

CONCLUSION

Le but de notre travail est de contribuer à l'extension de la base de données lexicale du programme de recherche SAMIA (Synthèse et Analyse Morphosyntaxique Informatisée de l'Arabe) dans un cadre syntactico-sémantique. Nous avons le devoir de trouver toutes les structures morphosyntaxiques des phrasèmes ou expressions figées arabes, de les étudier et de constituer un Arbre de leur mise en format.

Nous avons choisi d'étudier les expressions figées en nous inspirant de la conception de MEL'CUK dont la démarche s'appuie sur la théorie sens-texte. Cette théorie est un «modèle de la paraphrase dont la tâche est de faire correspondre à une représentation sémantique donnée toutes les phrases qui lui sont synonymes grâce au système de paraphrasage».(EZZAHID, 1996, p.466) Il «part d'une représentation sémantique pour construire des arbres syntaxiques à l'aide du lexique» (MEL'CUK, CLAS, POLGUERE, 1995, p.17) et décrit l'expression figée ou le phrasème comme étant «une locution prise dans une seule acception bien déterminée et munie de tous les renseignements qui spécifient totalement son comportement dans un texte». (MEL'CUK, CLAS, POLGUERE, 1995, p.57) Cette définition, contrairement à d'autres, prend en compte le figement sémantique et syntaxique des phrasèmes. En étudiant les phrasèmes à partir d'une représentation sémantique, nous nous rendons compte qu'il existe trois groupes de phrasèmes : ceux dont le sens est opaque appelés phrasèmes complets, ceux ayant un sens semi-opaque appelés semi-phrasèmes et ceux dont le sens est compositionnel avec un surplus imprévisible appelés quasi-phrasèmes. Chaque catégorie de phrasème appartient à l'une de ces trois classes.

A partir de cette définition, nous avons constitué une liste de phrasèmes arabes à partir des dictionnaires de **Hans wehr, al Mu^cjam al wasīT**, de la **thèse de Benkaddour** et des **locuteurs natifs algériens, marocains et tunisiens**. De cette liste, nous avons pu établir un classement de toutes les structures morphosyntaxiques des phrasèmes arabes (voir annexe). Dans notre démarche, tous les phrasèmes de l'arabe générale sont représentés dans ce classement.

Les études sur les phrasèmes les ont toujours classés selon la fonction syntaxique, nous avons alors décidé de traiter des phrasèmes à partir de leur structure morphosyntaxique et ainsi les classer selon leur premier constituant. Ce choix est conforté par le fait que dans un texte, la reconnaissance du phrasème, lors du traitement automatique, se passe d'abord à partir de sa structure. Nous avons donc rencontré quatre cent soixante dix (470) structures morphosyntaxiques de phrasèmes classés en cinq familles :

-La première est celle à initiale nominale qui comprend cinquante sept (57) structures.

-La seconde comprenant deux cent dix sept structures (217) est celle à initiale verbale.

-La troisième, avec cent cinquante neuf (159) structures est celle à initiale prépositionnelle.

-La quatrième, à initiale pronominale a vingt et quatre (24) structures.

-La cinquième avec treize (13) structures est celle à initiale adjectivale.

Toutes ces structures ont été étudiées à partir d'au moins deux exemples ; l'étude de chaque catégorie de phrasèmes se déroulant en trois étapes. Dans la première, nous décrivons la suite à étudier, dans la seconde nous étudions son degré de figement syntaxique et sémantique, dans la dernière, nous donnons la classe de phrasèmes à laquelle appartient cette catégorie tout en précisant pour qui ce phrasème peut être employé (Humain ou non-humain). Chaque famille de phrasèmes a des caractéristiques sémantico-syntaxiques définissant son degré de figement. Le phrasème à initiale nominale a pour caractéristiques :

- 1) *non compositionnalité du sens*
- 2) *substitution synonymique impossible*
- 3) *insertion d'un nouveau terme impossible*
- 4) *impossibilité d'actualiser les éléments de façon autonome*
- 5) *adjonction d'un modifieur impossible*

Le phrasème à initiale verbale a pour caractéristiques :

- 1) *non compositionnalité sémantique*
- 2) *blocage des transformations syntaxiques*
- 3) *substitution synonymique impossible*
- 4) *impossibilité d'insérer un nouveau terme*

Le phrasème à initiale prépositionnelle a pour caractéristiques :

- 1) *non compositionnalité sémantique*
- 2) *substitution synonymique impossible*
- 3) *détermination globale*
- 4) *insertion d'un nouvel élément impossible*

