

Que
sais-je ?

L'ARABE



Djamel E. Kouloughli

puf

QUE SAIS-JE ?

L'arabe

DJAMEL EDDINE KOULOUGHLI

Agrégé de l'Université

Directeur de recherche au CNRS



Introduction

Les grandes langues de culture ont, chacune, une aventure qui lui est propre. Celle de l'arabe est assez singulière. À l'origine, c'est la langue de tribus de pasteurs nomades qui sillonnent les steppes et les déserts d'Arabie, ainsi que de quelques groupes sédentaires installés dans des agglomérations dont bien peu méritaient le nom de ville. Mais cet état de choses, qui existait probablement depuis des siècles, est brusquement bouleversé à partir du second tiers du vii^e siècle de l'ère courante à la suite du succès rencontré par *Muhammad Ibn bAbdillah* de la tribu des *Qurayš* de La Mecque qui prêche une nouvelle religion monothéiste, l'islam. Porté par les conquêtes islamiques l'arabe va, en quelques décennies, s'étendre sur un immense empire couvrant le Proche-Orient et l'Afrique du Nord et débordant sur la péninsule ibérique à l'ouest et sur l'Asie centrale à l'est. Dans les deux ou trois siècles qui suivent, cet idiome de Bédouins primitifs qui n'excellaient qu'en poésie, intègre, à travers une gigantesque entreprise de traduction, l'ensemble de l'héritage culturel, scientifique et philosophique des vieilles civilisations de la région et devient le moyen d'expression écrite de milliers d'auteurs dans tous les domaines de la vie culturelle d'alors : théologie, droit, littérature, sciences, philosophie. Ayant recueilli et développé l'héritage de la culture grecque, il a joué un rôle décisif dans sa transmission à l'Occident et ainsi dans la préparation de la Renaissance européenne qui est le fondement de la culture moderne.

Comme langue parlée, l'arabe a donné naissance à une nombreuse descendance qui, recouvrant complètement les anciens domaines de l'araméen et du copte et partiellement celui du berbère, s'étend sur un territoire de plus de 13 millions de kilomètres carrés de la Mauritanie à l'Oman.

L'arabe est aujourd'hui la langue officielle de plus de 20 pays dont la population totale avoisinait les 300 millions au début du xxi^e siècle et pourrait atteindre les 500 millions en 2050. L'arabe est langue de travail de l'onu et des organisations internationales depuis le milieu des années 1970 du siècle dernier (la seule langue non indo-européenne avec le chinois) et est classée au 6e rang des langues du monde. Enfin, en tant que langue liturgique de l'islam, l'arabe

intéresse peu ou prou près d'un quart des êtres humains aujourd'hui.

Cet ouvrage se propose de retracer quelques-unes des étapes significatives de l'aventure historique de l'arabe telle que les connaissances actuelles permettent de se la représenter. Dans les limites de cet ouvrage, il ne pouvait être question de présenter un " état des connaissances " en la matière. On espère néanmoins esquisser un cadre global cohérent à l'intérieur duquel le lecteur intéressé pourra situer les recherches actuelles, recherches dont nous avons donné une bibliographie sommaire mais suffisante pour une première approche. Deux ouvrages de synthèse, relativement récents, ceux de Kees Versteegh (1997) et d'Ignacio Ferrando (2001), tous deux riches d'une abondante bibliographie, pourront apporter une information plus complète.

Mais même l'objectif modeste que nous nous proposons exige, pour pouvoir être visé de façon réaliste, de faire des choix, parfois difficiles, dans la masse des informations intéressantes. En général nous avons privilégié les renseignements relatifs à " l'histoire externe " de la langue, celle qui concerne son environnement social et culturel. Par contre, nous avons presque toujours évité de rentrer dans le détail de son " histoire interne " parce que celle-ci n'intéresse qu'un public limité de spécialistes et que les exposer de façon claire et complète aurait vite transformé la nature et les objectifs de l'ouvrage.

Par contre, pensant que le lecteur pourrait juger utile, voire indispensable, d'avoir une idée globale de la structure générale de l'arabe, nous avons présenté de façon synthétique l'essentiel de ce qu'il faut en savoir dans une annexe d'une douzaine de pages en fin de volume. Cette annexe peut être consultée de façon complètement indépendante et à n'importe quel moment du processus de lecture.

Chapitre I

L'arabe, langue sémitique

I. Les langues sémitiques

Les langues se laissent regrouper en « familles » sur la base de ressemblances systématiques entre, d'une part, leurs paradigmes grammaticaux, et, d'autre part, les éléments de leur lexique présumés « primitifs », c'est-à-dire les moins susceptibles d'avoir fait l'objet d'emprunts. Des ressemblances de cette nature, mises en évidence par des comparaisons méthodiques, systématisées à partir du xix^e siècle par la linguistique comparée, permettent de classer l'arabe dans la famille des langues sémitiques.

