

Faut-il réorganiser le lexique arabe ? Proposition pour une catégorisation du lexique arabe dans une perspective de lexicométrie

Zoubeir Mouelhi

▶ **To cite this version:**

Zoubeir Mouelhi. Faut-il réorganiser le lexique arabe ? Proposition pour une catégorisation du lexique arabe dans une perspective de lexicométrie. VIIème Colloque International de la Lexicologie (CIL) sur " Le Lexique entre Langue et Discours ", May 2017, Tunis, Tunisie. hal-01534162

HAL Id: hal-01534162

https://univ-sorbonne-nouvelle.hal.science/hal-01534162

Submitted on 7 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Faut-il réorganiser le lexique arabe ?
Proposition pour une catégorisation du lexique arabe dans une perspective de lexicométrie

Zoubeïr MOUELHI
(CEAO – Université Sorbonne Nouvelle)

L'approche lexicométrique s'inscrit dans la perspective plus générale de l'approche quantitative de l'étude des textes, à l'intersection de plusieurs disciplines, notamment la linguistique, l'informatique et la statistique.

Cette approche trouve plusieurs applications eu égard aux textes, qu'ils soient pris isolément (préoccupations d'ordre stylistique, didactique, historique, etc.), comparés entre eux (typologies de textes, approche contrastive, etc.), considérés dans leur relation aux auteurs (homogénéité d'auteur, attribution d'auteurs, etc.) ou dans leur relation au temps (séries textuelles chronologiques, spécificité chronologique, etc.).

Définie comme étant la science qui étudie l'organisation du vocabulaire dans le discours d'un point de vue quantitatif, la lexicométrie renferme un ensemble de méthodes permettant d'opérer « des réorganisations formelles de la séquence textuelle et des analyses statistiques portant sur le vocabulaire d'un corpus de textes »[1]. Ayant le texte au centre de son objet d'étude, la lexicométrie partage de ce fait, des zones plus ou moins vastes avec d'autres disciplines à savoir : la linguistique, l'analyse de discours, la recherche documentaire, l'analyse de contenu, l'informatique, l'intelligence artificielle, la statistique, etc.

De la norme lexicologique

Avec l'accroissement fulgurant de l'informatique depuis près de trois décennies d'un côté, et l'essor considérable que connaît la linguistique de corpus dont la manifestation la plus parlante est la multiplicité des banques de données textuelles, de l'autre côté, les études et les recherches lexicométriques accusent des évolutions capitales tant sur le plan des objectifs fixés qu'au niveau des principes méthodologiques adoptés. Parmi ces bouleversements sur le plan méthodologique, le recours inévitable à l'outil informatique pour pouvoir faire face à cette grande envergure des opérations de dépouillement et des calculs statistiques colossaux que nécessitent le traitement et l'analyse lexicométriques ; ce qui a pour conséquence que l'on doive constamment œuvrer à définir « de manière toujours plus précise l'ensemble des règles qui présideront au dépouillement automatique des textes stockés en machine »[2].

En traitement automatique des langues écrites et surtout dans les études quantitatives des textes, le point de départ est généralement les mots graphiques. Or le mot graphique ne correspond pas toujours, surtout dans les langues à systèmes d'écritures non segmentées, à l'unité lexicale de base, ce qui rend indispensable l'établissement d'une norme lexicologique qui a pour principale tâche de définir les unités de décompte, faciliter et harmoniser les opérations de dépouillement et préparer le corpus aux traitement et l'analyse statistiques. Cela se traduit par un certain nombre de décisions lexicographiques à prendre dont la principale caractéristique est qu'elles doivent être à la fois simples et durables.

En premier lieu, l'élaboration, pour l'arabe, de ce que dans le domaine de la lexicométrie on appelle une norme lexicologique, donnant une assise théorique et méthodologique aux travaux lexicométriques futurs sur les textes arabes, norme qui a fait défaut, jusque-là, aux rarissimes recherches, le plus souvent fragmentaires, dans ce domaine.

[1] André Salem, *Pratique des segments répétés. Essai de statistique textuelle*, 1987, p. 315
[2] *Ibid, idem*, p. 315

Quantifier un corpus, c'est en extraire les occurrences des unités lexicales définies comme unités de décompte en dressant la liste structurée et exhaustive de ces unités, associées à leurs fréquences, pour pouvoir les analyser par la suite et leur opérer les calculs statistiques voulus. Il s'agit donc du passage d'un texte à des listes de mots-formes, lemmes, catégories lexicales, segments répétés, contextes (concordances), etc. Ā l'issue de cette opération, la linéarité du texte cède sa place à la verticalité des listes et des tableaux. Cette quantification constitue la première étape nécessaire à cette phase du traitement et d'analyse du corpus en lexicométrie.

Parmi les impératifs méthodologiques que le développement de la discipline a pu imposer à tous les travaux s'inscrivant dans le domaine de la linguistique de corpus en général et des études quantitatives en particulier, c'est justement la nécessité d'établir et puis de suivre une norme lexicologique de dépouillement au moment du prétraitement des données textuelles du corpus étudié. Loin des différentes connotations que peut avoir, surtout en linguistique, le terme norme, la notion de norme de dépouillement doit être comprise comme « une exigence de standardisation provisoire des textes contenus dans un corpus. Cette standardisation est destinée avant tout à les rendre comparables, à les stabiliser le temps d'une expérience »[3].

Mais l'établissement d'une telle norme lexicologique de dépouillement ne va pas sans susciter des difficultés de nature différente. L'une de ces difficultés est de satisfaire aux exigences parfois contradictoires de la linguistique et de la statistique et pour lesquelles « la norme devrait être acceptable à la fois pour le linguiste, pour ses auxiliaires, et pour le statisticien. Mais leurs exigences sont souvent contradictoires. L'analyse linguistique aboutit à des classements nuancés […], la matière sur laquelle elle opère est éminemment continue […]. La statistique, dans toutes ses applications, ne va pas sans une certaine simplification des catégories ; elle ne pourra entrer en action que quand le continu du langage a été rendu discontinu »[4].

Il faut cependant prendre soin, au moment de l'acte interprétatif, de restituer ces éléments discontinus (unités lexicales, phrasèmes, co-occurrents, codes grammaticaux, …) dans leur cadre textuel, puis dans un cadre plus élargi, le corpus. En effet, si « le texte est pour une linguistique évoluée l'unité *minimale* »[5], c'est le corpus qui représente son unité maximale. Le cadre discursif de contextualisation par excellence est au-delà du texte, il est au niveau du corpus, comme nous le précise François Rastier : « tout texte placé dans un corpus en reçoit des déterminations sémantiques, et modifie potentiellement le sens de chacun des textes qui le composent »[6]. Ce parcours interprétatif se fait au sein du corpus où s'articulent textes et contexte, et qui marque les limites ultimes de la linguistique, comme le souligne Damon Mayaffre : « De la lettre au mot, du mot à la phrase, de la phrase au texte, du texte au corpus textuel… L'expansion de la linguistique semble devoir s'arrêter là. Ce qui se passe au-delà du corpus est, par définition, hors de portée du linguiste »[7].

Outre la rigueur scientifique exigée dans tout travail de recherche quant à la définition des unités utilisés, l'ensemble des choix et décisions formant la norme lexicologique vont pouvoir servir de base à toute étude contrastive lexicométrique ou stylométrique. Sans cette harmonisation des critères de dépouillement des textes, base de tout parcours analytique et interprétatif, il serait difficile sinon impossible de pouvoir comparer deux corpus sur la base de leurs indices lexicométriques ou stylométriques respectifs, des similitudes/dissimilitudes qui peuvent exister entre eux, ni même pouvoir juger des éventuelles corrélations/dispersions décelées entre les éléments de chacun d'entre eux.

Il convient donc de définir pour les études lexicométriques arabes une norme lexicologique comportant les règles qui servent à définir la nature ou les frontières des unités sur lesquelles porteront les décomptes, les lemmes ou les limites, plus ou moins claires, des catégories lexicales auxquelles seront rattachées les formes du texte étudié. C'est une norme de dépouillement incluant la définition, les critères et les règles de chacune des quatre grandes opérations qui composent ce que l'on appelle l'analyse lexicale : il s'agit du processus

[3] Habert B., Nazarenko A. et Salem A., *Les linguistiques de corpus*, 1997, p. 187
[4] Muller Ch., *Initiation aux méthodes de la statistique linguistique*, 1992, p. 113
[5] Rastier F., *Enjeux épistémologiques de la linguistique de corpus*, in G. Williams (éd.), *La linguistique de corpus*, 2005, p. 31
[6] Rastier F., *Arts et sciences du texte*. Paris, 2001, p. 2
[7] Mayaffre D., *Les corpus politiques : objet, méthode et contenu. Introduction*, 2005, note 5

qui, analysant le texte, va définir les limites des unités de décompte (segmentation), assigner le lemme (lemmatisation) et la catégorie lexicale (catégorisation) à chaque unité en fonction de ses voisins (désambiguïsation).

La segmentation ou itémisation ouvre la voie aux autres étapes principales puisqu'elle se charge de découper le texte en unités minimales d'un certain type : en lexicométrie, ce sont en général les mots-formes. La lemmatisation consiste à identifier pour chaque mot-forme son adresse lexicale, que l'on appelle le lemme. La catégorisation, quant à elle, consiste à choisir parmi les différentes catégories lexicales possibles du lemme, la bonne catégorie à assigner à ce dernier en fonction du contexte immédiat ou médiat du mot-forme ; la tâche de l'étape de catégorisation se voit nettement facilitée si l'on procède d'abord à une opération de désambiguïsation.

Catégorisation des unités lexicales en lexicométrie

Pouvant être considérée soit comme une opération indépendante de la lemmatisation au sens strict et succédant à celle-ci, soit comme le deuxième volet de la lemmatisation au sens large après un premier volet d'identification, la catégorisation est cette opération qui consiste à assigner à chaque lemme sa catégorie lexicale correspondante. Cette opération est d'autant plus aisée que les catégories lexicales sont clairement définies et nettement délimitées. Cependant, certains cas d'ambiguïtés mono- et polycatégorielles peuvent venir entacher le bon déroulement de cette entreprise. C'est pourquoi, la mise en place des filtres réducteurs d'ambiguïtés avant la segmentation ainsi que le recours aux procédés de désambiguïsation avant la lemmatisation peuvent être d'un grand secours à l'opérateur humain ou au programme informatique, préparant le terrain à une catégorisation réussie. Réciproquement, le fait d'affecter sa catégorie lexicale à chaque lemme, contribue d'une manière efficace à la levée d'ambiguïtés effective. Mais une bonne catégorisation commence toujours par la définition des catégories lexicales. Quelles sont donc les catégories lexicales de l'arabe que nous avons retenues et quel est le code fixé pour chacune d'entre elles ?

Cependant, le choix des critères de classification est certes important, mais ce qui l'est encore davantage dans cette entreprise de catégorisation, c'est la stabilité d'application de ces mêmes critères de classement.

Principes de catégorisation

« *Procéder à l'indexation d'un texte*, nous dit Charles Bernet, *c'est opter pour une certaine définition du lexique et de ses éléments* »[8], mais cette relation entre les éléments du texte, et ceux du lexique n'est pas une relation directe dans ce sens où l'objet de la lexicométrie n'est pas le lexique mais le(s) vocabulaire(s). Néanmoins, à une certaine étape du cheminement de l'analyse lexicométrique, l'étape de catégorisation, l'on est amené à faire correspondre aux unités du vocabulaire, dotées d'une dimension réelle, des unités virtuelles du lexique. Ce lexique qui est considéré, selon l'approche adoptée, comme :

- la somme, théorique, de tous les vocabulaires,
- l'ensemble des unités lexicales, c'est-à-dire des mots lexicaux et des mots-outils,
- l'ensemble des morphèmes d'une langue,
- l'ensemble des mots lexicaux et locutions.

