

Bohas Georges, *Matrices, Étymons, Racines : Éléments d'une théorie lexicologique du vocabulaire arabe*.

Peeters (coll. *Orbis / Supplementa*, tome 8),  
Leuven-Paris, 1997. 16 × 24 cm, 207 p.

Pour une grande majorité d'arabisants, ce qu'il est convenu d'appeler la « théorie bilitéraliste » n'a guère bonne presse. On l'associe le plus souvent à un débat périmé, ou à des efforts plus récents, mais trop marginaux pour avoir pu réellement remettre en question l'idée selon laquelle le lexique de l'arabe repose massivement sur des racines trilitères indécomposables, dont chacune est, en principe, porteuse d'une charge sémantique propre ; cette idée, héritée de la tradition linguistique arabe, a été de surcroît confortée par un certain nombre de facteurs qui ont tendu, de longue date, à l'institutionnaliser : l'organisation des dictionnaires, l'enseignement de la morphologie (celui-ci n'étant d'ailleurs pour une large part qu'une introduction à l'usage de ceux-là), et jusqu'à la tendance de certains de nos grands anciens à célébrer, avec un enthousiasme que l'on peut juger parfois excessif, la « rigueur mathématique » que l'existence de racines trilitères conférerait au lexique de l'arabe.

Parallèlement à cette conception, en quelque sorte officielle, on sait qu'il existe cependant des faits, assez nombreux du reste, qui tendraient à suggérer, à l'inverse, que « certains mots trilitères en arabe ont été manifestement amenés à la trilitarité d'une manière artificielle, en partant d'une base bilitère » (H. Fleisch, *Traité de philologie arabe*, vol. I, p. 256). Certains cas, particulièrement repérables, avaient déjà été relevés par les philologues arabes médiévaux ; d'autres ont pu être mis au jour dans le cadre du comparatisme sémitique ou chamito-sémitique. On sait également que, sur la base de ces observations, des tentatives récurrentes ont eu lieu pour établir le caractère foncièrement bilitère du lexique de l'arabe, et, plus généralement du sémitique, sans qu'aucune ait réellement réussi à s'imposer.

Tout en intégrant les acquis de ces travaux, et en s'appuyant sur les développements les plus récents de théorie générative en morpho-phonologie, l'ouvrage de G. Bohas renouvelle radicalement cette problématique ; le modèle théorique qu'il met en œuvre présente tout à la fois un degré d'abstraction plus élevé que les tentatives précédentes, et un pouvoir explicatif plus large, en cela qu'il permet de rendre compte, sur la base de principes identiques, d'un nombre important de faits considérés jusqu'à présent comme indépendants les uns des autres.

Ce modèle, exposé globalement dans l'« Introduction » (p. 1-9), est élaboré exclusivement à partir du lexique de l'arabe : bien qu'il présente, à l'évidence, des implications au niveau du sémitique et du chamito-sémitique, l'A. a préféré, pour des raisons de méthode, s'en tenir à une seule

langue connue pour l'ampleur de son vocabulaire, qui, de ce fait, constitue un terrain particulièrement propice. Ce lexique est abordé dans une perspective « panchronique » (au sens de Hagège et Haudricourt, *La phonologie panchronique*, Paris, 1979) : il ne s'agit pas de reconstituer un « état primitif » de l'arabe, mais de mettre en lumière le système sous-jacent qui organise son lexique, ce système n'étant pas spécifique à une étape particulière de son développement. Toutefois, précise l'A., ce système se laisse plus directement appréhender à travers les travaux des lexicographes arabes médiévaux et les dictionnaires bilingues qui en sont issus (notamment celui de Kazimirski), qui fournissent un corpus plus complet que les dictionnaires modernes. Que ce lexique soit, pour une part importante, constitué de vocables rares ou tombés de l'usage n'est pas, dans cette perspective, une objection réhabilitrice : selon l'A., l'organisation du lexique intériorisé par les locuteurs contemporains (et enregistré par les dictionnaires modernes) n'est pas en soi différente de celui qui se laisse dégager de « l'immense lexique littéraire de l'arabe » ; celui-ci en donne simplement une image plus complète.