Cette famille est attestée historiquement, c'est-à-dire par l'écriture, depuis près de cinq millénaires dans l'actuel Moyen-Orient. La plupart des spécialistes l'incluent aujourd'hui, à côté d'autres familles de langues, dans un ensemble plus vaste, la « superfamille » (ou « phylum ») afro-asiatique. En l'état actuel des connaissances, on regroupe sous ce nom, outre la famille sémitique, au moins quatre familles de langues qui sont toutes réparties sur le nord et l'est du continent africain, à savoir : *l'égyptien*, attesté depuis le iii^e millénaire avant l'ère courante (désormais ec) dans la vallée du Nil, et dont le descendant moderne, le copte, a survécu jusqu'au xix^e siècle ; *le berbère*, attesté depuis au moins le ii^e siècle avant ec par des inscriptions dites « libyques », et dont le domaine, aujourd'hui morcelé, s'étendait sur tout l'ouest de l'Afrique du nord ; *les langues tchadiques*, parlées en Afrique centrale, et dont la plus importante est le hausa ; *les langues couchitiques*, dont le domaine couvre la « corne » de l'Afrique et se prolonge plus au nord le long de la mer Rouge, et dont les langues les plus importantes sont l'oromo et le somali. Les langues de ces deux dernières familles sont très nombreuses, et les spécialistes les subdivisent en plusieurs sous-groupes. Pour certains, les langues de l'ouest du domaine couchitique formeraient en fait une famille linguistique distincte, l'omotique,

que les uns proposent de rattacher, comme branche distincte, à l'afro-asiatique, alors que d'autres contestent même ce rattachement.

Ces hésitations révèlent assez que les critères linguistiques sur lesquels repose l'intégration d'une langue dans une famille de l'ensemble afro-asiatique restent assez flous. Ce flou résulte, pour une bonne part, du fait que nombre de ces langues sont encore très imparfaitement décrites. Mais il tient aussi au fait que les traits linguistiques sur lesquels on se fonde pour faire des regroupements sont encore peu nombreux et peu systématiques, voire contradictoires, suggérant parfois des rapprochements entre deux langues que d'autres critères conduisent à éloigner l'une de l'autre. Certains auteurs doutent même que le phylum afro-asiatique ait jamais eu un ancêtre commun, et attribuent les ressemblances entre les langues sémitiques et les langues africaines du phylum à l'effet plus ou moins profond de « vagues de sémitisation » exercées par contacts sur ces langues.

En comparaison avec les traits linguistiques qui rapprochent les langues afro-asiatiques, ceux qui unissent les membres de la famille sémitique sont beaucoup plus nombreux et systématiques, et ont été perçus bien avant la naissance de la linguistique comparée. En particulier, la parenté de l'hébreu, de l'araméen et de l'arabe était déjà parfaitement perçue par les savants juifs et musulmans au Moyen Âge. Elle fonde par exemple chez le grand philologue juif maghrébin Yehuda Ibn Qurayš (x^e siècle), l'intérêt porté à la connaissance des langues sœurs de l'hébreu pour expliciter certains termes bibliques. De même, le jurisconsulte musulman andalou 'Alī Ibn Hazm (xi^e siècle) n'hésite pas à affirmer [\[1\]](#) que « quiconque examine l'arabe, l'hébreu et le syriaque acquière la certitude [...] qu'il s'agit d'une seule et même langue à l'origine ».

En Occident, c'est l'orientaliste allemand Job Ludolf qui, en 1702, établit solidement la parenté des langues sémitiques connues à l'époque, en se fondant sur une analyse systématique de leurs ressemblances à la fois lexicales et grammaticales. Mais l'appellation de « langues sémitiques » est due à un autre allemand, August Schlözer, qui l'emploie pour la première fois en 1781.

Les premières langues reconnues comme sémitiques furent l'hébreu, l'araméen (et sa variante syriaque) et l'arabe. Au xix^e siècle, la famille s'élargit d'une part au phénicien, reconstruit grâce à l'hébreu dont il est proche, et d'autre part à l'éthiosémitique constitué par le ge'ez, langue classique d'Éthiopie, et les langues apparentées. Puis s'y ajoutât le sud arabe ancien, déchiffré en partie

grâce à l'éthiosémitique. Avec le déchiffrement de l'écriture cunéiforme par Henry Creswicke, entre 1833 et 1845, on découvrit la plus ancienne langue sémitique écrite, l'akkadien, et ses descendants, le babylonien et l'assyrien. Au ^{xx}^e siècle, la famille sémitique s'agrandit encore par la reconstitution (approximative) de l'amorite, connu seulement à partir de noms propres figurant dans des textes akkadiens, mais surtout par la découverte en 1928 de l'ougaritique, puis celle de l'éblaïte en 1975. À chacune de ces étapes, l'image que l'on se faisait de la famille sémitique, des relations entre ses membres et des caractéristiques linguistiques probables du « proto-sémitique », ancêtre commun supposé de ces langues, s'est trouvée plus ou moins profondément modifiée. Aujourd'hui le consensus est loin de régner sur l'ensemble de ces questions.

