\), and so the analyzer does not obtain the analysis in which the sequence

(9) apporté la preuve que ... à la suite

is taken as similar to apporter un chapeau à Marie.

Sentence JM-3 Remarks

a) The sequence

(10) le découvritre de l'ARN messager

is analyzed as a "compound noun". This term was meant for such groups as

(11) lycée d'Etat; homme de l'espace; bateau à vapeur

\(^5\) When the mechanism for treating complex prepositions like à la suite de has been added to the program, it will be specified in the grammar that the prepositional phrase à la suite de \(N\) must be analyzed as a sentence adjunct.
in which the prepositional phrase de N or à N cannot be pronominalized to en or y:

(12)  
   a) Je vois un lycée d'Etat \[\Rightarrow\] *J'en vois un lycée  
   b) Je vois un bateau à vapeur \[\Rightarrow\] *J'y vois un bateau

But this pronominalization is possible for (10):

(13)  
   On décrira la découverte de l'ARN messager \[\Rightarrow\] On en décrira la découverte

and so should not be analyzed as a compound noun. However, some technical difficulties with repeated adjuncts made it convenient to retain this analysis temporarily.

Rejected Analysis.

c) The analyzer rejects the analysis in which the object of confirmé is

(14)  
   a) Les ... hypothèses ... et les protéines-enzymes

i.e., the analysis in which the sentence is understood as

b) (Les découvertes ...) ont confirmé les hypothèses ... et (les découvertes ont confirmé) les protéines-enzymes.

The verb confirmer cannot take a concrete noun as its object:

(15)  
   a) *[Pierre / l'hypothèse] a confirmé les protéines

except for a few special cases such as

b) Pierre a confirmé [sa place (dans l'avion) / le rendez-vous/ ...]

Sentences JM-4 Remarks

In the analysis shown (Fig. 9), the relative clause qui caractérisent une cellule is analyzed as a right adjunct of activités (cf. line 17.), although it should be an adjunct of the entire subject group, i.e. an adjunct of les propriétés, les structures, les activités. There is at present no provision in the grammar for attaching an adjunct to a sequence of conjoined structures.
Rejected analysis.
The analyzer rejects the analysis in which the object of démontré is

\[(16) \quad \text{que } C1 \text{ et } * \text{ } N_a; \text{ } C1 = \text{les propriétés, les ... rapportées à la structure}\]

\[\quad * = \text{à l'activité};\]

\[\quad N_a = \text{des protéines que ...}\]

For one, protéines is not a correct object of démontré:

\[(17) \quad \text{*Pierre a démontré les protéines}\]

so that the restriction mentioned above in JM-3 in connection with confirmé disallows the analysis. However, even if \(N_a\) happened to be a licit noun object of démontré there is still a question whether the conjunction of a nominalized sentence and a noun phrase yields a grammatical object:

\[(18) \quad ? \text{Pierre a démontré que le problème est difficile et l'impossibilité de le résoudre}\]

Such a sentence is difficult or impossible to accept; in the present grammar, it is taken as ungrammatical, and a restriction prevents the conjoining of such dissimilar objects.

Sentence JM-5

The prepositional phrase \textit{par un segment génétique} is taken as a right adjunct of the verb affirmer, instead of being analyzed as an adjunct of the participle définie.

Sentence JM-6A

The two analyses show the alternatives for conjoining the sequence \textit{ou d'une lignée cellulaire}. In the first analysis (Fig. 11) it is conjoined to \textit{d'une cellule}, which yields the meaning intended by the author: \textit{l'ADN d'une cellule ou (l'ADN) d'une lignée cellulaire}. In the second reading, it is attached to \textit{la structure}: \textit{la structure de l'ADN ... ou (la structure) d'une
lignée cellulaire. The general problem of the correct conjoining of strings headed by a conjunction has not yet been solved.

Sentence JM-6B

Because the mechanism for treating idioms has not yet been incorporated in the analyzer, it was not possible to treat en fonction de as a complex preposition. It is therefore analyzed as a prepositional phrase en fonction modified by the right adjunct de signaux ... (cf. note 2).

Sentence JM-7A  Remarks.

a) The sequence la conversion de son système excrétoire is analyzed as a compound noun; remark (a) on sentence JM-3 applies here too. Because of this, the relative clause qui, de semblable ... is incorrectly attached to conversion, instead of being analyzed as the right adjunct of système.

b) The exact status of the sequence de semblable à celui d'un poisson is not clear. It could be analyzed as a verb adjunct (that can appear at the beginning of a clause or sentence) for a subclass of verbs like devenir, se transformer, ...; or it can be treated as part of the object for these verbs. I have chosen the second solution, and it appears as the first part of a split object of deviendra in an inverted center string (line 13, Fig. 14).

c) The adjective excrémentaire, since it is neither definitely masculine nor definitely feminine in form, could modify either conversion or système. The analysis shows it modifying système (line 7), and further analyses in which it would modify conversion are not printed by the analyzer, since such an ambiguity is predictable from the first analysis. The same remark applies to thyroidienne, which can modify either injection or hormone. The printing of these ambiguities (and of many others), which can be predicted from the form of the structures involved, is suppressed by the program.

