

Approche linguistique et statistique de la variation apophonique du verbe trilitère sain en arabe standard

A Linguistic and statistical approach of the apophonic variation of the sound trilateral verb in standard Arabic

Hicham OUARDI *

Université Mohamed V de Rabat,

FLSH, Maroc,

hicham.ouardi@flsh.um5.ac.ma

Reçu le : 23/09/2021 Accepté le : 12/12/2021 Publié le : 31/12/2021

Résumé :

Le présent papier se dresse comme objectif de quantifier le phénomène de la variation apophonique. Une question épineuse qui occupe depuis des siècles une place centrale dans les annales portant sur la morphophonologie verbale .

En effet, par le biais d'une approche statistique et linguistique, nous essayerons de montrer la représentativité des variables contenues dans un corpus de 1816 verbes trilitères sains et de présenter en suite une interprétation linguistique reposant sur les prémisses de la théorie de la marque en phonologie. Notre hypothèse de base soutient le fait qu'une quantification limpide de la distribution de ces formes trilitères est susceptible de révéler tout un ensemble de corrélats sous-tendant leurs structures morphophonologiques.

Mots clés : Morphophonologie ; Corpus ; voyelle lexicale ; arabe standard ; matrice des traits

Abstract:

The present paper sets out to quantify the phenomenon of apophonic variation. This is a thorny issue that has occupied a central place in the annals of verbal morphology for centuries .

* Auteur correspondant.

Indeed, through a statistical and linguistic approach, we will try to show the representativeness of the variables contained in a corpus of 1816 sound trilateral verbs and to present a linguistic interpretation based on the premises of the markedness theory in phonology. Our basic hypothesis is that a clear quantification of the distribution of these trilateral forms is likely to reveal a set of correlates underlying their morphophonological structures..

Key words: Morphophonology; Corpus; lexical vowel; standard Arabic, feature matrix

1/ Introduction

Il est connu dans la littérature très abondante de la morphologie flexionnelle de l'arabe classique qu'il est des questions délicates se rapportant essentiellement à l'étude de formes verbales dupliquées dites « rédupliquées », des verbes hamzés* et des alternances apophoniques liées aux conversions temporelles et modales [1], [2], [3], [4] . Pour des raisons méthodologiques, nous concentrerons notre intérêt sur la question de la variation apophonique tout en limitant le champ de notre intervention autour des verbes dits forts. Ce choix se veut inévitable vu la complexité qui surgirait dès que l'on tente d'appréhender la classe verbale. Pour ce faire et en vue de mieux saisir la logique qui sous-tend le phénomène de l'apophonie dans le cadre de cette classe, nous essayerons de puiser les résultats statistiques que l'on a pu collecter sur la base d'un corpus implémenté en fouillant les matériaux linguistiques dans le fameux dictionnaire LISSAN AL ARABE d'IBN MANDOUR. Puis, nous tenterons de présenter une interprétation linguistique guidée par les prémisses théoriques de la théorie de la marque en phonologie.

2/ Matériels et méthodologie

La modélisation puis la mise en exergue du corpus occupent une place centrale dans notre réflexion. Nous estimons que l'étude de la variation de voyelle lexicale ne peut passer qu'à travers une analyse des données intelligiblement échantillonnées, classées et enfin traitées. La constitution d'un corpus de données attestées est conçue d'être la tâche la plus fatale dans la quête linguistique. En réalité, un corpus bien ficelé doit permettre avant tout de représenter de façon exhaustive les variables nécessaires à la description et à l'explication du phénomène en question

* Initié par le coup de glotte /ʔ/ tel qu'il est illustré dans les exemples : /ʔalifa/ (Prendre l'habitude de) , /saʔala/ (interroger) , /bada ʔa/ (Initier), entre autres.

et de témoigner aussi la validité des résultats obtenus. En effet, pour mieux appréhender la question des alternances apophoniques, il est jugé inévitable de saisir et les différentes observables constituant la charpente du phénomène et la tendance des processus phonologiques les régissant.