Longtemps l'opinion dominante fut que l'arabe était la langue sémitique la plus typique, la plus proche du proto-sémitique. Le fait qu'elle ait émergé le plus tardivement sur la scène de l'histoire ne constituait pas, aux yeux des tenants de cette thèse, une objection. Au contraire, postulant que le berceau primitif de tous les peuples sémitiques ne pouvait être que l'Arabie, ils considéraient que la dernière langue à avoir quitté ce sol originaire, n'en avait que mieux gardé les caractéristiques les plus « pures » de la famille. Les autres peuples sémites auraient quitté l'Arabie par vagues successives échelonnées dans le temps en direction de l'est, du nord et du sud, et se seraient sédentarisés dans l'ensemble du « croissant fertile », du Yémen et de l'Afrique orientale. Leurs parlers s'y seraient progressivement diversifiés, subissant les effets de substrats linguistiques divers et s'éloignant d'autant du prototype sémitique, conservé par l'arabe.

À l'appui de ce scénario, on évoquait un certain nombre d'arguments linguistiques dont les principaux sont les suivants : 1 *au niveau phonologique, l'arabe aurait conservé pratiquement intact le stock primitif des consonnes sémitiques, en particulier les interdentes, les emphatiques et les gutturales que de nombreuses langues sœurs (celles du nord notamment) auraient perdues ou confondues ; de même, l'inventaire vocalique de l'arabe, réduit aux trois voyelles cardinales a/ i et u (longues et brèves) refléterait mieux le système sémitique primitif que les systèmes plus complexes de certaines autres langues, souvent corrélatifs de la perte de consonnes gutturales.* 2 *Au niveau morphologique, l'arabe exhiberait dans toute sa pureté le système sémitique de formation des mots à partir de racines généralement triconsonantiques et de schèmes constitués de voyelles et éventuellement de consonnes ; la régularité de sa morphologie serait aussi une preuve de la vitalité des mécanismes hérités de*

la langue mère. 3 Au niveau syntaxique, il aurait conservé un système complet de marques casuelles (des déclinaisons) dont il ne resterait dans la plupart des autres langues que des vestiges ; corrélativement la structure, essentiellement synthétique, des énoncés dans cette langue correspondrait mieux au « génie » sémitique. 4 / Au niveau lexical, l'arabe, peu en contact avec d'autres langues, aurait conservé un fonds lexical sémitique primitif beaucoup plus riche que celui de ses sœurs périphériques.

La découverte de l'akkadien ébranla quelque peu ces convictions. L'impressionnante profondeur historique des textes qui représentaient cette langue et ses descendantes et l'importance de la culture matérielle sur laquelle elles reposaient ne pouvait pas ne pas en imposer. Des auteurs comme Alfred Von Kremer en Allemagne ou Ignazio Guidi en Italie soutinrent, dans le dernier quart du xix^e siècle, que les Sémites étaient originaires de Mésopotamie. Leurs arguments reposaient sur des faits de vocabulaire. D'après ces auteurs, le lexique primitif commun à toutes les langues de la famille ne reflétait pas l'environnement écologique ou géographique du désert, mais pointait plutôt vers des régions situées à l'extérieur de la péninsule arabique, quelque part plus au nord-est. Les adversaires de la thèse de l'Arabie berceau des Sémites ajoutaient que les conditions naturelles de la péninsule rendaient très improbable l'idée qu'elle ait pu, au moins avant la domestication du dromadaire, qui date seulement du ii^e millénaire avant ec, abriter assez d'habitants pour être le réservoir humain des vagues d'émigration successives qui auraient sémitisé ses marges. Le débat sur l'origine géographique des Sémites connut d'autres rebondissements au xx^e siècle, à la suite du développement des recherches sur l'histoire du peuplement humain de la planète. La plupart des auteurs qui abordent encore aujourd'hui ces questions pensent plutôt, maintenant, à une origine africaine, mais des scénarios contradictoires sont proposés à partir de cette hypothèse.

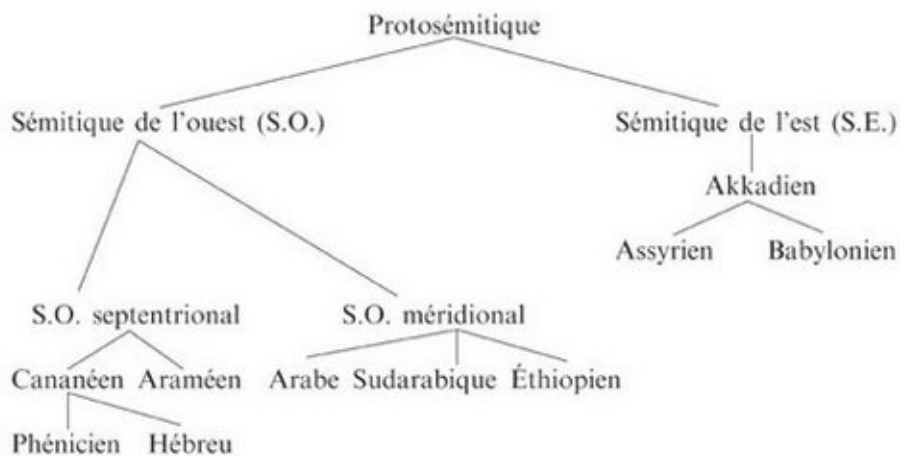
II. La classification des langues sémitiques