Cette organisation, selon l'A., est constituée de trois niveaux hiérarchiquement ordonnés : au-delà de la racine, l'A. met en lumière l'existence d'« étymons » (noté e), combinaisons non ordonnées (*i. e.* réversibles) de deux consonnes, associés à une charge sémantique ; ainsi, l'étymon *bt* se réalise dans *batta*, *barata*, *batara*, *balata*, *tabba*, qui contiennent tous la notion de « couper ». Toutefois, le niveau de l'étymon apparaît encore trop peu abstrait pour rendre compte de certaines séries, par exemple celle que forment *šabara*, *ḡabba*, *ḡafrun*, *ṭaffa*, *ṭafana*, *zaffa*, et *rabaṭa*, qui expriment tous l'idée de « lier, entraver » ; bien que ces verbes ne comportent pas deux consonnes communes, ils ont tous en commun de comporter une labiale (*b* ou *f*) et une dentale emphatique (*ṣ*, *ḡ*, *ṭ* ou *ẓ*). Ceci amène l'A. à postuler l'existence d'un niveau supérieur, celui de la « matrice » (notée m), définie dans un premier temps comme une combinaison de deux points d'articulation, également non ordonnée.

Quant à la racine proprement dite, elle n'a pas de statut à proprement parler dans ce modèle : selon l'A., qui se réfère aux développements récents en morpho-phonologie dans le cadre génératif, le point de départ (*input*) des règles morphologiques et phonologiques est non la racine, mais le radical, défini comme une séquence ordonnée de trois consonnes (les deux consonnes de l'étymon élargies par la reduplication de l'une d'entre elles ou par l'ajout d'une autre consonne), portant au moins une voyelle qui lui est attribuée par le lexique. Autrement dit, dans ce modèle, la racine « n'est pas assez abstraite pour le niveau 2 [celui de l'étymon – JPG] et l'est trop pour le niveau 3 [celui du radical – JPG] » (p. 9). Un tel dédain pour la racine pourra sembler choquant aux yeux de certains, même si une telle position rejoint en partie les idées des grammairiens arabes médiévaux. Il faut

toutefois signaler que ce point n'est pas réellement crucial, en ce sens que les conclusions essentielles de l'A. resteraient pertinentes, me semble-t-il, même dans le cas où l'on tiendrait à conserver son statut à la notion de racine trilitère.

Le chapitre I « Matrices et étymons », le plus long de l'ouvrage (p. 9-141), contient l'essentiel de la discussion. L'A. s'attache tout d'abord à mettre en lumière (p. 9-52) les processus d'étoffement qui, à partir d'un étymon bilitère, permettent de dériver un schème trilitère, la « diffusion » (qui produit un schème à redoublement, de type *madda*), et l'« incrémentation », c'est-à-dire l'ajout d'une consonne en première, deuxième ou troisième position ; un troisième processus, la « reduplication » (e.g. *bašša* / *bašbaša* « briller »), n'est qu'évoqué ici ; il sera abordé plus en détail au chapitre III.

Sur le premier de ces processus, la diffusion, l'A. s'attache à rendre compte de l'anomalie selon laquelle, en arabe comme dans la plupart des langues sémitiques, elle ne peut se faire que de la seconde à la troisième position ; autrement dit, à partir d'un étymon *md*, on peut avoir *madad*, mais jamais *mamad*. La discussion, hautement technique, n'en reste pas moins accessible au lecteur peu familiarisé avec les modèles théoriques mis en œuvre, grâce à la clarté de l'exposé et au grand sens pédagogique dont fait preuve l'A.

En ce qui concerne l'incrémentation, il établit que celle-ci met en jeu non seulement les glides *w* et *y* (e.g. *ḡāsa* / *ḡassa*, « espionner » ; *ḡālā* / *ḡalla*, « être élevé, imposant » ; *baṭṭa* / *wabaṭa* « être affaibli »), ainsi que l'avaient remarqué de nombreux arabisants et sémitisants, mais aussi les gutturales (e.g. *dakka* / *dahaka*, « piler, broyer » ; *šamma* / *šamaḥa* / *šama'a*, « frapper avec une pierre / un fouet / un bâton » ; *ṭarra* / *ṭaraḥa* « pousser devant soi, repousser ») et les sonantes (e.g. *ḡalla* / *raḡala* « être bas, vil » ; *ṭalla* / *naṭala*, « curer un puits »). Ces deux derniers cas sont toutefois plus complexes que celui des glides. D'une part, en effet, ils sont dans certains cas « liés à des modifications parfois importantes de la charge sémantique de l'étymon » (p. 41), ce qui laisserait à penser qu'il s'agirait, dans ces cas, non pas de simples incréments phonologiques dépourvus d'incidence sémantique, mais d'anciens marqueurs morphologiques dont la valeur se serait au moins en partie effacée ; à cet égard, l'A. établit un parallèle avec l'assyrien moderne, où ce type de phénomène apparaît de façon particulièrement visible. En second lieu, lorsqu'une sonante ou une gutturale apparaît dans une base triconsonantique, il ne s'agit pas nécessairement d'un incrément ; elle peut aussi bien faire partie de l'étymon. Ainsi, note l'A., le *m* de l'arabe peut soit faire partie de l'étymon, comme *tamma* (« achever »), soit être un extenseur d'origine morphologique comme dans *mašaḡa* (« mélanger »), soit encore d'un pur incrément phonologique comme dans *fašama* (« séparer », cf. *fašša*, *fašala*...).