À cet effet, nous nous intéresserons dans la section suivante à une description brève des outils engagés dans notre étude empirique ainsi que les procédés méthodologiques mis en exergue.

2.1/ Les outils de sauvegarde et d'analyse statistique

Les matériaux linguistiques constituant les variables d'observation sont stockés dans une base de données usant du gestionnaire MySQL. Les tables sont élaborées en respectant certains principes et contraintes méthodologiques rigoureux. Ainsi, chaque table réunit les informations correspondant à un seul type d'objet. Par exemple, la table des traits phonétiques étiquetant chacun des sons radicaux collecte toutes les informations relatives aux sons de la parole sans que ses données soient reproduites dans une autre table. De surcroît, chaque rubrique de cette table correspond à une propriété unique de cet objet. On ne peut saisir qu'une seule fois une donnée dans la rubrique et l'information y correspondant n'est disponible que si cette table destinée à l'annotation existe. Le cas échéant, l'étude statistique des cooccurrences prenant comme critère les données phonétiques ou phonologiques semblent impossible en l'absence de cette table. D'autant plus, la même information ne doit jamais être redondante. Dans la table des radicaux verbaux, nous optons pour la segmentation des racines trilitères en vue de faciliter l'importation et l'exportation des données et pour focaliser sur les verbes initiés par le coup de glotte, il suffit de se rendre à la table des radicaux et sélectionner le critère de tri pour extraire facilement les variables manifestant. Cette opération même est possible via le gestionnaire MySQL en introduisant le code d'extraction requis .

Tablant sur le volet numérique, nous opterons pour les traitements à l'appui du logiciel SPSS 17.0 dédié aux analyses statistiques les plus complexes. Ce choix nous permettra, en fait, de saisir les caractéristiques quantitatives et qualitatives de la variation apophonique et de relever les résultats ce qui donnera lieu à la possibilité de construire des graphiques plus représentatifs.

2.2/ *Nature des données collectées*

Dans le cadre de ce travail, il est question de constitution et d'analyse d'un large corpus de structures verbales trilitères. Pour ce faire, la collecte des données a puisé différentes ressources lexicales du dictionnaire monumental d'IBN-MANDHOUR ce qui nous a permis de mettre la main sur un thésaurus d'usage polyvalent et nous avons pu collecter 1816 racines verbales trilitères saines. En réalité, le macro-corpus fait état de 19 455 entités appartenant à différentes classes lexicales (verbe, nom, adjectif, préposition, entre autres). Après l'étape de vérification et de brassage des données, nous avons gardé 10 011 entités réparties selon leurs sources textuelles et leurs origines linguistiques. Puis, nous avons échantillonné les verbes selon le nombre la nature des consonnes formant leurs radicaux. Cette opération nous a fourni une liste de 1816 racines verbales trilitères saines.

2.3/ *Classification des données*

Les entrées collectées sont classées dans une matrice comme suit : les lignes représentent les entrées lexicales trilitères arabes et empruntées et les colonnes indiquent respectivement l'origine, le nombre de caractères, les traits phonologiques et phonétiques ainsi que d'autres paramètres linguistiques utiles tel le degré de sonorité. Subséquemment, cette matrice fait état du statut de chaque segment par rapport aux traits ainsi que d'autres informations susceptibles de contextualiser l'entrée lexicale ; tel qu'il est montré dans les tables N°1 et N°2.

Table N°1. Illustration graphique du mode de classification du corpus.