La conception des relations existant entre les diverses langues sémitiques a, elle aussi, beaucoup évolué à mesure que l'on en a su plus sur ces langues. La découverte de l'akkadien avait révélé que cette langue ne connaissait que des

conjugaisons à préfixes personnels, alors que toutes les langues sémitiques connues jusque-là avaient un système verbal reposant sur l'opposition d'un paradigme à préfixes de personnes (à valeur d'inaccompli) et d'un paradigme à suffixes de personnes (à valeur d'accompli). L'akkadien utilisait en outre une forme nominale, exprimant l'état, qui pouvait être spécifiée par des pronoms personnels suffixés. On ne tarda pas à faire l'hypothèse que cette forme, dite « stative », était l'ancêtre probable du paradigme à suffixes des langues sémitiques de l'ouest. Du coup était ébranlé le mythe de l'arabe prototype du sémitique ancien. L'akkadien, malgré les traits indéniables d'évolution qu'il présentait, notamment la perte de nombreuses consonnes gutturales et emphatiques, apparaissait comme ayant conservé des traits linguistiques proto-sémitiques perdus par les langues sémitiques « occidentales », y compris l'arabe. On se mit même à soupçonner que la très grande régularité morphologique de l'arabe, loin de refléter une exemplaire conservation du proto-sémitique, pouvait résulter d'un processus tardif de régularisation des paradigmes hérités, les éloignant ainsi d'états plus anciens que des langues moins idéalement régulières auraient mieux conservé. C'est notamment l'opinion de Robert Hetzron (1976).

La classification des langues sémitiques qui a prévalu pendant la plus grande moitié du ^{xx}e siècle a été essentiellement typologique. On a procédé par regroupements sur la base de traits linguistiques qui permettaient de constituer de grands groupes, qui se sont trouvés recouper la localisation spatiale des langues, d'où les dénominations géographiques de ces groupes. La première grande division opposait le « sémitique oriental », constitué par l'akkadien et ses langues filles, à toutes les autres langues sémitiques alors connues, regroupées en « sémitique occidental ». Cette division se fondait sur la différence d'organisation du système verbal, déjà évoquée, mais aussi sur un certain nombre de spécificités phonologiques et morphologiques de l'akkadien. Une seconde division opposait, au sein du sémitique occidental, le sémitique septentrional, regroupant en gros les langues sémitiques du Croissant fertile (Syrie, Liban, Palestine) et le sémitique méridional avec l'arabe, le sudarabique et les langues éthiopiennes. Cette division correspondait, selon les spécialistes, à l'existence de traits linguistiques propres au groupe méridional, par exemple la réalisation /f/ du phonème que toutes les autres langues sémitiques réalisent /p/, l'infexion d'une voyelle longue /ā/ pour former certaines classes de verbes « dérivés », ou encore l'usage étendu de pluriels nominaux dits « internes » (cf. Annexe p. 115). Dans le groupe sémitique du nord-ouest, on distinguait aussi deux ensembles : les langues « cananéennes », représentées par le phénicien et l'hébreu, et l'araméen avec ses diverses variétés. La classification traditionnelle

se présentait donc comme dans le tableau ci-contre.



Classification traditionnelle des langues sémitiques

Cette classification foncièrement géographique n'établit de relations de filiation qu'entre les variétés historiques d'une même langue. Elle a reçu cependant une interprétation historique qui, sans fonder une classification généalogique des diverses langues de la famille, permettait à tout le moins de se représenter globalement les grandes étapes de la constitution du domaine. Selon cette interprétation, la première grande division qu'aurait connue le proto-sémitique serait antérieure au ⁱⁱⁱ^e millénaire avant ec, et aurait vu se constituer la branche orientale du sémitique avec l'akkadien. Vers le ⁱⁱ^e millénaire avant ec, le sémitique occidental aurait connu une nouvelle séparation entre nord et sud. Le groupe du nord se serait, à son tour, fractionné un millénaire plus tard en deux groupes, les langues cananéennes et l'araméen. À peu près dans le même temps, le groupe du sud aurait connu une division comparable entre l'arabe et le sudarabique, lequel aurait essaimé sur le continent africain, donnant naissance au groupe éthiopien.

Mais ce scénario a été remis en cause à la fois par la découverte de nouvelles langues de la famille et par une analyse plus fine de celles qui étaient déjà connues.

La langue découverte sur le site d'Ougarit, non loin de l'actuelle Lattaquié en Syrie, et dont la documentation remonte au ^{xiv}^e avant ec, a révélé des traits la rapprochant de l'arabe : inventaire consonantique riche, avec en particulier les

consonnes interdentes, devenues occlusives, ou chuintantes dans les langues plus jeunes de la région, et flexion casuelle pour les noms, du moins au singulier, alors que les langues plus tardives n'en gardent que des vestiges. Les spécialistes sont d'avis que l'ougaritique est une langue sémitique du nord-ouest, mais sont obligés d'admettre qu'en remontant dans le temps certains traits qui opposaient ce groupe et celui du sud-ouest semblent s'estomper.