Seulement, la statistique lexicale depuis Pierre Guiraud et Charles Muller, a toujours favorisé la deuxième définition du lexique.

Deux démarches différentes parfois même contradictoires sur certains points, s'opposent quant à la classification des unités lexicales de la langue arabe : la Tradition grammaticale arabe aspirant à

[8] Bernet Charles, *Le vocabulaire des tragédies de Jean Racine. Analyse statistique*, 1983, p.17

l'exhaustivité et donnant la prééminence du sens sur la forme, et la linguistique moderne soucieuse de systématisation et insistant sur l'analyse de la forme qui est « une démarche contemporaine, issue du positivisme logique et de la pensée structurale »[9].

Mis devant les problèmes d'ordre méthodologique qu'il rencontre très souvent dans son domaine et qui sont, entre autres, la définition des unités, le repérage des occurrences relevant d'une même unité, la levée d'ambiguïtés polycatégorielles, etc., le chercheur en lexicométrie se trouve confronté à une double obsession : d'un côté, faire des inventaires les plus exhaustifs possible des unités du lexique pour qu'elles puissent correspondre aux unités du (des) vocabulaire(s) étudié(s) et de l'autre côté, tenir compte des exigences du TAL et de ses formalismes avec toute la rigueur que cela implique notamment au niveau des critères de catégorisation pour obtenir en fin de compte que chaque unité lexicale puisse être affectée à une seule et unique catégorie lexicale.

- **Mots lexicaux *vs* mots-outils**

Notons d'abord que c'est le poète Zhang Yan, au XIII^{ème} siècle, qui « avait inventé les notions de *mot plein* (*Shi Zi*) et de *mot vide* (*Xu Zi*) – pour opérateur grammatical – qui ont fait dans nos grammaires contemporaines une si belle carrière »[10]. Même si, à la suite de B. Pottier, « nous nous refusons à croire que la langue puisse posséder des mots vides »[11], qu'on les appelle *mots vides*, *mots de relation*, *mots grammaticaux* ou *mots-outils*, ces unités sont en quelque sorte comparables à du ciment qui lie entre eux les *mots pleins*, *mots sémantiques*, *mots lexicaux*. Ce sont des mots « dont le rôle est plus syntaxique que sémantique et qui servent davantage à établir une relation qu'à définir une substance »[12].

Les mots-outils servent donc à faire fonctionner le discours, à construire des phrases et à les organiser en un texte cohérent. Même s'ils contribuent certainement à la signification générale du contexte discursif dont ils font partie, les mots-outils n'ont cependant pas de signification en dehors de ce contexte. Contrairement aux mots lexicaux, ils ne font pas penser à une image mentale correspondant à une part du réel. Le contenu de ces unités à fonction syntaxique, ne peut être appréhendé conceptuellement hors contexte. Ils sont « sémantiquement déterminés par leur fonction, qui reflète la structure immanente de la langue. Leur sens est entièrement investi par le code syntactique ou par la situation de communication, comme on le voit pour les pronoms personnels, pour les adverbes de lieu et de temps »[13].

Cette dépendance des mots-outils vis-à-vis de la structure syntaxique et du contexte discursif, les mots lexicaux eux, ne la subissent pas. Au contraire, « les signifiés des mots "lexicaux" sont relativement indépendants du système abstrait de la langue et sont sémantiquement analysables hors contexte, d'où la possibilité de périphrase synonymique : la définition du dictionnaire »[14]. Ces mots lexicaux sont doués d'une valeur dénominative leur permettant de référer aux choses du monde.

Au niveau de la répartition des unités lexicales entre ces deux ensembles de classes, la langue arabe ne fait pas exception et ses unités lexicales se prêtent également à cette opposition entre *mots lexicaux* ayant une signification en soi et *mots-outils* servant de liant au discours et recevant leur signification de cette fonction syntaxique qu'ils sont censés remplir.

- **Qu'appelle-t-on mot-outil en arabe ?**

[9] Joseph Dichy, *L'écriture dans la représentation de la langue : la lettre et le mot en arabe*, 1990, p. 494
[10] Henri Meschonnic, *Des mots et des mondes. Dictionnaire, encyclopédies, grammaires, nomenclatures*, 1991, p. 20
[11] Bernard Pottier, *Systématique des éléments de relation : étude de morphosyntaxe structurale romane*, 1962, p. 95
[12] Etienne Brunet, *Le vocabulaire français de 1789 à nos jours*, p. 361
[13] Alain Rey, *Le lexique : Images et modèles. Du dictionnaire à la lexicologie*, 1977, p. 165
[14] A. Rey, *idem*, p. 165

Souvent, dans un contexte linguistique arabe, le terme أداة *ʾadât* (pl. أدوات *ʾadawât*) [outil] est utilisé comme simplification de أداة نحويّة *ʾadât naḥwiyya* [outil grammatical]. Contrairement à ce qu'on trouve dans certains écrits linguistiques, commentaires de textes de grammaire ou traductions de citations de livres de grammaire arabe, le terme d'*outil grammatical*, ou de *mot-outil*, n'est en aucun cas synonyme de *particule*, chère à la Tradition grammaticale arabe. Le terme *particule* est proposé par la tradition orientaliste pour traduire le terme حرف *ḥarf* [lettre, mot[15], litt. : extrémité].

Dans la division tripartite opérée par la Tradition grammaticale, la *particule* est définie par opposition au nom et au verbe ; tout ce qui n'est ni nom ni verbe est *particule*. Le *mot-outil*, lui, est défini par rapport à la relation syntaxique dont il acquiert sa signification ; il est dépendant du contexte. Le critère qui permet de définir le *mot-outil*, et donc de le distinguer du *mot lexical*, est cette opposition : fonction syntaxique *vs* valeur dénominative. Le *mot-outil* n'est donc pas la *particule*. Le premier englobe la deuxième. Toute *particule* est *mot-outil* mais tout *mot-outil* n'est pas *particule*. Cette vision du *mot-outil*, souvent implicitement adoptée ou appliquée, n'a été que rarement précisée ou énoncée d'une manière explicite.

Il est frappant de voir que, sur le plan de la catégorisation classique, les unités qui sont dépourvues de toute valeur dénominative et vouées à des relations syntaxiques, et qui auraient dû être regroupées dans un seul et même ensemble, celui des mots-outils, ont été éparpillées entre les trois classes, la classe des verbes (pour les verbes figés), celle des noms (pour les noms-outils) et celle des particules. Les deux figures suivantes montrent la classe des verbes dans laquelle on trouve les verbes figés (ou fonctionnalisés), et celle des noms où l'on trouve un ensemble de mots-outils (relatifs, démonstratifs, interrogatifs, ...) formant ce que l'on appelle les noms-outils.

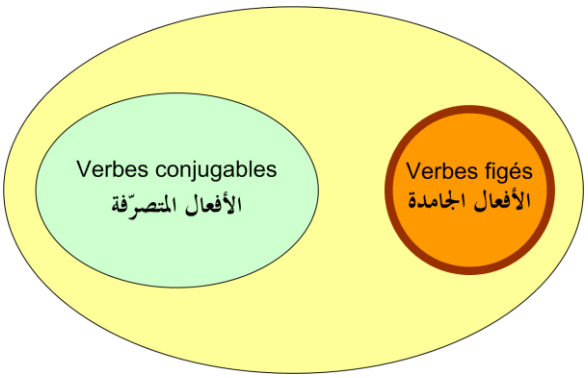


Figure 1 : Catégorisation classique des Verbes contenant les Verbes figés

[15] Voir ces notions dans (Dichy, 1990).

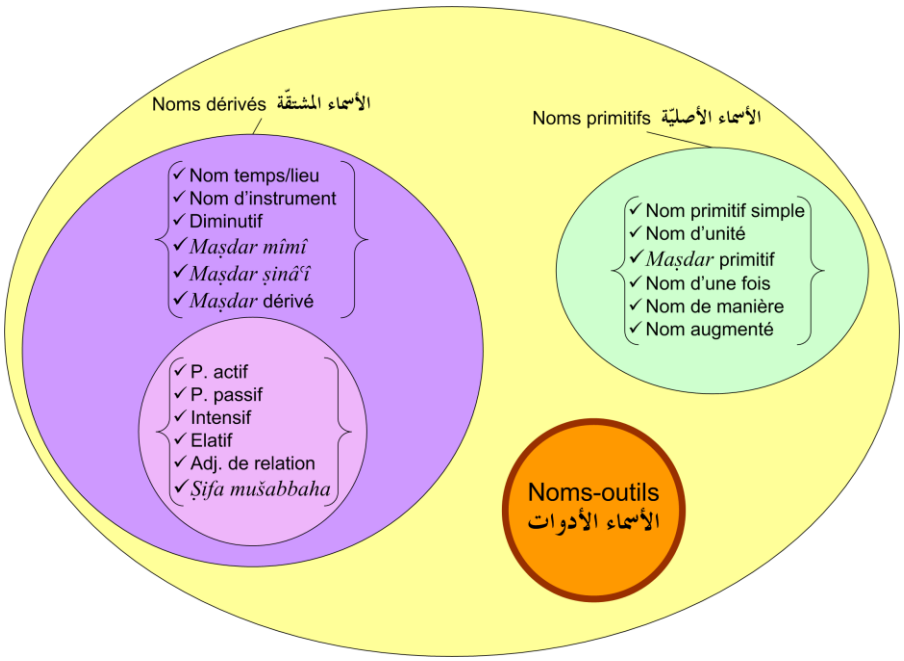


Figure 2 : Catégorisation classique des Noms contenant les Noms-outils

Nous présentons dans ce qui suit, quelques définitions, usages ou indications qui corroborent l'idée que, malgré la confusion qui a accompagné les notions de *mot-outil* et de *particule*, les deux termes ne sont aucunement synonymes et que la *particule* n'est qu'un élément parmi d'autres du *mot-outil* :

- Dans *Miftâḥ as-saʿâda wa-miṣbâḥ as-siyâda* مفتاح السعادة ومصباح السيادة , Ṭâš Kubrâ Zâda (1495-1561) définit les mots-outils comme suit :

« والمراد بالأدوات : الحروف ، وما شابهها من الأسماء والأفعال والظروف »[16]

« Par *mots-outils*, il faut entendre : les particules et leurs semblables parmi les noms, les verbes et les adverbes » (la traduction est de nous).

- De la même manière, Georges Mitrî ʿAbd al-Masîḥ et Hânî Georges Tâbrî, dans leur *Lexique des termes de la grammaire arabe*, *Al-Ḫalîl : Muʿjam muṣṭalaḥât an-naḥw al-ʿarabî* معجم الخليل : مصطلحات النحو العربي , définissent le mot-outil comme suit :

« الأداة : الحرف ، وما تضمّن معناها من الأسماء والأفعال والظروف ، نحو : مِنْ - سِوَى - حاشا - أمس »[17]

« Mot-outil : c'est la particule et tout ce qui en a la même signification [syntaxique] parmi les noms, les verbes et les adverbes, comme *min* [de, à partir de] - *siwâ* [autre que] - *ḥâšâ* [sauf, en dehors de] - *ʾamsi* [hier] » (la traduction est de nous).