Le système des étymons permet également d'isoler certains « noms-bases », qui ne lui sont pas intégrés ; tel est le cas de *kalb* (« chien »), qu'il apparaît impossible de rattacher aussi bien à l'étymon *k/* qu'à l'étymon *kb*. À ce propos, l'A., reprenant une hypothèse émise par Diakonoff (*Afrasian Languages*, Moscou, 1998) suggère que le *b* final pourrait être corrélé à la notion d'« animal sauvage et dangereux » (cf. notamment *di'b*, « loup » ; *dubb*, « ours » ; *lubb*, « lion » ; *ta'lab*, « renard »). Sur la base d'un sondage partiel dans Kazimirski, l'A. fait état d'une vingtaine de cas du même genre, qui s'ajoutent aux huit relevés par Diakonoff, tous étant extérieurs au système des étymons. L'idée selon laquelle le *b*, dans tous ces noms, serait le vestige d'un système très ancien de classificateurs nominaux peut surprendre ; il n'en reste pas moins que le nombre relativement important de cas qui viennent la conforter donne à réfléchir.

Quoi qu'il en soit, l'existence de noms primitifs extérieurs au système des étymons permet de comprendre certains faits de polysémie repérés de longue date : ainsi, *di'b*, nom primitif, n'a rien à voir avec *da'aba* (« chasser, éloigner ») qui est, quant à lui, une réalisation de l'étymon *db* (cf. *ḡabba*, « éloigner »), bien que le système traditionnel des racines les classent tous deux ensemble (et séparent *da'aba* de *ḡabba*).

Est abordée ensuite (p. 53-64) la question du non-ordonnancement des étymons ; ainsi, l'étymon *bt* se réalise aussi bien par *batta* (« couper, retrancher ») que par *tabba* (même sens). Sur les 325 paires possibles en arabe, l'A. relève que 135 présentent une liaison sémantique forte quel que soit l'ordre de réalisation (un inventaire en est donné en appendice au chapitre, p. 95-141) ; à cela s'ajoutent les paires présentant une relation sémantique lâche, dont l'inventaire fait l'objet d'une thèse en cours sous la direction de l'A.

Les pages suivantes (64-70) s'attachent à montrer que le système ainsi mis au jour permet de résoudre un certain nombre d'anomalies apparentes à la contrainte bien connue selon laquelle les trois éléments d'une même racine doivent appartenir à des points d'articulation différents. En simplifiant un peu les choses, on dira que cette contrainte, selon l'A. (qui s'appuie sur les développements récents de la phonologie plurilinéaire) ne s'applique qu'entre éléments de même statut. Ainsi l'anomalie apparente constituée par *fašama* (« couper, retrancher »), où *f* et *m* sont tous deux des labiales, se dissipe si l'on considère que cette forme est une réalisation de l'étymon *fš* (voir *fašša* « extraire une chose d'une autre, séparer » ; *tafašfaša* « se séparer » ; *fašala* « séparer »), et que *m* est un incrément, donc de statut différent. Ce principe est ensuite testé systématiquement (p. 70-79) à propos des paires incluant *b*, ce qui permet également à l'A. de préciser les relations sémantiques entre les différentes réalisations d'un même étymon, ou plus précisément les diverses modalités selon lesquelles chacune d'entre elles développe et spécifie la charge sémantique propre à l'étymon.