Avec l'éblaïte, dont la documentation découverte au sud d'Alep, en Syrie, nous fait remonter au III^e millénaire avant ec, la situation devient encore plus compliquée : cette langue, en effet, présente des traits qui la rapprochent de l'akkadien (notamment une conjugaison à préfixes pour l'accompli), et d'autres qui la rapprochent des langues sémitiques de l'ouest (par exemple l'ordre Verbe-Sujet-Objet, alors que l'akkadien a un ordre Sujet-Objet-Verbe). Elle a aussi un système consonantique beaucoup plus conservateur que celui de l'akkadien, en particulier au niveau des consonnes gutturales. En outre, l'éblaïte a conservé, à l'instar de l'akkadien, de l'ougaritique et de l'arabe, un système de déclinaison du nom. Certains spécialistes proposent d'ajouter cette langue au sémitique de l'est, d'autres au sémitique de l'ouest. Mais surtout, ici encore, une remontée dans le temps semble estomper des distinctions qui fondaient la classification ancienne.

Le réexamen des langues sémitiques plus anciennement connues a, lui aussi, apporté son lot d'objections à la classification traditionnelle. En particulier, on a été amené à reconsidérer le rattachement de l'arabe au groupe sémitique du sud, en faisant valoir qu'il partageait avec le sémitique occidental du nord des traits non attestés dans les langues sudarabiques et éthiopienne [2]. Ces traits linguistiques partagés par l'arabe et les langues sémitiques du nord-ouest ont conduit Hetzron (1976), à proposer un réaménagement global de la classification traditionnelle, en créant notamment, à la place du groupe sémitique occidental, un « sémitique central » avec un embranchement « arabo-cananéen » et un embranchement araméen, ne laissant dans le sémitique sud que l'éthiopien, le sudarabique ancien et le sudarabique moderne (ce dernier n'étant d'ailleurs plus du tout considéré comme un descendant du premier).

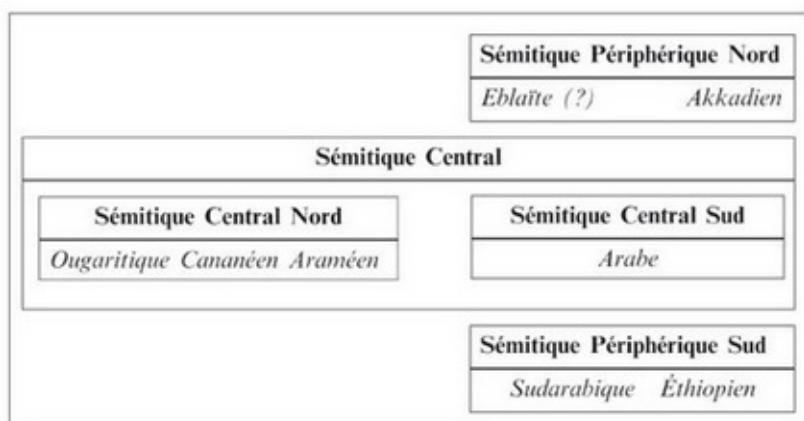
Cependant, ce réaménagement ne répond pas à la question fondamentale suivante : comment se fait-il, si l'on se situe dans une perspective généalogique, que l'arabe, rattaché à l'une des branches de l'arbre sémitique (peu importe laquelle), exhibe des traits linguistiques que les autres langues du même embranchement ne présentent pas, mais qui sont partagés par des langues

appartenant à un rameau différent, parfois non adjacent ? À cette question, le sémitisant italien Giovanni Garbini (1984) apporte une réponse radicale : pour lui, le concept théorique sous-jacent à l'usage des arbres généalogiques en linguistique, concept qui a donné sa pleine mesure dans l'étude des langues indo-européennes, est tout simplement inadapté aux réalités des langues sémitiques. Ces dernières, en effet, sont restées sur une très longue période en contact suivi les unes avec les autres, sur un territoire relativement fixe, et donc avec toute latitude pour s'emprunter les unes aux autres du matériel linguistique à tous les niveaux (phonologie, morphologie, syntaxe, lexique). En outre, certaines de ces langues, notamment l'akkadien, puis l'araméen, puis l'arabe, ont fonctionné comme *lingua franca*, langue véhiculaire, dans l'ensemble de la région favorisant encore les emprunts et les calques. Par suite, face à un trait linguistique partagé par deux langues sémitiques voisines, il est très difficile de dire s'il est hérité de l'ancêtre commun (conclusion la plus plausible pour les langues indo-européennes, séparées les unes des autres et sans contacts suivis), d'un emprunt de l'une à l'autre, voire des deux à une troisième, ou encore d'un « développement parallèle », indépendant, et dû seulement à des tendances typologiques partagées. Garbini n'en déduit pas qu'une histoire linguistique des langues sémitiques est impossible, mais pense qu'il faut, pour ce faire, renoncer à toute conception généalogique. Le scénario qu'il propose consiste à poser un « centre » initial de développement du sémitique, centre qu'il situe dans les plaines de Syrie, et qui pourrait bien être le point d'arrivée des protosémites au Moyen-Orient. Ce centre, en bordure du grand désert d'Arabie, a toujours été un lieu de contact entre les populations sédentaires et nomades de toute la région, et a vu, au cours du temps, des processus alternatifs de sédentarisation et de bédouinisation. Ces mouvements incessants de population ont été, selon Garbini, le moteur de la diffusion d'innovations linguistiques de ce « centre attracteur » vers les périphéries sémitisées, dans toutes les directions. Toujours selon lui, les premiers émigrants, notamment les ancêtres des locuteurs de l'akkadien et de l'éblaïte se sont trouvés isolés, dans une large mesure, des courants migratoires ultérieurs, et leurs langues ont donc conservé tous les archaïsmes qui les caractérisent. Les locuteurs du sudarabique ancien représentent une seconde vague d'émigration qui se serait produite après que certaines innovations aient eu lieu en sémitique occidental. Garbini considère que le point de départ de toutes les innovations linguistiques qui vont caractériser le sémitique occidental est l'amorite, langue très peu connue, mais dont le peu que l'on a pu en reconstruire suggère qu'il contient en germe beaucoup des innovations qui se développeront par la suite dans les langues sémitiques occidentales. En fait, il est probable que Garbini se représente l'amorite plutôt comme un type linguistique