Par la suite, les auteurs de ce lexique présentent plusieurs sous-groupes de mots-outils en précisant à chaque fois, la nature des éléments qui les composent. Ils présentent, par exemple, les mots-outils d'exception أدوات الاستثناء *ʾadawât al-ʾistiṯnâʾ* comme étant composés de *particules* حروف *ḥurûf* comme إلاّ *ʾillâ* (...), de *noms* أسماء *ʾasmâʾ* comme غيْر - سِوَى *ġayr - siwâ* (...), de *verbes* أفعال *ʾafʿâl* comme لا يكون - ليس *laysa - lâ yakûnu* (...), de *verbes*/*particules* comme عدا - خلا - حاشا *ʿadâ - ḫalâ - ḥâšâ* (...) et en fin, de لا سيّما *lâ siyyamâ* et ses sœurs لا مثل ما *lâ miṯla mâ* (...).

[16] Ṭâš Kubrâ Zâda, *Miftâḥ as-saʿâda wa-miṣbâḥ as-siyâda fî mawḍûʿât al-ʿulûm*, p. 379
[17] ʿAbd al-Masîḥ Georges Mitrî et Tâbrî Hânî Georges, *Al-Ḫalîl : Muʿjam muṣṭalaḥât an-naḥw al-ʿarabî*, p. 38

- Dans le livre d'Az-Zajjâjî[18] (m. 340/951), كتاب حروف المعاني *Kitâb ḥurûf al-maʿânî*, sur les 137 mots-outils recensés (ou 146 si l'on sépare les quelques entrées groupées), près de la moitié, 62 (ou 68) mots-outils, ne sont point des *particules* au sens de la Tradition grammaticale arabe. Nous y repérons des verbes (أمام , تحت , ثَمَّ , بين , بعد , الآن ...), des adverbes (هات , تعال , ظلّ , بات , أمسى , أضحى , أصبح , ليس ...), des interjections (حَنَانَيْك , لعمرك , صددك , رويداً , حيّهل , ويْكأنّ , صه ...), des démonstratifs (أولى), des circonlocutions (كَأَيِّن), des noms (غَيْر , كُلّ , بعض.), voire même des expressions figées (سبحان الله , معاذ الله , لبيك وسعديك , غفرانك لا كفرانك ...).

- Parmi les 105 mots-outils qu'Ibn Hišâm al-ʾAnṣârî (1309-1360) recense et analyse dans son livre, مغني اللبيب عن كتب الأعاريب *Muġnî al-labîb ʿan kutub al-ʾaʿârîb*[19], consacré en grande partie au sujet des mots-outils, une trentaine ne sont pas des particules au sens de la Tradition grammaticale arabe. Parmi eux, on trouve des adverbes (عَوْضُ , عَلُ ...), des pronoms personnels (هُوَ ...), des verbes figés (عَسَى , لَيْسَ ...), des circonlocutions (كَذَا , كَأَيِّ ...), des interjections (بَلْهَ , بَجَلْ ...) ou d'autres noms comme غَيْر , كُلّ , كِلاَ...

- Muhammad At-Tûnjî inclut dans les 156 mots-outils que comporte son dictionnaire des mots-outils[20] معجم الأدوات النحويّة *Muʿjam al-ʾadawât an-naḥwiyya* [litt. : Dictionnaire des outils grammaticaux], une soixantaine de mots-outils autres que particules. Ce sont des relatifs (الأُلى), des adverbes (عَوْضُ , الآنَ , عَلِ ...), des pronoms personnels (هُوَ , إيّاك...), des verbes figés (لَيْسَ ...), des circonlocutions (كَذَا , كَأَيِّ ...), des interjections (بَلْهَ , هَيَّا , هَيْهَاتَ , مَهْ ...) ou d'autres noms comme غَيْر , كُلّ , كِلاَ...

- Dans son livre الجنى الداني في حروف المعاني *Al-Janâ ad-dânî fî ḥurûf al-maʿânî*[21], Al-Ḥasan ʾIbn Al-Qâsim Al-Murâdî (m. 749/1348) recense quant à lui, 105 mots-outils ; parmi eux nous repérons une trentaine qui ne sont pas des particules au sens de la Tradition grammaticale arabe. Nous y trouvons tous les pronoms personnels isolés (أَنَا ... هُنَّ), des interrogatifs (كَمْ , مَنْ...), des adverbes (إذَا , إذ , مُذْ ...), des démonstratifs (ذا), des verbes figés (لَيْسَ , عَسَى ...), des interjections (بَجَلْ , بَلْهَ ...), etc.

- Comme beaucoup d'autres, Donnat Vernier tombe lui aussi, dans la confusion entre *particules* et *mots-outils*. Il déclare que « les particules en arabe se divisent, comme dans les autres langues, en adverbes, prépositions, conjonctions et interjections »[22].

- Ā plusieurs endroits de son livre *Jâmiʿ ad-durûs al-ʿarabiyya* où il traite des mots-outils, Muṣṭafâ Al-Ġalâyînî fait systématiquement la distinction entre ceux qui sont des particules et ceux qui sont des noms, des adverbes ou des verbes, comme en témoignent les citations suivantes :

« حرفا الاستفهام هما : (هل والهمزة) وبقيّة أدوات الاستفهام أسماء »[23]

« Les deux particules interrogatives sont *hal* et la *hamza*. Le reste des mots-outils interrogatifs sont des noms » (la traduction est de nous) ;

Ou :

« أدوات الشرط : منها ما هو حرف، وهما: "إنْ وإذْ ما" [...]. ومنها ما هو اسم [...] وهي "مَنْ وما ومهما وأيُّ وكيفما" ومنها ما هو ظرفُ زمانٍ [...] وهي "أينَ وأنَّى وأيَّانَ ومتى وإذ". ومنها ما هو ظرفُ مكانٍ [...] وهي "حيثما" »[24]

[18] Az-Zajjâjî, *Kitâb ḥurûf al-maʿânî*, 1986
[19] ʾIbn Hišâm Al-ʾAnṣârî, *Muġnî al-labîb ʿan kutub al-ʾaʿârîb*, 1997
[20] Muhammad at-Tûnjî, *Muʿjam al-ʾadawât an-naḥwiyya*, 1995
[21] Al-Ḥasan ʾIbn al-Qâsim al-Murâdî, *Al-Janâ ad-dânî fî ḥurûf al-maʿânî*, 1992
[22] Donat Vernier S. J., *Grammaire arabe, composée d'après les sources primitives*, 1892, t.1, p. 393
[23] Muṣṭafâ al-Ġalâyînî, *Jâmiʿ ad-durûs al-ʿarabiyya*, t. 1, p.12, note 4.

« Les mots-outils de condition sont : soit une particule, "*ʾin* et *ʾiḏ mâ*" [...] ; soit un nom [...], "*man*, *mâ*, *mahmâ*, *ʾayyu* et *kayfamâ*" ; soit un adverbe de temps [...], "*ʾayna*, *ʾannâ*, *ʾayyâna*, *matâ* et *ʾiḏ*" ; soit un adverbe de lieu [...], "*ḥayṯumâ*" » (la traduction est de nous) ;

Ou encore :

« وبقيّة الأدوات التي تجزم فعلَيْن أسماءٌ لا حروف، كمَن وما ومهما ومتى وأخواتها »[25]

« Le reste des mots-outils qui régissent l'apocopé pour deux verbes successifs, sont des noms[-outils] et non des particules comme *man*, *mâ*, *mahmâ*, *matâ* et leurs sœurs » (la traduction est de nous).

- Même s'il n'échappe pas, lui non plus, à cet amalgame entre *particule* et *mot-outil*, Henri Fleish se pose tout de même la question des frontières de la notion de "particule" *ḥarf* et ce du fait de la définition fourre-tout qu'on a voulu donner à ce terme : « On traduit habituellement par "particule" *ḥarf* ainsi défini. Mais jusqu'où fallait-il étendre l'extension de ce *ḥarf* ? Beaucoup d'instruments grammaticaux (ces *adawât* [sing. *adât*], terme d'*al-Farrâʾ*, [...]) se présentaient d'eux-mêmes. Des cas étaient moins clairs [...] »[26]

- En complément à la célèbre concordance du Coran *Al-Muʿjam al-mufahras li-ʾalfâẓ al-Qurʾân al-karîm* المعجم المفهرس لألفاظ القرآن الكريم faite par Mohammad Fuʾâd ʾAbd Al-Bâqî[27] et dans laquelle, à part quelques rares mots-outils, il n'est principalement question que des noms et des verbes contenus dans le Coran, ʾIsmâʿîl ʾAḥmad ʿAmâyra et ʿAbd al-Ḥamîd Muṣṭafâ as-Sayyid ont réalisé la concordance des mots-outils et des pronoms personnels utilisés dans le Coran : معجم الأدوات والضمائر في القرآن الكريم *Muʿjam al-ʾadawât wa-ḍ-ḍamâʾir fi-l-Qurʾân al-karîm*. Même si les auteurs de cette concordance placent les pronoms personnels en dehors des mots-outils (l'on remarque ici qu'en dépit de la distinction faite entre les deux groupes, le fait de les traiter ensemble traduit bien leur grande similitude fonctionnelle), ils insistent bien dans l'introduction, sur la définition et les raisons du choix du terme *mots-outils* أدوات et non de celui de *particules* :

« ونستميح القارئ عذرًا عن استخدام كلمة الأدوات بدلاً من المصطلح الشائع "حروف المعاني"، فهذه الكلمة أوفى بالحاجة من المصطلح المركّب من كلمتين "حروف المعاني"؛ فإنّ من الحروف ما هو خالص في الحرفيّة كالباء والفاء وبل... ومنها ما يجمع بين الاسميّة والحرفيّة والفعليّة كـ"ما" و"حاشا" و"عدا". وهو على أيّ حال مصطلح كوفيّ قديم فضلاً عن تجدّد استعماله لدى المحدثين »[28]

« Le lecteur est prié d'excuser l'utilisation du mot [mots-]outils au lieu du terme courant "particules" *ḥurûf al-maʿânî*. Ce mot est plus adéquat que le terme composé "*ḥurûf al-maʿânî*" ; aussi certains de ces *ḥurûf* sont-ils exclusivement de véritables particules comme le "*bâ*", le "*fâ*", "*bal*"... D'autres en revanche, peuvent être des noms, des particules ou des verbes comme "*mâ*", "*ḥâšâ*", ou "*ʿadâ*". Quoi qu'il en soit, c'est un ancien terme coufique qui de plus, est réactualisé par les [linguistes] contemporains ». (La traduction est de nous)

Par ailleurs, parmi les mots-outils pour lesquels les auteurs établissent la concordance, l'on trouve des adverbes (حَيْثُ , ثَمَّ , إِذْ , إِذَا ...), des relatifs (الذِي , التِي , الذِينَ ...), des démonstratifs (ذَا , أُولاَءِ , ذَانِكَ ...), des circonlocutions (كَأَيِّنْ), des interrogatifs (أَنَّى , كَيْفَ , مَاذَا , مَتَى ...), etc.

---

[24] *Ibid*, *idem*, t. 2, p.203.

[25] *Ibid*, *idem*, t. 3, p.253, note 1.

