MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

ANNALES ISLAMOLOGIQUES

en ligne en ligne en ligne en ligne en ligne en ligne en ligne en ligne en ligne en ligne

AnIsl 22 (1987), p. 217-256

Claude Audebert, André Jaccarini

À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.

Conditions d'utilisation

L'utilisation du contenu de ce site est limitée à un usage personnel et non commercial. Toute autre utilisation du site et de son contenu est soumise à une autorisation préalable de l'éditeur (contact AT ifao.egnet.net). Le copyright est conservé par l'éditeur (Ifao).

Conditions of Use

You may use content in this website only for your personal, noncommercial use. Any further use of this website and its content is forbidden, unless you have obtained prior permission from the publisher (contact AT ifao.egnet.net). The copyright is retained by the publisher (Ifao).

Dernières publications

9782724711264	*La chapelle d'Osiris Ounnefer Neb-djefaou*	Laurent Coulon
9782724711752	*Guide du Caire Islamique*	Hadrien Collet
9782724711660	*Mélanges de l'Institut dominicain d'études orientales 41*	Adrien Candiard (éd.)
9782724711745	*Bulletin archéologique des Écoles françaises à l'étranger (BAEFE)*	
9782724711271	*Du côté de chez Zeus*	Emmanuel Botte (éd.), Hélène Cuvigny (éd.), Michel Reddé (éd.)
9782724711677	*La Description de l'Égypte*	Paul-Marie Grinevald
9782724711714	*La pensée et la pratique pharmacologiques d'Avicenne*	Sylvie Ayari
9782724711899	*BCAI 40*	

© Institut français d'archéologie orientale - Le Caire

À LA RECHERCHE DU *ḤABAR*

OUTILS EN VUE DE L'ÉTABLISSEMENT D'UN PROGRAMME D'ENSEIGNEMENT DE L'ARABE ASSISTÉ PAR ORDINATEUR

Claude AUDEBERT et André JACCARINI

I

Nous essayons de définir ici un ensemble d'outils nécessaires à la conception d'un programme d'enseignement assisté par ordinateur. Cette réflexion part d'une expérience d'enseignement et de l'observation d'un apprenant lorsqu'il cherche à décoder un texte arabe non vocalisé. L'analyse de la démarche de l'étudiant et les correctifs à lui apporter devront constituer la logique du programme.

Le niveau de l'apprenant observé est celui d'un débutant ayant fait 200 à 250 heures d'arabe dans un laps de temps continu — une année scolaire par exemple —. En conséquence celui-ci aura fait un tour d'horizon assez complet de la grammaire : acquis les conjugaisons, les formes dérivées et l'essentiel des règles de syntaxe avec les possibilités d'erreur que cela implique à ce niveau. L'apprenant n'est donc pas ignorant et dispose de certaines connaissances non encore assises.

Le niveau de langue des textes à étudier sera celui d'un arabe contemporain, de moyenne difficulté excluant tout archaïsme et toute préciosité. Ce type de langue est représenté en arabe contemporain dans la presse non spécialisée, nombre d'œuvres à caractère littéraire, articles etc... A défaut d'études statistiques suffisantes, nous aurons recours à une appréciation intuitive de ce niveau de langue [1].

Ces données une fois fixées il convient d'analyser les difficultés de l'apprenant, les procédures généralement mises en œuvre par l'enseignant dans la classe et l'objectif fixé par celui-ci. En d'autres termes les considérations qui suivent se placeront du triple point de vue de la démarche de l'apprenant, de celle de l'enseignant analysant sa propre démarche (le choix des textes qu'il fait du point de vue du niveau de langue, des objectifs qu'il fixe à l'étudiant, et de la démarche qu'il cherche à lui inculquer) et des exigences linguistiques du texte même.

[1] Voir plus bas p. 233-240 (Les règles illustrées par les diagrammes).

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

La principale difficulté que rencontre l'apprenant face à un texte consiste à repérer des unités de sens dans un flux d'information qui peut sembler assez flou. Le texte est représenté dans une *graphie* d'où sont absentes les voyelles brèves et dont le découpage en « mots » n'apparaît pas toujours évident. Il arrive par exemple parfois à l'apprenant de ne pas identifier un même mot répété à quelques lignes de différence du fait de l'agglutination d'un pronom, d'un suffixe ou de la préposition. Parfois en dépit de la ponctuation — souvent défectueuse il est vrai — il ne se représente pas clairement où commence la phrase et où elle finit. Les anaphoriques lui échappent. Il a enfin tendance à se jeter dans le dictionnaire, espérant que la signification d'un mot le mettra sur la voie. Ceci dans le cas le plus favorable où il s'est montré capable de dégager la « bonne » racine et où le sens qu'il choisit, souvent au hasard, sera celui qui cadre avec la phrase. Il serait trop long d'énumérer ici les mille et un chemins où peut s'égarer l'apprenant malgré son acquis grammatical qui est souvent considérable mais mal organisé.

Si l'on cherche à comprendre les raisons des déboires de l'apprenant, on se rendra vite compte qu'elles tiennent à la manière anarchique dont il travaille, — à la fois par manque de savoir faire mais aussi parce que sa démarche lui est imposée par la nature même de l'arabe où s'imbriquent morphologie et syntaxe. Manière anarchique et tout à la fois mécanique et à courte vue. En effet, la pratique habituelle consiste à faire progresser l'apprenant *mot à mot* en opérant un décorticage morphologique sans l'amener à prendre une vue d'ensemble du terrain sur lequel il avance. Pourtant à y bien regarder, le terrain est balisé par quantité de points de repère qui doivent permettre à un étudiant bien guidé de dégager des *ensembles* dans un premier temps et à laisser le décorticage morphologique rigoureux pour un deuxième temps et dans la mesure où cela est nécessaire [1]. L'apprenant a donc besoin d'acquérir une *démarche* de travail, de mettre en jeu toutes les connaissances dont il dispose tant syntaxiques que morphologiques afin d'être en mesure d'avoir accès au *dictionnaire* de la manière la plus fructueuse possible. Ceci implique son aptitude à découvrir la racine, mais aussi à repérer à quelle classe de mot il a affaire : nom, verbe, maṣdar, ce qui veut dire à son tour que plus il aura établi de relations entre les groupes de mots dans la phrase, plus il sera à même d'y parvenir. En effet la manière la plus fructueuse de faire le bon choix dans le dictionnaire consiste à avoir déjà formulé des hypothèses sur ces relations et sur la direction de la phrase. Une dernière remarque : l'apprenant, bien guidé, en possession d'un nombre relativement

[1] En effet cela dépend de l'objectif visé : la lecture à haute voix imposera un décodage morphologique rigoureux. Une lecture qui a pour but une compréhension du sens ne l'exigera pas forcément.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

restreint de règles tant morphologiques que syntaxiques, peut arriver à résoudre un texte de difficulté moyenne. C'est dire qu'il peut faire beaucoup avec un minimum de règles, à condition de savoir les mettre en jeu au bon moment.

De cette brève analyse on peut tirer les conclusions suivantes en termes de *démarche* à proposer à l'apprenant, démarche qui sera reflétée dans la logique du programme :

On amènera l'apprenant : a) à avoir une vue d'ensemble de la phrase, par exemple à commencer prosaïquement par se rendre compte où elle commence et où elle finit; b) à prendre la *syntaxe* comme moniteur c'est-à-dire à rechercher dans le texte des points de *repère,* d'où la notion de « structurants » qui serviront au repérage de groupes de mots susceptibles de s'agencer ensemble. Ceci à son tour met en jeu la notion de *globalité* et de *localité* ainsi que celle de *hiérarchie* de règles que nous clarifierons plus bas.

Pendant ces opérations, il est évident que les règles morphologiques auront leur rôle à jouer. L'expérience d'enseignement démontre en effet qu'il se produit lors du décodage d'un texte arabe non vocalisé un *chassé croisé* continuel de la morphologie à la syntaxe (ceci reste partiellement vrai d'un texte vocalisé, l'expérience montrant que l'étudiant omet souvent les voyelles même si elles sont notées!).

Une fois le balayage du texte opéré, les structurants repérés, le problème revient à distinguer les groupes nominaux (GN) et leur hiérarchie, les verbes (V) et leurs dépendants qui viennent s'insérer dans les armatures de la phrase constituées par les structurants qui servent de vecteur à la proposition.

Comment repérer rapidement GN et V en partant d'éléments directement observables et en ayant recours à divers registres de connaissances? Pour prendre un exemple trivial il existe pour déterminer le nom (N) des marques directement observables : la présence de l'article ال, le ة, marque dominante du féminin, ou qui signale un pluriel non doué de raison; à un autre niveau : la distribution, par ex. une préposition entraîne GN ou pronom et ne saurait précéder un verbe.

Les indications sur la direction de la phrase, les attentes après certaines occurrences auront amené l'apprenant à formuler des hypothèses. Que celles-ci soient contradictoires voire erronées est moins grave que de n'en point formuler du tout, ce qui est fréquemment le cas dans la procédure du mot à mot. L'essentiel est qu'elles soient formulées sur l'ensemble de la phrase. Les connaissances morphologiques rigoureuses viendront alors dans un deuxième temps lever les doutes.

Le programme tenant compte de la démarche proposée consistera à la *simuler.* Il devra donc tenir compte du *chassé-croisé* syntaxe-morphologie, de manière à faire intervenir les deux niveaux ensemble *parallèlement.* Ceux-ci interagissent de manière de plus en

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

plus forte au fur et à mesure que la compétence de l'apprenant augmente, ainsi que le démontre l'expérience.

C'est cette constatation qui nous a conduits à chercher un formalisme rigoureux : a) qui fournisse une véritable description logique des processus d'appréhension et non simplement une description du texte de la phrase étudiée; b) qui permette de mieux cerner les passages d'information du niveau de la syntaxe au niveau de la morphologie.

D'où le recours à la notion d'un processeur abstrait ou idéal que l'apprenant mettrait en marche dans sa tête lorsqu'il cherche à décoder une phrase. L'apprenant bien formé se pose en effet implicitement des questions et fournit des réponses qui entraînent une cascade de questions et de réponses jusqu'à ce qu'il ait trouvé — ou non — la solution. Au cours de ce processus il garde en mémoire des définitions correspondant à chacune des options et n'opérera en définitive que des choix compatibles entre eux en fonction de sa compétence.

Ce point de vue nous amènera à esquisser dans les pages qui suivent une grammaire des attentes.

II

LOGIQUE DU PROGRAMME

Nous disposons d'un programme morphologique qui ouvre l'accès au dictionnaire du fait qu'il permet l'extraction de la racine. Il fournit d'autre part tous les schèmes possibles attachés à une forme non vocalisée. Dans certains cas les ambiguïtés du schème peuvent être importantes comme le montre l'exemple فعل [1]. Toutefois ce programme fournit un grand nombre d'informations sur les catégories lexicales sans avoir recours au dictionnaire : repérages de verbes, noms etc... Les catégories grammaticales qui apparaissent à ce niveau sont considérables.

Un grand nombre d'ambiguïtés peuvent être levées par le recours au contexte d'où la notion de *moniteur syntaxique* qui guidera le programme morphologique. Ce moniteur syntaxique permettra :

1) de lever les ambiguïtés qui affectent l'entrée dans le dictionnaire (ambiguïté de la racine). Ex. فقد décodable en فَـقَـدَ (V) ou فَـقَـدْ;

[1] En l'absence de contexte فعل peut représenter dans le domaine verbal فَعَل ، فَعِلْ ، فَعَّل avec leurs formes passives et dans le domaine nominal فَعَل ، فُعّل ، فَعّل ، فِعّل etc.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

2) de lever les ambiguïtés des schèmes (à l'intérieur du dictionnaire);
3) de définir une stratégie de recherche pour l'apprenant comme on le verra dans la structure du programme exposé ci-dessous;
4) d'optimiser le programme morphologique.

En effet, du point de vue de l'optimisation, le programme morphologique est très peu effectif lorsqu'il est appliqué à un texte réel du fait que les opérations de décodage deviennent très répétitives.

Au-delà donc de l'optimisation de l'analyse morphologique locale il faut mettre au point un autre programme qui utilise cette analyse et l'étende au texte dans son ensemble. Ainsi à la procédure mot à mot nous substituerons des traitements en bloc qui permettent par exemple un traitement unique des mots identiques ou qui ne font appel qu'à un nombre restreint d'opérations pour ceux d'une longueur déterminée etc... Cela suppose donc un premier balayage du texte en vue d'une préclassification des mots. Ces derniers sont ordonnés en fonction de la complexité de leur analyse morphologique. Cependant cette préclassification peut être considérablement optimisée si l'on tient compte de la syntaxe. En effet cette dernière représente des contraintes sur les enchaînements de mots à l'intérieur de la phrase de même que la morphologie représente les contraintes d'enchaînement de lettres à l'intérieur du mot. Nous nous sommes donc attachés à dégager les règles minimum qui régissent les agglomérations de mots.

Nous avons cherché ensuite un formalisme qui puisse refléter cette stabilité syntaxique et qui soit compatible avec le programme morphologique. Ces remarques précisent la signification que nous avons donnée au terme « moniteur syntaxique ».

Le programme morphologique permet de dégager des éléments qui échappent totalement ou en partie aux règles de formation morphologiques du mot et ont le statut de mots « vides ». On a donc intérêt à les repérer en premier lieu afin d'éviter que le programme ne se bloque [1].

Ces éléments au nombre de 150 environ, que nous nommerons *token*-s, appellent des définitions élémentaires que nous pouvons classer dans un *tableau* auquel se réfère constamment le programme morphologique.