La discussion concernant les matrices (p. 79-85) est relativement brève ; toutefois, de nombreux éléments ont déjà été fournis dans les pages précédentes, ou le seront par la suite. L'A. base essentiellement son argumentation sur la matrice constituée d'une coronale (*t, ṭ, d, s, z, š, ʒ, ʃ, ɖ, ʈ*) et d'une uvulaire (*q* et *ğ*, lui-même issu d'un ancien phonème \*G, correspondant voisé de *q*), et correspondant à la charge sémantique « couper, trancher ». En s'en tenant simplement à l'ordre uvulaire - coronal, il ne relève pas moins de 45 formes (verbes ou noms primitifs) présentant indiscutablement cette charge sémantique, avec des modalités diverses. Après une brève « Conclusion » (p. 85-86), le chapitre se clôt par une « Discussion » (p. 86-94), qui, reprenant un certain nombre de travaux récents s'inscrivant dans le cadre de l'hypothèse bilittère, s'attache à montrer que le modèle proposé par l'A. permet un traitement plus efficace des faits. Les p. 94-141, enfin, donnent une liste détaillée des étymons réversibles, comme on l'a dit plus haut.

Reprenant un problème classique de la phonologie historique de l'arabe, le chapitre II, « Les sources du *ǧīm* » (p. 143-152), développe l'une des implications de la théorie des matrices : selon l'A., le *ǧīm* arabe proviendrait tantôt d'un phonème noté *G*, qui serait l'équivalent voisé du *q*, tantôt d'un *g*, équivalent voisé de *k*. Cette hypothèse permettrait de rendre compte de nombreux cas de « polysémie » comme celui que présente le verbe *ǧadda*, qui signifie à la fois « s'appliquer, travailler avec ardeur » et « couper, tailler » ; en fait, il y aurait là deux verbes différents à l'origine, le premier étant issu d'un étymon *gd*, réalisation d'une matrice [vélaire, coronal] à laquelle on peut également rattacher, par exemple, *kadda* (« s'épuiser à force de travailler »), alors que le second proviendrait d'un étymon *Gd*, réalisation de la matrice [uvulaire, coronal], porteuse de la charge sémantique « couper, retrancher », à laquelle se rattachent également *qadda*, *qaṭṭa*, *qaṭa'a* etc. Bien que les faits, plus complexes, conduisent en définitive l'A. à nuancer quelque peu ce tableau, l'arabisant généraliste retiendra surtout de ce chapitre la réponse particulièrement ingénieuse et élégante qu'il apporte à la question de la polysémie de certaines racines.

Le chapitre III, « Quelques aspects des quadrilittères » (p. 153-163) reprend en le développant l'argument fondé sur la liaison sémantique forte entre certaines racines à redoublement et les quadrilittères correspondantes (comme *bazza* et *bazbaza*, « enlever, emporter »), déjà présenté au chapitre I. Sont également envisagées, dans ce cadre, les formes présentant une troisième consonne identique à la première, ainsi *zalla* (« glisser, broncher »), *zaliza* (« s'agiter continuellement ») et *zalzala* (« secouer, agiter »).

Le chapitre IV, enfin, « Les trilittères sains » (p. 165-187), aborde les racines dites « saines », (c'est-à-dire celle qui ne sont pas redoublées et ne comportent pas de sonantes), qui constituent l'argument traditionnel en faveur de la notion de racine triconsonantique. L'A. s'attache à montrer que ces racines, dans de nombreux cas, peuvent

se ramener à des étymons biconsonantiques, non seulement par adjonction d'une obstruante en troisième position, mais aussi par adjonction d'un *s* ou d'un *t* en première ; dans ce cas, il s'agirait, selon l'A., d'anciens préfixes. Dans d'autres cas, il s'agit de trilittères formés par croisement, comme *batta* / *takka* / *bataka* qui ont tous trois le sens de « couper, retrancher » ; l'A. montre à ce propos que ce genre de croisements fonctionne également au niveau des étymons, et constitue un autre procédé d'étoffement, à côté de l'incréméntation.

Cette analyse, nécessairement succincte, aura à tout le moins permis - du moins l'espère-t-on - de mettre en lumière le caractère profondément novateur du modèle proposé par G. Bohas, ainsi que le nombre et la richesse des pistes nouvelles qu'il ouvre à la réflexion. Sans doute, et l'A. en convient volontiers, de nombreuses implications de ce modèle restent-elles à préciser ou à élaborer, aussi bien dans le cadre d'une description du lexique de l'arabe (inventaire systématique des matrices et des étymons), qu'à celui d'un élargissement au sémitique, mais également des questions plus générales touchant au statut et à la nature mêmes du signe linguistique. Il n'en reste pas moins que l'abondance et la solidité des arguments empiriques apportés ici à l'appui du modèle suffisent amplement à en établir la pertinence, et conduisent à une révision radicale de ce qu'il faut bien appeler, désormais, la *doxa* trilittéraliste.

Jean-Patrick Guillaume  
Université Paris III