جاء الجاء الجوار فقط من غير الوضيمز وأ. يميز. والجمع خولوا. وكان عظيم جالسا ج																											
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table N°2. Illustration graphique du mode d'annotation des radicaux verbaux.

verbe	C1	C1_phon1	C1_phon2	C2	C1_phon1	C1_phon2	C3	C1_phon1	C1_phon2	NATURE	FORMEflex	symetrie
أَب	ء	Glottale	occlusive	ب	Labiale	occlusive	ب	Labiale	occlusive	مهموز الءاء و مضصف	فعل-يُفَعِّلُ	c2c3
أَنْب	ء	Glottale	occlusive	د	Dentale	occlusive	ب	Labiale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَنْب	ء	Glottale	occlusive	د	Dentale	occlusive	ب	Labiale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَرْب	ء	Glottale	occlusive	ر	Dentale	Liquide	ب	Labiale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَرْب	ء	Glottale	occlusive	ر	Dentale	Liquide	ب	Labiale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَشَب	ء	Glottale	occlusive	ش	Palatale	Fricative	ب	Labiale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَلَب	ء	Glottale	occlusive	ل	Dentale	Liquide	ب	Labiale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَث	ء	Glottale	occlusive	ث	Interdentale	Fricative	ث	Interdentale	Fricative	مهموز الءاء و مضصف	فعل-يُفَعِّلُ	c2c3
أَث	ء	Glottale	occlusive	ث	Interdentale	Fricative	ث	Interdentale	Fricative	مهموز الءاء و مضصف	فعل-يُفَعِّلُ	c2c3
أَح	ء	Glottale	occlusive	ح	Palatale	occlusive	ح	Palatale	occlusive	مهموز الءاء و مضصف	فعل-يُفَعِّلُ	c2c3
أَرْج	ء	Glottale	occlusive	ر	Dentale	Liquide	ج	Palatale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَزج	ء	Glottale	occlusive	ز	Sifflante	Fricative	ج	Palatale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَز	ء	Glottale	occlusive	ز	Sifflante	Fricative	ج	Palatale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0
أَغ	ء	Glottale	occlusive	غ	Pharyngale	Fricative	ع	Pharyngale	Fricative	مهموز الءاء و مضصف	فعل-يُفَعِّلُ	c2c3
أَزح	ء	Glottale	occlusive	ز	Sifflante	Fricative	ح	Pharyngale	Fricative	مهموز الءاء	فعل-يُفَعِّلُ	0
أَله	ء	Glottale	occlusive	ل	Dentale	Nasale	ع	Pharyngale	Fricative	مهموز الءاء	فعل-يُفَعِّلُ	0
أَد	ء	Glottale	occlusive	ب	Labiale	occlusive	د	Dentale	occlusive	مهموز الءاء	فعل-يُفَعِّلُ	0

Cette représentation détaillée a le mérite de dévoiler systématiquement le contraste qui existe entre chacune des classes de segments et de mettre la main sur les dissemblances structurales perçues dans le cadre de la même classe de sons. Aussi, elle permet de distinguer les différentes classes verbales sur la base de la nature et la position des segments formants le radical et le chemin apophonique suivi lors de la flexion (/faʕala/-/jaʕʕalu/, /faʕala/-/jaʕʕulu/, etc.)

2.4/ Méthode de travail

Dans le cadre de ce papier, nous adopterons une approche stochastique des formes verbales appartenant à la classe du verbe trilitère et des schèmes output, jadis connues dans la littérature morphologique de l'arabe classique sous l'appellation «schèmes morphophonologiques». À notre sens, ce choix judicieux semble utile à détecter les occurrences ce qui permettrait d'identifier les corrélats morphophonologiques pesant sur la structure des verbes trilitères dits forts. Cette alternative méthodologique, à notre égard, est jugée fructueuse tantôt sur l'axe théorique que sur l'axe empirique .

D'autant plus, l'annotation des entités phono-lexicales repose sur l'exploitation des tables élaborées en vue de traiter la structure des segments des racines empruntées et celles qui sont arabes. Nous avons établi quinze tables permettant l'étiquetage des matrices conçues pour le traitement des racines arabes ou empruntées. Ces mêmes tables sont importées à la base de données. La table des traits phonétiques et phonologiques reste l'épine dorsale de l'annotation de nos corpus. Cette matrice regroupe les graphèmes arabes ainsi que les sons correspondants bien que d'autres informations d'ordre linguistiques (les traits phonétiques, le degré de sonorité du segment, entre autres) et organisationnelles (l'identifiant du graphème, le niveau d'abord du trait selon l'architecture proposée par la géométrie des traits, parmi d'autres informations permettant la bonne gestion et l'exploitation optimale de cette matrice exhaustive).