Or on remarque que la plupart de ces éléments ont une répercussion syntaxique importante soit qu'ils engagent la globalité de la phrase, soit qu'ils aient une incidence sur l'environnement immédiat. Ceci pose le problème de la hiérarchie de ces éléments

[1] Cette opération coïncide avec l'expérience d'enseignement qui consiste à conduire l'étudiant à rechercher d'emblée les points de repère.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

structurants. Les critères retenus seront ceux du degré de *globalité* que nous tenterons de définir dans des règles.

Les globalités ne sont pas toutes de même nature. C'est en réfléchissant sur la nature de ces dernières qu'on est amené à poser des règles qui opèrent sur des chaînes entières et à développer de même les rapports entre certains types de localité et les globalités.

Nous devrons donc représenter dans notre tableau les critères: 1) de positionnement du token; 2) du niveau de globalité qu'il engage; 3) de son environnement. Nous reviendrons sur ces notions plus bas.

Il est normal, lorsqu'on décode un message, de repérer d'abord l'élément de ce message qui engage le plus grand environnement. La logique commande de dégager en premier lieu les plus forts structurants qui réduisent d'un coup le « niveau entropique » (le degré d'indétermination du flux informationnel, ceux qui introduisent les contraintes les plus importantes. Ce sont ces considérations qui nous ont amenés à adjoindre à notre programme morphologique le tableau des opérateurs qui doit refléter sous forme de relations logiques toute notre information syntaxique.

En effet, ces opérateurs peuvent selon les cas représenter de véritables automates descriptifs [1] de la structure syntaxique.

Enfin ce tableau pourra être aussi conçu comme une classification de ces automates.

Pour fixer les idées considérons la phrase :

ان الاتفاقية الاردنية الفلسطينية التى تستهدف السعى من أجل تسوية سلميه هى بمثابة تطور عظيم ينبغى الا نقلل من شأنه او نجنبه

1) Le balayage donne la liste des tokens :

ان — التى — من — هى — الا — من — او

2) On notera que certains de ces tokens sont *ambigus graphiquement*, c.-à-d. qu'ils peuvent représenter sous le même signe graphique divers *opérateurs* grammaticaux.

Ainsi ان peut représenter : أنْ — أنَّ — إنَّ — إنْ

De même الا peut représenter : إلَّا — ألَّا — ألا (interro-négatif) et ألا exclamatif; من peut représenter مِنْ préposition et مَنْ pronom relatif, مَنْ interrogatif. Par contre les *tokens* التى — هى — أو sont reconnaissables en première lecture.

[1] Sur la définition des automates, voir p. 231, le paragraphe intitulé « Les Diagrammes ».

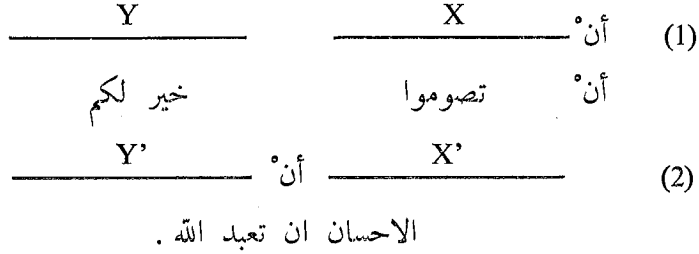
AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Les éléments ambigus figurent sous une seule entrée dans le tableau ce qui pose le problème d'une décision à prendre lors du décodage.

* Attachons-nous au symbole ان. Selon qu'il représente une des 4 possibilités, il engage différemment le processus du décodage. Chacun de ces signes a, en effet, une incidence différente sur l'ensemble des occurrences de la phrase. Il conviendra donc de les classer en fonction de leur influence plus ou moins forte sur la structure de la phrase.

CAS A

Ainsi l'on peut trouver ان comme structurant du noyau prédicatif dans les configurations suivantes :



où X et X' représentent respectivement un ensemble d'occurrences en fonction de thème et Y et Y' un ensemble d'occurrences en fonction de prédicat.

On voit que cet opérateur engage la globalité de la phrase. Si nous envisageons maintenant le positionnement immédiat des occurrences environnant ان on remarque que celui-ci doit être *strictement* suivi (occurrence n + 1) d'un verbe. Il peut d'autre part être précédé d'une préposition éventuellement agglutinée comme dans بأنْ (Le programme morphologique aura préalablement replacé les éléments agglutinés en leur occurrence véritable).

ان commande en n + 1 un verbe; on dira qu'il s'agit là d'une contrainte *stricte.*

X représente le groupe verbal. Ce dernier occupe un ensemble d'occurrences dont la longueur n'est pas déterminée. Cet ensemble possède néanmoins une occurrence privilégiée n + 1 qui la gouverne. La contrainte qui commande le positionnement de X par rapport à ان est dite, par opposition à la précédente, une contrainte de *position lâche* car elle commande un ensemble de positions non fixées.

En ce qui concerne Y des critères de globalité et de position entrent en jeu. En effet Y apparaît nécessairement dans la structure globale YX ان et son positionnement dépend de l'expansion de X [1].

[1] Remarque : on appellera *contrainte stricte* une contrainte portant sur une *position déterminée.*

ex. V أن

L'occurrence n + 1 doit être un v. (n désignant la position de l'opérateur).

On appellera *contrainte lâche* une contrainte

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

— Les relations qui touchent X et Y en tant que noyau prédicatif seront évoquées plus bas.

Dans la configuration (2) ان n'implique aucune contrainte de position stricte en < n mais uniquement une contrainte de position lâche. Effectivement X' représente un groupe d'occurrences de longueur non déterminée commandées par la première occurrence nominale; en effet cet X' ne peut être qu'un groupe nominal. Y' doit représenter un groupe verbal (GV) commandé par ان et qui possède une occurrence privilégiée n + 1 = V (contrainte de position stricte).

CAS B

Dans la configuration suivante :

$$\frac{\overset{V}{\text{أَنْ}}}{Y} \quad \frac{V}{X} \qquad (3)$$

$$\frac{\overset{V}{\text{أَنْ}}}{Y'} \quad \frac{\text{على}}{X'} \qquad (4)$$

— On retrouve la même contrainte *stricte* de position relative à أن c.-à-d. n + 1 est un verbe.

— Y représente un GV initialisé par l'occurrence n + 1.

— X est un GV où V occupe l'une des occurrences inférieures à n. Il s'agit donc d'une contrainte de position lâche. (Ce V marqué sémantiquement gouverne en fait le أن).

Nous avons donc à loger un GV en X, duquel Y est entièrement dépendant.

* Il est également utile de considérer de part et d'autre de l'opérateur les extensions de X et de Y.

Si X minimum, alors X = V, ce qui donne :

$$\frac{V}{Y} \text{ أن } V$$

portant sur un *ensemble* d'occurrences dont la longueur n'est pas nécessairement déterminée.

ex. $\frac{\text{الاتفاقية الاردنية الفلسطينية}}{X}$ أن

X représente l'ensemble des occurrences n + 1, n + 2, n + 3

Cet ensemble d'occurrences doit être celui d'un groupe nominal GN. Sa longueur n'est pas nécessairement déterminée. Dans ce cas il existe cependant une contrainte stricte car n + 1 doit être égal à un nom ou un pronom.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Ce X n'est extensible que par ajouts au verbe ce qui amènera à étudier ceux-ci [1]. Toutefois on peut noter qu'avec un X minimum on est obligé de chercher pour certains verbes du type ينبغى le sujet en Y. A la différence du cas A les groupes X et Y sont des propositions qui ne reçoivent pas leur fonction grammaticale par l'opérateur ان. On remarque, dans le cas A, que ان commande deux ensembles d'occurrences X et Y qui coïncident avec des fonctions grammaticales qui sont celles du thème et du prédicat et sont induites par l'opérateur lui-même. Tandis que dans B, les fonctions des deux ensembles ne sont plus induites par l'opérateur. Les fonctions des ensembles X et Y ne sont pas définies. Le cas A engage a priori une globalité celle du noyau prédicatif) et se trouve à un niveau supérieur de structuration. Le cas B engage également une globalité mais de niveau moindre.

Le cas B (4) est de même nature que B (3) à cette différence près que X' est un groupe nominal commandé par على ou ل dont l'extension possible est celle de ce groupe lui-même. On n'a donc ici aucune contrainte de position stricte. Mais il y aura lieu d'expliciter le groupe nominal initialisé par une préposition.

CAS B'

Dans la configuration $\underset{Y}{\text{———}}$ أنّ $\underset{X}{\text{———}}$ V (5)

illustrée par l'exemple : الولد مريض أنّ علم

X représente un GV et Y une phrase nominale.

L'opérateur إن entraîne une contrainte stricte de position : en n + 1 on doit avoir un nom ou pronom en fonction de thème suivi d'un prédicat. Les deux éléments étant susceptibles de recevoir des ajouts. Comme dans le cas B c'est la sémantique de V qui entraîne le choix de l'opérateur.

CAS C

Dans la configuration $\underset{Y}{\text{———}}$ $\underset{X}{\text{———}}$ إنّ (6)

illustrée par l'exemple : مريضٌ الوَلَدَ إنّ

[1] Ce type d'étude peut faire l'objet d'une grammaire en chaîne de l'arabe. En effet, le sujet du verbe peut être représenté par un N dont les extensions sont celles du GN ou GND dans toute leur généralité (diagramme XXV et XXVI). Ces GN peuvent être extensibles à leur tour, ils sont récursifs. Le verbe admet également comme possibilité d'ajouts des groupes adverbiaux, récursifs également etc...

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

إنّ introduit une contrainte de position stricte sur n + 1 qui doit être égal à un nom ou un pronom et une contrainte de position lâche sur X et Y qui doivent coïncider avec les fonctions de thème et de prédicat.

X doit être un groupe nominal déterminé (GND). La longueur de ce groupe n'est fixée que par le fait qu'il doit être suivi d'un autre ensemble d'occurrences Y. Y peut être un GN, un GV ou un GAD.

Remarque : Si l'on supprimait أنْ dans la configuration du cas A (1) illustrée par أنْ تصوموا خير لكم la phrase ne tiendrait pas. Par contre la suppression de إنّ dans la configuration C n'entraîne pas l'effondrement de la phrase qui peut se réduire à : الولد مريض

إنّ se situe cependant à un niveau qui engage la globalité de la phrase puisque son existence implique l'existence des occurrences X et Y et entraîne une modification de la vocalisation. Dans la hiérarchie des globalités, il sera tenu compte de ce fait.

CAS D

Nous ne nous étendrons pas ici sur la configuration :

$$\frac{}{Y}\ \text{إنّ}\ \frac{}{X}\ V \qquad (7)$$

illustrée par : الولد مريض إنّ قال

qui constitue un cas particulier du cas B'.

CAS E

Dans la configuration :

$$\frac{}{Y}\qquad \frac{}{X}\ \text{إنْ} \qquad (8)$$

illustrée par les phrases :

إنْ تجتهدْ تنجحْ et إنْ اجتهدت نجحت

X et Y constituent chacun un noyau avec la contrainte de position fixe n + 1 = verbe. X est donc une phrase verbale, alors que Y peut également être nominale, auquel cas elle sera introduite par ف. Nous reviendrons sur ce cas (cf. diagramme IX).

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Dans les considérations précédentes nous avons été exclusivement guidés par la logique des attentes induites par les opérateurs, laissant de côté tout autre type de remarque. Examinons maintenant les contraintes induites par les autres *tokens* repérés dans la phrase étudiée.

Parmi les opérateurs ambigus graphiquement nous avons noté الا et من que nous aborderons brièvement :

* الا

1 — أَلَا. Au stade actuel de notre programme nous considérerons الا (exclamatif) comme ajout à droite d'une phrase quelconque. Il sera défini par une contrainte de début de phrase. Nous parlons d'ajouts car il peut être enlevé sans modifier le statut de la phrase à la différence d'un opérateur comme ان (cas C). Ex. : ألا فاسقنى خمراً

2 — أَلَا (*interro-négatif*). De par sa nature interrogative, il a une position stricte au début de la phrase. Il induit en général une contrainte de position fixe sur n + 1 qui initialise un GV. Cet élément décomposé en أ et لا commande du fait de la négation, un inaccompli qui constitue un élément important de reconnaissance puisque le temps du verbe est reconnaissable par le programme morphologique. Ex. : ألا تريد ان تذهب إلى القاهرة ؟

3 — أَلَّا peut également se décomposer en أنْ et لا que le programme morphologique aura replacé dans leurs occurrences respectives ce qui donnera la configuration :

$$\frac{\qquad\qquad}{Y}\ \text{لا}\ \text{أنْ}\ \frac{\qquad\qquad}{X}\ V$$

ce qui nous ramène au cas B. Il est important de noter que أَلَّا ne peut pas figurer en position initiale.

4 — إِلَّا particule d'exception. Nous ne décrivons pas les contraintes attachées à إِلَّا car elles recouvrent un ensemble très vaste. Nous retiendrons que cette particule ne peut intervenir en position initiale. On notera simplement qu'elle joue le rôle d'un restricteur de fonction grammaticale au sein d'une proposition et non celle d'un joncteur entre deux phrases noyau; que d'une manière générale la particule d'exception est suivie d'un groupe nominal et que, si elle est suivie d'un verbe, la négation s'impose dans le premier membre de la proposition exceptive comme dans : ما سمع هذا الا تذكر . . .

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

* مَنْ / مِنْ

Pour lever l'ambiguïté entre مِنْ préposition et مَنْ relatif, le critère porte sur l'occurrence n + 1. En effet si l'on trouve un nom déterminé ou non déterminé ou un pronom après من on sait qu'on est en présence d'une préposition. Tandis que si من est suivi d'un verbe l'on est en présence du relatif, ou d'une particule introduisant une phrase double. Dans le diagramme IV nous avons esquissé les attentes que peut entraîner la préposition مِنْ au cas où elle est suivie d'opérateurs tels que : ما, مَنْ, أَنْ, أَنَّ qui engagent des structures complètes et fréquentes du type : من الناس من يقول آمنا

Nous ne développerons pas ici la structure comparative.