[26] Fleisch Henri, *Encyclopédie de l'Islam*, article ḤARF, p. 210

[27] Mohammad Fuʾâd ʾabd al-Bâqî, *Al-Muʿjam al-mufahras li-ʾalfâẓ al-Qurʾân al-kaîm*, 1987

[28] ʾIsmâʿîl ʾAḥmad ʿAmâyra et ʿAbd al-Ḥamîd Muṣṭafâ as-Sayyid, *Muʿjam al-ʾadawât wa-ḍ-ḍamâʾir fi-l-Qurʾân al-kaîm*, 1988, p. 10 de l'introduction.

- Fidèle à son engagement de réorganiser la classification des unités lexicales de l'arabe et de les définir sur la base de la complémentarité de la forme et du sens (voir *infra*), Tammâm Ḥassân définit clairement le *mot-outil* sous cet angle de vue :

« الأداة مبنى تقسيمي يؤدّي معنى التعليق والعلاقة التي تعبّر عنها الأداة إنّما تكون بالضرورة بين الأجزاء المختلفة من الجملة »[29]

« Le mot-outil est une catégorie [grammaticale] dont le rôle est d'établir une relation. La relation [syntaxique] exprimée par le mot-outil est forcément [établie] entre les différents constituants de la phrase » (la traduction est de nous).

Ā la suite de cette définition claire et en parfait accord avec la vision, quasi-universelle, qu'on a aujourd'hui du *mot-outil*, Tammâm Ḥassân présente le *mot-outil* comme étant composé de deux types : 1) le **mot-outil primitif** *al-ʾadât al-ʾaṣliyya* الأداة الأصليّة qui correspond aux *particules* ; et 2) le **mot-outil converti** *al-ʾadât al-muḥawwala* الأداة المحوّلة correspondant aux adverbes, aux noms-outils (interrogatifs, démonstratifs, relatifs, noms de condition...), aux verbes (figés ou incomplets) et aux pronoms personnels.

Faut-il réorganiser le lexique arabe ?

Même si, depuis Sîbawayhi et jusqu'au début du VIII/XIV^{e} siècle, la tendance générale était de considérer que le discours tout entier est Verbe, Nom et Particule, cette division tripartite du discours n'a pas toujours fait l'unanimité entre les grammairiens arabes classiques ou plus tardifs, et encore moins chez les linguistes contemporains. Il y a toujours eu, tout au long de l'histoire de la grammaire arabe, des remises en question de la rigidité de cette classification, et une insistante dénonciation de son hétérogénéité. Plusieurs études critiques autour des sciences du langage arabes le signalent[30].

Il n'est pas dans notre propos ici d'exposer toutes les tentatives qui ont été faites pour sortir de cette rigidité qui caractérise la division tripartite des classes de mots de la langue arabe. Mais, en plus de ce qui a été cité plus haut concernant les mots-outils, nous évoquons ici à titre d'exemple quelques propositions allant dans le sens de la dénonciation de la rigidité et de l'hétérogénéité de la division tripartite :

- Celle par exemple, de ʾAḥmad Ibn Ṣâbir, grammairien du VIII/XIV^{e} siècle, d'ajouter une quatrième classe, celle du *nom de verbe* اسم الفعل *ʾism al-fiʿl* [interjection].
- ʿAḍud ad-Dîn al-ʾḥjî (m. 756/1355) a proposé, quant à lui, une typologie en neuf classes.
- ʾIbn as-Sarrâj (m. 316/928-9), quant à lui, donne, dans son livre *Kitâb al-ʾuṣûl fi n-naḥw*, une classification qui trouble la typologie tripartite.
- Parmi les linguistes contemporains, nous citons Tammâm Ḥassân qui déclare, suite à une étude critique des fondements théoriques et méthodologiques de la classification des grammairiens arabes classiques, que :

« التقسيم الذي جاء به النحاة بحاجة إلى إعادة النظر ومحاولة التعديل بإنشاء تقسيم آخر جديد مبنيّ على استخدام أكثر دقّة لاعتباري المبنى والمعنى »[31]

« La classification opérée par les grammairiens [arabes] a besoin d'être revue et remaniée de façon à aboutir à une nouvelle classification basée sur un emploi plus rigoureux des notions de forme et de sens » (la traduction est de nous)

[29] Tammâm Ḥassân, *Al-luġa l-ʿarabiyya. Maʿnâhâ wa mabnâhâ*, p.123.
[30] Voir entre autres, les travaux de Tammâm Ḥassân, d'André Roman et (Gabr 1980)
[31] Tammâm Ḥassân, *Al-luġa l-ʿarabiyya*, op. cit., p.88

C'est dans la continuité de son analyse des unités lexicales de l'arabe et de sa ligne de pensée insistant sur la complémentarité de la forme et du sens (complétées d'ailleurs et approfondies dans son livre *Manâhij al-baḥṯ fî al-luġa*), que Tammâm Ḥassân propose sept catégories lexicales pour classer les éléments du lexique arabe : Nom, Adjectif, Verbe, Pronom, *Ḫâlifa*[32], Adverbe et Mot-outil.

- André Roman, qui se place, quant à lui, dans une perspective de rupture totale avec la Tradition grammaticale arabe, met l'accent sur « l'échec » des grammairiens arabes classiques à définir les unités lexicales sans pour autant que cet échec ne les incite à repenser la classification tripartite : « Les grammairiens, inlassablement, s'attacheront, sans les reclasser, à définir ces données traditionnelles. Jamais ils n'aboutiront. Jamais les quelques définitions proposées par eux ne seront opératoires, régulièrement. Jamais, cependant, ils ne concluront de leur échec millénaire à la nécessité de rompre avec la Tradition »[33].

Que l'on adopte la position appelant à la rupture avec la division tripartite du discours considérée trop rigide, peu cohérente ou hétéroclite, ou que l'on considère que cette division tripartite a fait tout au long de son histoire l'objet d'un consensus quasi-total et que les différentes propositions, dans le passé, voulant la réformer « restent en deçà d'une rupture théorique avec Sîbawayhi. Ils [les changements] présentent, si l'on peut dire, une diversité dans l'unité et un changement dans la continuité »[34], il est évident, du point de vue du TAL arabe, que la classification de la Tradition grammaticale fait fonctionner un système de catégories, par endroits, lexicalement hétérogènes ne permettant aucun gain de clarté pour la répartition catégorielle de certaines unités lexicales.

Certains "réaménagements" s'imposent donc pour avoir une organisation des éléments du lexique s'adaptant aux exigences de l'analyse lexicométrique du discours et du TAL arabe en général. L'opposition, par exemple, entre lexicalité et fonctionnalité des unités du lexique est devenue une pratique inéluctable, pour toutes les langues, dans les domaines de l'ingénierie linguistique et du traitement de l'information textuelle.

Le but visé dans la proposition d'une telle organisation des unités lexicales n'est évidemment pas l'originalité. Nous ne prétendons pas proposer une typologie nouvelle et définitive du Lexique arabe ; tel n'est aucunement notre propos. Nous proposons simplement une organisation ponctuelle des unités lexicales, dictée par les exigences de l'approche adoptée, et appliquant, loin d'une démarche purement éclectique, certains changements et aménagements proposés par des linguistes arabes et arabisants (certaines de ces propositions, comme nous l'avons vu plus haut, ont même été faites par des grammairiens classiques).

Nous présentons dans la page suivante le schéma général de l'organisation des unités lexicales de l'arabe avec les ramifications sur trois niveaux (quatre si l'on considère l'opposition mots lexicaux/mots-outils) ; suite à quoi, nous présentons en détail, toutes les catégories lexicales de base, les catégories lexicales, les sous-catégories et les sous sous-catégories que nous avons retenues pour la *norme lexicologique* que nous proposons dans une perspective de lexicométrie arabe. Norme que nous avons déjà appliquée à différents corpus, littéraires et politiques.

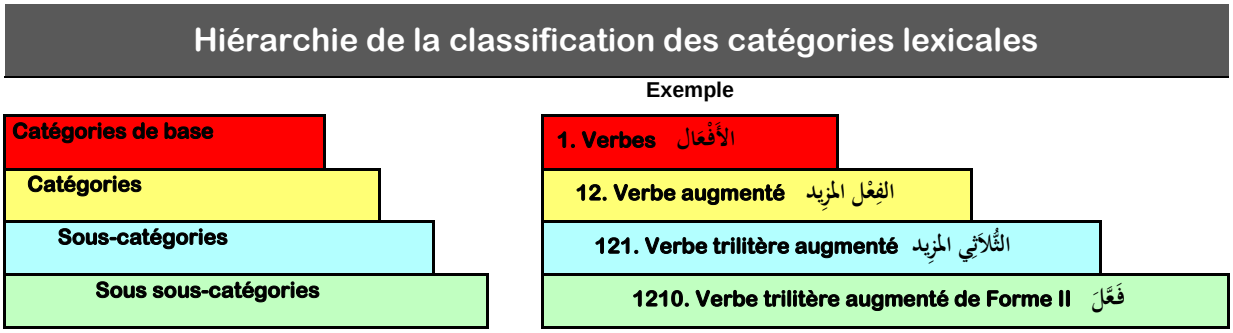
[32] Dans cette catégorie des *ḫawâlif* (pl. de *ḫâlifat*), Tammâm Ḥassân classe les interjections (أسماء الأفعال), les onomatopées (أسماء الأصوات) et les deux formes d'exclamation (مَا أَفْعَلَ - أَفْعِلْ بِـ...).
[33] André Roman, *Genèse et typologie des unités de la langue arabe*, 1994, p. 118
[34] Hassan Hamzé, *Les parties du discours*, 1994, p. 97


Organisation du **lexique arabe** dans une perspective de lexicométrie
Mots lexicaux
Mots-outils
Verbes
Noms primitifs
Noms dérivés
Adjectifs
Particules
Noms-Outils
Verbes fonctionnalisés
Mots-outils composés
Verbes simples
Trilitères simples
Quadrilitères simples
Verbes augmantés
Trilitères augmentés
Quadrilitères augmantés
Primitifs simples
Trilitères simples
Quadrilitères simples
Pentalitères simples
Noms d'unité
Maṣdar primitif
Noms d'une fois
Noms de manière
Noms primitifs augmentés
Noms de temps et de lieu
Noms d'instrument
Diminutifs
Maṣdar mîmî
Maṣdar ṣinâ'î
Maṣdar dérivé
Maṣdar trilitère augmenté
Maṣdar quadrilitère simple
Maṣdar quadrilitère augmenté
Participes actifs
Participes passifs
Intensifs
Elatifs
Ṣifa mušabbaha
Adjectifs de relation
Article défini
Prépositions
Coordonnants
Particules de négation
Particules de comparaison
Particules d'attente ou de certitude
Particules du futur
Particules de condition
Particules de corroboration
Particule d'espérance et d'apitoiement
Particules de maṣdarité
Particules d'appel
Particules de réponse
Particules interrogatives
Particules d'exception
Particules d'ouverture
Particules d'incitation et de remords
Particules assimilées au verbe
Particules d'alternatives
Particules d'explication
Démonstratifs
Interrogatifs
Noms de condition
Relatifs
Noms de verbes
Onomatopées
Noms de nombres
Nombres cardinaux
Nombres ordinaux
Nombres fractionnaires
Circonlocutions
Adverbes
Pronoms personnels
Annectifs
Verbes figés à l'accompli
Verbes figés à l'impératif
Verbes figés à l'inaccompli

Les catégories lexicales retenues

Nous présentons ci-après en détail, toutes les catégories lexicales retenues. Elles sont classées hiérarchiquement : les catégories de base (verbes, noms primitifs, noms dérivés, adjectifs et noms composés pour les mots lexicaux, et mots-outils ; sans oublier les noms propres qui sont eux, à la périphérie du lexique), les catégories (verbes simples, noms primitifs simples, diminutif, participe actif, etc. pour les mots lexicaux, et particules, noms-outils, verbes fonctionnalisés, etc. pour les mots-outils ; noms propres de personnes, etc.), les sous-catégories (verbe trilitère simple, nom trilitère simple etc. pour les mots lexicaux et prépositions, démonstratifs, verbes figés à l'accompli, etc. pour les mots-outils) et les sous sous-catégories (verbe augmenté de forme II, etc. pour les mots lexicaux et nombres cardinaux, etc. pour les mots-outils). Devant chaque catégorie lexicale nous donnons le code que nous avons choisi pour elle et sous lequel elle doit enregistrée dans un souci d'harmonisation.