Rejected analyses.

d) In fig. 14, analogue is taken as the adjective object of deviendra. However, analogue is also a noun:

(19)  Cette situation est en effet l'analogue (de telle autre)
but it cannot appear as the noun object of être without an article. Only a certain subclass of nouns, called N17, can appear without the article:

(20) \[ \text{Pierre est [patron / ambassadeur / professeur/ ...]} \]
\[ \ast \text{Pierre est rocher} \]

Since analogue, as a noun, does not belong to this subclass N17, the analysis is eliminated.

c) The analyzer rejects the analysis in which the object of provoque is

(21) \[ N_1 P N_2; N_1 = \text{la conversion} \ldots \]
\[ P N_2 = \text{à celui d'un mammifère} \]

This is analogous to the use of provoque in sentences like

(22) \[ \text{On a provoqué Pierre [à la violence / à un acte désespéré] \ldots} \]

But provoque belongs to a subclass of verbs, V15, which require a human or human-like noun in the position of N1:

(23) \[ \ast \text{On a provoqué la table à une chute brutale} \]

A restriction then disallows la conversion... in the position N1 of (21).

2. The semantic data.

From the analyses presented above, we see that two distinct types of semantic data are incorporated in the grammar.

(1) A subclass of some major grammatical class cannot appear in a given position. This was the case for the verb subclasses V15 and V16, which require a human noun in one of the positions of their NP object; for the noun object of être, which can drop the article only if the noun is in the subclass N17; and for the object of verbs like confirmer, démontrer, ... which cannot be a concrete noun.

(2) Some sequence of conjoined strings is not possible for given values of one of the strings or of some subclass appearing in the strings. Thus, the two conjoined nouns in the sequence (sentence JM-1) depuis N1 et N2 may both be nominalizations, time nouns, or concrete nouns, but not one of each kind.
Another example was seen in the sentence JM-4 where the conjunction of dissimilar objects of a verb is ungrammatical. These two examples illustrate the problem of the conjunction of classes and strings which is not solved for the general case.

The interesting point is that these two types of semantic restrictions are of the same nature as the syntactic restrictions and are incorporated in the grammar in the same way as the latter. This means that there is no need for a semantic component, or for semantic considerations that are completely separate from the usual grammatical procedures. The definition of subclasses is required in any case by the grammar, e.g. for the syntactic subclasses (singular, plural, ...); and the specification of the conjoinability of given sequences (strings) is required in order to treat conjunctions. In this way, the semantic component becomes part of the syntax and is incorporated without any special mechanism. The same type of restriction as that which forbids *Pierre sont ici or *L′homme est courageuse is used to prevent the analyzer from presenting a sequence such as *provoquer la table à la violence.

Hence I extend this use to semantic subclasses. Some of these subclasses can be defined syntactically, e.g., N17, but in any case they are sometimes used to exclude sequences that are not necessarily syntactically forbidden. These sequences do not violate any rules of the grammar, but violate what is usually called a selection rule or a semantic constraint.

To the extent that these semantic constraints, or selection rules, can in fact be formulated, the formulation can probably always be stated in one of the ways (1) or (2) given above. If this is true, then the syntactic analyzer based on string grammar which I have presented here can incorporate semantics as well as syntax. Two difficulties in this formulation via subclasses immediately present themselves:

a) The subclass itself is difficult to define, e.g., the subclass human or human-like which is needed to define V15 and V16, or the subclass «concrete» noun.

b) The decision as to whether a given word does in fact belong to some subclass is not always easy to make. In the discussion of JM-2, I said that apporter belongs to the subclass V15, and this is how apporter is presently coded in the dictionary. There are however well-formed sentences in which the N1 P N2 object of apporter has a non-human noun in the position N2:

\[(24) \quad J′apporterai une solution au problème\]
but this seems to be limited to pairs of nouns standing in some relationship to each other, as *solution* and *problème*. This relationship is very hard to define, since other pairs of nouns, seemingly related in a similar fashion, do not yield well-formed sentences with *apporter*:

(25) *J'ai apporté un pied à la chaise*

But these difficulties are not specific to the analyzer nor to the string grammar that I use; rather, they are independent of the parsing strategy — no matter what the analyzer — and will be solved, if indeed they can be solved, by more detailed research into the linguistic problems involved. From a practical point of view, this uniform treatment of the semantic and the syntactic data leads to a more compact grammar and a simpler analyzer than one containing separate semantic and syntactic components. Only semantically and syntactically correct analyses are furnished by such an analyzer, as is desired, and this is of prime importance for later applications of the analyzer to the problems of automatic translation or automatic documentation.
MÉCANISMES BIOCHIMIQUES ET GÉNÉTIQUES
DE LA RÉGULATION DANS LA CELLULE BACTÉRIENNE

par François Jacob et Jacques Monod.