Pour ce qui est de مَنْ, il est le plus fréquemment suivi d'un verbe; cependant il peut être suivi d'un pronom isolé introduisant une phrase nominale.

Ex. : مَن هو بَرٌّ

Il est à noter qu'il ne peut être précédé d'un GND.

L'ambiguïté entre من suivi d'un élément nominal et d'un pronom (ex. : منه) sera levée par l'obligation que مَنْ relatif, s'il est suivi d'un élément nominal, l'est par un pronom *isolé*. D'où la non confusion possible entre منه et مَنْ هو.

Pour le مَنْ introduisant une phrase double, voir diagramme IX plus bas.

Subsiste le cas de مَنْ interrogatif qui peut être suivi d'un verbe ou d'un pronom nominal (GND ou GN). La distinction entre l'interrogatif et le relatif se fera à partir de la position de l'opérateur. L'interrogatif se trouvera nécessairement en première position dans la phrase à la ponctuation près. Ce qui laisse ouverte l'ambiguïté مَنْ (interrogatif) et مِن (préposition). Une analyse syntaxique sera nécessaire pour lever l'ambiguïté.

* الذي et sa classe

الذي appelle une structure noyau mais qui est gouvernée par une structure supérieure. Cette structure noyau peut être un GV ou une phrase nominale que nous n'avons pas détaillée. Le diagramme III rendra cependant compte d'un grand nombre d'attentes liées à ce token. Ce diagramme couvre les deux types possibles de relatives : celui qui est introduit par un pronom relatif et celui où ce dernier n'apparaît pas.

Dans ces structures apparaissent des pronoms liés. Ils peuvent jouer un rôle essentiel dans le repérage de la relative et à plus forte raison en l'absence du token. Or, le programme

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

initial de balayage de la phrase les repère du premier coup. Cependant l'automate ne peut décider, dans la succession de ces pronoms, lequel peut constituer un pronom de rappel proprement dit. Il peut se trouver en présence, par exemple, de deux pronoms dont l'un est de rappel et l'autre un simple ajout au verbe. Pour cette raison nous avons prévu tous les arcs pronoms liés (sans distinguo) et des arcs ajouts au verbe. Pour mieux tenir compte de la présence des pronoms liés nous avons été conduits à exprimer un arc N + pronom lié plutôt que de le traduire par un GND (qui figure bien comme possibilité dans le diagramme mais qui aurait supposé un rappel récursif du sous-automate GND).

D'autre part nous nous sommes livrés à des tests de compatibilité entre les diagrammes I, II et III, d'où la soustraction de certains arcs figurée par l'écriture GND* ou GND- { arcs / exclus }

On peut remarquer à ce sujet que l'explicitation des arcs pronoms est compatible avec la logique du programme qui commande de repérer en premier les structures (les automates) de plus haut niveau qui couvrent el maximum d'occurrences et de n'appeler ensuite les sous-automates que pour les faire agir comme accepteur (structure de contrôle). Nous avons donc intérêt à faire apparaître le pronom dans l'automate directeur (le pronom étant considéré comme révélateur de cette structure) et non dans le GND.

Nous pouvons examiner l'exemple الرجل الذى ضرب اخاه où اخاه peut être analysé soit comme N + pronom ce qui nous permet de rester au niveau de l'automate III, soit comme un GND (diagramme II, arc. 6 + arc pronom lié) ce qui suppose donc de suspendre l'attente de l'automate directeur et d'appeler récursivement un sous-automate GND. Dans ce cas, le pronom apparaît dans le sous-automate et non dans la structure relative qui nous intéresse. L'analyseur, privilégiant les chemins les plus simples, emprunte l'arc N du diagramme III qui est étiqueté par un symbole terminal plutôt que l'arc GND qui provoque une mise en pile de mémoire.

Signalons enfin que, dans l'état de notre recherche et en dépit de la précision du programme morphologique, nous n'avons pas tiré parti des informations que l'on peut obtenir quant à la nature du sujet du verbe de la relative et aux conséquences qu'il entraîne en ce qui concerne la présence ou l'absence d'un pronom de rappel. En effet $S_1 = S_2$ (S_1 = sujet de la relative. S_2 = antécédent) entraîne l'absence d'un pronom de rappel :

Ex. : الرجل الذى خرج أخى . . .

Par contre si $S_1 \neq S_2$ le pronom de rappel est de rigueur.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Si, en l'absence de tout pronom, on peut déduire $S_1 = S_2$, en revanche, à partir du moment où apparaîtra un pronom dans la phrase, on ne pourra pas prédire $S_1 = S_2$ quand bien même ce serait le cas :

Ex. : $S_1 = S_2$: L'homme qu'il frappa
الرجل الذى ضربه $S_1 \neq S_2$: L'homme qui le frappa.

En fonction des remarques précédentes, on se rend compte qu'on ne peut pas faire figurer l'arc GV en tant que tel car on ne serait plus en mesure d'expliciter les pronoms de rappel; ce qui entraîne à son tour la nécessité d'arcs AjV (ajouts au verbe).

* هو *et sa classe :*

Le token هو est non ambigu graphiquement mais peut remplir des fonctions diverses. Ce pronom isolé est porteur d'une grande information. En effet tout comme les autres pronoms isolés, il peut remplir la fonction de thème mais en outre il peut servir de pronom de séparation entre le thème et le prédicat (sous sa forme masculine ou féminine). En ce sens il est un indicateur de fonction grammaticale.

On sait aussi qu'il peut se présenter, au même titre que les autres pronoms isolés, comme pronom de corroboration. Ce dernier cas n'étant pas pertinent du point de vue d'une grammaire des attentes ne sera pas isolé, pour le moment, du cas du pronom de séparation. Comme pour les autres tokens, nous envisageons les critères de position tout en notant cependant qu'en ce qui concerne le pronom de séparation il conviendrait d'étudier plus en détail la notion d'extension des deux termes qu'il sépare.

En effet les règles d'occurrence du pronom de séparation sont floues. En règle dominante, on peut dire, qu'ayant pour fonction de lever les ambiguités entre le thème et le prédicat, il apparaîtra dans les cas suivants :

1) GND GND
   Thème prédicat
2) GND très étendu GN
   Thème prédicat

Ce dernier cas est illustré par notre phrase où هى apparaît comme séparateur vu l'occurrence du thème.

On notera que le pronom de séparation est obligatoire si l'on a GND الذى (....)
Thème Prédicat

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Anisl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 Anisl en ligne https://www.ifao.egnet.net

TABLEAU

Symbole	Operateur	Position opérateur = n		Contraintes				Nombre d'ensembles	Positionnement des ensembles	Nature des ensembles	Hiérarchie	Possibilité de suppression de l'opérateur
		Début	Milieu	n + 1	n − 1	< n	> n					
أَنْ	أَنْ	+		+ (V)			+ (GV)	2	CONTIGUS x > n; y > n; x < y	x = GV; y = OUVERT	++	−
	أَنْ		+	+ (V)		+ (GND)	+ (GV)	2	NON CONTIGUS x < n; y > n	x = GND; y = GV	+	−
	أَنْ		+	+ (V)		+ (GV)	+ (GV)	2	NON CONTIGUS x < n; y > n	x = GV; y = GV	−	−
	أَنْ		+	+ (V)		+ (prep GND)	+ (GV)	2	NON CONTIGUS x < n; y > n	x = GN; y = GV (PREP)	−	−
أَنَّ	أَنَّ		+	+ (≠ V)		+ (GV)	+ (GN)	2	NON CONTIGUS x < n; y > n	x = GV; y = PHN	−	−
إِنَّ	إِنَّ	+		+ (≠ V)			+ (GND)	2	CONTIGUS x > n; y > n; y > x	x = GND; y = GN (x, y) = PHN	+	+
إِنْ	إِنْ	+		+ (V)			+ (ف* GV)	2	CONTIGUS x > n; y > x	x = GV; y = GV	++	−
الذى	الذى		+	+ (V)		+ (GND)	+ $(Rel)_D$	⩾ 2	NON CONTIGUS x < n; y > n	x = GND; y = GV	−	
مَنْ	مَنْ	+		+ $(V)_D$			+ (ف* GV)	2	CONTIGUS x > n; y > x	x = GV; y = GV	++	−
	مَنْ		+	+ (V, PRO) ISOLÉ$_D$	≠ GND		+ (GV)	2	NON CONTIGUS x < n; y > n	x = OUVERT; y = PH	+	−
	مَنْ int	+		+ (V)			+ (GV)	1		x = GV من x = PHN	++	+
	مَنْ int	+		+ (N) + (ART)			+ (GN) + (GND)	1		x = { GN / GND } من x = PHN	++	−
مِنْ	مِنْ		+	+ (N) + (ART) + (PRO LIÉ) + (REL)			(GN) (GND)	2	NON CONTIGUS	x = OUVERT y = GN; GND	−	−
هو	هو	+			(ف ; و)		(GN) (GND)	1		x = GN; GND = PRÉDICAT	++	−
	هو		+			+ (GND)	(GN) (GND) (GV)	2		x = THÈME y = PRÉDICAT		(+)
	هو		+			+ (GN)				x = GN y = Rel		+
او	او		+					2	NON CONTIGUS x < n; y > n	x = y (STR)		−
الا	الا excl	+						1		x = OUVERT	+	+
	الا int	+		+ $(V)_D$ inaccompli			+ (GV)	1		x = GV	+	
	الا		+	+ (V)		+ (GV)	+ (GV)	2		x = GV; y = GV	−	−

Enfin هو et هى peuvent introduire une phrase nominale en fonction de relative marquée ou non par الذى

* او

او est un coordonnant qui fait apparaître deux ensembles X et Y situés de part et d'autre de l'opérateur. Ce token impose une relation très forte entre X et Y. En effet la structure syntaxique de X doit être de même niveau que la structure syntaxique de Y (ce qui veut dire que X et Y doivent être acceptés par un automate de même niveau). Par exemple dans notre phrase شأنه (n + pronom) ne peut être équivalent à نجنبه (V + pronom). Voir plus bas l'analyse de la phrase.

On voit apparaître dans le tableau, des GN ou des GND dans la colonne correspondant aux contraintes de positions lâches. Ces groupes occupent un ensemble d'occurrences dont la longueur n'est pas déterminée à l'avance. Pour cette raison nous sommes amenés à considérer les ajouts possibles au nom (ajouts à droite et ajouts à gauche).

A cette fin nous avons dessiné les diagrammes complémentaires au tableau principal qui englobent également les attentes induites par certains tokens que nous avons discutés.

LES DIAGRAMMES

Les diagrammes représentent des successions d'attentes. On passe d'un état d'attente représenté par un nœud à un nouvel état au moyen d'une transition qui est symbolisée par un arc. Ces arcs sont étiquetés soit par des symboles terminaux soit par des états. Dans ce dernier cas le processus de contrôle effectue un saut. Il empile l'état suivant sur lequel il pointe et fait appel à un sous-processus initialisé par l'état cité sur l'arc. Il peut éventuellement s'appeler lui-même comme dans le cas du diagramme IX, à la sortie de Q 4.

Ces diagrammes sont en fait des représentations de machines abstraites qui peuvent accepter (ou produire) des chaînes de mots qui appartiennent au langage. Ces machines possèdent un mécanisme d'entrée qui permet à la machine d'accéder à un seul symbole d'entrée à la fois, et un mécanisme de contrôle qui peut lire et tester cette entrée et provoquer le passage de la machine d'un état donné à un nouvel état. De plus, ces machines peuvent avoir une « mémoire » à laquelle on peut accéder par l'unité de contrôle. Ce sont les registres où l'on peut emmagasiner l'information et effectuer au besoin des tests. Une telle machine est aussi appelée automate abstrait, ou plus simplement automate. Nous verrons que la démarche mise en œuvre lors de l'analyse d'une phrase correspond bien au fonctionnement d'une telle machine.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Les règles illustrées par les diagrammes ne sont de toute évidence pas exhaustives. Elles se fondent également sur l'idée de *règle dominante* entraînant les attentes les plus probables. En d'autres termes, elles rendent compte de l'usage actuel le plus répandu. Il est toujours possible de trouver des exceptions du fait de la variété des niveaux de langue en cours aujourd'hui dans les pages d'un même journal. Pour vérifier nos intuitions nous nous sommes livrés à des sondages dans la presse et certains écrits littéraires. Une étude statistique plus poussée des fréquences sera utilisée lors de l'étude définitive de la compatibilité des diagrammes.

Commentaires sur les diagrammes

Les états marqués par une barre transversale sont ceux qui offrent une possibilité de sortie. S'ils ne sont suivis d'aucun arc, l'automate s'arrête nécessairement.

— Lorsque sur un arc un élément du type ذو, الذى apparaît, ce dernier représente en fait toute sa classe.

— L'étude des compatibilités entre I, II et III nous a amenés à prévoir des restrictions dans l'application des diagrammes. GND (— 10) représente par exemple le diagramme GND d'où l'on a supprimé l'arc numéro 10.

GND* = GN — { arcs contenant ART }

— Le diagramme III ne comporte pas d'état initial. Il doit être précédé d'un groupe nominal. Il représente les formes de relatives marquées ou non par un relatif de la classe de الذى. Dans le cas de la relative marquée on mentionne ce diagramme par Rel, dans le cas où la relative est non marquée, le diagramme est appelé Relind.

Ce diagramme accepte des phrases du type de celles qui figurent ci-dessous.

Devant chaque phrase figure une suite de nombres qui représente les arcs des diagrammes que l'automate doit emprunter pour l'accepter. Le nombre 1 est précédé d'un tiret qui représente les arcs non numérotés correspondant au segment précédant le pronom relatif.