Pour examiner les racines collectées, nous avons opté pour un test de la productivité portant sur les formes arabes uniquement. L'idée est explicitée comme suit : si les trois segments de la racine arabe ne sont pas sujets à des contraintes de localité et/ ou de nature de la consonne, ils n'engendreraient que des formes arabes. Autrement dit, une racine arabe génère plus de racines arabes que de racines empruntées et vice versa .

3/ Résultats

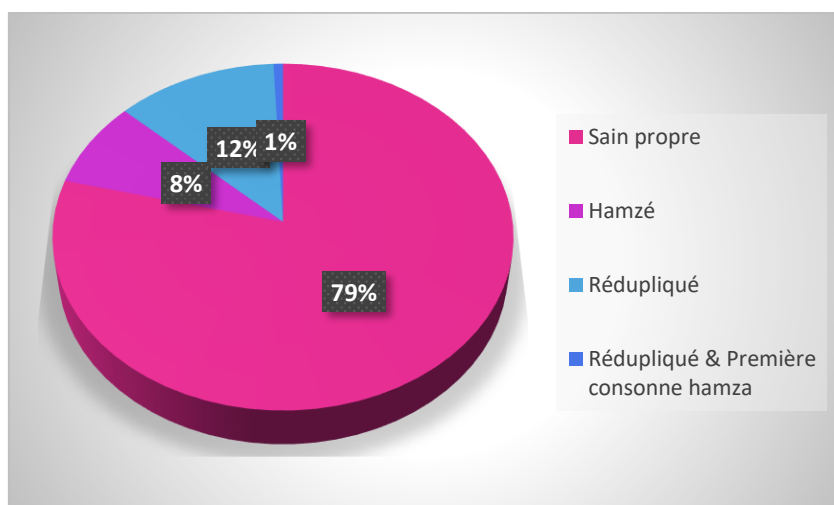
La grammaire de l'arabe classique manifeste une forte attraction aux formes non marquées. La considération de la classe du verbe trilitère sain affirme cette réalité (Tableau N°3).

3.1/ Distribution des sous-classes du verbe trilitère sain

Table N°3. Illustration numérique de la fréquence des sous-classes du verbe trilitère sain en arabe classique.

Sous-classes du verbe trilitère		Somme	Fréquence	
Sain propre		1433	78,91%	
Hamzé	Première consonne est hamza	76	146	4,19%
	Deuxième consonne est hamza	26		1,43%
	Troisième consonne est hamza	44		2,42%
Dupliqué		224	12,33%	
Rédupliqué & Première consonne hamza		13	0,72%	
		1816	100,00%	

Figure 1. Représentation graphique des sous-classes du verbe trilitère sain en arabe classique.



À bien fouiller les données collectées de ces statistiques, il paraît que les verbes trilitères sains propres dominent incontestablement cette catégorie (78,91%). La sous-classe du verbe redupliqué (12,33%) paraît moins contrainte par rapport aux sous-classes hamzé (8,04%) et redupliqué et hamzé en même temps (0,72%). Ceci dit que la morphologie de cette langue autorise la succession de deux segments consonantiques plus qu'elle tolère les formes hamzées. Aussi, le cumul