Services de Génétique microbienne et de Biochimie cellulaire, Institut Pasteur, Paris.

1. INTRODUCTION

Les connaissances acquises depuis quelques années sur la structure des macromolecules biologiques essentielles, acides nucléiques et protéines, permettent de comprendre, au moins dans ses grandes lignes, le rapport entre les fonctions de ces macromolecules et leur structure chimique. L'étude de la réplication de l’ADN \textit{in vivo et in vitro} a apporté la preuve que le mécanisme chimique fondamental de l'hérédité est bien celui qu'avaient proposé Watson et Crick (1953) à la suite de leur découverte de la structure de l'ADN. La découverte de l'ARN messager et de son rôle dans la biosynthèse des protéines, l'étude des processus de transcription, les recherches sur le déterminisme génétique des structures primaires des protéines ont entièrement confirmé, en les renouvelant, les anciennes hypothèses sur les relations entre les déterminants génétiques et les protéines-enzymes.

Les progrès de la biochimie réalisés depuis 50 ans ont, en outre, démontré que les propriétés, les structures, les activités qui caractérisent une cellule doivent, en définitive, être rapportées à la structure et à l'activité des protéines que cette cellule est capable de synthétiser. Or, la structure de chacune de ces protéines est intégralement définie, on peut l'affirmer aujourd'hui, par un segment génétique. Mais alors que la structure de l'ADN d’une cellule ou d’une lignée cellulaire est invariante, les propriétés biochimiques réalisées et exprimées par cette cellule pourront être profondément différentes; en outre, ces propriétés sont modifiables en fonction de signaux chimiques spécifiques venus de l’extérieur. L'injection d’hormone thyroïdienne à un têtard provoque, avant même toute manifestation morphologique, la conversion de son système excrétoire qui, de semblable à celui d’un poisson, deviendra chimiquement analogue à celui d’un mammifère: quelques heures après l’injection, les enzymes spécifiques du cycle de l’urée augmenteront en proportion de 50 à 100 fois et l’animal excrètera de l’urée alors que, jusque-là, il n’avait excrété que de l’ammo-

BULL. SOC. CHIM. BIOL., 1964, 46, N° 12.
LES CONNAISSANCES ACQUIS ÉS DEPUIS QUELQUES AN NÉES SUR LA STRUCTURE D'-ES MACROMOLÉCULES BIOLOGIQUES ESSENTIELLES, ACIDES NUCLEIQUES ET PROTEINES, PERMETTENT DE COMPRENDRE, A- -U MOINS DANS SES GRANDES LIGNES, LE RAPPORT ENTRE LES FONCTIONS DE CES MACROMOLÉCULES ET LEUR STRUCTURE CHIMIQUE.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN
2.

2. CHAINE D'ASSERTION = • (PART.) SUJET •(PART.) VERBE •
3. CONNAISSANCES 4. PERMETTENT

OBJET DV •

5.

3. GN = ARTICLE QUANT ADJ
LES

4. DN = VE OMEGA-PASSIF

6.

5. A S DE V OMEGA = CHAINE D'ASSERTION

7.

6. VE OMEGA-PASSIF = VE • OM-PASS DV •
ACQUISÉS 8.

7. CHAINE D'ASSERTION = • (PART.) SUJET •(PART.) VERBE • 0

DE COMPRENDRE , 9. 1

B JET DV •

0. RAPPORT 11.

8. P N = G.P. PREPOSITION D.P. N

DEPUIS 12. ANNEES 13.

9. P N = G.P. PREPOSITION D.P. N
A- -U MOINS DANS 14. LIGNES

, 15.

10. GN = ARTICLE QUANT ADJ
LE

11. P N = G.P. PREPOSITION D.P. N
ENTRE 15. FONCTIONS 16.

ET 17.