— الرجل الذى خرج [—, 1]
— الرجل الذى خرج أمس [—, 1, 11]
— الرجل الذى ضرب الكلب [—, 1, 11]
— الرجل الذى ذهب الى المحطه [—, 1, 10, 27]
— الرجل الذى ذهبت اليه [—, 1, 8, 14]

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

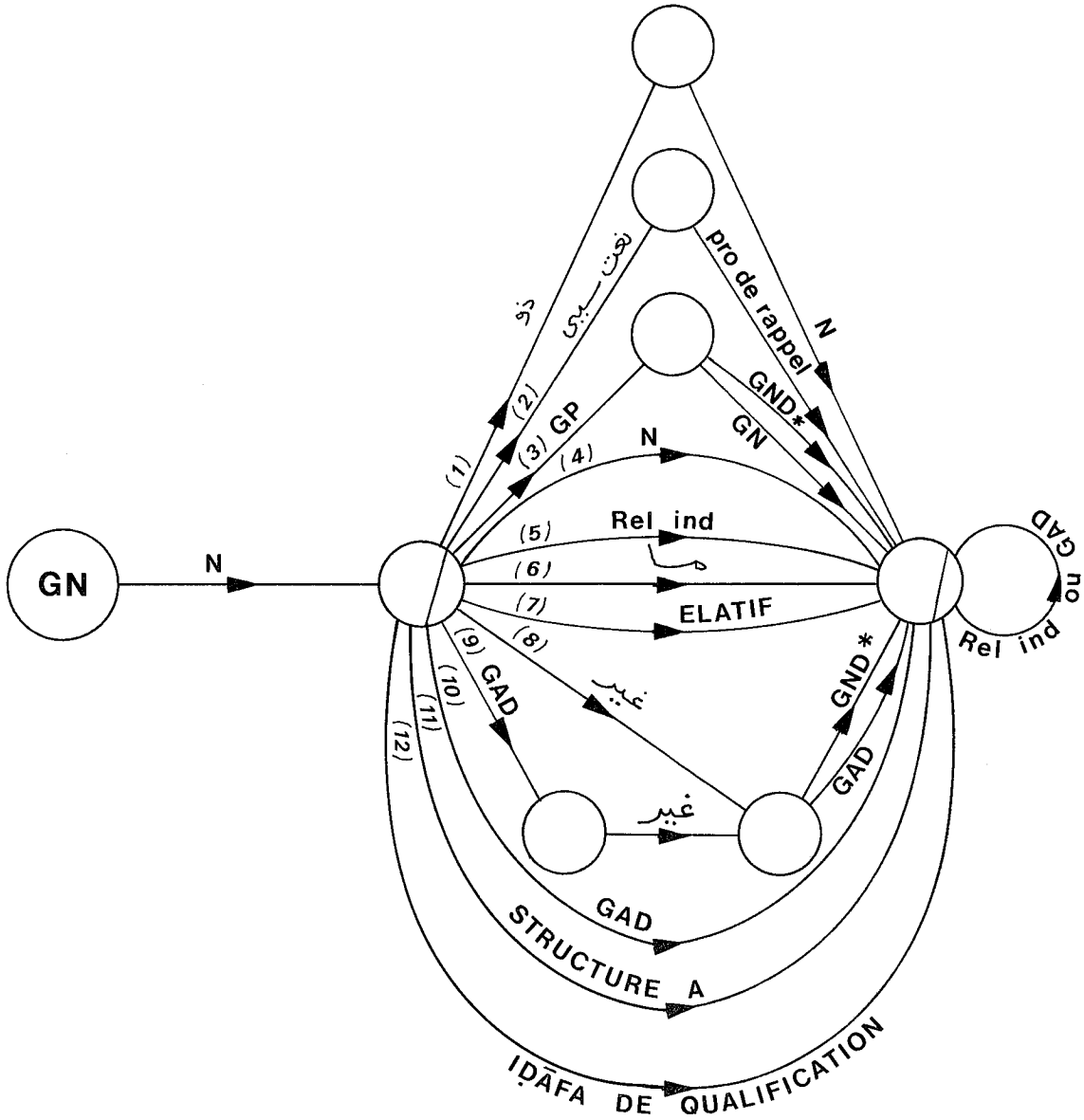


Diagramme I.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

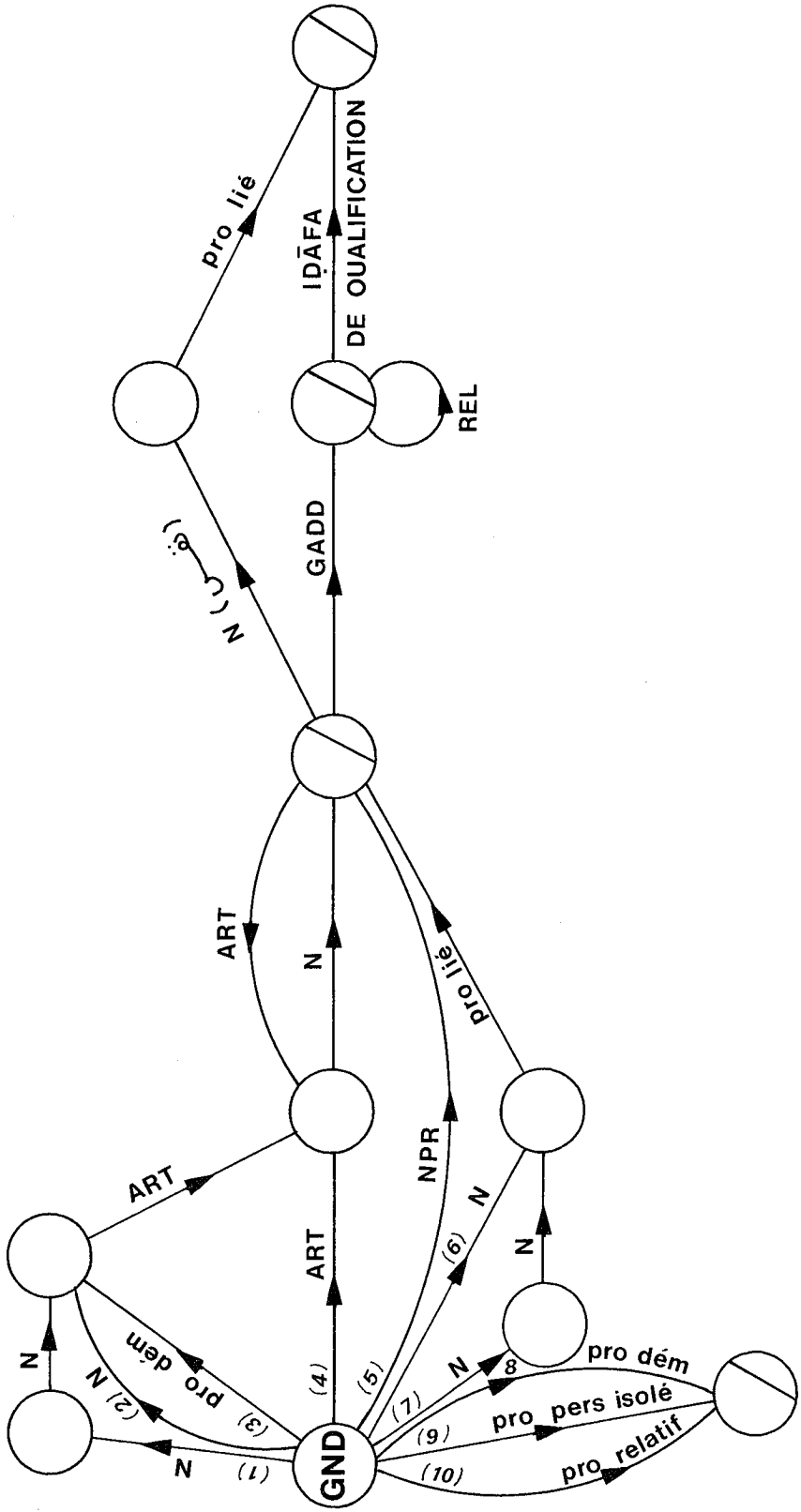


Diagramme II.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

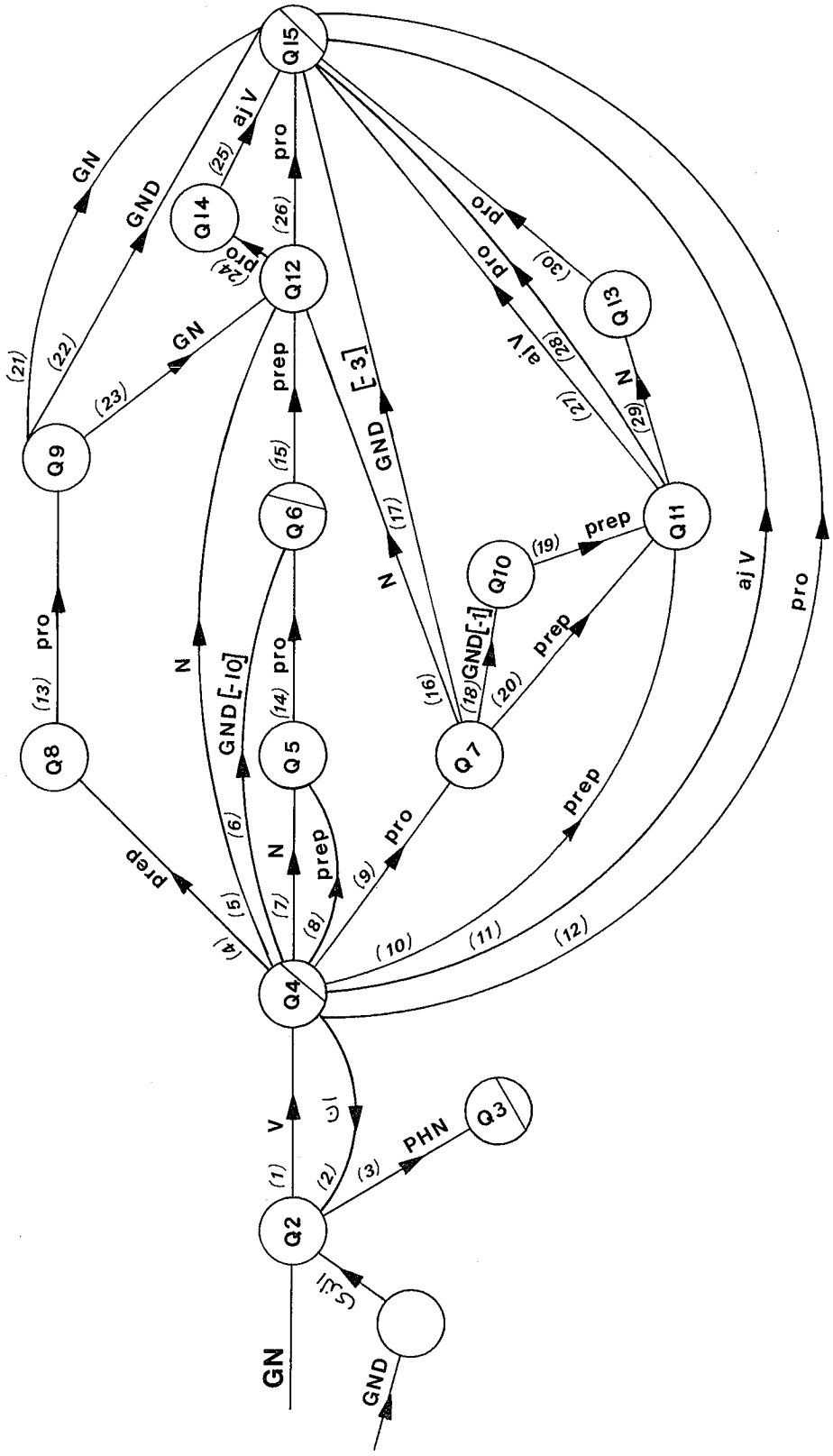


Diagramme III.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

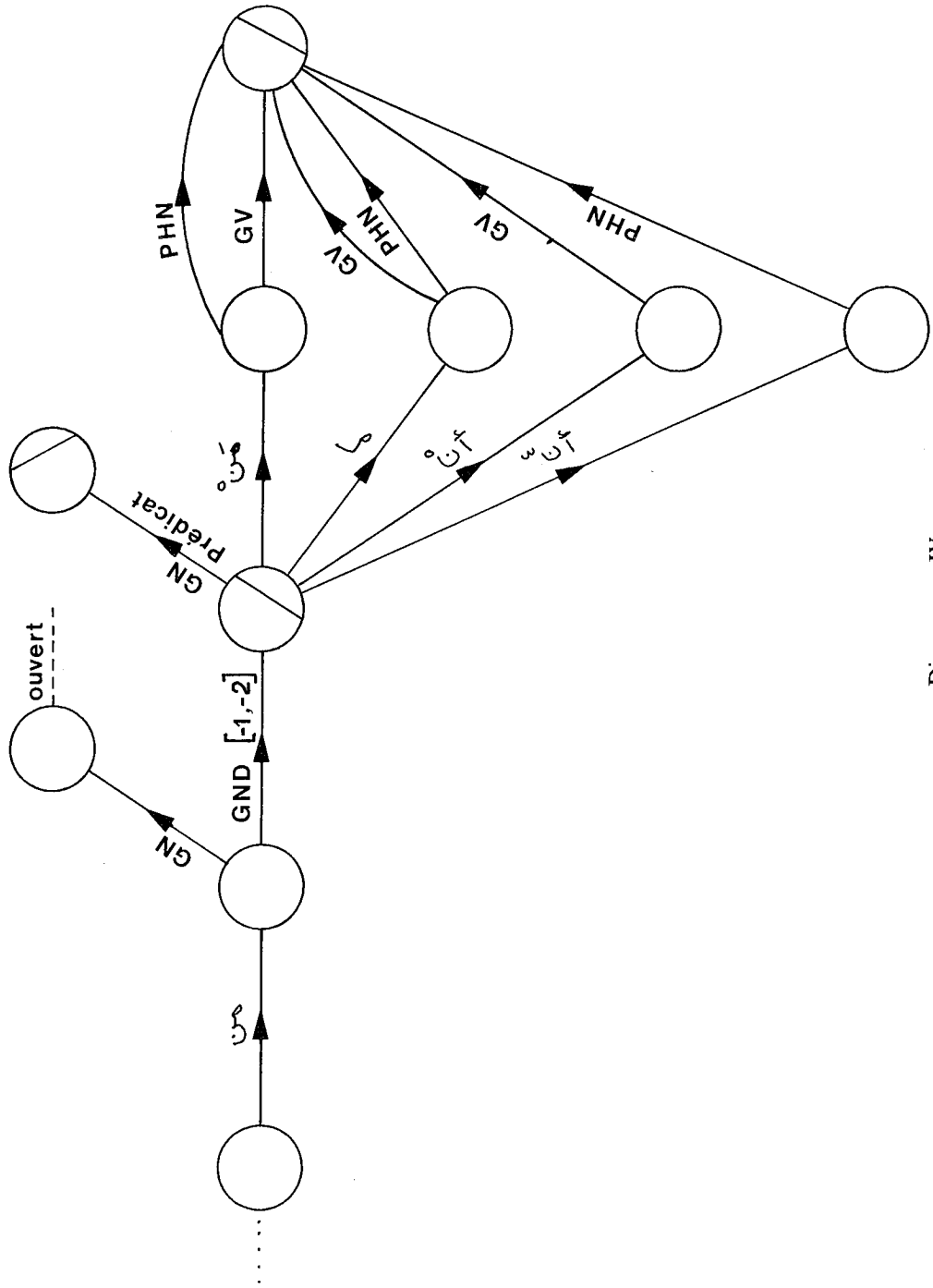


Diagramme IV.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

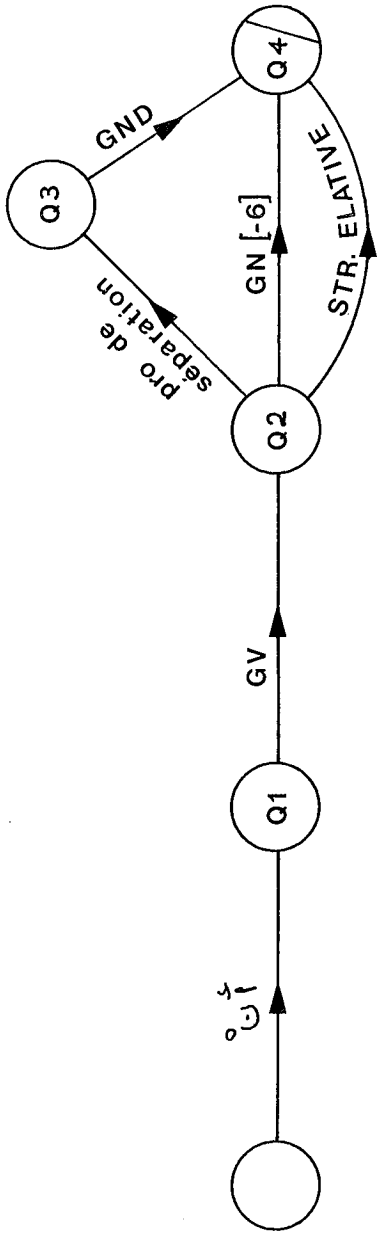


Diagramme V.

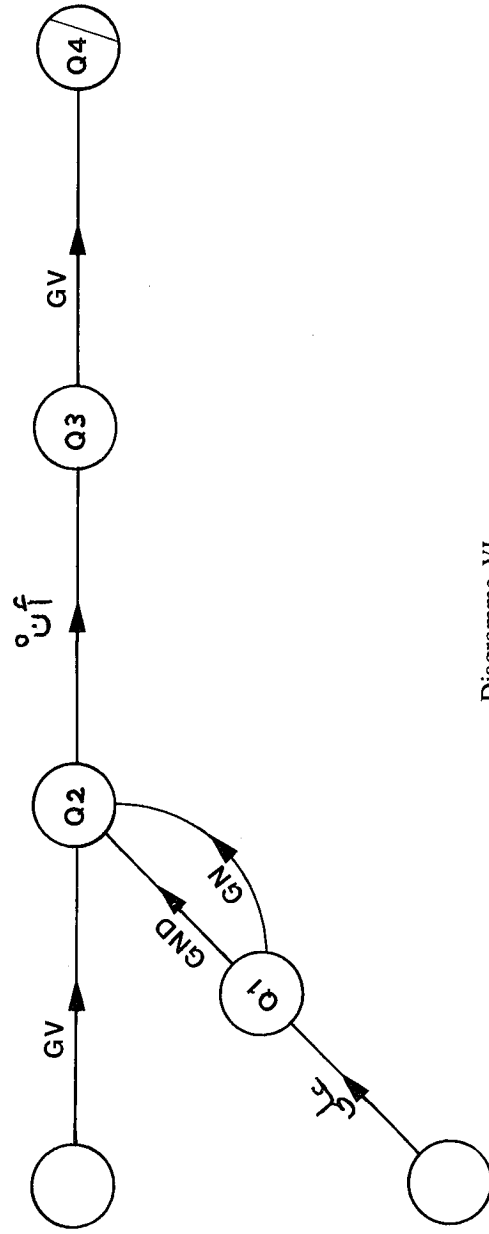


Diagramme VI.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

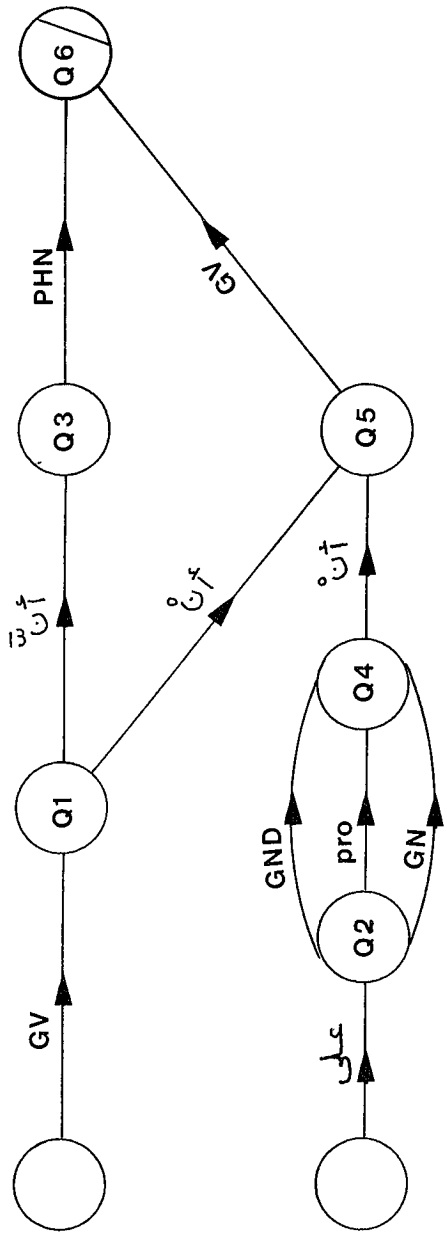


Diagramme VII.

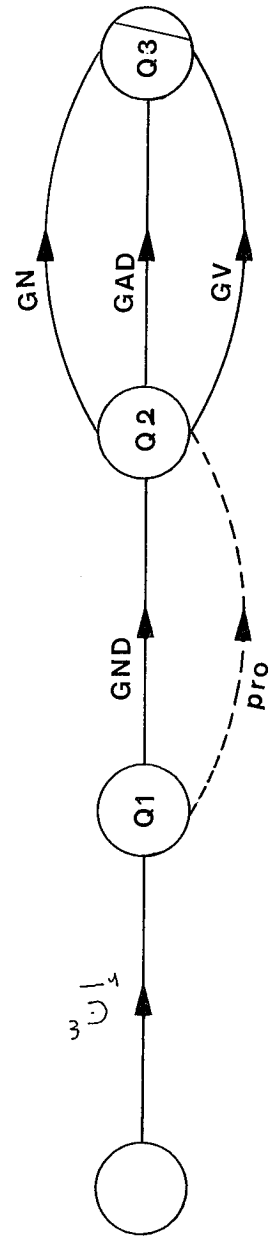


Diagramme VIII.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

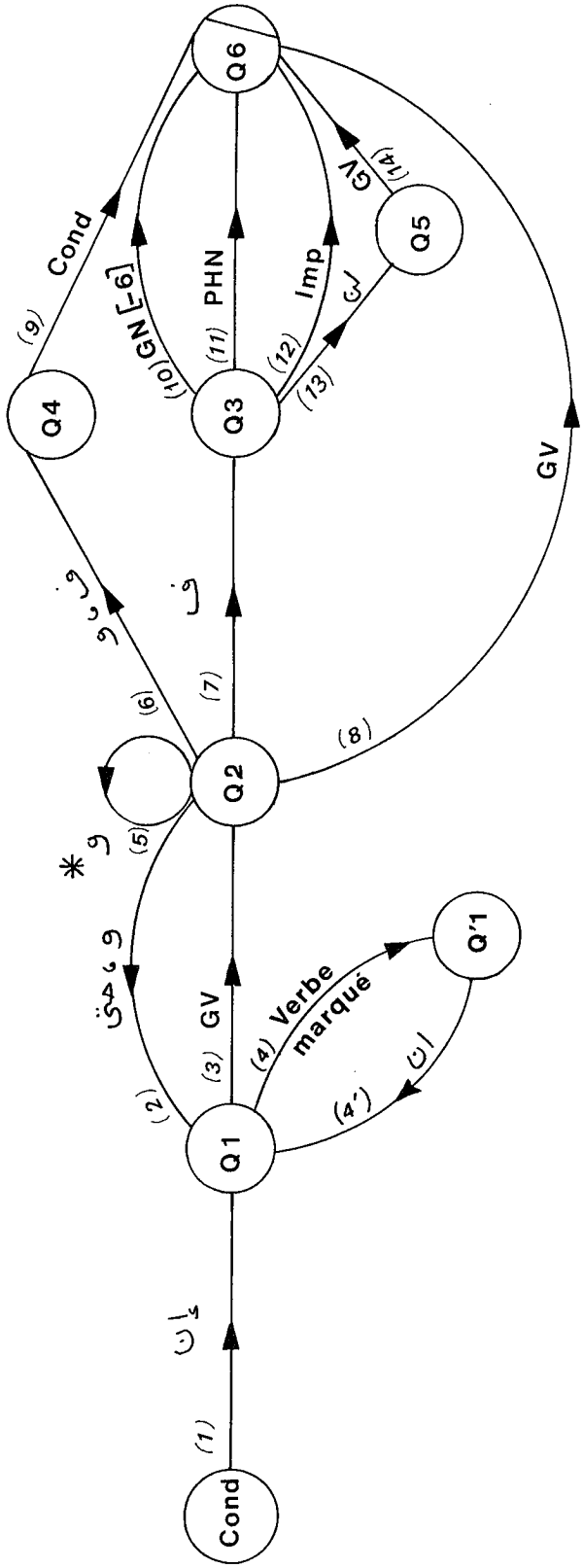