innovateur que comme une langue restructurable.

Les traits linguistiques que l'arabe partage avec les langues sémitiques du nord-ouest datent de la période où ses locuteurs vivaient encore près du centre innovateur. Quant aux traits que l'arabe ne partage qu'avec le sudarabique, ils résulteraient d'influences réciproques entre les Arabes bédouins et les sédentaires d'Arabie du sud, influences rendues possibles après la colonisation par les Arabes de la plus grande partie de l'Arabie.

La théorie de Garbini rend caduque toute classification des langues sémitiques fondée sur des arborescences prétendant à une interprétation généalogique. Elle est par contre compatible avec une classification aréale « large » du type de celle que propose Igor Diakonoff (1994) et qui s'articule selon deux paramètres : une opposition entre langues sémitiques centrales et langues sémitiques périphériques d'une part, et une autre entre nord et sud. Selon cette classification, le sémitique central nord comprendrait l'ougaritique, le cananéen et l'araméen, le sémitique central sud serait réservé à l'arabe en ses diverses variétés (dont nous discuterons au prochain chapitre), le sémitique périphérique nord serait constitué par l'akkadien et ses langues filles, mais on pourrait y ajouter (ce que ne fait pas Diakonoff) l'éblaïte, enfin le sémitique périphérique sud comprendrait le sudarabique et l'éthiopien. À cette classification peut se superposer une relation d'ordre dans le temps mettant en évidence le passage des formes les plus archaïques aux formes les plus modernes des langues sémitiques en passant par les formes intermédiaires.



Classification aréale des langues sémitiques

Retenons de tout cela la position de l'arabe, « intermédiaire » à plusieurs titres :

entre le nord et le sud, le centre et la périphérie, l'archaïque et le novateur.

Notes

[1] Ibn Hazm , *Al-'Ihkām fī 'usūl al-'ahkām* chap. 4, Le Caire, Šakir Éd., (n.d.).

[2] Les principaux de ces traits sont : l'usage d'un suffixe /t/ pour la 2^e personne de la conjugaison là où les langues du sud ont un /k/ ; une base verbale d'inaccompli en $R_1R_2VR_3$ alors que les langues du sud ont un schème $R_1VR_2VR_3$ (avec possible gémination de R_2) ; un pronom de 3^e personne à initiale /h/ alors que les langues du sud ont un /s/, sauf le sabéen qui a aussi un /h/ qu'il pourrait bien avoir emprunté à l'arabe ; enfin, seules les langues sémitiques du nord et l'arabe ont développé une forme pausale des mots féminins en /a/. Plus surprenant encore, l'arabe partage avec le seul cananéen certains traits qui les opposent à l'araméen, en particulier l'usage d'un suffixe de féminin pluriel à initiale /-n/ et celui d'un article défini préfixé (/ha-/ en hébreu, /al-/ en arabe).

Chapitre II

L'émergence de l'arabe

I. L'Arabie avant l'islam

La péninsule Arabique, territoire où commence l'histoire de l'arabe, est un vaste espace au climat aride où l'on peut distinguer, en gros, deux grandes zones écologiques : celle de la steppe, qui soude l'Arabie au Croissant fertile et se prolonge vers les zones côtières de la mer Rouge, le long de la chaîne du Ḥiǧāz, et celle du désert, qui commence au centre du plateau oriental du *Naǧd*, avec le *Nafūd*, et se poursuit sur tout le sud-est de la péninsule avec le *Rub' al-Ḳālī*. Au sud-ouest, cependant, une zone montagneuse bénéficie des pluies de mousson. Cette région, dite « Arabie heureuse » (en gros l'actuel Yémen), a développé, depuis les temps les plus anciens et grâce à une gestion évoluée des ressources hydrauliques, une agriculture florissante, base d'une vie sédentaire, et même citadine remarquable. Mais précisons d'emblée que l'Arabie du Sud n'est pas, à l'origine, un territoire de langue arabe : c'est là en effet que se sont installées à une époque ancienne, probablement antérieure au I^{er} millénaire avant ec, des populations parlant diverses formes du sudarabique ancien, notamment le sabéen (langue de la tribu de *Saba'*) et le minéen (langue du royaume de *Ma'in*). Ces langues sont connues seulement par des données épigraphiques, qui utilisent un alphabet propre, quoique dérivé d'une écriture sémitique du nord.