12. GN = ARTICLE QUANT ADJ
QUELQUES
Fig. 2.
LES CONNAISSANCES ACQUISES DEPUIS QUELQUES ANNEES SUR LA STRUCTURE DES MACROMOLECULES BIOLOGIQUES ESSENTIELLES, ACIDES NUCLEIQUES ET PROTEINES, PERMETTENT DE COMPRENDRE, A- U MOINS DANS SES GRANDES LIGNES, LE RAPPORT ENTRE LES FONCTIONS DE CES MACROMOLECULES ET LEUR STRUCTURE CHIMIQUE.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN
2.
3. CHAINE D’ASSERTION = *(PART.) SUJET *(PART.) VERBE *(PART.) OBJET DV *
4. GN = ARTICLE QUANT ADJ LES
5. DN = VE OMEGA-PASSIF
6. A S DE V OMEGA = CHAINE D’ASSERTION ET
7. ET 8.
8. VE OMEGA-PASSIF = VE * OM-PASS DV * ACQUISES 9.
9. CHAINE D’ASSERTION = *(PART.) SUJET *(PART.) VERBE *( ) DE COMPRENDRE , 10.
10. OBJET DV *
11. RAPPORT 12.
12. CONJONCTION = CHAINE D’ASSERTION
13.
14. P N = G.P. PREPOSITION D.P. N DEPUIS ANNEES 15.
15. P N = G.P. PREPOSITION D.P. N A- U MOINS DANS LIGNES 16.
16. LE
17. GN = ARTICLE QUANT ADJ
18. P N = G.P. PREPOSITION D.P. N ENTRE FONCTIONS 17.
19. CHAINE D’ASSERTION = *(PART.) SUJET *(PART.) VERBE *( ) OBJET *( ) RUCTURE CHIMIQUE
20.
Fig. 3.
ON USING SEMANTIC DATA IN AUTOMATIC SYNTACTIC ANALYSIS

TEMPS D'ANALYSE = 1359 SEC/100

JM-1

*************************** ANALYSE NO 3 ********************

LES CONNAISSANCES ACQUISES DEPUIS QUELQUES ANNEES SUR LA STRUCTURE DES MACROMOLECULES BIOLOGIQUES ESSENTIELLES, ACIDES NUCLEIQUES ET PROTEINES, PERMETTENT DE COMPRENDRE, A MOINS DANS SES GRANDES LIGNES, LE RAPPORT ENTRE LES FONCTIONS DE CES MACROMOLECULES ET LEUR STRUCTURE CHIMIQUE.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN

2. CHAINE D'ASSERTION = • (PART.) SUJET VIRGULE • (PART.) CONNAISSANCES 4, 5. PERMETTE VERBE • OBJET DV • NT 6.

3. GN = ARTICLE QUANT ADJ LES

4. DN = VE OMEGA-PASSIF 7.

5. CONJONCTION = (PART.) SUJET ET ACIDES NUCLEIQUES ET 8.

6. A S DE V OMEGA = CHAINE D'ASSERTION 9.

7. VE OMEGA-PASSIF = VE • OM-PASS DV • ACQUISES 10.

8. CONJONCTION = (PART.) SUJET PROTEINES ,

9. CHAINE D'ASSERTION = • (PART.) SUJET • (PART.) VERBE • ( ) DE COMPRENDRE , 11.

10. P N = G.P. PREPOSITION D.P. N DEPUIS 14. ANNEES 15.

11. P N = G.P. PREPOSITION D.P. N A MOINS DANS 16. LIGNES ,

12. GN = ARTICLE QUANT ADJ LE

13. P N = G.P. PREPOSITION D.P. N ENTRE 17. FONCTIONS 18.

ET 19.
MORRIS SALKOFF

412

14. GN

15. PN

16. GN

17. GN

18. PN

19. CONJONCTION

20. GN

21. NOM COMP

22. GN

23. GN

24. GN

= ARTICLE QUANT ADJ
QUELOUES

= G.P. PREPOSITION D.P. N
SUR

= ARTICLE QUANT ADJ
SES GRANDES

= ARTICLE QUANT ADJ
LES

= G.P. PREPOSITION D.P. N
DE

= N

= ARTICLE QUANT ADJ
LA

= PREPOSITION GROUPE NOM
D-

= ARTICLE QUANT ADJ
CES

= ARTICLE QUANT ADJ
LEUR

= ARTICLE QUANT ADJ
-ES

= STRUCTURE CHIMIQUE

= STRUCTURE 21.

= 22. MACROMOLECULE

= 22. MACROMOLECULES BIOLOGIQUES ESSENTIELLES

Fig. 4.
TEMPS D'ANALYSE = 34 SEC/100

JM-1

************************ ANALYSE NO 4 ************************

LES CONNAISSANCES ACQUISES DEPUIS QUELQUES ANNEÉES SUR LA STRUCTURE DES MACROMOLÉCULES BIOLOGIQUES ESSENTIELLES, ACIDES NUCLEIQUES ET PROTEINES, PERMETTENT DE COMPRENDRE, A- JU MOINS DANS SES GRANDES LIGNES, LE RAPPORT ENTRE LES FONCTIONS DE CES MACROMOLÉCULES ET LEUR STRUCTURE CHIMIQUE.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN

2. CHAINE D'ASSERTION = • (PART) SUJET VIRGULE (PART) VERBE • OBJET DV •

3. CONNAISSANCES 4. , 5. PERMETTE

4. ARTICLE QUANT ADJ

5. VE OMEGA-PASSIF

6. CHAINE D'ASSERTION ET

7. VE OMEGA-PASSIF

8. (PART) SUJET PROTEINES

9. CHAINE D'ASSERTION = • (PART) SUJET (PART) VERBE ( ) DE COMPRENDRE , 12.

10. CHAINE D'ASSERTION

11. G.P. PREPOSITION D.P. N

12. G.P. PREPOSITION D.P. N

13. ARTICLE QUANT ADJ

14. G.P. PREPOSITION D.P. N

15. CHAINE D'ASSERTION

16. ANNEÉES

17. CHAINE D'ASSERTION

18. LIGNES

19. FONCTIONS 20.
15. CHAINE D'ASSERTION = * (PART.) SUJET * (PART.) VERBE * OBJET
   ( ) (I) 21. ST
   DV *
   RUPTURE CHIMIQUE