Diagramme IX.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

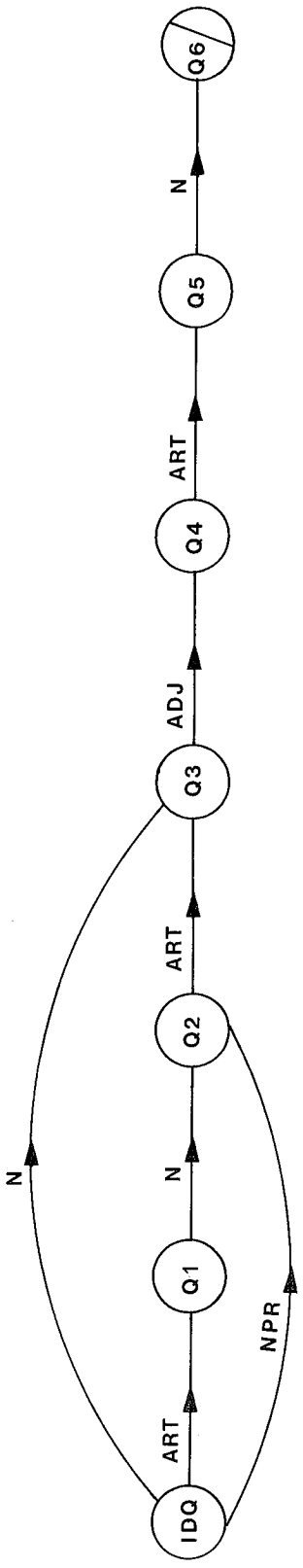


Diagramme X.

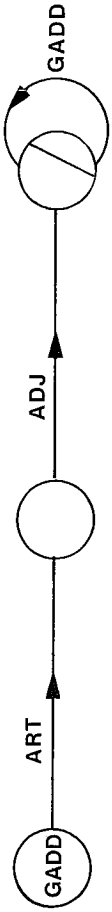




Diagramme XI.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

— الرجل الذى خرج أخوه [—, 1, 7, 14]
— الرجل الذى ضرب ابوه الكلب [—, 1, 5, 24, 25]
— الرجل الذى وصلت اليه الرسالة [—, 1, 4, 13, 22]
— الرجل الذى زارته أمه [—, 1, 9, $\underset{17}{16, 26}$]
— الرجل الذى ضربه أبوه [—, 1, 9, $\underset{17}{16, 26}$]
— الرجل الذى وصلت اليه أمه [1, 4, 13, $\underset{22}{23, 26}$]
— الرجل الذى ذهب أخوه إليه [—, 1, 7, 14, 15, 26]
— الرجل الذى ضربه الشرطى [—, 1, 9, 17]
— أعمال العنف التى قامت بها قلة [—, 1, 4, 13, 21]
— العناصر التى تم إلقاء القبض عليها [—, 1, 6, 15, 26]
— الإتفاقية التى ينبغى أن نعتبرها [—, 1, 2, 1, 12]

— En ce qui concerne le diagramme IV, nous devons noter que dans une étude plus détaillée, la nature du GND devrait être précisée afin de tenir compte de la structure très courante أنْ/أنّ . . . من الـ. En effet, seuls certains arcs du GND sont compatibles avec les arcs أنْ/أنّ

Ex. : من المهم أن . . . ، من المعروف أن . . .