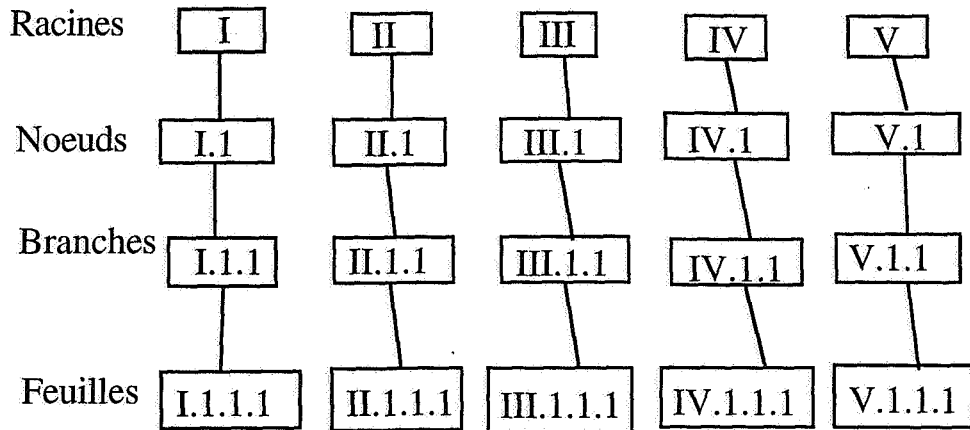
Le phrasème à initiale pronominale a pour caractéristiques :

- 1) *non compositionnalité sémantique*
- 2) *détermination contrainte*
- 3) *substitution synonymique impossible*
- 4) *adjonction d'un modifieur impossible*

Le phrasème à initiale adjectivale a pour caractéristiques :

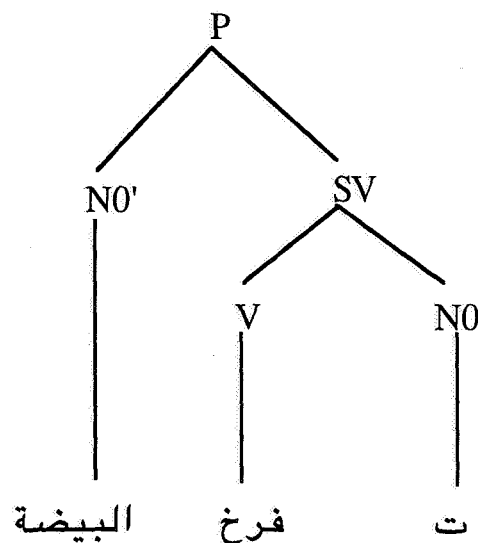
- 1) *non compositionnalité du sens*
- 2) *détermination contrainte*
- 3) *insertion d'un nouveau terme impossible*
- 4) *substitution synonymique impossible*
- 5) *nominalisation de l'adjectif impossible.*

L'étude de ces structures nous a permis de réaliser l'Arbre de Mise en Format des Phrasèmes que nous avons nommé A.MI.FO.P. Cet arbre comprend toutes les structures morphosyntaxiques des phrasèmes arabes et se présente sous forme d'arbres reliés les uns aux autres avec pour racines les cinq familles de phrasèmes. Ces racines comportent des noeuds desquels partent des branches aux bouts desquelles nous avons des feuilles. La représentation générale de l'A.MI.FO.P est la suivante :



Un phrasème peut donc être un noeud, une branche ou une feuille de l'arbre. Pour exemple, nous avons le noeud II.1 à la structure : *verbe à l'achevé + nom déterminé* : فرخت البيضة (*faraxat il bayDatu*) : *la peur s'empare de* ; la branche II.1.2 à la structure : *verbe à l'achevé + nom déterminé au cas accusatif* : رفع الأستئناف (*rafaCa l 'isti'nâfa*) : *faire appel* ; la feuille : II.1.2.1 à la structure : *verbe à l'achevé + nom déterminé (cas accusatif) + pronom complément* : سند معدته (*sana miCdatahu*) : *casser la croûte*.

Pour utiliser cet arbre, nous avons créé un système d'analyse basé d'une part sur l'arbre élémentaire pour la représentation morphosyntaxique et d'autre part sur l'arbre de dérivation pour le figement sémantique. Les phrasèmes étant des unités sémantiques, l'arbre élémentaire nous donne des informations sur l'ordre et les caractéristiques des mots composant ce phrasème, exemple l'arbre élémentaire du phrasème complet فرخت البيضة (*faraxat il bayDatu*) : *la peur s'empare de lui* nous donne la structure du phrasème : *verbe à l'achevé (troisième personne du féminin singulier) + nom déterminé (cas sujet)*.