D'une manière générale, les classes lexicales sont hiérarchisées en quatre niveaux allant de la catégorie de base à la sous sous-catégorie, passant par la catégorie et la sous-catégorie ; c'est ce que nous schématisons dans le tableau suivant :



- Sept catégories de base ont été retenues ; cinq catégories de base pour les mots lexicaux et qui représentent des listes ouvertes (Verbes, Noms Primitifs, Noms dérivés, Adjectifs et Noms Composés) et une catégorie de base pour les mots-outils qui est une liste fermée. Comme dans toutes les langues, les Noms Propres constituent en fait une classe à la périphérie du lexique ; pour la nécessité des études lexicométriques, nous les avons retenus et nous en avons fait une catégorie de base.
- 29 catégories se répartissent entre les 7 catégories de base : 2 catégories pour les Verbes, 6 pour chacune des catégories de base des Noms Primitifs, des Noms dérivés et des Adjectifs, 4 pour les Mots-Outils et 5 pour les Noms Propres.
- Les sous-catégories quant à elles, sont au nombre de 44 dont la part du lion (20 sous-catégories) revient à la catégorie des Particules qui appartient à la catégorie de base des Mots-Outils.
- Nous avons retenu enfin, 16 sous sous-catégories : 10 dans la sous-catégorie des Verbes Trilitères Augmentés, 3 dans celle des Verbes Quadrilitères Augmentés et 3 dans celle des Noms-Outils.

- Verbes الأفعال → Code : 1
  - Verbe simple الفعل المجرّد → Code : 11
  - Verbe augmenté الفعل المزيد → Code : 12
- Noms primitifs الأسماء الأصليّة → Code : 2
  - Nom primitif simple الاسم الأصليّ المجرّد → Code : 21
  - Nom d'unité اسم الوَحدة → Code : 22
  - Maṣdar primitif المصدر الأصلي → Code : 23
  - Nom d'une fois اسم المرّة → Code : 24
  - Nom de manière اسم الهيئة → Code : 25

- Nom augmenté الاسم المزيد → Code : 26

- Noms dérivés الأسماء المشتقّة → Code : 3
  - Noms de temps et de lieu اسما الزمان والمكان → Code : 31
  - Nom d'instrument اسم الآلة → Code : 32
  - Diminutif اسم التصغير → Code : 33
  - Maṣdar mîmî المصدر الميمي → Code : 34
  - Maṣdar ṣinâʿî المصدر الصناعي → Code : 35
  - Maṣdar dérivé المصدر المشتقّ → Code : 36
- Adjectifs الصفات → Code : 4
  - Participe actif اسم الفاعل → Code : 41
  - Participe passif اسم المفعول → Code : 42
  - Intensif اسم المبالغة → Code : 43
  - Elatif اسم التفضيل → Code : 44
  - Ṣifa mušabbaha الصفة المشبّهة → Code : 45
  - Adjectif de relation الاسم المنسوب → Code : 46
- Mots-Outils الأدوات → Code : 5
  - Particules الحروف → Code : 51
  - Noms-Outils الأسماء الأدوات → Code : 52
  - Verbes fonctionnalisés الأفعال الجامدة → Code : 54
  - Mots-Outils composés الأدوات المركّبة → Code : 55
- Noms propres أسماء الأعلام → Code : 6
  - Noms Propres de personnes أسماء الأشخاص → Code : 61
  - Noms Propres de lieux أسماء الأماكن → Code : 62
  - Noms de tribus, groupes et nations أسماء القبائل والأمم والفرق → Code : 63
  - Théonymes الأسماء الدينيّة → Code : 64
  - Noms des œuvres أسماء الكتب → Code : 65
- Noms composés الأسماء المركّبة → Code : 7

Détail des catégories lexicales, définitions, schèmes, exemples et codes

Verbes الأفعال → Code : 1

Cette catégorie de base regroupe tous les éléments du lexique en relation au temps et dont les formes fléchies sont conjuguées, autrement dit, elles varient en personne, en "temps" (aspect), en mode et en voix[35]. De cette définition, il est clair que les "verbes" qui ne remplissent pas cette condition ne sont pas admis dans cette classe ; c'est pour cette raison que les verbes fonctionnalisés (verbes figés الأفعال الجامدة dans le Tradition grammaticale arabe) ne sont pas classés parmi les verbes de cette catégorie de base, mais plutôt dans les mots fonctionnels (mots-outils).

Verbe simple الفعل المجرّد → Code : 11

Verbe trilitère simple الفعل الثلاثي المجرّد **(فَعَلَ\فَعُلَ\فَعِلَ) → Code : 111**

Verbe quadrilitère simple الفعل الرباعي المجرّد **(فَعْلَلَ) → Code : 112**

[35] Une présentation claire et bien structurée des paradigmes de conjugaison des verbes arabes est donnée dans l'introduction de : Joseph Dichy et Sam Ammar, *Les verbes arabes*, 1999, p. 11-20

Verbe augmenté الفعل المزيد → Code : 12

Verbe trilitère augmenté الفعل الثلاثي المزيد **→ Code : 121**

Forme **II** (*faʿʿala* - فَعَّلَ) **→ Code : 1210**

Forme **III** (*fâʿala* - فَاعَلَ) **→ Code : 1211**

Forme **IV** (*ʾafʿala* - أَفْعَلَ) **→ Code : 1212**

Forme **V** (*tafaʿʿala* - تَفَعَّلَ) **→ Code : 1213**

Forme **VI** (*tafâʿala* - تَفَاعَلَ) **→ Code : 1214**

Forme **VII** (*ʾinfaʿala* - اِنْفَعَلَ) **→ Code : 1215**

Forme **VIII** (*ʾiftaʿala* - اِفْتَعَلَ) **→ Code : 1216**

Forme **IX** (*ʾifʿalla* - اِفْعَلَّ) **→ Code : 1217**

Forme **X** (*ʾistafʿala* - اِسْتَفْعَلَ) **→ Code : 1218**

Autres **→ Code : 1219**

اِفْعَالَّ - اِفعَوْعَلَ - اِفْعَوَّلَ - اِفْعَنْلَى - اِفْتَعْلَى

Verbe quadrilitère augmenté الفعل الرباعي المزيد **→ Code : 122**

(*tafaʿlala* - تَفَعْلَلَ) **→ Code : 1220**

(*ʾifʿanlala* - اِفْعَنْلَلَ) **→ Code : 1221**

(*ʾifʿalalla* - اِفْعَلَلَّ) **→ Code : 1222**

Noms primitifs الأسماء الأصليّة → Code : 2

Les noms primitifs sont des noms qui ne sont pas dérivés d'un verbe. Ils désignent des êtres animés ou des choses, entités concrètes ou abstraites. En font partie les collectifs et les singulatifs (individués).

Nom primitif simple الاسم الأصليّ المجرّد → Code : 21

Les noms primitifs simples peuvent être composés de trois consonnes initiales (trilitères) de quatre (quadrilitères) ou de cinq consonnes initiales (pentalitères).

Nom trilitère simple الاسم الثلاثي المجرّد **→ Code : 211**

فَعْل (شَمْس) \ **فَعَل** (بَصَل) \ **فَعِل** (كَبِد) \ **فَعُل** (رَجُل) \ **فُعَل** (صُرَد) \ **فِعْل** (رِجْل) \ **فِعَل** (عِنَب) \ **فِعِل** (إِبِل) \ **فُعْل** (قُفْل) \ **فُعُل** (عُنُق)

Nom quadrilitère simple الاسم الرباعي المجرّد **→ Code : 212**

فَعْلَل (جَعْفَر) \ **فِعْلِل** (زِبْرِج) \ **فِعْلَل** (دِرْهَم) \ **فَعْلُل** (بَرْشُن) \ **فِعَلّ** (سِبَطْر) \ **فُعْلُل** (جُحْدُب)

Nom pentalitère simple الاسم الخماسي المجرّد → **Code : 213**

فَعَلَّل (سَفَرْجَل) \ **فَعْلَلِل** (جَحْمَرِش) \ **فُعَلِّل** (خُزَعْبِل) \ **فِعْلَلّ** (زِنْجَفْر)

Nom d'unité اسم الوَحْدة → **Code : 22**

C'est un singulatif (individué) formé par l'ajout d'un ـة at à un collectif pour former un individu de l'espèce ou de la masse désignée par le collectif. Exemple, du collectif سمك *samak* [du poisson], nous obtenons le nom d'unité سمكة *samaka* [un poisson].

...**ـة** (سمك : سمكة)

***Maṣdar* primitif** المصدر الأصلي → **Code : 23**

En plus du ***maṣdar*** [nom d'action] primitif lui-même, nous classons dans cette catégorie le **nom de *maṣdar*** اسم المصدر comme وضوء *wuḍû'* [ablutions] ainsi que le ***maṣdar* intensif** مَصْدَر المُبَالَغَة construit sur le schème تَفْعَال taf'âl (تَضْرَاب \ تَعْدَاد \ تَطْوَاف \ تَسْآل).

فَعْل (أَكْل) \ **فَعَل** (فَرَح) \ **فُعُول** (جُلُوس) \ **فِعَال** (جِماح) \ **فَعَلان** (غَلَيان) \ **فُعَال** (سُعال) \ **فَعِيل** (رَحِيل) \ **فِعَالة** (إِمارة) \ **فُعْلة** (سُمْرة) \ **فُعُولة** (سُهُولة) \ **فَعَالة** (فَصَاحة) \ **تَفْعَال** (تَعْدَاد)

Nom d'une fois اسم المرّة → **Code : 24**

C'est un nom formé par l'ajout d'un ـة at au nom d'action des verbes simples (il aura, dans ce cas, le schème فَعْلة) et dérivés (sauf, pour ces derniers ceux qui se terminent déjà par ـة).

فَعْلة (وَقْفة) \ **انْفِعالة** (انطلاقة) \ **افْتِعالة** (ابتسامة)

Nom de manière اسم الهيئة → **Code : 25**

Appelé aussi en arabe, اسم النوع ***ism an-naw'***, ce nom est construit exclusivement, à partir du verbe trilitère simple, sur le schème فِعْلة et indique la manière d'accomplir l'action.

فِعْلة (جِلْسة)

Nom primitif augmenté الاسم الأصلي المزيد → **Code : 26**

(حِصَان – قِنْدِيل – إنسان – عصفور –...)

Noms dérivés الأسماء المشتقّة → Code : 3

Contrairement aux noms primitifs, les noms dérivés eux, sont soit des noms ou maṣdars dérivés d'un verbe (déverbaux), soit des noms dérivés du nom primitif (comme le diminutif). Les formes déverbales sont

« celles des formes non conjuguées qui dénotent un procès, c'est-à-dire une action ou une actualisation, exprimé dans son déroulement »[36].