Le domaine de l'arabe se situe plus au nord, dans les steppes et les déserts de l'Arabie du centre et du nord, où règne essentiellement le pastoralisme nomade, même si les oasis permettent la culture du palmier et de quelques céréales et légumes. On ignore quand les ancêtres des Arabes ont commencé à coloniser l'Arabie, mais on suppose que ce processus n'a connu toute son ampleur qu'après la domestication du dromadaire, laquelle ne serait pas antérieure à la première moitié du II^e millénaire avant ec. Selon Garbini (1984) l'arabe s'est détaché du sémitique du nord au cours du processus de bédouinisation rendu possible par la domestication des camélidés. Certains spécialistes ajoutent que

l'invention, au cours des derniers siècles avant ec, d'un nouveau type de selle spécialement conçue pour le dromadaire, et la mise au point, quelques siècles plus tard, d'armes adaptées à ce type de monture, ont permis le développement de tribus de guerriers méharistes capables d'effectuer en une journée d'importants déplacements et de transporter des charges considérables. Les conditions techniques pour que les Arabes contrôlent le commerce caravanier étaient ainsi réunies.

II. Les premières attestations de l'arabe

Les premières mentions du terme « Arabes » apparaissent dans des textes assyriens du ^{viii}^e siècle avant ec et font référence à des populations nomades que les Assyriens eurent à combattre ou à fixer aux marges de l'empire, s'en servant éventuellement comme troupes auxiliaires. Des textes sudarabiques appliquent le terme « Arabes » aux populations nomades exogènes. C'est aussi, semble-t-il, le sens de ce mot dans la Bible. Les Romains puis les Byzantins feront à leur tour mention des Arabes comme de peuples nomades. En langue arabe même, le mot *'a'rāb* désigne spécifiquement les nomades, alors que *'arab* peut désigner les Arabes en général ou les Bédouins en particulier.

Les toutes premières traces de la langue arabe sont des noms propres qui se rencontrent dans des inscriptions en langues sudarabiques. Mais ces noms ne nous apprennent pas grand-chose sur l'état ancien de la langue arabe. Plus intéressantes à cet égard sont des inscriptions que l'on nommera, en suivant la suggestion de Christian Robin (1992), « nordarabiques ». Il s'agit de milliers d'inscriptions parfois simples graffitis, laissées par les nomades, durant des siècles, au long de leurs pérégrinations à travers la péninsule, souvent dans les oasis d'étape des grandes voies caravanières. Ces inscriptions sont regroupées en quatre grandes familles en fonction de leurs affinités formelles et des circuits de transhumance qu'elles semblent avoir suivis. On reconnaît ainsi un groupe « *ṭamūdéen* », situé le long d'un arc d'oasis allant du nord au nord-ouest de l'Arabie, avec des prolongements jusqu'au nord du Yémen, et qui couvre une période allant du ^{vi}^e siècle avant ec au ^{vi}^e siècle de l'ère courante ; un groupe « *liḥyānite* », localisé dans l'oasis de *Didān* (nom antique d'*Al-'Ulā* au nord de Médine), sur une période allant du ^{iv}^e au ⁱ^{er} siècle avant ec ; un groupe «

ṣafā'itique », du nom de Ṣafā' au sud-ouest de la Syrie, où elles ont d'abord été trouvées, mais qui sont réparties au long des confins du Croissant fertile et couvrent la période qui va du i^{er} siècle avant ec au iii^e siècle ; enfin un groupe « ḥasā'itique », localisé dans la région du Golfe, et dont les inscriptions sont situées entre le v^e et le i^{er} siècle avant ec.

Toutes ces inscriptions utilisent une forme de l'alphabet sudarabique et sont en général courtes et stéréotypées : dédicaces, généalogies, invocations à divers dieux du panthéon arabe. Quelques-unes, un peu plus longues évoquent des événements politiques, marquent la présence de campements ou évoquent des défunts. Il est, bien sûr, impossible, avec si peu de données, d'esquisser un vrai profil de la variété linguistique représentée. Mais un trait linguistique se retrouve dans toutes ces inscriptions, qui permet de les caractériser sans hésitation comme du nordarabique : l'usage d'un article défini préfixé, /h-/ ou /hn-/, ce qui les exclut du domaine sudarabique, qui ne connaît qu'un article suffixé. Trois autres indices linguistiques d'appartenance au nordarabique se retrouvent dans certaines de ces inscriptions : l'usage d'un préfixe verbal factitif ' -, là où le sudarabique aurait /h-/ ou /s-/, l'usage des deux sifflantes /s/ et /š/ seulement, alors que le sudarabique en a une troisième (transcrite /ś/), et celui d'un pronom suffixe de 3^e personne à initiale h, alors que le sudarabique a /s/ (sauf le sabéen dont la forme, analogue à celle de l'arabe pourrait bien être un emprunt). On trouve aussi, dans les inscriptions safā'itiques, une alternance entre suffixes féminins /(a)t/ et /(a)h/, propre aux langues sémitiques du nord, et un suffixe de masculin pluriel /n/, qui, en l'absence de notation des voyelles, pourrait correspondre au suffixe de pluriel /ūn~īn/ de l'arabe.