16. GN = ARTICLE QUANT ADJ QUELQUES

17. P N = G.P. PREPOSITION D.P. N SUR 22. STRUCTURE 23.

18. GN = ARTICLE QUANT ADJ SES GRANDES

19. GN = ARTICLE QUANT ADJ LES

20. P N = G.P. PREPOSITION D.P. N DE 24. MACROMOLECULE

21. GN = ARTICLE QUANT ADJ LEUR

22. GN = ARTICLE QUANT ADJ LA

23. NOM COMP = PREPOSITION GROUPE NOM D- 25. MACROMOLECULES BIOLOGIQUES ESSENTIELLES

24. GN = ARTICLE QUANT ADJ CES

25. GN = ARTICLE QUANT ADJ -ES

TEMPS D'ANALYSE = 2659 SEC/100

PLUS D'ANALYSES POUR CETTE Phrase

Fig. 5.
L'étude de la réplication de l'ADN in-vivo et in-vitro a apporté la preuve que le mécanisme chimique fondamental de l'hérité dit est bien celui qu'avaient proposé Watson et Crick (1953) à la suite de leur démonstration de la structure de l'ADN.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN
2. CHAINE D'ASSERTION = *(PART.) SUJET *(PART.) VERBE * OBJET
3. ÉTUDE 4. A APPORT
    DV *
    É1 5. PREUVE 6.
4. NOM COMP = ARTICLE QUANT ADJ LA
5. GN = ARTICLE QUANT ADJ LA
6. DN = PHRASE NOMINALISÉE
7. GN = ARTICLE QUANT ADJ LA
8. NOM COMP = PREPOSITION GROUPE NOM DE 7. RÉPPLICATION 8. IN-VIVO ET
9. CONJONCTION = ADJECTIF IN-VITRO
10. QUE C1/C15 = QUE CHAINE CENTRALE QUE 12.
11. GN = ARTICLE QUANT ADJ L'
12. CHAINE D'ASSERTION = *(PART.) SUJET 13. MÉCANISME CHIMIQUE FONDAMENTAL 14
    *(PART.) VERBE * OBJET DV *
    EST BIEN CELUI 15.
13. GN = ARTICLE QUANT ADJ LE
14. NOM COMP = PREPOSITION GROUPE NOM DE 16. HÉRITÉ DITE
15. DN = QU-C1,N-OMIS
16. GN = ARTICLE QUANT ADJ L'
17. QU-C1,N-OMIS = QU- CHAINE CENTRALE QU' 18.
18. C1 INVERSE = * OBJET * (PART, VERBE * OBJET * (PA ( ) AVAIENT PROPOSE1 WAT RT.) SUJET ET DV * SON ET 19. 20.
19. CONJONCTION = (PART, SUJET CRICK ( 1953 )
20. P N = G.P. PREPOSITION D.P. N A2 21. SUITE 22.
21. GN = ARTICLE QUANT ADJ LA
22. P N = G.P. PREPOSITION D.P. N DE 23. DE1COUVERTE 24

23. GN = ARTICLE QUANT ADJ LEUR
24. NOM COMP = PREPOSITION GROUPE NOM DE 25. STRUCTURE 26.
25. GN = ARTICLE QUANT ADJ LA
26. NOM COMP = PREPOSITION GROUPE NOM DE 27. ADN
27. GN = ARTICLE QUANT ADJ L'