Noms de temps et de lieu اسما الزمان والمكان → Code : 31

مَفْعِل (مَجْلِس – مَوْقِع – مَبِيت) \ **مَفْعَل** (مَكْتَب – مَرْمَى – مَقام) \ **مَفْعَلة** (مَزرعة – مَنامة – مَزَلّة) \ **مُفْعَل** (مُقام) \ **مُنْفَعَل** (مُنْصَرَف) \ **مُسْتَفْعَل** (مُسْتَقْبَل)

Nom d'instrument اسم الآلة → Code : 32

C'est un nom généralement dérivé du verbe trilitère simple transitif pour désigner un instrument utilisé pour accomplir l'action du verbe. Il y a quelques cas où le nom d'instrument est dérivé d'un verbe augmenté (comme مِزْمار mizmâr [trompette], du verbe de la forme II), d'un verbe intransitif (comme مِدْخَنَة *midẖana* [cheminée], du verbe intransitif دَخَنَ) voir même, et c'est rare, d'un nom primitif (comme مِحْبَرَة *miḥbara* [encrier] du nom primitif حِبْر *ḥibr* [encre]).

مِفْعال (مِزْمار) \ **مِفْعَل** (مِصْعَد) \ **مِفْعَلة** (مِلْعَقة) \ **فاعِلة** (ساقِية) \ **فاعُول** (ساطُور) \ **فَعَّالة** (ثلاّجة) \ **فِعَّال** (إِرَّاث)

Diminutif اسم التصغير → Code : 33

فُعَيْل (قُلَيْم) \ **فُعَيِّل** (كُتَيِّب) \ **فُعَيْعِل** (جُعَيْفِر) \ **فُعَيْعِيل** (عُصَيْفِير)

***Maṣdar mîmî* المصدر الميمي → Code : 34**

Le ***maṣdar mîmî*** est un *maṣdar* d'un type particulier. Sa particularité est qu'il commence par la lettre مـ m. Outre les schèmes que nous présentons dans l'encadré suivant correspondant à des *maṣâdir* de verbes trilitères simples, le ***maṣdar mîmî***, quand il correspond à des verbes trilitères augmentés, a le même schème que le participe passif. Exemple : le *maṣdar* مُعْتَقَد *muʿtaqad* [croyance, conviction] de schème مُفْتَعَل *mufta'al* et qui correspond au verbe de Forme VIII اِعْتَقَدَ *'iʿtaqada* [croire] de schème اِفْتَعَلَ *'ifta'ala*.

مَفْعَل (مَدْخَل) \ **مَفْعِل** (مَوْعِد) \ **مَفْعِلة** (مَعْرِفة) \ **مَفْعَلة** (مَفْسَدة)

***Maṣdar ṣinâʿî* المصدر الصناعي → Code : 35**

Le ***maṣdar ṣinâʿî*** est obtenu par l'ajout d'un ـة at final à un adjectif de relation, lui-même obtenu par l'adjonction à un nom, d'un ـيّ *iyy* final. Exemple : إنسان *'insân* [Homme] → إنسانـيّ *'insâniyy* [humain] → إنسانـيّـة *'insâniyya* [humanité].

...يّة (إنسانيّة)

***Maṣdar dérivé* المصدر المشتقّ → Code : 36**

***Maṣdar* trilitère augmenté** المصدر الثلاثي المزيد **→ Code : 361**

[36] André Roman, *Etude de la phonologie et la morphologie de la Koinè arabe*, 1983, p. 1002

تَفْعِيل – تَفْعِلة (تقسيم – تسمية) \ **فِعَال – مُفَاعَلة** (خِصام – مُخاصَمة) \ **إِفْعال – إِفَالة** (إحسان – إعانة) \ **تَفَعُّل** (تحسُّن) \ **تَفَاعُل** (تَخاصُم) \ **اِنْفِعَال** (انطلاق) \ **اِفْتِعَال** (اقتراب) \ **اِفْعِلال** (احمرار) \ **اِسْتِفْعَال – اِسْتِفَالة** (استقبال – استقامة)

***Maṣdar* quadrilitère simple** المصدر الرباعي المجرّد **→ Code : 362**

فَعْلَلة (دحرجة) \ **فِعْلال** (زلزال)

***Maṣdar* quadrilitère augmenté** المصدر الرباعي المزيد **→ Code : 363**

تَفَعْلُل (تَزَلْزُل) \ **اِفْعِنْلال** (احرنجام) \ **اِفْعِلّال** (اقشعرار)

Adjectifs الصفات → Code : 4

Il est à noter que les formes déverbales telles que les adjectifs, sont sujettes à une éventuelle nominalisation et constituent de ce fait un réservoir de formes nominales.

Participe actif اسم الفاعل → Code : 41

Le participe actif est un adjectif qui exprime celui qui fait l'action du verbe.

فاعِل \ مُفَعِّل \ مُفاعِل \ مُفْعِل \ مُتَفَعِّل \ مُتَفاعِل \ مُنْفَعِل \ مُفْتَعِل \ مُفْعَلّ \ مُسْتَفْعِل \ مُفَعْلِل \ مُتَفَعْلِل \ مُفْعَنْلِل \ مُفْعَلِلّ

Participe passif اسم المفعول → Code : 42

Le participe passif est un adjectif exprimant celui qui subit l'action.

مَفْعُول \ مُفَعَّل \ مُفاعَل \ مُفْعَل \ مُتَفَعَّل \ مُتَفاعَل \ مُنْفَعَل \ مُفْتَعَل \ مُفْعَلّ \ مُسْتَفْعَل \ مُفَعْلَل \ مُتَفَعْلَل \ مُفْعَنْلَل \ مُفْعَلَلّ

Intensif اسم المبالغة → Code : 43

Il sert à exprimer l'intensité ou la répétition. L'un des schèmes de l'intensif (فَعَّال) est notamment utilisé pour la formation d'un grand nombre de noms de métier.

فَعَّال (كذّاب) \ **مِفْعَال** (مِفْضال) \ **فَعُول** (ضَرُوب) \ **فَعِيل** (عَلِيم) \ **فِعّيل** (سِكّير) \ **فُعُّول** (قُدُّوس) \ **فَعَّالة** (علاّمة) \ **مِفْعِيل** (مِعْطِير) \ **فَيْعُول** (قَيُّوم) \ **فُعَّال** (كُبّار) \ **فَاعُول** (فاروق)

Elatif اسم التفضيل → Code : 44

Servant à exprimer la comparaison, l'élatif est un adjectif construit sur le schème أَفْعَل *ʾafʿal* (فُعْلَى *fuʿlâ* pour le féminin), dérivé du verbe trilitère conjugable actif. Il a la valeur d'un comparatif quand il est

indéfini suivi de la préposition مِن *min* [de], et d'un superlatif quand il est défini par l'article ou premier terme d'une annexion (que le deuxième terme de l'annexion soit défini ou non).

أفْعَل (أحْسَن)

Ṣifa mušabbaha الصفة المشبّهة → Code : 45

La ***Ṣifa mušabbaha*** (*bi-sm al-fâʿil*) (الصفة المشبّهة (باسم الفاعل [litt. l'adjectif ressemblant ou assimilé (au participe actif)] est un adjectif dérivé du verbe intransitif pour exprimer une qualité durable, comme أعمى *ʾaʿmâ* [aveugle].

أفعل – فَعْلاء (أعمى – عمياء) \ **فَعْلان – فَعْلى** (شَبعان – شَبْعى) \ **فَعِل** (فَرِح) \ **فَعِيل** (كَرِيم – عَلِيّ) \ **فَعْل** (شَهْم) \ **فُعْل** (صُلْب) \ **فَعَل** (حَسَن) \ **فُعَال** (شُجاع) \ **فَعَال** (جَبان) \ **فَاعِل** (طاهر) \ **فَيْعِل** (طَيِّب) \ **فَيْعَل** (فَيْصَل)

Adjectif de relation الاسم المنسوب → Code : 46

L'adjectif de relation est obtenu par l'ajout d'un يّ- *iyy* final à un nom. Il sert à exprimer la relation, la matière ou l'origine.

يّ... (عربيّ – فرنسيّ)

Mots-Outils الأدوات → Code : 5

Les mots-outils regroupent les particules, les noms-outils et les verbes fonctionnalisés. Vu leur spécificité et pour permettre, le cas échéant, de les étudier isolément, nous avons décidé de classer dans une catégorie à part, les mots-outils composés.

Particules الحروف → Code : 51

Article défini حرف التعريف → Code : 510

اَلْ

Prépositions حروف الجرّ → Code : 511

إِلَى – بِ – تَ – حَتَّى – رُبَّ – عَلَى – عَنْ – فَ – فِي – كَ – لِ – مُذْ – مِنْ – مُنْذُ – وَ (واو رُبَّ)

Coordonnants حروف العطف → Code : 512

أَمْ – أَوْ – بَلْ – ثُمَّ – حَتَّى – فَ – وَ – لَكِنْ

Particules de négation حروف النفي → Code : 513

إِنْ – لاَ – لاَتَ – لَمْ – لَمَّا – لَنْ – مَا

Particules de comparaison حرفا التشبيه **→ Code : 514**

كـ – كَأَنَّ

Particule d'attente ou de certitude حرف التوقّع أو التحقيق **→ Code : 515**

قَدْ – لَقَدْ

Particules du futur حرفا الاستقبال **→ Code : 516**

سَ – سَوْفَ

Particules de condition حروف الشرط **→ Code : 517**

إِذْمَا – أَمَّا – إِنْ – لَمَّا – لَوْ – لَوْلاَ – لَوْمَا – كَيْفَمَا

Particules de corroboration حروف التوكيد **→ Code : 518**

أَنَّ – إِنَّ – قَدْ – لَ – نَّ – نْ

Particule d'espérance et d'apitoiement حرف الترجّي والإشفاق **→ Code : 519**

لَعَلَّ

Particules de "*maṣdarité*" حروف المصدريّة **→ Code : 5191**

أَ – أَنَّ – أَنْ – كَيْ – لَوْ – مَا – كَمَا

Particules d'appel حروف النداء **→ Code : 5192**

آ – أَ – أَيْ – أَيَا – أَيَّتُهَا – أَيُّهَا – هَيَا – وَا – يَا – اللَّهُمَّ

Particules de réponse حروف الجواب **→ Code : 5193**

أَجَلْ – إِنَّ – إِيْ – بَجَلْ – بَلَى – جَلَلْ – جَيْرِ – كَلاَّ – لاَ – نَعَمْ – إِذًا

Particules interrogatives حرفا الاستفهام **→ Code : 5194**

أَ – هَلْ

Particules d'exception حروف الاستثناء **→ Code : 5195**

إِلاَّ – حَاشَا – خَلاَ – عَدَا – سِيَّمَا – لاَسِيَّمَا

Particules d'ouverture حرفا الاستفتاح **→ Code : 5196**

أَلاَ – أَمَا

Particules d'incitation et de remords حروف التحضيض والتنديم → **Code : 5197**

أَلاَ – أَلاَّ – لَوْلاَ – لَوْمَا – هَلاَّ

Particules assimilées au verbe الحروف المشبّهة بالفعل → **Code : 5198**

أَنَّ – إِنَّ – كَأَنَّ – لَعَلَّ – لَكِنَّ – لَيْتَ

Particules d'alternative حرفا التفصيل → **Code : 5199**

أَمَّا – إِمَّا

Particules d'explication حرفا التفسير → **Code : 51991**

أَنْ – أَيْ

Noms-Outils الأسماء الأدوات → Code : 52

Nous avons proposé, en premier temps, cette appellation sans savoir qu'elle avait déjà été utilisée auparavant ; nous avons découvert par la suite que Régis Blachère l'avait déjà proposée dans sa Grammaire de l'arabe classique. Il définit les noms-outils comme suit : « Sous cette appellation, seront désignés des thèmes nominaux ou autres qui ont perdu leur valeur primitive ou qui, s'ils l'ont conservée, en ont pris parallèlement une autre qui permet de les utiliser comme de véritables outils grammaticaux »[37].