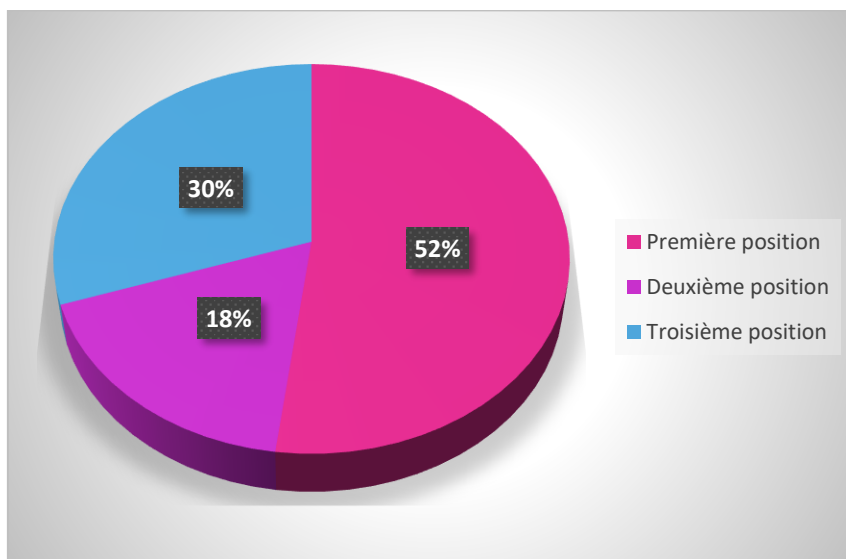
des deux anomalies morphologiques incarnées par la conjonction « redupliqué&hamzé » se passe beaucoup moins récurrente par rapport à toutes les autres classes dérivées ; même les sous-classes du verbe hamzé (dont la première consonne est un hamza (4,19%), dont la troisième est un hamza (2,42%), dont la deuxième consonne est un hamza (1,43%)).

Une lecture rapide des données dévoile sans conteste la dominance des verbes ne manifestant ni le coup de glotte ni la reduplication d'une consonne dans l'une ou l'autre position du schème verbal. Par ailleurs, l'abondance des schèmes manifestant une consonne redupliquée par rapport à ceux affichant un coup de glotte se veut significative sur le plan phonologique. Le système morphophonologique a tendance à neutraliser les formes complexes.

3.2/ Distribution du coup de glotte (Hamza) dans la classe du verbe trilitère sain.

Pour mieux comprendre l'impact de l'émergence du coup de glotte dans la structure du verbe trilitère sain, nous tenterons de retracer l'occurrence de la forme verbale en prenant en considération la distribution du hamza dans le schème verbal (Figure 2).

Figure 2. Distribution du coup de glotte dans la classe du verbe trilitère sain.

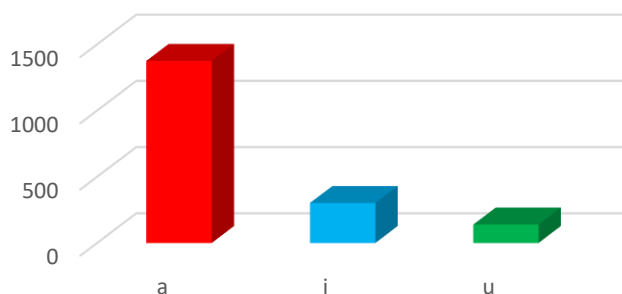


En survolant les données, il semble que la position d'attaque est moins marquée par rapport aux autres positions du fait que la moitié des verbes manifestant un coup de glotte font état du hamza dans la position initiale (52% des verbes recensés). Par ailleurs, la syllabe contenant la voyelle lexicale paraît plus contrainte que les deux autres positions. Cela s'explique par le fait que cette syllabe rend compte des données linguistiques d'ordre phonologique et lexical (seulement 18% des radicaux verbaux manifestent un coup de glotte dans cette localité). Quant à la troisième position, elle manifeste le tiers des candidats contenus dans le corpus (30% des radicaux trilitères sains).