Fig. 6.
LA DÉCOUVERTE DE L’ARN MESSAGER ET DE SON RÔLE DANS LA BIOSYNTÉHÈSE DES PROTEINES, L’ÉTUDE DES PROCESSUS DE TRANSCRIPTION, LES RECHERCHES SUR LE DÉTERMINISME GÉNÉTIQUE DES STRUCTURES PRIMAIRES DES PROTEINES ONT ENTièrement CONFIRMé1, EN LES RENOUVELANT, LES ANCIENNES HYPOTHÈSES SUR LES RELATIONS ENTRE LES DÉTERMINANTS GÉNÉTIQUES ET LES PROTEINES-ENZYMES.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN
2. CHAINE D’ASSERTION = *(PAR.) SUJET VIRGULE *(PART.)
3. DÉCOUVERTE 4. 5. 6. ONT VERBE OBJET DV
7. 
3. GN = ARTICLE QUANT ADJ LA
4. NOM COMP = PRÉPOSITION GROUPE NOM ET DE 8. ARN MESSAGER ET 9.
5. PN = G.P. PRÉPOSITION D.P. N DANS 10. BIOSYNTÉHÈSE 1
1. 6. CONJONCTION = *(PAR.) SUJET VIRGULE 12. ÉTUDE 13. , 14.
7. VE OMEGA = VE OBJET ENTièrement CONFIRMé1, 15. 16. HYPOTHE DV
23 17.
8. GN = ARTICLE QUANT ADJ L’
9. CONJONCTION = PRÉPOSITION GROUPE NOM DE SON RÔLE
10. GN = ARTICLE QUANT ADJ LA
11. NOM COMP = PRÉPOSITION GROUPE NOM D-18. PROTEINES
12. GN = ARTICLE QUANT ADJ L’
13. NOM COMP = PRÉPOSITION GROUPE NOM D-19. PROCESSUS 20.
14. CONJONCTION = *(PAR.) SUJET 21. RECHERCHES 22.
15. CS3 VANT OMEGA = CS3 VANT OMEGA EN 23.
Fig. 7.
La découverte de l'ARN messager et de son rôle dans la biosynthèses des protéines, l'étude des processus de transcription, les recherches sur le déterminisme génétique des structures primaires des protéines ont entièrement confirmé, en les renouvelant, les anciennes hypothèses sur les relations entre les déterminants génétiques et les enzymes.

1. Phrase = INTRO chaîne centrale marque de fin 2.
2. Chaîne d'assertion = * (PART.) sujet virgule * (PART.)
   3. Découverte 4. 5. 6. ont verbe objet dv *
   7. 
3. GN = ARTICLE quant ADJ la
4. NOM COMP = PREPOSITION groupe nom et de 8. ARN messager et 9.
5. PN = G.P. PREPOSITION D.P. N dans 10. BIOSYNTHESE 1 1.
6. CONJONCTION = (PART.) sujet virgule 12. étude 13. 14.
7. VE OMEGA = VE * objet ENTIEREMENT CONFIRMÉ 15. 16. HYPOTHE
   DV 2SES 17.
8. GN = ARTICLE quant ADJ l'
9. CONJONCTION = PREPOSITION groupe nom de son rôle 10. GN = ARTICLE quant ADJ l'
11. NOM COMP = PREPOSITION groupe nom d- 18. PROTEINES 12. GN = ARTICLE quant ADJ la
13. NOM COMP = PREPOSITION groupe nom d- 19. PROCESSUS 20. 14. CONJONCTION = (PART.) sujet 21. RECHERCHES 22.
15. CS3 VANT OMEGA = CS3 VANT OMEGA en 23.
TEMPS D'ANALYSE = 1194 SEC/100

PLUS D'ANALYSES POUR CETTE PHRASE

Fig. 8.
Les progrès de la biochimie réalisés depuis 50 ans ont, en outre, démontré que les propriétés, les structures, les activités qui caractérisent une cellule doivent, en définitive, être rapportées à la structure et à l'activité des protéines que cette cellule est capable de synthétiser.

1. Phrase = INTRO CHAÎNE CENTRALE MARQUE DE FIN
2. CHAÎNE D'ASSERTION = (PART.) SUJET (PART.) VERBE
3. PROGRÈS 4. 5. ONT, EN O
   OBJET D.V.
4. UTRE, 6.
5. 3. GN = ARTICLE QUANT ADJ
   LES
6. 4. NOM COMP = PREPOSITION GROUPE NOM
   DE 7. BIOCHIMIE
7. 5. DN = VE OMEGA-PASSIF
   8.
8. 6. VE OMEGA = VE OBJET D.V.
   DE1MONTRE1 9.
9. 7. GN = ARTICLE QUANT ADJ
   LA
10. 8. VE OMEGA-PASSIF = VE OM-PASS D.V.
    DE1ALISE1S 10.
11. 9. QUE C1/C15 = QUE CHAÎNE CENTRALE
    QUE 11.
12. 10. P N = G.P. PREPOSITION D.P. N
    DEPUIS 12. ANS
13. 11. CHAÎNE D'ASSERTION = (PART.) SUJET VIRGULE (PART.) VER
    13. PROPRIÉTÉS, 14. DOIVENT,
    OBJET D.V. * EN DEFINITIVE, ÊTRE 15.
14. 12. GN = ARTICLE QUANT ADJ
    50
15. 13. GN = ARTICLE QUANT ADJ
    LES
16. 14. CONJONCTION = (PART.) SUJET VIRGULE
    16. STRUCTURES, 17.
17. 15. VE OMEGA PASSIF = VE OM PASS ET D.V.
    RAPPORTÉES 18. ET 19.
18. 16. GN = ARTICLE QUANT ADJ
    LES
19. 17. CONJONCTION = (PART.) SUJET
    20. ACTIVITÉS 21.
20. 18. P N = G.P. PREPOSITION D.P. N
    A2 22. STRUCTURE
19. CONJONCTION = OM-PASS
20. GN = ARTICLE QUANT ADJ
   LES
21. DN = QU-C1,N-OMIS
   24.
22. GN = ARTICLE QUANT ADJ
   LA
23. PN = G.P. PREPOSITION D.P. N
   A2  25. ACTIVITE1 26.
   27.
24. QU-C1,N-OMIS QU CHAINE CENTRALE
   QUI 28.
25. GN ARTICLE QUANT ADJ
   L'
26. NCM COMP PREPOSITION GROUPE NOM
   D  29. PROTEINES
27. DN QU-C1,N-OMIS
   30.
28. CHAINE D'ASSERTION = (PART.) SUJET (PART.) VERBE OBJET
   ( ) CARACTERE1RISENT 31. C
   D.V.
   ELLULE
29. GN ARTICLE QUANT ADJ
   -ES
30. QU-C1,N-OMIS QU CHAINE CENTRALE
   QUE 32.
31. GN ARTICLE QUANT ADJ
   UNE
32. CHAINE D'ASSERTION = (PART.) SUJET (PART.) VERBE
   OBJET
   33. CELLULE EST CAPABL
   E 34.
   D.V.
33. GN = ARTICLE QUANT ADJ
   CETTE
34. DA PHRASE NOMINALISEE
   35
35. DE V OMEGA = DE V OMEGA
   DE 36.
36. V OMEGA = V
   OBJET D.V.
   SYNTHE1TISER ( )

