

Quel corpus pour la TME ?

Bruno Paoli

▶ **To cite this version:**

Bruno Paoli. Quel corpus pour la TME ?. Langues et Littératures du Monde Arabe, 2018. halshs-02092137

HAL Id: halshs-02092137
https://shs.hal.science/halshs-02092137

Submitted on 7 Apr 2019

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.

Quel corpus pour la TME ?

Bruno Paoli, ICAR (UMR 5191)

Il est incontestable que la théorie des matrices et des étymons (TME), depuis bientôt trois décennies, a permis des progrès considérables dans le domaine de la lexicologie arabe, en montrant, d'une part, que les racines trilitères, loin de constituer le stade ultime de l'organisation du lexique de la langue, étaient elles-mêmes la résultante de l'expansion de noyaux binaires constitués de deux phonèmes, les étymons, et, à un niveau plus abstrait encore, de la combinaison de deux matrices de traits distinctifs, elle-même appelée matrice ; et, d'autre part, que les matrices, en nombre limité, étaient phonétiquement motivées, avec pour conséquence de battre en brèche le postulat structuraliste de l'arbitraire du signe. Ainsi, la matrice {[nasal], [+continu]} est associée à l'invariant notionnel « le nez » et englobe les champ sémantique du souffle et des odeurs ; et il n'est pas besoin d'un dessin pour saisir le rapport « mimophonique » entre le nez et la nasalité, d'une part, et entre le souffle et la continuité, d'autre part.

La TME a permis de comprendre et d'expliquer de nombreux cas d'homonymie (Bohas et Saguer 2014) ; mais elle bute encore sur le problème de la synonymie. Dans mon étude sur le lexique arabe des odeurs (Paoli 2015), j'avais par exemple identifié pas moins de soixante-dix-sept verbes signifiant « sentir mauvais », sans pour autant pouvoir dégager des principes permettant d'établir entre eux des nuances sémantiques qui seraient en rapport, par exemple, avec la source, la cause ou l'intensité de l'odeur. Si de telles nuances ont pu exister, elles ont de longue date été enfouies sous des siècles, sinon des millénaires, de changements linguistiques. La question du « sens » des

augments, les consonnes ajoutées aux étymons pour former des radicaux triconsonantiques, fait toujours débat : jouent-ils le rôle d'affixes spécifiant d'une manière ou d'une autre le sens de l'étymon bilitère ? Ou bien sont-ils de simples « bouche-trous » qui n'apportent rien de plus sémantiquement ? Dans ce même numéro, Salem Khchoum montre de manière convaincante que le *ʿayn* final a souvent valeur d'intensif ; mais Jean-Claude Rolland montre pour sa part que le *mīm* final, dans la plupart des cas, ne modifie pas le sens de l'étymon qu'il développe. Le débat est ouvert... et loin d'être clos !

J'en reviens aux soixante-dix-sept verbes signifiant « sentir mauvais ». Pourquoi une telle profusion ? La première idée qui vient à l'esprit est de considérer une partie au moins de ces verbes comme des variantes dialectales, propres aux dialectes de telles ou telles tribus arabes ; l'arabe « classique » serait, de ce point de vue, la somme des dialectalismes acceptés par les lexicographes arabes médiévaux et consignés dans leurs dictionnaires. Cette possibilité permettrait à l'évidence de résoudre, au moins en partie, le problème, à condition, bien sûr, de pouvoir identifier les variantes pouvant à coup sûr être considérées comme dialectales.

Il arrive que celles-ci soient explicitement présentées comme telles par les lexicographes arabes, généralement au moyen de l'expression « *X luġatun fī Y* », comme dans ce passage du *Lisān al-ʿarab* qui dit : « *saniḫa d-duhnu wa-ṭ-ṭaʿāmu wa-ġayru-humā sanaḫan : taġayyara, luġatun fī zaniḫa yaznaḫu iḏā fusida wa-taġayyarat rīḥu-hu* » ; autrement dit, *saniḫa* est une variante de *zaniḫa* dans le sens d'« être gâté et changer d'odeur », en parlant de la graisse (*duhn*) ou de la nourriture (*ṭaʿām*). Plus rarement, la tribu qui utilise cette variante est mentionnée, comme dans cet autre passage du *Lisān al-ʿarab* : « *wa-l-sūdadu al-šaraf, maʿrūf, wa-qad yuhmazu wa-tuḍammu ad-dāl, ṭāʾiyya. Al-Azharī : as-suʾdudu bi-ḍamm ad-dāl al-ʾūlā luġatu Ṭayyʾ* » ; autrement dit, le mot *sūdad*, qui signifie « noblesse » (*šaraf*) peut être prononcé avec une *hamza* et avec une voyelle *u* sur le premier *d*, soit *suʾdud*, ce qui, d'après al-Azharī, est le cas dans le dialecte de la tribu des Ṭayyʾ.