— Le diagramme IX représente le إنْ conditionnel, il offre un exemple intéressant de récursivité. En effet l'arc de transition entre Q_4 et Q_6 (n° 8) est étiqueté par Cond c'est-à-dire que l'automate peut s'appeler lui-même (possibilité d'imbrication des conditionnelles). L'arc *و (n° 5) représente un automate que nous n'avons pas explicité car nous n'avons pas étudié les problèmes des coordonnants dans leur généralité. Il nous suffit de noter qu'il peut accepter une structure complète, l'important est de savoir à quel état on aboutit dans l'automate directeur après avoir balayé cette structure. L'automate Cond peut accepter par exemple une phrase du type :

إن قامت الساعة وفى يد أحدكم فسيلة فان استطاع ألا يقوم حتى يغرسها فليغرسها [1, 3, 5, 6, 9]
Le dernier arc 9 se décomposant en [1, 4, 4', 3, 7, 12].

La clôture de la sous-conditionnelle entraîne la clôture de la conditionnelle. L'élément ف introduisant la 2e partie de la conditionnelle apparaît dans le sous-automate.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

ABRÉVIATIONS. :

AjV	: Ajouts au verbe (compléments d'objet direct, adverbes ...)
Cond	: Conditionnelle
GAD	: Groupe Adjectival
GADD	: Groupe adjectival déterminé
N	: Nom
GN	: Groupe Nominal
GND	: Groupe Nominal déterminé
GP	: Groupe prépositionnel représentant un ensemble de prépositions
GV	: Groupe verbal
Imp	: Impératif
NPR	: Nom propre
PHN	: Phrase nominale
Pro	: Pronom
Prep	: Préposition

STRUCTURE A : Représente un groupe adverbial. Ex. : وحده / خارج البيب

STRUCTURE ÉLATIVE : N'a pas été figurée par un diagramme car elle suppose une étude des compatibilités qui reste à faire.

III

ANALYSE DE LA PHRASE

Après avoir représenté un noyau minimal des contraintes attachées aux opérateurs et établi les quelques règles d'ajouts nécessaires pour définir les possibilités et les modes d'extension des ensembles d'occurrences X et Y de notre tableau, il serait utile maintenant, pour fixer les idées, de nous représenter dans ses grandes lignes la démarche générale d'un analyseur syntaxique qui progresse mot à mot et qui correspond donc à l'opération de lecture. Nous rendrons mieux compte ainsi de l'évolution de la construction de l'arbre parallèlement au flux de surface, des différentes tentatives de constructions d'arbres partiels et des opérations de va-et-vient de l'analyseur qu'elles impliquent.

Soit la phrase :

ان الاتفاقية الاردنية الفلسطينية التى تستهدف السعى من اجل تسوية سلمية هى بمثابة تطور عظيم ينبغى الا نقلل من شأنه او نجنبه .

L'analyseur pointe sur la première occurrence. Il trouve le signe ان qui est ambigu puisqu'il correspond à 4 entrées principales dans le tableau. L'ensemble des contraintes attachées à chacune de ces entrées (relatives aux positions strictes, lâches, les types de relations qui régissent les ensembles d'occurrences X et Y qui à leur tour obéissent à des

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

règles d'expansion bien déterminées) ont été regroupées et exprimées par les diagrammes que nous représentons (pages 233 à 240) auxquels nous nous référons dorénavant systématiquement.

Nous nous trouvons donc en face de 4 diagrammes principaux. Parmi ces 4 diagrammes (qui, nous le rappelons, ne représentent que des regroupements de règles) trois seulement admettent ان comme état initial. Nous avons donc à ce niveau trois possibilités d'attente représentées par les diagrammes V, VIII, IX.

L'analyseur se met en position [2] et trouve الاتفاقية (que le programme d'analyse du mot reconnaît aux marques extérieures الـ et ة comme nom déterminé sans avoir à chercher la racine et la forme de cette occurrence. Cette reconnaissance permettra à l'analyseur d'écarter les successions d'attentes exprimées par les diagrammes V et IX qui font bien apparaître la nécessité d'un verbe au niveau de cette occurrence). Il retient donc les attentes exprimées par le diagramme VIII.

L'analyseur pourra donc revenir sur l'occurrence 1 et lever l'ambiguïté graphique de ان que l'on notera dorénavant إنّ.

إنّ *induit l'attente d'une phrase nominale* dont les principaux constituants sont GND suivi d'un GN ou d'un GV. Ce انّ introduit deux fonctions grammaticales thème et prédicat à répartir entre les deux groupes (entre GND et GN ou entre GND et GV).

Nous discuterons au cours de l'analyse la question du pronom de séparation.

Nous retenons donc l'arc ان ——→ GND puisque nous avons reconnu en l'occurrence 2 un nom déterminé qui initialise un groupe nominal déterminé dont les possibilités d'ajouts sont exprimées par le diagramme II. Nous aurons donc à vérifier si les occurrences qui suivent satisfont aux règles d'ajouts. Il faut noter toutefois à ce niveau — le programme morphologique opérant au minimum — que si la première marque extérieure ال a été reconnue, la deuxième (*tā' marbūṭa*) demeure ambiguë dans la mesure où elle peut indiquer le féminin ou être révélatrice d'un schème pluriel. Pour lever l'ambiguïté il est nécessaire d'avancer dans le groupe nominal.

A ce niveau l'analyseur a la possibilité de clore le groupe nominal déterminé en fonction du THÈME qu'il vient de reconnaître et de *s'orienter vers la recherche du PRÉDICAT,* ce qui correspondrait à un parcours minimal du diagramme VIII : aucun ajout au GND.

L'analyseur se positionne en [3]. Les marques ال et ة qu'il reconnaît tout de suite (grâce au programme de décomposition du mot) lui permettent *d'écarter l'attente du prédicat* puisque l'on tient compte de la règle dominante.

THÈME	PRÉDICAT
GND	GN indéterminé

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Sans analyser complètement الاردنية il peut accepter cette entrée comme ajout au GND (diagramme II).

Il y a lieu de noter ici que la probabilité du ة = marque du féminin singulier augmente du fait de la succession de deux marques terminales ة, ة.

En effet ⁝FEM entraîne nécessairement ⁝FEM dans la perspective des ajouts alors qu'une forme pluriel en ة. n'entraîne pas nécessairement une forme pluriel en ة

⁝FEM ⟶ ⁝FEM
⁝PLURIEL —/⟶ ⁝PLURIEL

En [4] l'analyseur effectue la même analyse qu'en [3]. La probabilité du féminin augmente.

En [5] *l'analyseur reconnaît* التى : pronom relatif, féminin et singulier, ce qui lui permet : 1) de clore l'automate groupe adjectival déterminé (GADD); 2) de lever définitivement l'ambiguïté féminin-pluriel; il place donc les occurrences [2] [3] [4] dans un registre avec les définitions GND = THÈME FÉMININ SINGULIER. Il sera en mesure de construire l'arbre partiel (Fig. 1); 3) *de différer la Recherche du prédicat* ce qui revient à mettre dans une pile de mémoire l'état Q_3 du diagramme VIII. En effet التى agit comme un indicateur d'appel d'un sous-automate qui doit être rattaché à une structure précédemment balayée.

Les attentes induites par التى sont partiellement représentées par le diagramme III.

Ce diagramme indique que التى doit être précédé par un GND auquel il se rapporte et qu'il introduit l'attente d'un V ou d'une PHN.

L'analyseur se positionne donc en [6], تستهدف

On peut faire appel ici au processeur MORPHOLOGIQUE qui lui indique que تستهدف est bien un verbe. Il emprunte donc l'arc (1) V qui le conduit à l'état Q_4 (du diagramme III) qui lui offre la possibilité soit de sortir de la structure relative et de ramener l'état Q_3 (du diagramme VIII) de l'automate directeur au niveau courant (on désempile), c'est-à-dire de se remettre à la recherche du prédicat, soit de compléter la structure relative.

L'analyseur avance d'un cran, il se trouve en position [7].

Le mot est السعى. On peut noter que le programme morphologique ne peut pas lever l'ambiguïté portant sur le signe extérieur ال [1] qui est cependant réduite par les attentes attachées à cette occurrence). En effet d'une part l'automate directeur (THÈME-PRÉDICAT) exclut la présence d'une forme interrogative d'autre part l'automate III n'offre

[1] Celui-ci peut en effet être interprété comme أ interrogatif, et لَ préposition ou أ interrogatif et لِ de corroboration etc...

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

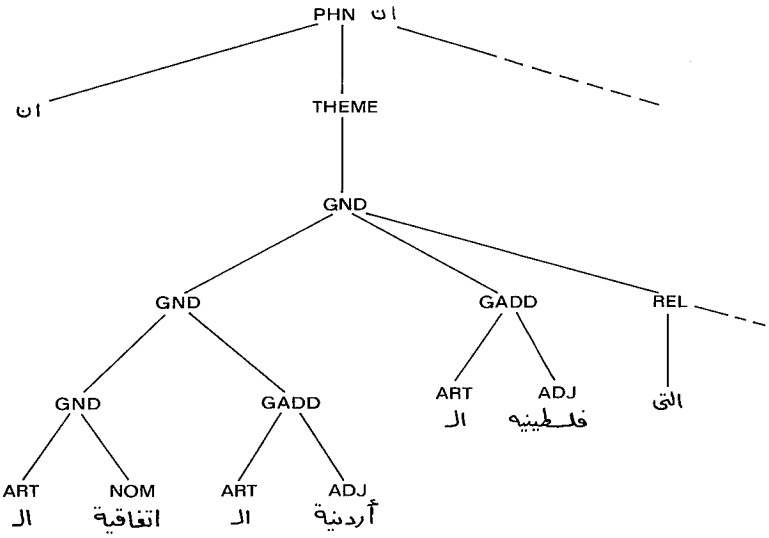


Fig. 1.

comme possibilités de transition que deux arcs GND (6) et AjV (11). L'occurrence [7] est donc reconnue avec un minimum d'analyse morphologique comme un nom déterminé de la forme الفعل .

L'automate directeur exclut en règle dominante la possibilité de retenir cette occurrence comme PRÉDICAT.

L'analyseur gardera donc *l'attente du PRÉDICAT en suspens* et complètera la structure relative. En ce point l'analyseur peut emprunter deux arcs, soit GND (6), soit AjV (11). Dans le premier cas, il est amené à l'état Q_6 et attend donc une préposition qu'il trouve bien à l'occurrence [8]. Pour passer à l'état final, il doit transiter par un pronom. En [9] le mot est اجل qu'il ne peut accepter. Il revient donc à l'état Q_4 et *emprunte l'arc AjV* (11) *qui provoque l'appel d'un nouveau graphe décrivant les ajouts possibles au verbe.* Or le verbe peut être complété par un GND dont le premier élément vient d'être reconnu (occurrence 7).

Nous nous situons donc à présent dans un *automate de plus bas niveau GND* décrit par le diagramme II. (*L'état final de l'automate AjV est mis en pile* en attendant d'arriver à un état de sortie du sous-automate GND). Or cet automate permet une sortie après ART NOM. Nous avons donc ici la possibilité de clore le GND, ce qui entraîne à son tour la clôture du AjV qui le gouverne. L'analyseur peut donc *désempiler l'état correspondant au prédicat,* c'est-à-dire se remettre en attente du prédicat.

A l'occurrence suivante, l'automate rencontre en [8] من qui est un signe ambigu.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Il existe en effet dans le tableau deux entrées principales correspondant aux éléments مِن et مَن (voir entre autres le diagramme IV).

Il écarte les entrées correspondant à مَن en début de phrase. Il restera donc les possibilités suivantes : °مَن suivi d'un verbe en règle dominante ou °مِن suivi de GND, GN, d'un pronom lié ou d'un pronom relatif. Il doit donc avancer d'un cran et se trouve en présence de اجل qu'il peut analyser soit comme un verbe, soit comme un nom indéterminé ce qui ne lui permet pas de lever l'ambiguïté entre مِن et °مَن. Cependant dans le cas de مَن + $\sqrt{}$ il trouve dans la colonne (n — 1) correspondant à la contrainte de position que cette occurrence doit être différente d'un GND, ce qu'il vérifie en retournant à l'occurrence numéro 7 = السعى = GND. Il opte donc pour مِن préposition et analyse [9] اجل comme un nom qui initialise la structure d'un groupe nominal [1]. Ce groupe peut être un GN ou un GND. Voir diagrammes I et II [2]. *En clôturant AjV après* السعى — *ce qui correspond à l'une des analyses possibles, l'autre étant de continuer AjV — nous avons du même coup désempilé l'état* Q_3 *de l'automate directeur* qui se retrouve donc *dans l'attente d'un GN* en fonction de *prédicat* (règle dominante). Or cette attente est corroborée par le résultat de l'analyse de من اجل qui nous a fourni la définition اجل = NOM que l'analyseur après avoir écarté GND, comprendra comme le premier élément de GN. Une fois ce choix effectué l'analyseur aura intérêt à étendre le GN jusqu'à la fin de la phrase pour sauver la structure noyau THÈME ET PRÉDICAT. Il fera donc appel au diagramme numéro I et contrôlera la bonne formation du GN. Aucun problème ne se pose jusqu'en [12] où l'analyseur reconnaît un token هى. La consultation du tableau lui indique la *possibilité* d'emprunter *un arc correspondant à une structure relative à ce niveau.* Il appelle donc récursivement le graphe représenté par le diagramme numéro III, Relind (relative indéterminée).