PLUS DE PLACE DANS LES ARCHIVES
TEMPS D'ANALYSE = 1459 SEC/100
PLUS D'ANALYSES POUR CETTE PHRASE

Fig. 9.
OR, LA STRUCTURE DE CHACUNE DE CES PROTEINES EST INTEGRALE- 
MENT DEFINIE, ON PEUT L'AFFIRMER AUJOURD'HUI, PAR UN SEGMENT 
GENETIQUE.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN
2. 
3. CHAINE D'ASSERTION = * (PART.) SUJET * (PART.) VERBE * OBJET
4. OR, 3. STRUCTURE 4. EST
5. BJET
6. NTEGRALEMENT DEFINIE, 5.
7. 
8. GN = ARTICLE QUANT ADJ LA
9. 
10. PN = G.P. PREPOSITION D.P. N
11. DE CHACUNE 6.
12. 
13. PROP. INCISE = CHAINE D'ASSERTION
14. 7.
15. 
16. PN = G.P. PREPOSITION D.P. N
17. DE 8. PROTEINES
18. 
19. CHAINE D'ASSERTION = * (PART.) SUJET * (PART.) VERBE * OBJET
20. ON PEUT
21. 9.
22. 
23. GN = ARTICLE QUANT ADJ CES
24. 
25. V OMEGA = V
26. OBJET DV
27. L' AFFIRMER AUJOURD'HUI, 10.
28. 
29. PN = G.P. PREPOSITION D.P. N
30. PAR 11. SEGMENT GENETIQUE
31. 
32. GN = ARTICLE QUANT ADJ UN
33. 

TEMPS D'ANALYSE = 662 SEC/100

PLUS D'ANALYSES POUR CETTE Phrase

Fig. 10.
TEMPS D'ANALYSE = 313 SEC/100

JM-6A

ANALYSE NO 1

MAIS ALORS-QUE LA STRUCTURE DE L'ADN D'UNE CELLULE OU D'UNE LIGNEÉE CELLULAIRE EST INVARIANTE, LES PROPRIÉTÉS BIOCHIMIQUES RÉALISÉES ET EXPRESSÉES PAR CETTE CELLULE POURRONT ÊTRE PROFONDEMENT DIFFÉRENTES.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN MAIS 2.
2. CHAINE D'ASSERTION = (PART.) SUJET (PA 3. 4. PROPRIÉTÉS BIOCHIMIQUES 5. POU RT.) VERBE • OBJET PRONT 6. STRUCTURE 7. EST INVARIANTE,

3. CS1 C1 = CS1 CHAINE D’ASSERTION ALORS-QUE 6.
4. GN = ARTICLE QUANT ADJ LES
5. DN = VE OMEGA-PASSIF 7.
6. CHAINE D’ASSERTION = (PART.) SUJET (PART.) VERBE • OBJET T DV * ENTES

7. VE OMEGA-PASSIF = VE ET CM-PASS DV * RÉALISÉES ET 10.
8. GN = ARTICLE QUANT ADJ L'
9. NOM COMP = PREPOSITION GROUPE NOM DE 11. ADN 12.
10. CONJONCTION = VE 13. EXPRESSÉES
11. GN = ARTICLE QUANT ADJ LA
12. NOM COMP = PREPOSITION GROUPE NOM OU 14. CELLULE OU 15. D.
13. PN = G.P. PREPOSITION D.P. N PAR 16. CELLULE
14. GN = ARTICLE QUANT ADJ UNE
15. CONJONCTION = PREPOSITION GROUPE NOM D. 17. LIGNEÉE CELLULAIRE
16. GN = ARTICLE QUANT ADJ CETTE
17. GN = ARTICLE QUANT ADJ UNE