Démonstratifs أسماء الإشارة → **Code : 520**

أُولاَءِ – أُولاَئِكَ – أُولاَكَ – أُولاَلِكَ – أُولَى – أُولَى لِكَ – تَا – تَاكِ – تَانِ – تَانِكَ – تَانِّكَ – تَلْكِ – تِلْكَ – تِلْكِ – تِهِ – تِهْ – تِي – تِيكَ – تِيكِ – تَيْنِ – تَيْنِكَ – تَيْنِّكَ – ذَا – ذَاكَ – ذَانِ – ذَانِكَ – ذَانِّكَ – ذَلِكَ – ذِلْكِ – ذِهِ – ذِهْ – ذِي – ذِيكِ – ذَيْنِ – ذَيْنِكَ – ذَيْنِّكَ – هَؤُلاَءِ – هَاتَانِ – هَاتِهِ – هَاتِيكَ – هَاتَيْنِ – هَذَا – هَذَاكَ – هَذَانِ – هَذِهِ – هَذَيْنِ – هُنَا – هُنَاكَ – هُنَالِكَ – هَهُنَا

Interrogatifs أسماء الاستفهام → **Code : 521**

أَنَّى – أَيُّ – أَيَّانَ – أَيْنَ – كَمْ – كَيْفَ – مَا – مَاذَا – مَتَى – مَنْ – مَنْ ذَا

Noms de condition أسماء الشرط → **Code : 522**

أَنَّى – أَيُّ – أَيَّانَ – أَيْنَ – أَيْنَمَا – حَيْثُمَا – كَيْفَمَا – مَا – مَتَى – مَنْ – مَهْمَا

Relatifs الأسماء الموصولة → **Code : 523**

[37] R. Blachère, *Grammaire de l'arabe classique*, 1952, p.277.

أَيُّ – اَلْ – الأُلاَءِ – الأُلَى – التِي – الذِي – الذِينَ – اللاَّءِ – اللاَّئِي – اللاَّتِ – اللاَّتِي – اللَّتَانِ – اللَّتَيْنِ – اللَّذَانِ – اللَّذَيْنِ – اللَّوَاتِي – ذَا – مَا – مَنْ

Interjections (Noms de verbes) أسماء الأفعال **→ Code : 524**

أُفٍّ – إِلَيْكَ – أَمَامَكَ – آمِينَ – آهِ – أَوَّهْ – إِيهِ – بَجَلْ – بَخٍ – بُطْآنَ – بَلْهَ – حَذَارِ – حَيَّ – حَيَّهَلَ – دُونَكَ – رُوَيْدَ – زِهْ – سُرْعَانَ – شَتَّانَ – شُكَانَ – صَهْ – عَلَيْكَ – مَكَانَكَ – مَهْ – نَزَالِ – هَا – هَاءَ – هَاءِ – هَاؤُمُ – هَاؤُمَا – هَاؤُنَّ – هَيَّا – هَيْهَاتَ – وَا – وَاهًا – وَرَاءَكَ – وَيْ

Onomatopées أسماء الأصوات **→ Code : 525**

سَأْ – طَقْ – عَدَسْ – غَاقِ – قَبْ – كَخْ – كِخْ – نِخْ – هَلاَ – وَيْهِ

Noms de nombre أسماء العدد **→ Code : 526**

Nombres cardinaux الأعداد الأصليّة **→ Code : 5261**

عدد مفرد : أحد – واحد – واحدة – اثنان – اثنين – اثنتان – اثنتين – ثلاثة – ثلاث – أربعة – أربع – خمسة – خمس – ستّة – ستّ – سبعة – سبع – ثمانية – ثمانٍ – ثمانيَ – تسعة – تسع – عشرة – عشر – مئة – مائة – ألف – مليون – مليار – بليون – بليار – ترليون

عدد مركّب : أحدَ عشر – إحدى عشرة – اثنا عشر – اثنيْ عشر – اثنتا عشرة – اثنتيْ عشرة – ثلاثَ عشرة – ثلاثةَ عشر – أربعَ عشرة – أربعةَ عشر – خمسَ عشرة – خمسةَ عشر – ستَّ عشرة – ستّةَ عشر – سبعَ عشرة – سبعةَ عشر – ثمانيَ عشرة – ثمانيةَ عشر – تسعَ عشرة – تسعةَ عشر

عقود : عشرون – عشرين – ثلاثون – ثلاثين – أربعون – أربعين – خمسون – خمسين – ستّون – ستّين – سبعون – سبعين – ثمانون – ثمانين – تسعون – تسعين

عدد معطوف : واحد وعشرون – … – اثنيْن وعشرين – … – ثلاثة وثلاثون – … – أربعة وثلاثين – … – خمسة وأربعون – … – ستّة وخمسين – … – سبعة وثمانون – … – ثمانية وتسعين – …

Nombres ordinaux الأعداد الترتيبيّة **→ Code : 5262**

أَوَّلٌ – أُولَى – ثَانٍ – ثَانِيَةٌ – ثَالِثٌ – … – ثَامِنٌ – … – عَاشِرٌ – … – حَادِي عَشَر – حَادِيَة عَشْرَة – … – تَاسِعَ عَشَر – تَاسِعَة عَشْرَة – عِشْرُونَ – … – مِئَة – … – أَلْف – …

Nombres fractionnaires الأعداد الجزئيّة **→ Code : 5263**

ثُلُثٌ – ثُلُثٌ – رُبْعٌ – رُبُعٌ – خُمْسٌ – خُمُسٌ – … – عُشْرٌ – عُشُرٌ

Circonlocutifs أسماء الكناية **→ Code : 527**

ذَيْتَ – ذَيَّةَ – كَيْتَ – كَيَّةَ – كَذَا – هَكَذَا – كَمْ – فُلاَن – بِضْع – كَمِنٍ – كَأَيٍّ – كَأَيِّنْ – كَأَيْنْ – كَأْيِنْ – كَائِنْ

Adverbes الظروف → **Code : 528**

Sans entrer dans les détails des divergences terminologiques de la Tradition grammaticale arabe concernant les adverbes[38] (*mafʿûl fîh, ẓarf* pl. *ẓurûf, maḥall* pl. *maḥâll*, ou *ṣifa* pl. *ṣifât*), nous avons adopté une position d'ordre morpho-lexical. En effet, par adverbe, *ẓarf*, nous entendons tout nom-outil non déclinable (figé) *ẓarf mabniyy* et ayant une utilisation exclusivement adverbiale *lâ yufâriqu ẓ-ẓarfiyya* [ne pouvant être utilisé qu'adverbialement] et ce qu'il soit strictement figé au cas direct *mabniyy ʿalâ n-naṣb* (أَنَّى - عَوْضُ - قَطُّ -...) ou pouvant être précédé des prépositions مِن ou إِلَى comme (فَوْقَ - بَعْدَ - قَبْلَ - مَتَى -...).

Adverbes de temps :

إِذْ – إِذَا – أَيَّانَ – مَتَى – أَمْسِ – الآنَ – مُذْ – قَطُّ – عَوْضُ – بَيْنَا – بَيْنَمَا – رَيْثَ – رَيْثَمَا – لَمَّا – خِلاَلَ – كُلَّمَا

Adverbes de lieu :

ثَمَّ – ثَمَّةَ – حَيْثُ – أَمَامَ – قُدَّامَ – بُحَاهَ – قُبَالَةَ – إِزَاءَ – وَرَاءَ – خَلْفَ – تَحْتَ – دُونَ – فَوْقَ – حِينَ – عَلُ – عِنْدَ – نَحْوَ

Adverbes communs :

بَيْنَ –لَدَى – لَدُنْ – أَنَّ – قَبْلَ – قَبْلُ – بَعْدَ – بَعْدُ

Adverbes composés :

بَيْنَ بَيْنَ – صَبَاحَ مَسَاءَ – لَيْلَ نَهَارَ

Pronoms personnels الضمائر → **Code : 529**

« La classe des pronoms, nous dit Henri Fleisch, est un domaine particulier, en dehors du comportement normal de la langue et dans la morphologie nominale et dans la morphologie verbale »[39]

أَنَا – ي – إِيَّايَ – أَنْتَ – كَ – إِيَّاكَ – أَنْتِ – كِ – إِيَّاكِ – هُوَ – هُ – إِيَّاهُ – هِيَ – هَا – إِيَّاهَا

نَحْنُ – نَا – إِيَّانَا – أَنْتُمْ – كُمْ – إِيَّاكُمْ – أَنْتُنَّ – كُنَّ – إِيَّاكُنَّ – هُمْ – هُمْ – إِيَّاهُمْ – هُنَّ – هُنَّ – إِيَّاهُنَّ

أَنْتُمَا – كُمَا – إِيَّاكُمَا – هُمَا – هُمَا – إِيَّاهُمَا

Annectifs الأسماء الملازمة للإضافة → **Code : 530**

En arabe, la relation de subordination entre deux noms (un nom et son complément déterminatif) s'exprime par l'adjonction de l'élément subordonné (le complément de nom) à l'élément qu'il complète ; on dit alors que les deux termes sont en état d'annexion. Il existe en arabe un petit nombre de noms qui ne peuvent être utilisés que premier terme d'une annexion (il faut rappeler ici que la suffixation, pour les noms, est considérée comme un cas particulier de l'annexion). Les grammairiens arabes appellent ce type de noms الأسماء الملازمة للإضافة *al-ʾasmâʾ al-mulâzima li-l-ʾiḍâfa* [les noms faisant toujours partie d'une annexion]. Nous proposons d'appeler ***annectifs***, les noms ayant cette particularité.

كِلاَ – كِلْتَا – سِوَى – ذُو – أُولُو – وَحْدَ – كُلّ – بَعْض – غَيْر

Verbes fonctionnalisés الأفعال الجامدة → Code : 54

[38] Voir à cet effet, Moussaoui M., *Le circonstant de temps et de lieu dans le Coran*, Lyon, 2005, p. 25-42.

[39] Henri Fleisch, *Sur les pronoms personnels en arabe classique*, 1968, pp. 65-73, p. 70.

Verbes figés à l'accompli أفعال ملازمة لصيغة الماضي → **Code : 541**

Ce sont des verbes qui sont figés à l'accompli et ne peuvent être conjugués ni à l'inaccompli ni à l'impératif.

بِئْسَ – تَبَارَكَ – حَبَّ – حَبَّذَا – سَاءَ – سُقِطَ – شَدَّمَا – طَالَمَا – عَسَى – قَصُرَمَا – قَلَّمَا – كَثُرَمَا – لَيْسَ – نِعْمَ – نِعِمَّا – هَدَّ

Verbes figés à l'impératif أفعال ملازمة لصيغة الأمر → **Code : 542**

Ces verbes sont figés à l'impératif.