Bref, la phrase entière est acceptée jusqu'à l'occurrence [22] et l'arbre produit est le suivant : إن الاتفاقيه من أجل (Fig. 2).

L'analyseur cependant ne peut s'arrêter là, puisqu'il avait laissé en suspens au niveau des occurrences [7] [8] *la possibilité d'accepter les occurrences suivantes comme complément*

[1] Il est à remarquer qu'on aurait eu intérêt ici à reconnaître من اجل comme une expression figée qui gouverne un GN. Ceci reviendrait à prévoir une ligne supplémentaire correspondant à une classe d'expressions de ce type.

[2] En effet il existe dans le diagramme GND des arcs initiaux correspondant à N. Cependant aucun chemin induit par ces arcs ne pourra conduire l'automate à un état terminal comme le lecteur pourra s'en rendre compte aisément. Il retiendra donc la structure GN (diagramme III).

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

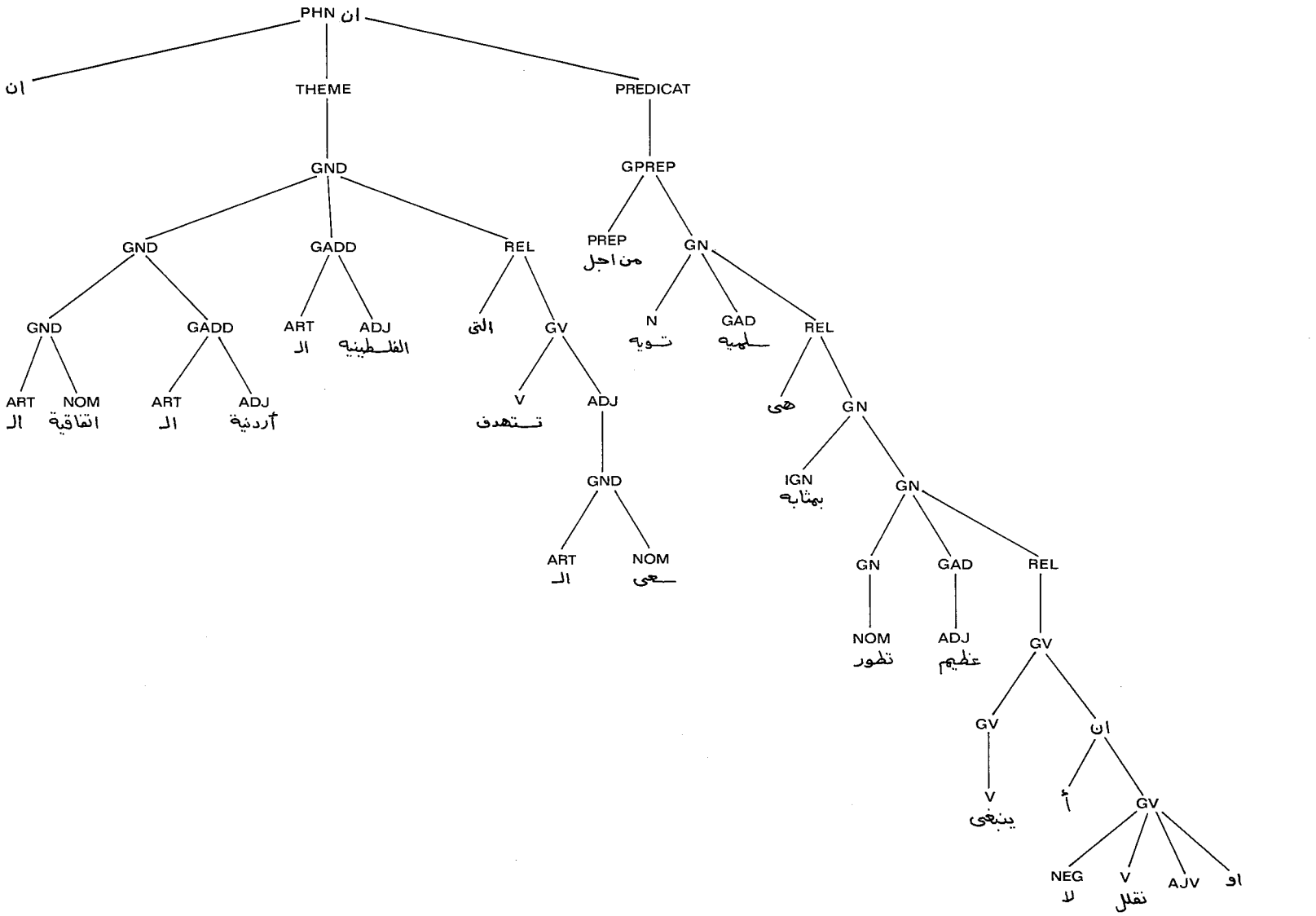


Fig. 2.

à AjV. Il revient donc à l'occurrence [7], rappelle le sous-automate AjV, *remet donc l'état Q_3 de l'automate directeur en pile* et essaye d'accepter ce qui suit conformément aux attentes exprimées par AjV. Nous rappelons que l'analyse de من اجل a déjà été faite dans le cadre de la tentative précédente et que les résultats locaux (ceux qui touchent aussi bien la morphologie que les contraintes de liaisons locales), ne sont pas affectés par le nouveau choix effectué par l'analyseur. Les définitions sont gardées en mémoire. Nous savons donc que اجل = N et initialise un GN. Toutefois un problème de segmentation apparaîtra à ce niveau. En évitant d'entrer dans le détail de l'analyse, nous signalons les possibilités suivantes :

GN1 = من اجل / من فعْل est considéré comme un groupe nominal en position d'ajout au verbe et = GN2 ... تسوية سلمية comme un groupe nominal en fonction de prédicat (GN1 = ajout au verbe; GN2 = prédicat) ce qui fournit l'analyse : الاتفاقية تسوية

ou bien :

GN1 = من اجل تسوية en position d'ajout au verbe

GN2 = سلمية en fonction du prédicat

(GN1 = ajout au verbe; GN2 = prédicat) ce qui fournit l'analyse الاتفاقية سلمية

Toutefois en testant les attentes attachées à chacune de ces deux possibilités l'analyseur se bloquera, car, arrivé au niveau de هى تطور il analysera ces deux occurrences comme un GND qu'il sera obligé de considérer comme ajout à un GN précédent — (le thème ayant été déjà initialisé) ce qui est rigoureusement impossible. L'analyseur est donc en mesure d'écarter ces deux dernières possibilités. — Il essaiera donc de compléter GN jusqu'en [11], ce qui ne pose aucune difficulté.

Arrivé en [12] = هى *l'analyseur peut retenir la définition pronom de séparation entre THÈME et PRÉDICAT*. Il clôt donc le 1er GND = thème, *désempile* le Q_3 de l'automate directeur et se met dans l'*attente du prédicat*.

En [13] il trouve بمثابة , mot vide qui introduit l'attente d'un nom.

Il arrive à l'occurrence [14] تطور qui initialise un automate GN (diagramme I); ce GN lui permet d'accepter l'occurrence suivante [15] عظيم comme ajout. A l'occurrence [16] il se trouve en présence de ينبغى qu'il peut analyser sans difficulté en appelant le programme morphologique comme un verbe. Or l'état terminal de l'automate GN lui offre *la possibilité d'une transition Relind* qui est représentée sur le diagramme I par une

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

boucle [1]. Le diagramme III de la relative accepte, après la première transition qui l'amène en Q_2, deux transitions possibles : PHN (pour Phrase Nominale) (arc 3) et un arc V (arc 1) qui correspond bien à notre situation. L'analyseur explorera donc l'arc Relind du GN précédent qui a été initialisé en [14] (تطور).

Avançant d'une position l'analyseur se trouvera en présence d'un الا [17] qui ne peut en tant que tel être accepté par l'ensemble des arcs suivants. Mais comme il s'agit d'un token ambigu, il doit faire appel au tableau des structurants et tâcher de lever l'ambiguïté en testant les compatibilités avec l'ensemble des arcs précédents et suivant l'état dans lequel il se trouve (et qui est l'état Q_4 du diagramme III). Or une seule option parmi les quatre possibilités d'analyse de الا lui permet de rester dans l'automate Relind et le ramène à l'état Q_2 : il s'agit de الا = لا أن . Il emprunte donc l'arc ان (2) qui lui permet de progresser d'une occurrence dans le texte. Il se positionne en [18] qui est bien un verbe نقلل compatible avec l'attente suivante (transition $Q_2 \rightarrow Q_4$). L'analyseur avance d'un cran, il se trouve en [19] en présence de من qui est un token ambigu. Il fait donc appel au tableau des structurants et teste toutes les possibilités compatibles avec les attentes de l'état du diagramme III. Il choisira PREP suivi des arcs N et pronom. Une seule possibilité sera donc retenue, celle qui correspond à la définition : من préposition induisant une attente de groupe nominal. L'analyseur avance d'une occurrence; il se trouve en [20] en présence de شأنه qu'il décompose en شأن et o : N + pronom. Cette reconnaissance lui permet de passer en Q_{13} et d'emprunter l'arc pronom de rappel qui l'amène en Q_{15} (état terminal de la relative) (10, 29, 30).

Ayant clôturé le sous-automate Relind, il désempile l'état terminal du GN. Comme la phrase continue, il est obligé de réemprunter l'arc Relind (*choix que le token* او *de l'occurrence 21 lui permet de faire*) et de retester, à partir de [22], les attentes exprimées par le diagramme III.

En [22] il analyse نجنبه comme V + pronom. Il emprunte donc l'arc V qui le fait passer en Q_4 et après avoir testé les possibilités attachées aux différents arcs, il emprunte l'arc Pro et arrive en Q_{15} qui lui permet de sortir du sous-automate Relind et de clore du même coup l'automate directeur GN qui se situait dans le registre prédicat. L'analyse fournit l'arbre représenté sur fig. 3.

[1] Cet arc est récursif puisqu'il est étiqueté par un état et (non par un symbole terminal). Il pointe donc sur un sous-automate que nous avons esquissé dans le diagramme III. Nous avons essayé dans ce graphe de capter les régularités qui apparaissent lorsqu'on examine plusieurs cas de relatives indéterminées (plusieurs exemples sont fournis pour illustrer le diagramme). Voir plus haut.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

Anlsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 Anlsl en ligne https://www.ifao.egnet.net

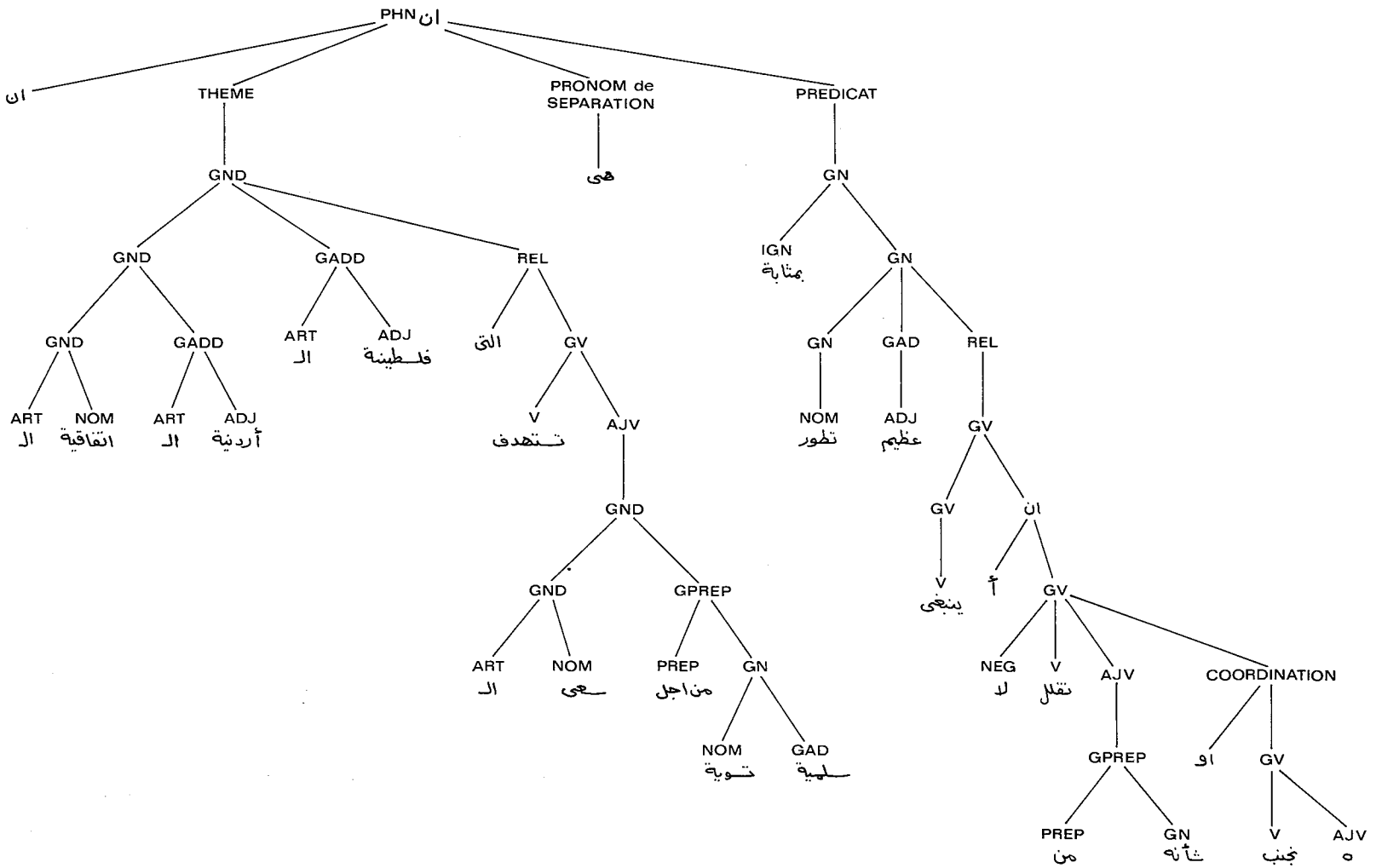


Fig. 3.