Fig. 11.
MAIS ALORS-QUE LA STRUCTURE DE L'ADN D, UNE CELLULE OU D, UNE LIGNEÉE CELLULAIRE EST INVARIANTE, LES PROPRIÉTÉS BIOCHIMIQUES RÉALISÉES ET EXPRIMÉES PAR CETTE CELLULE POURRONT ÊTRE PROFONDEMENT DIFFÉRENTES.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN MAIS 2.
2. CHAINE D'ASSERTION = *(PART.) SUJET *(PA 3, 4. PROPRIÉTÉS BIOCHIMIQUES 5. POU RT.) VERBE • OBJET PRONT ÊTRE PROFONDEMENT DIFFÉRENTES DV •
3. CS₁ C₁ = CS₁ CHAINE D'ASSERTION ALORS-QUE 6.
4. GN = ARTICLE QUANT ADJ LES
5. DN = VE OMEGA-PASSIF 7.
6. CHAINE D'ASSERTION = *(PART.) SUJET *(PART.) VERBE • OBJE 8. STRUCTURE 9. EST INVANTA T DV •
7. VE OMEGA-PASSIF = VE ET • OMEGA-PASSIF • REALISÉES ET 10.
8. GN = ARTICLE QUANT ADJ LA
9. NOM COMP = PREPOSITION GROUPE NOM OU DE 11. ADN 12. OU 13.
10. CONJONCTION = VE EXPRIMÉES 14.
11. GN = ARTICLE QUANT ADJ L'
12. NOM COMP = PREPOSITION GROUPE NOM D, 15. CELLULE
13. CONJONCTION = PREPOSITION GROUPE NOM D, 16. LIGNEÉE CELLULAIRE
14. PN = G.P. PREPOSITION D.P. N PAR 17. CELLULE
15. GN = ARTICLE QUANT ADJ UNE
16. GN = ARTICLE QUANT ADJ UNE
17. GN = ARTICLE QUANT ADJ CETTE

*** PLUS DE PLACE DANS LES ARCHIVES ***
TEMPS D'ANALYSE = 5131 SEC/100

***** PLUS D'ANALYSES POUR CETTE PHRASE *****

Fig. 12.
TEMPS D'ANALYSE = 200 SEC/100

JM-6B

***************ANALYSE NO 1 ***************

EN OUTRE, CES PROPRIÉTÉS SONT MODIFIABLES EN FONCTION DE SIGNAUX CHIMIQUES SPÉCIFIQUES VENUS DE L'EXTERIEUR.

1. PHRASE = INTRO CHAINE CENTRALE MARQUE DE FIN
2. CHAINE D'ASSERTION = *(PART.) SUJET *(PART.) VER
   EN OUTRE, 3. PROPRIÉTÉS SONT
   BE * OBJET DV *
   MODIFIABLES 4. 5.
3. GN = ARTICLE QUANT ADJ
   CES
4. PN = G.P. PREPOSITION D.P. N
   EN FONCTION
5. PN = G.P. PREPOSITION D.P. N
   DE SIGNAUX CHIMIQUES
   SPECIIFIQUES 6.
6. DN = VE OMEGA
7.
7. VE OMEGA = VE * OBJET DV
   VENUS 8.
8. PN = G.P. PREPOSITION D.P. N
   DE 9. EXTERIEUR
9. GN = ARTICLE QUANT ADJ
   L'

TEMPS D'ANALYSE = 260 SEC/100

***** PLUS D'ANALYSES POUR CETTE PHRASE *****

Fig. 13.
L'INJECTION D'HYPERNORME THYROIDIENNE À UN TESTARD PROVOQUE, AVANT TOUTE MANIFESTATION MORPHOLOGIQUE, LA CONVERSION DE SON SYSTÈME EXCRETORIÈRE QUI, DE SEMBLABLE À CEUX D'UN POISSON, DEVIENDRA CHIMIQUEMENT ANALOGUE À CEUX D'UN MAMMAL.
15. P N = G.P. PREPOSITION D.P. N A2 CELUI 17.
16. P N = G.P. PREPOSITION D.P. N A2 CELUI 18.
17. P N = G.P. PREPOSITION D.P. N D, 19. MAMMIFE2RE
18. P N = G.P. PREPOSITION D.P. N D, 20. POISSON ,
19. GN = ARTICLE QUANT ADJ UN
20. GN = ARTICLE QUANT ADJ UN

*** PLUS DE PLACE DANS LES ARCHIVES ***
TEMPS D'ANALYSE = 1428 SEC/100

***** PLUS D'ANALYSES PO UR CETTE PHRASE *****

Fig. 14.