Un certain nombre de ces variations dialectales ont été nommées par les auteurs arabes médiévaux, comme la *ʿanʿana* (ʾ > ʿ), la *kaskasa* (k > s) ou encore la *kaškaša* (k > š). Dans un article à paraître, Bohas et Rolland ont identifié un grand nombre de témoins lexicalisés de ce que les grammairiens arabes ont identifié sous le terme de *kaškaša*, autrement dit, de couples de termes de même sens qui ne diffèrent que par l'alternance entre *kāf* et *šīn*. Ces doublons témoignent aussi de la dimension pluri-dialectale de l'arabe des dictionnaires classiques. D'autres, nombreuses et variées, et qui n'ont pas nécessairement été

nommées, sont également attestées dans les dictionnaires. Je vais bientôt y revenir.

La TME s'est construite grâce au dépouillement d'une base de données, baptisée *Kazimiro*, qui reproduit le corpus du dictionnaire arabe-français de Kazimirski (1860). Si ce dictionnaire est sans aucun doute le meilleur dictionnaire bilingue de l'arabe classique, il a cependant aussi ses limites, en ce sens qu'il ne reproduit pas, loin de là, l'ensemble des données et des informations fournies par les dictionnaires arabes classiques sur lesquels il est basé. C'est le cas, notamment, de ces informations dialectales que je viens d'évoquer ; mais aussi de certaines entrées ou racines. Je n'en donnerai qu'un exemple, relatif aux mauvaises odeurs.

Les principaux étymons olfactifs peuvent être rattachés à la matrice {[nasal], [+ continu]}. C'est le cas de {m, š} (*šamma, šamšama, šamā, rašima*), {n, š} (*našara, našaqa, našā, nušūʾ*), {n, s} (*ʾasina*, *saniḫa*), {n, ṣ} (*ʾaṣanna, ṣaniqa*), {n, l} (*laḫina*), {n, ṯ} (*ṯanita, ṯatina, naṯita, ṯadina, ʿaṯina*), {m, ḫ} (*ḫamma, ḫašama, ḫaṭama, ḫamiǧa, taḫammara, ḫamaṭa, zaḫima, saḫḫama, ʾašḫama*), {n, ḫ} (*istaḫanna, ḫaniza, ḫazana, daḫana, zaniḫa, saniḫa*) et bien d'autres encore, au sujet desquelles je renvoie à l'inventaire exhaustif dressé dans Paoli (2015).

Mais il est un fait que le verbe couramment utilisé aujourd'hui pour dire « sentir mauvais », *natuna*, comporte deux nasales et une consonne non-continue, ce qui ne permet pas de le rattacher, sous cette forme, à la matrice {[nasal], [+continu]}. Une solution consisterait à considérer le verbe *natuna* comme une variante dialectale d'un hypothétique **naṯuna*. La transformation des interdentales en dentales est bien attestée dans les dialectes arabes anciens et modernes. D'après Jean Cantineau (1960, p. 44), elle est caractéristique des parlers sédentaires, tandis que dans les parlers nomades, les interdentales sont généralement conservées.

Or, si l'on s'en tient au dictionnaire de Kazimirski, qui n'a pas d'entrée [nṯn], ce verbe n'existe pas. Voulant en avoir le coeur net, je suis retourné aux dictionnaires arabes et, pour ma plus grande satisfaction, j'ai découvert que le verbe *naṯana* existe bel et bien. Ibn Manẓūr, dans le *Lisān al-ʿarab*, en donne la définition suivante : « *naṯana al-laḥmu naṯnan wa-naṯanan : taġayyara* » ; autrement dit *naṯana* se dit de la viande pour dire qu'elle a « changé », sous-entendu que « son odeur a changé » (*taġayyarat rāʾiḥatu-hu*). L'existence de ce verbe rend donc plausible l'hypothèse selon laquelle l'occlusive dentale médiane de *natuna* serait issue de l'interdentale correspondante, rendant possible son rattachement à un étymon {n, ṯ} sous-jacent.