3.3/ *Distribution des segments vocaliques dans la syllabe dite lexicale.*

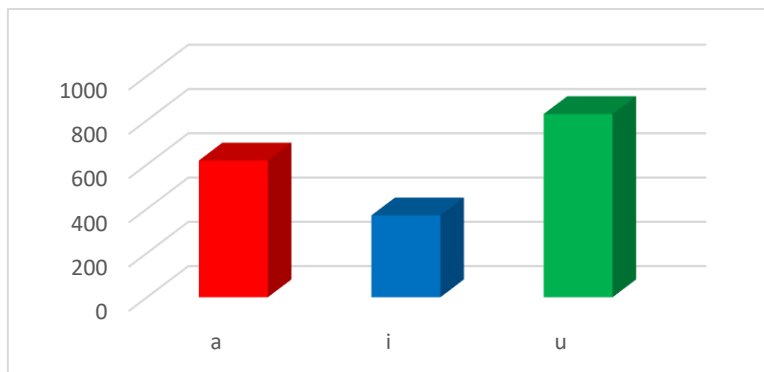
D'autre part, le chemin apophonique reste en étroite relation avec la distribution des trois segments vocaliques [a], [i] et [u]. Nous essayerons dans le cadre de cette section de quantifier le nombre des observables selon l'occurrence des de ces trois voyelles.

Figure 3. *Distribution des segments vocaliques dans la syllabe lexicale. Cas de l'imperfectif.*



Les formes trilitères contenant la voyelle [a] semblent dominantes dans le cas de l'imperfectif (1375 verbes). Une fréquence moins abondante pour les cas des segments [i] (302) et [u] (139 formes verbales). Ce résultat prévisible paraît en parfaite adéquation avec le chemin apophonique explicité par

Figure 4. Distribution des segments vocaliques dans la syllabe lexicale. Cas du perfectif.



Du côté de la flexion au perfectif, la distribution des segments vocaux démontre sans aucune réticence que le choix d'un tel ou tel itinéraire phonologique n'est pas lié uniquement aux propriétés intrinsèques des trois voyelles. D'autres paramètres d'ordre linguistiques interviennent dans la formation du schème verbal ([a] (618 verbes), [i] (370 verbes) et [u] (828 verbes)).

4/ Discussion

Avant de procéder au traitement des différents paradigmes vocaliques, il incombe de survoler les propriétés idiosyncrasiques des segments qui sont à la base du système vocalique de l'arabe classique.

Table N°4. Propriétés intrinsèques du triangle vocalique de l'AC

	/a/	/i/	/u/
Haut	-	+	+
Bas	+	-	-
Antérieur	-	+	-
Arrondi	-	-	+
Rond	-	-	+

[1] Source: *Brandão de Carvalho*, 1997, 16.

Le segment vocalique /u/ semble le plus affecté par ces traits en contraste aux autres éléments vocaliques de la matrice. La hiérarchie /a/ < /i/ < /u/ émerge avec une netteté frappante. En effet, la spécification de ces propriétés intrinsèques est faite conformément à la notion de marque telle qu'elle est conçue de la littérature du SPE. En revanche, la théorie de la marque présente une nouvelle optique selon laquelle la spécification des traits et aussi le traitement des structures morphologiques répondent aux exigences d'une hiérarchisation universelle telle qu'elle est illustrée en infra.

Formule 1. *Hiérarchie des contraintes de marquage portant sur les segments vocaliques de l'arabe classique.*

*[+rond], *[+arrière], *[+haut] >> *[-rond], *[-arrière], *[-haut]

L'avantage d'une telle hiérarchisation est qu'elle permette tout d'abord de rendre compte l'inégalité segmentale illustrée ci-dessous.:

Formule 2. *Hiérarchie partielle des contraintes de marquage portant sur les segments vocaliques de l'AC.*

*u >> *a, *i

Cet état de choses est prédictible en nous basant, tout simplement, sur une hiérarchisation partielle des contraintes. Il importe de noter que la mise en exergue d'une telle hiérarchisation est basée uniquement sur les traits phonétiques de V-Place. D'autre part, elle est en mesure d'établir une hiérarchisation entre les deux segments /a/ et /i/. Ainsi, le conflit des contraintes paraît sensible à la hiérarchisation illustrée infra.