تَعَالَ – تَعَالَيْ – تَعَالَيَا – تَعَالَوْا – تَعَالَيْنَ – هَاتِ – هَاتِي – هَاتِيَا – هَاتُوا – هَاتِينَ – هَبْ – هَلُمَّ – هَلُمَّا – هَلُمُّوا – هَلُمِّي – هَلُمِّينَ – هَيْتَ/هِيتَ

Verbes figés à l'inaccompli أفعال ملازمة لصيغة المضارع → **Code : 543**

Il n'y a qu'un seul verbe dans cette catégorie et qui est figé à inaccompli.

يَهِيطُ

Mots-Outils composés الأدوات المركّبة → Code : 55

La composition peut être avec ou sans agglutination (ou amalgame). Dans cette catégorie, nous avons donc regroupé trois ensembles de mots-outils composés ; un groupe sans agglutination réunissant les deux formes d'exclamation مَا أَفْعَلَ - أَفْعِلْ بِـ... , et deux groupes de mots-outils composés avec agglutination c'est-à-dire des composés dont les composants sont graphiquement soudés :

- le premier groupe comporte les mots-outils résultant de l'agglutination de certains mots-outils simples avec la particule مَا empêchant la rection مَا الكَافَّة عَنِ العَمَل *mâ al-kâffa ʿan al ʿamal* (أَيُّمَا - رُبَّمَا - كَأَنَّمَا - أَنَّمَا - إِنَّمَا)
- le deuxième groupe comporte les interrogatifs composés résultant de l'agglutination de certaines prépositions avec l'interrogatif simple مَا qui, de ce fait, subit une transformation en perdant son *ʾalif* final (لِمَ - بِمَ - مِمَّ - عَمَّ - فِيمَ - عَلاَمَ ...)

أَفْعِلْ بِـ... – مَا أَفْعَلَ
أَيُّمَا – رُبَّمَا – كَأَنَّمَا – أَنَّمَا – إِنَّمَا
لِمَ – لِمَهْ – بِمَ – بِمَهْ – مِمَّ – مِمَّهْ – عَمَّ – عَمَّهْ – فِيمَ – فِيمَهْ – عَلاَمَ – عَلاَمَهْ – حَتَّامَ – حَتَّامَهْ – إِلاَمَ – إِلاَمَهْ – كَيْمَ – كَيْمَهْ

Noms propres أسماء الأعلام → Code : 6

Le nom propre est un mot qui désigne autoréférentiellement un individu donné sous une appellation spécifique qui n'est pas commune aux individus de la même espèce.

Noms Propres de personnes أسماء الأشخاص → Code : 61

Le nom propre de personne est traditionnellement composé, en arabe, de cinq éléments : كُنْية *kunya* [nom patronymique], اسْم *ʾism* [prénom], نَسَب *nasab* [filiation], نِسْبَة *nisba* [adj. de relation] et لَقَب *laqab* [surnom]. La *kunya* est composée de *ʾabû* أبو [père de] ou de *ʾumm* أمّ [mère de] et d'un nom propre, généralement celui de la fille ou du fils aîné. L'*ʾism* est le prénom de la personne proprement dite, il peut être composé de *ʿabd* عبد [serviteur, esclave] suivi de l'un des 99 qualificatifs de Dieu. Le *nasab* est la partie dans laquelle sont mentionnés le père et les aïeux qui se suivent liés par le mot ابن *ʾibn* [fils de] (quand il y en a plusieurs, à partir du deuxième, ابن *ʾibn* est écrit بن *bn*). La *nisba* est un adjectif de relation se terminant par ي *î* et qui indique le pays, la ville ou le village, origine de la personne ; il peut aussi indiquer le métier. Enfin, le *laqab* qui se place généralement après la *nisba* est un surnom tiré d'un trait physique ou d'un titre honorifique.

Exemple : أبو جعفر محمّد بن أحمد بن محمّد الصَيْمَرِيّ

Noms Propres de lieux أسماء الأماكن → Code : 62

Exemple : سُرّ مَن رَأَى (سَامَرَّاء)

Noms de tribus, groupes et nations أسماء القبائل والأمم والفرق → Code : 63

Exemple : بنو تميم

Théonymes الأسماء الدينيّة → Code : 64

Exemple : الله – الكريم

Noms des œuvres أسماء الكتب → Code : 65

Ce sont des noms simples ou composés de plusieurs unités qui correspondent à des entités culturelles ou à des productions littéraires, théologiques, scientifiques, philosophiques ou autres.

Exemple : إصلاح المنطق

Noms composés الأسماء المركّبة → Code : 7

Exemple : شقائق النعمان

Références bibliographiques

ʿAbd Al-Masîḥ Georges Mitrî et Tâbrî Hânî Georges (1990), *Al-Ḫalîl : Muʿjam muṣṭalaḥât an-naḥw al-ʿarabî*, Maktabat Lubnân, Beyrouth

ʿAmâyra ʾIsmâʿîl ʾAḥmad et as-Sayyid ʿAbd al-Ḥamîd Muṣtafâ (1988), *Muʿjam al-ʾadawât wa-ḍ-ḍamâʾir fi-l-Qurʾân al-kaîm*

Bâqî (al-) Mohammad Fuʾâd ʾabd (1987), *Al-Muʿjam al-mufahras li-ʾalfâẓ al-Qurʾân al-kaîm*

Bernet Charles (1983), *Le vocabulaire des tragédies de Jean Racine. Analyse statistique*, Slatkine/Champion, Genève/Paris.

Blachère Régis (1997), *Eléments de l'arabe classique*, Maisonneuve et Larose, Paris, 4e édition (1ère édition 1946)

Brunet Etienne (1981), *Le vocabulaire français de 1789 à nos jours*, Slatkine/Champion, Genève/Paris.

Dichy Joseph (1990), *L'écriture dans la représentation de la langue : la lettre et le mot en arabe*, Thèse pour le Doctorat d'Ētat (ès Lettres), Université Lumière-Lyon2

Dichy Joseph et Ammar Sam (1999**),** *Les verbes arabes*, collection Bescherelle, Paris, Hatier, décembre 1999 (édition bilingue arabe-français)

Donat Vernier S. J. (1892), *Grammaire arabe, composée d'après les sources primitives*

Fleisch Henri, *Encyclopédie de l'Islam*, article ḤARF

Ġalâyînî (al-) Muṣṭafâ (1992), *Jâmiʿ ad-durûs al-ʿarabiyya*, 26e édition

Hamzé Hassan (1987), *Les théories grammaticales d'Az-zajjâjî* , Thèse de Doctorat d'Etat (és- lettres), Université Lumière-Lyon 2

Hamzé Hassan (1994), "Les parties du discours dans la tradition grammaticale arabe", dans : *Les classes de mots. Tradition et perspectives*, sous la direction de Louis Basset et Marcel Pérennec, Presses Universitaires de Lyon, pp. 93-115

ʾIbn Hišâm Al-ʾAnṣârî (1997), *Muġnî al-labîb ʿan kutub al-ʾaʿârîb*

Mayaffre Damon (2005), « Les corpus politiques : objet, méthode et contenu. Introduction », dans : *Corpus* [En ligne], 4 | 2005, mis en ligne le 05 septembre 2006, consulté le 8 mai 2017. URL : http://corpus.revues.org/292

Meschonnic Henri (1991), *Des mots et des mondes. Dictionnaire, encyclopédies, grammaires, nomenclatures*, Hatier, Paris

Muller Charles (1992-a), *Initiation aux méthodes de la statistique linguistique*, Slatkine/Champion, Genève/Paris.

Muller Charles (1992-b), *Principes et méthodes de statistique lexicale*, Collection « Unichamp » n° 33, Honoré Champion Ğditeur, Paris, Réimpression de l'édition Hachette de 1977.

Murâdî (al-) Al-Ḥasan ʾIbn al-Qâsim (1992), *Al-Janâ ad-dânî fî ḥurûf al-maʿânî*

Pottier Bernard (1962), *Systématique des éléments de relation : étude de morphosyntaxe structurale romane*, Klincksieck

Rastier F. (2001). *Arts et sciences du texte*. Paris, Puf

Rastier F. (2005). "Enjeux épistémologiques de la linguistique de corpus", dans : G. Williams (éd.), *La*

linguistique de corpus. Rennes : Pur, pp. 31-45. [En ligne sur *Texto !* (http://www.revue-texto.net/Inedits/Rastier/ Rastier_Enjeux.html)]

Rey Alain (1977), *Le lexique : Images et modèles. Du dictionnaire à la lexicologie*, Armand Colin, Paris

Roman André (1983-a), *Etude de la phonologie et la morphologie de la Koinè arabe*, 2 vol, Publication de l'Université de Provence, Aix-en-Provence, Jeanne Laffite, Marseille.

Roman André (1983-b), "Sur la constitution des unités de la langue arabe", dans : *Travaux du Cercle linguistique d'Aix-en-Provence*, vol. I, pp. 113-144.

Roman André (1983-c), "La genèse de la langue arabe et la constitution de ses formes verbales", dans : *La linguistique appliquée à la langue arabe*, C.E.R.E.S - Université de Tunis, Série Linguistique N° 5, pp. 29-45, Tunis.

Roman André (1988), "Les formes infinitives de l'arabe dans l'organisation générale de la langue arabe" dans *: L'infinitif*, sous la direction de Sylvianne Rémi-Giraud, Presses Universitaires de Lyon, pp. 211-242.

Roman André (1994), "Genèse et typologie de la langue arabe" dans : *Les classes de mots. Tradition et perspectives*, sous la direction de Louis Basset et Marcel Pérennec, Presses Universitaires de Lyon, pp. 117-147.

Roman André (1999), *La création lexicale en arabe, ressources et limites du système de nomination d'une langue humaine naturelle*, Presses Universitaire de Lyon, Lyon.

Roman André (2001), *Systématique de la langue arabe*, Université Saint-Esprit de Kaslik, Kaslik-Liban.

Salem André (1982), "Analyse factorielle et lexicométrie : synthèse de quelques expériences", dans : *Mots*, N°4, Mars, pp. 147-168.

Salem André (1987), *Pratique des segments répétés. Essai de statistique textuelle*, Klincksieck, Paris.

Suyûṭî (Al-) (1998), *al-Muzhir fî ʿulûm al-luġat wa ʾanwâʿihâ*, édition scientifique de Fuʾâd ʿAlî Mansûr, Dâr al-kutub al-ʿilmiyya, Beyrouth.

Tammâm Ḥassân (sans date), *Al-luġa l-ʿarabiyya. Maʿnâhâ wa mabnâhâ*, Dâr aṯ-Ṯaqâfa, Casablanca.

Ṭâš Kubrâ Zâda (ʾAḫmad Ibn Muṣṭafâ) (sans date), *Miftâḥ as-saʿâda wa-miṣbâḥ as-siyâda fî mawḍûʿât al-ʿulûm*, Dâr al-Kutub al-ʿilmiyya, Beyrouth

Tournier Maurice (1980), "D'où viennent les fréquences de vocabulaire", dans : *Mots*, N°1, octobre, pp. 189-209.

Tournier Maurice (1981), "Spécificité politique et spécificité lexicale", dans : *Mots*, N°2, mars, pp. 5-10.

Tûnjî (al-) Muhammad (1995), *Muʿjam al-ʾadawât an-naḥwiyya*

Zajjâjî (al-) ʾAbû l-Qâsim ʿAbd Ar-Raḥmân Bn ʾisḥâq (1986), *Kitâb ḥurûf al-maʿânî*, Muʾassasat ar-Risâlat et Dâr al-ʾAmal, Beyrouth et ʾIrbid (Jordanie), deuxième édition