On voit que cet analyseur est de nature non déterministe puisqu'il fournit deux analyses possibles :

(1) إن الاتفاقية من أجل . . .

et (2) إن الاتفاقية تطور . . .

Dans le cas de la deuxième analyse, c'est le token هى qui a marqué le début de la recherche du prédicat, en vertu de la règle dominante située au niveau de l'automate directeur qui veut que l'on ait GND هى GND ou GND étendu هى GN.

Dans la première analyse qui est fautive, هى apparaît dans un sous-automate (Relind). L'analyseur privilégie celle où هى apparaît au niveau de l'automate directeur. La raison pour laquelle cette décision n'a pas été prise tout de suite est la suivante : l'analyseur tend à retrouver l'automate directeur au plus tôt. L'état Q_3 de cet automate directeur qui avait été mis en pile afin de pouvoir achever l'analyse des constituants du thème a été désempilé aussitôt que la possibilité s'en est offerte, à savoir, après السعى. Tout le reste de la phrase a donc automatiquement été mis dans le registre prédicat, dans ce cas هى ne pouvait être analysé que comme pronom introduisant une relative non marquée.

REMARQUES :

1 — Le diagramme III ne doit pas être considéré comme l'expression d'une règle générale. En effet, il pourrait donner à penser que chaque fois que l'on a N + V = Relind. Or ceci peut être inexact notamment dans le cas où N + V peut être interprété comme THÈME ET PRÉDICAT. Toutefois cette analyse est impossible dans notre exemple, le thème ayant déjà été repéré.

2 — Il faut noter que les diagrammes exprimés ci-dessus ne tiennent pas compte des cas $S_1 = S_2$ et $S_1 \neq S_2$ qui méritent d'être considérés dans les cas de relatives particulièrement des non marquées car ils sont en relation avec la présence ou l'absence du pronom de rappel.

Une ambiguïté secondaire subsiste quant au sujet du verbe ينبغى qu'on pourrait identifier à تطور. Dans le cas réel cette affectation est écartée par la sémantique dont on ne tient pas compte ici.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

IV

ANALYSE DE LA PHRASE VIDE

Nous avons simulé l'évolution d'un analyseur syntaxique procédant mot à mot. Cet analyseur est commandé essentiellement par les règles attachées aux opérateurs (lesquels, nous le rappelons, ont été choisis en fonction du programme morphologique). Nous avons figuré ces règles sous forme de graphes pour mieux illustrer les possibilités de parcours et mieux rendre compte de la construction parallèlement au flux de lecture.

Ces graphes en fait induisent une méthode d'analyse qui n'est pas la seule possible. On aurait pu par exemple écrire ces règles sous la forme :

GND ⟶ ART N
GND ⟶ GND ADD
ADD ⟶ Art Adj

et l'analyseur aurait pu systématiquement tester toutes les possibilités de parenthésage attachées à chaque règle ce qui serait revenu à tenter de construire l'arbre syntaxique à partir du bas en examinant de proche en proche toutes les combinaisons permises par les règles.

Notre méthode d'analyse, par contre, tente de repérer d'abord une structure (c'est-à-dire un arbre ou un ensemble d'arbres) par un élément révélateur qui apparaît dans le texte. On essaye ensuite d'appliquer cet arbre au flux de symboles terminaux. C'est une méthode descendante. C'est en ce sens que nous avons parlé d'un programme guidé par la syntaxe.

Les diagrammes décrits, qui ne sont que les organigrammes du moniteur syntaxique, ont pour rôle d'agir alors comme structure de contrôle. On voit là que les oppositions SYNTAXE-MORPHOLOGIE ne font que refléter celles beaucoup plus générales des analyses de reconnaissance montantes et descendantes et les problèmes d'optimisation qui leur sont attachés.

Les structurants ou opérateurs jouent le rôle d'indicateurs d'arbres. Il faut donc rechercher l'indicateur de l'arbre qui a le plus de chance de coiffer l'ensemble de la phrase (ou le maximum de flux), ensuite les arbres de moindre portée.

Pour tester la pertinence de notre hypothèse qui a consisté à considérer les tokens morphologiques comme structurants syntaxiques, on peut reprendre l'analyse de notre phrase en la vidant de toutes ses occurrences morphologiques, ce qui permettra d'annuler les informations morphologiques et les contraintes de liaison locale (celles qui sont liées à un processus d'analyse montante).

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

L'analyse qui suit a pour but de mesurer l'information apportée par les tokens et donc de justifier notre tableau. On notera qu'elle ne correspond pas nécessairement à la démarche précédente où l'on ne se privait pas systématiquement du recours au programme morphologique.

Soit la phrase vide :

$\overline{22}$ او $\overline{20}$ من $\overline{18}$ الا $\overline{16}$ $\overline{15}$ $\overline{14}$ $\overline{13}$ هى $\overline{11}$ $\overline{10}$ من أجل $\overline{7}$ $\overline{6}$ التى $\overline{4}$ $\overline{3}$ $\overline{2}$ ان

Si l'on cherche le token qui implique le plus de contraintes de positions fixes, التى apparaîtra comme le candidat privilégié. Bien que ان soit un plus fort structurant au niveau du noyau prédicatif, il est à écarter du fait de son ambiguïté graphique en l'absence de recours aux informations morphologiques.

التى doit être précédé d'un GND qui doit être logé entre les occurrences [2], [3], [4], et suivi, en règle dominante, d'un V. Il impose en outre la recherche d'un noyau prédicatif dont il dépend. D'où les hypothèses suivantes en tenant compte du fait de la contrainte de ان. Ceci revient à déterminer la position du noyau prédicatif en fonction de التى.

D'où les hypothèses suivantes :

I — GND est initialisé en [4], il reste les occurrences [2] et [3] où on peut loger en [2] un V. L'opérateur est donc إنْ ou أنْ ; à ces deux hypothèses peuvent correspondre les attentes représentées par les diagrammes IX et V.

Commençons par le cas IX (إنْ conditionnelle). Ce diagramme montre qu'après le premier GV l'automate ne peut continuer que par un autre GV dans le cas où il ne trouve pas de ف dans la phrase, ce qui correspond à notre situation. Ce deuxième √—, s'il n'est pas en [3], ne peut se loger qu'en [7], [11], [14], [15], [16] [1] car l'occurrence [4] est prise par un GND [2], [6] est prise par un verbe dépendant de la relative, l'occurrence [10] par un GN, la [13] éventuellement par une relative indéterminée initialisée par هى. Les occurrences [18] et suivantes sont à écarter car incompatibles avec le diagramme IX [3].

2e *cas de* أنْ introduisant une *phrase nominale,* diagramme V : le prédicat peut se situer en [7], [11], [14], [15], [16], en effet [13] est à écarter car incompatible avec le GND

[1] Donc dans les occurrences 7, 11 14, 15 16 il ira chercher un V (le recours au programme morphologique est ici minimum). Cette recherche n'est effectuée que dans le cas où il existe un V en 2.

[2] Voir diagramme III.

[3] Il ressort de ceci que l'on peut utiliser cette analyse en vue d'une optimisation du décodage. En effet l'analyseur dirigera le programme morphologique vers les occurrences 7, 11, 14, 15, 16, avec la tâche précise de rechercher un verbe. Toutefois tel n'est pas le but de la démarche décrite dans cette partie.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

nécessaire en [4], [6] est occupé par la relative, [10] par le GN dépendant de من أجل, [13] par une relative possible, quant aux occurrences qui suivent, elles sont incompatibles avec la structure du prédicat.

II — *Si GND est initialisé en* [3] a) dans l'hypothèse de la conditionnelle, le deuxième verbe ne peut se situer qu'en [7], [11], [14], [15], [16] comme précédemment; b) dans l'hypothèse de أنْ la recherche du prédicat se fera en [7], [11], [14], [15], [16].

III — *Si GND est initialisé en* [2], ceci impose إن (car les diagrammes IX et V sont inapplicables) et conduit à rechercher le prédicat soit en [3], [4], [7], [8] ou [11] ou immédiatement après هى ([13] ou [14]); le هى apparaît comme un très fort indicateur de structure, notamment dans le cas de phrases longues où apparaissent des relatifs du type التى qui impliquent des expansions de constituants. Le pronom personnel isolé apparaît alors en règle dominante comme un séparateur nécessaire entre le thème et le prédicat. En termes de programme cette option sera testée en premier car l'analyseur aura en connaissance le nombre d'occurrences de la phrase, entre le premier signe ان et هى.

Nous avons simplement voulu donner une idée des informations que l'on peut tirer après un premier balayage et qui peuvent orienter fortement l'analyseur vers la recherche d'une structure. Cette démarche représente une optimisation en ce sens que nous privilégions les attentes les plus probables. Toutefois nous devons signaler que dans le cas de notre phrase, la méthode qui consiste à se priver de tout recours à la morphologie apporte d'assez maigres résultats. Le lecteur pourra néanmoins vérifier par lui-même qu'il n'en est pas toujours ainsi. Nous en donnons pour preuve la phrase suivante sur laquelle il pourra appliquer sa perspicacité [1] :

. — . — فى — لها 14 لا 12 10 9 التى 7 ان 5 التى 3 هذه 1

V

CONCLUSION

L'idée directrice de ce travail a été de montrer la possibilité de décoder un texte arabe avec un minimum de connaissances syntaxiques et morphologiques. La notion de structurants repérables nous a amenés à esquisser les attentes attachées à un certain nombre

[1] وحاولت هذه المصادر التى تؤكد ان التدريبات التى تجرى الآن لا علاقة لها بالتطورات الأخيرة فى العلاقات الامريكية الليبية .... (الاهرام).

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

d'entre eux. L'analyse de ces attentes et la tentative de les classer par ordre d'importance en tenant compte des critères de globalité nous a conduits à les représenter sous forme d'automates abstraits. En ayant recours à ce formalisme, nous avons essayé ensuite lors de l'analyse d'une phrase réelle de faire ressortir la nature du processus de décodage idéal.

Cette première étape nous semble nécessaire avant d'aborder le problème de la conception d'un outil d'EAO proprement dit. Ce dernier devrait s'appuyer sur une banque d'automates hiérarchisés, que l'on pourrait appeler en cours de décodage.

Toutefois ces automates peuvent être appréhendés de manière différente selon le degré de connaissance de l'utilisateur. Nous avons cherché à faire ressortir l'importance de la notion de stratégie de décodage. A chaque état de connaissance correspond en effet une stratégie appropriée. Le lecteur a pu se rendre compte par exemple que la stratégie n'était pas la même lorsqu'on faisait travailler l'analyseur sur une phrase vide ou lorsque ce dernier travaillait en ayant recours de temps à autre au programme morphologique. Entre deux cas extrêmes : celui de la phrase vide et celui de la phrase entièrement résolue morphologiquement on peut concevoir tous les états intermédiaires possibles qui correspondraient à autant d'états de connaissance de l'utilisateur. La mise en œuvre des automates de contrôle dans chacun des cas est différente. Dès lors, on voit clairement la nécessité de mettre au point une bonne description de la grammaire sous forme de diagrammes qui sont en fait des graphes récursifs que l'on peut éventuellement augmenter par des « actions » (ou programmes) pour rendre compte par exemple des aspects de la langue qui échappent à une description que produirait une grammaire indépendante du contexte. En effet un graphe récursif est formellement équivalent à une telle grammaire.

Enfin nous voulons signaler que c'est à partir d'une description minimum de ces automates, dont le présent article ne vise qu'à donner une idée, que l'on pourra envisager sérieusement un analyseur d'erreurs.

Il nous reste donc à :

1) parfaire la description des attentes;
2) classer les automates.

Par contre ce qui a déjà été réalisé est un programme morphologique qu'on a cherché à optimiser par un moniteur syntaxique qui permet l'accès au dictionnaire et permet de dégager le noyau de la phrase et la détermination des ajouts (Pour dégager le noyau d'une phrase on n'a qu'à parcourir les chemins minimaux dans les automates).

Il faut noter que le programme fournira un grand nombre de définitions en recourant le moins possible au dictionnaire. Nous avons en effet été guidé tout au long de notre

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḥabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net

travail par le principe du dictionnaire vide ce qui revient à raisonner sur des phrases où l'on a remplacé toutes les occurrences réelles par celles de leurs schèmes non vocalisés et qui peuvent donc être ambigus. La phrase de notre article devenant par exemple :

ان الافتعاليه الفعليه الفعليه التى تستفعل

Il est également intéressant de noter que la nature de l'analyseur morphologique est la même que celle du moniteur syntaxique, le mot pouvant être aussi considéré comme un automate abstrait qui reflète les contraintes d'enchaînement des lettres entre elles.

L'analyseur syntaxique, au lieu d'interroger constamment le lexique pour mettre en œuvre ses automates pourra dans ce cas faire simplement appel récursivement à des sous-automates morphologiques. Ce type de programmation aura l'avantage de présenter une grande unité de méthode.

AnIsl 22 (1987), p. 217-256 Claude Audebert, André Jaccarini
À la recherche du ḫabar. Outils en vue de l'établissement d'un programme d'enseignement de l'arabe assisté par ordinateur.
© IFAO 2026 AnIsl en ligne https://www.ifao.egnet.net