De cet arbre, nous avons les informations suivantes : marque du sujet dans le verbe (N0) représenté par le *t* de la troisième personne du féminin singulier ; le verbe (V) à l'achevé ; le sujet à proprement parler (N0').

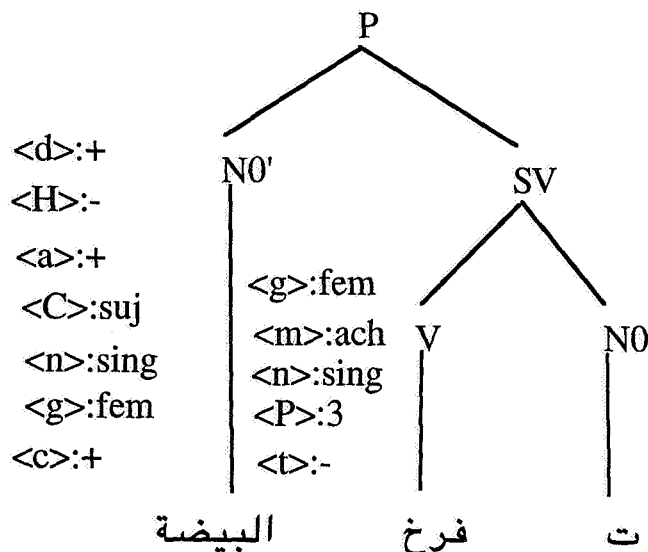
L'arbre de dérivation, quant à lui, nous renseigne sur le degré de figement sémantique du phrasème. Il nous donne, en particulier, le noyau sémantique du phrasème qui n'est autre que sa partie figée. Ce qui nous permet de différencier le phrasème d'une suite libre car dans une séquence ordinaire, on ne parle pas de noyau sémantique ni d'élément non figé mais plutôt de prédicat et arguments. L'arbre de dérivation du phrasème complet que nous avons représenté plus haut sera :

فرخت البيضة (noyau sémantique)

0 (élément non figé)

Le noyau sémantique de cette suite est le phrasème en entier car c'est un phrasème complet, ce qui induit un élément non figé nul car son sens est opaque.

Toutes les informations fournies par l'étude morphosyntaxique nous permettent de former les arbres élémentaires ou les arbres dérivés (un arbre dérivé est la combinaison de plusieurs arbres élémentaires, il est utilisé pour représenter certains semi-phrasèmes et les quasi-phrasèmes) selon que nous sommes en présence d'un phrasème complet, d'un semi-phrasème ou d'un quasi-phrasème. Ces arbres élémentaires ou dérivés, incluant les traits morphosyntaxiques et sémantiques caractérisant le phrasème, associés aux arbres de dérivation (qui nous donnent le degré de figement sémantique du phrasème) simplifient la reconnaissance des phrasèmes lors du traitement automatique. Les traits sémantiques et morphosyntaxiques que possède chaque noeud de l'arbre élémentaire constituent des informations supplémentaires. Ils sont d'ordre morphologique, syntaxique ou sémantique et figurent obligatoirement dans le même ordre dans un arbre dérivé. Ces traits, que le système enregistrera seront utilisés lors du traitement automatique. Leur présence dans un arbre élémentaire se passe de la manière suivante :



A partir des traits sémantiques et morphosyntaxiques, nous avons les détails suivants : *verbe à l'achevé* (<m>) ; à la troisième personne (<P>) du féminin (<g>) singulier (<n>) ; intransitif (<t>), son sujet est un nom déterminé (<d>) au cas sujet (<C>) est un non-humain (<H>), concret (<c>) et animé (<a>).

Le choix que nous avons fait d'adopter une représentation arborescente se justifie donc par le fait qu'elle permet une meilleure représentation des phrasèmes.

Dans la logique de la démarche que nous proposons, nous aboutissons à l'affirmation que le dictionnaire est la cible et le produit final des recherches lexicologiques comme l'est le dictionnaire informatisé de l'arabe (DIINAR). Toutefois, la conception d'un dictionnaire exhaustif des phrasèmes arabes dans tous les domaines de discours demeure un horizon presque inaccessible. Mais avec la liste des structures morphosyntaxiques des phrasèmes arabes, il devient plus aisé d'établir un dictionnaire de phrasèmes arabes appartenant à l'arabe générale. Ne figureront dans ce dictionnaire que les phrasèmes arabes courants et fréquemment utilisés à propos desquels ne subsiste aucun doute sur leur caractère figé. Ensuite, il faudra rajouter à ce dictionnaire les phrasèmes les plus caractéristiques de chaque domaine de discours. Pour cela, la parfaite connaissance du domaine de spécialité en question est nécessaire.

Voilà résumée la tâche à laquelle nous allons nous atteler dès la fin de notre recherche.