Formule 3. *Hiérarchie partielle des contraintes de marquage portant sur les segments vocaliques /a/ et /i/*

a >> *i*

La discrimination de ces deux segments est à la base d'une opposition construite sur la dimension de la hauteur. Le segment vocalique /a/ étant affecté par le trait [-haut] est conçu comme marqué par la contrainte *[+haut]. Aussi le segment /i/ est marqué par la

contrainte *[-haut] puisqu'il est affecté par le trait *[-haut]. La hiérarchie établie est, alors, facilement inductible grâce à la relation de dominance stricte conçue entre ces deux traits : *[+Haut] >> *[- Haut].

Ces différentes hiérarchies partielles permettent entre autres de prédire l'occurrence de la voyelle lexicale dans le schème verbal. Ainsi, le segment vocalique [a] tant qu'il est moins marqué paraît plus prédictible que les deux autres segments. Dans le cas de l'imperfectif, ce résultat semble harmonieux avec les données empiriques. Toutefois, l'abondance du segment [u] dans le corpus lié aux verbes fléchis au perfectif dévoile sans réserve l'intervention de paramètres d'ordre non phonologique à l'occurrence des formes verbales trilitères.

5/ Conclusion

À l'issu de ce travail, nous avons tenté aussi de quantifier le phénomène de la variation apophonique. À notre sens, l'analyse empirique a l'avantage non plus de cerner le champ de validité des résultats ; mais aussi de nous montrer à quel point la généralisation est-elle adéquate. Autrement dit, les schèmes de variations apophoniques modélisés et discutés depuis les premiers exégètes arabes jusqu'aux phonologues contemporains retracent six probabilités énigmatiques mettant en jeu des matériaux d'ordre phonétique (la classe phonétique de la consonne lexicale), syntaxique (la nature grammaticale du verbe transitif Vs intransitif), sémantique (verbes décrivant un état constant ou éphémère, entre autres). Par ailleurs, à travers la quantification il semblerait évident qu'il est bien des classes verbales restreintes qui puissent prendre un itinéraire morphologique confus (faʕila>jafʕulu ou jafʕalu).

En dépit de ce que la hiérarchie des contraintes que nous avons pu établir soit partielle et restreinte, son efficacité à rendre des questions telles que l'inventaire syllabique et le statut du hamza dans les formes impératives des verbes trilitères sains reste salubre. Par ailleurs, il serait judicieux de continuer à cet égard pour mieux peaufiner cette réflexion

enfin de couvrir les différents paradigmes du verbe trilitère et redresser une hiérarchie totale de l'arabe classique. Une fois celle-ci est établie, l'abord des autres phénomènes morphologiques et phonologiques connexes serait aussi systématique que fluide.

6/ *Références*

- M. Taki, «structure syllabique de l'arabe classique et représentation des voyelles longues,» chez *recherches en linguistique arabe*, Casablanca, publications de la faculté des lettres et des sciences humaines Ben'Msik, 1995.
- T. Scheer et A. Chekayri, «La provenance apophonique des semi-voyelles dans les formes verbales en arabe classique,» *Langues et Linguistique*, pp. 15-54, 1998.
- H. Ouardi, «Le degré de complexité syllabique : une apostille apodictique,» chez *La lecture syllabique : Fondements théoriques*, 2021, ISBN : 978-9920-230-9 .
- H. Ouardi, «Pour une reconnaissance automatique des formes trilitères allogènes en arabe,» *INTERNATIONAL CONFERENCE ON LANGUAGE AND CURRENT ISSUES - Issue 1 : "Language and Education"*, Laboratoire Valeurs, Société et développement, Université Ibn Zohr, Agadir., pp. 272-304, 2018.
- J. Brandão de Carvalho, «PRIMITIVES ET NATURALITE,» *Langages*, vol. 125, p. 14-34, 1997.
- [6] J. Prince, *Generative Grammar*, New york: Press, 1993.
- B. Bruce, «The Crossing Constraint and 'Backwards Languages,» *Natural Language & Linguistic Theory*, vol. Vol. 7, n° 4, pp. pp. 481-549, 1989.