Une paire de verbes semblable pourrait recevoir une même analyse : il s'agit des verbes *ḏafira* et *dafira*, qui ont respectivement le sens de « sentir fort (aisselles) » et « sentir mauvais ». Mais il arrive aussi que les interdentales soient transformées en dentales continues (ṯ > s et ḏ > z). Les deux alternatives peuvent aussi coexister, comme c'est le cas dans les dialectes syro-libanais : on dit par exemple *sawra*, de *ṯawra* (révolution), mais *tōr*, de *ṯawr* (taureau) ; et cette alternance ne paraît pas prédictible. L'adjectif orphelin *zifir*, « qui sent mauvais », pourrait être un exemple de cette seconde alternative, et les trois racines ([*ḏfr*], [*dfr*] et [*zfr*]) pourraient être considérées comme des instanciations d'un même étymon {f, ḏ}, issu de la matrice {[labial], [continu]}.

Les lexicographes arabes sont avares d'informations « dialectales ». Mais il se trouve que l'article consacré à la racine [dfr] dans le *Lisān al-ʿarab* en contient une, qui ne concerne pas la dentale initiale, mais la troisième consonne radicale : « *al-dafr*, dit Ibn Manẓūr, c'est [comme] *al-dafʿ*, [en] yéménite (*yamaniyya*) ». À la racine [ḏfr], on trouve ceci : « *al-ḏafr* [désigne] la mauvaise odeur (*natn*) [...] Ibn Sīdah a dit : nous avons mentionné [dans ce même sens ?] *al-dafr*, avec un *dāl* »... Sans plus de précisions. Quant à Ibn Fāris, l'auteur du *Maqāyīs al-luġa*, il nous informe qu' « ils disent une fois *ṯatinat* et une fois *ṯanitat* ». De qui parle-t-il ? Des Arabes, sûrement ; mais s'agit-il alors d'une variation libre ou d'une variation dialectale ? Nous n'en saurons pas plus. Nous en retiendrons tout de même l'idée que, d'après Ibn Fāris, le verbe *ṯatina* n'est pas « fondamental » (*al-ṯāʾ wa l-tāʾ wa l-nūn laysa aṣlan*) et donc, implicitement, qu'il est issu de la permutation des « lettres » de *ṯanita*, tout comme *naṯita*, lui aussi oublié par Kazimirski, mais attesté chez al-Ǧawharī (*Ṣiḥāḥ*), Ibn Manẓūr (*Lisān al-ʿarab*) et surout al-Fayrūzabādī (*Al-qāmūs al-muḥīṭ*), qui dit : « *naṯita al-laḥmu qalbu ṯanita* », autrement dit : « *naṯita* [se dit de] la viande et est [issu de] la permutation de *ṯanita* ».

Salloum (1986, p. 79-112) a dressé l'inventaire des alternances phonologiques explicitement qualifiées de dialectales par les lexicographes arabes. Il en a identifié pas moins de soixante-quinze, dont certaines paraissent naturelles (ṯ/t, ḏ/d, q/k, b/f, b/m, s/z, s/ṣ, l/r, etc.), d'autres moins (q/f, z/n, s/t, etc.). Un examen critique approfondi de ces données reste à faire ; et il ne fait guère de doute qu'il permettra d'expliquer de nombreux cas de synonymie.

Les dictionnaires arabes permettent donc sur bien des points de compléter les données recueillies dans le dictionnaire de Kazimirski, soit en ajoutant des racines qui ont été oubliées par ce dernier ; soit parce qu'ils contiennent des informations utiles sur les dialectes qui permettent de comprendre que nombre de cas de synonymie peuvent

s'expliquer par des variations dialectales. Autrement dit, si la base Kazimiro constitue un bon outil de travail, il est nécessaire, pour aller plus loin, de la compléter avec les données bien plus riches et complètes que contiennent les dictionnaires arabes.

RÉFÉRENCES

BOHAS, Georges et SAGUER, Abderrahim, 2014, *The Explanation of Homonymy in the Lexicon of Arabic*, Lyon, ENS Editions.

KAZIMIRSKI Arthur de Biberstein, 1860, *Dictionnaire arabe-français*, Paris, Maisonneuve.

PAOLI, Bruno, 2015, « Le lexique arabe des odeurs », *Bulletin d'études orientales* 64, p. 63-97.

Les dictionnaires arabes classiques consultés sont disponibles en ligne, par exemple sur le site d'al-bāḥiṯ al-ʿarabī : http://www.baheth.net. Cette base de données permet de faire des recherches dans les ouvrages suivants :

Ibn Fāris, *Maqāyīs al-luġa*.

Fayrūzabādī (al-), *Al-Qāmūs al-muḥīṭ*.

Ǧawharī (al-), *Al-ṣiḥāḥ fī al-luġa*.

Ibn Manẓūr, *Lisān al-ʿarab*.

Saġānī (al-), *Al-ʿubāb al-zāḫir*.