À partir du vi^e siècle avant ec de plus en plus de tribus de la péninsule nomadisent avec leurs troupeaux dans les steppes du nord. Certaines finissent par se sédentariser, et se dotent même d'une organisation sociale et politique qui permet de parler de véritables états. Un des premiers, le royaume nabatéen, se développe entre le iv^e siècle avant ec et le i^{er} siècle ec, dans l'actuelle Jordanie, avec Pétra pour capitale. Les Nabatéens pratiquent l'agriculture dans les steppes jordaniennes, mais doivent l'essentiel de leur richesse à leur position stratégique sur la grande route caravanière de l'ouest en provenance du Yémen. Le royaume nabatéen est annexé en 106 par les Romains qui en font leur *Provincia Arabia*.

Les Nabatéens, arabophones, utilisaient pour leurs inscriptions l'araméen, sans doute considéré alors comme plus prestigieux pour ce type de fonction. Cet

araméen, cependant, présente des interférences avec l'arabe, et constitue donc une source d'information indirecte sur l'état ancien de la langue. Le premier fait intéressant, dans ces textes, est que l'article défini arabe, quand il apparaît, est toujours <'l->, à lire /'al-/, ce qui marque une étape décisive vers l'arabe que nous connaissons.

Un des points qui a le plus retenu l'attention des spécialistes dans les textes nabatéens, et qui a soulevé de nombreuses discussions, est que certains noms propres arabes y sont écrits avec un w ou parfois un y final : on trouve ainsi <yzydw> pour *yazyd(u)* ou <'bd'lh> pour *'abd 'allāhi*. Certains auteurs, K. Versteegh (1997) par exemple, ont fait l'hypothèse que ces terminaisons pouvaient correspondre à la forme pausale des noms en arabe. Cela laisserait penser que cette forme d'arabe utilisait des marques casuelles analogues à celles de l'arabe classique, mais que les règles de pause auraient à l'époque été plutôt du type allongement des voyelles brèves finales, alors qu'en arabe classique, on efface ces mêmes voyelles à la pause. Pour sa part, Werner Diem (1973), soulignant l'absence de systématisme de ces marques et leur incohérence contextuelle, conclut que si ces graphies sont le reflet de marques de cas, elles manifestent que le système casuel était en crise et avait perdu toute valeur fonctionnelle.

Après l'annexion de Pétra, c'est l'oasis de Palmyre (*Tadmur*), dans le désert syrien, qui devient un grand centre du commerce caravanier, jusqu'à sa conquête par les Romains en 273. Comme les Nabatéens, les palmyréniens étaient arabophones, mais écrivaient en araméen. Leurs inscriptions ne nous apprennent pas grand-chose de nouveau par rapport à celles des Nabatéens.

Une étape nouvelle dans l'étude de l'émergence historique de l'arabe est constituée par l'apparition de textes en écriture non arabe mais dont la langue est en tout ou en partie une forme de nordarabique.

Une transition vers ce type de texte semble être constituée par ce que l'on appelle les inscriptions en « pseudo-sabéen » ou « qahṭānique ». Il s'agit de textes qui proviennent du site saoudien de *Qaryat al-Fāw*, rédigés en écriture sabéenne mais dont la langue, bien qu'elle « imite » le sabéen, s'en distingue par des traits tout à fait significatifs, notamment la présence des deux seules sifflantes /s/ et /š/ que connaît le nordarabique. Les textes en pseudo-sabéen présentent en outre des traits de lexique et de phraséologie qui rappellent le minéen mais aussi « quelques tournures typiquement arabes » (Robin, 1992,

120). Il n'est pas impossible que la langue parlée des auteurs d'inscriptions pseudo-sabéennes ait été le *ḥimyarite*, sorte de langue mixte entre l'arabe et le sudarabique dont l'auteur arabe *al-Hamdānī* parlera au x^e siècle (Robin, 1992, 101-108). Pour les Arabes le terme *Ḥimyar* désignera globalement le pays et les populations yéménites.

Dans la même région, on trouve des inscriptions en écriture sudarabique mais dont les textes peuvent désormais être qualifiés de « proto-arabes » car ils exhibent, entre autres, l'article défini préfixé arabe <'l->. L'une d'entre elles (présentée dans Robin, 1992, 114), l'inscription de Ya'mar, datée de la fin du iii^e siècle avant ec ou du début du ii^e, est la dédicace d'une construction à la déesse arabe 'Allāt. On y trouve la plus ancienne attestation, à ce jour, de l'article défini arabe /'al-/ ainsi que la séquence graphique <l-lt> à lire /li-llāt/ « pour 'Allāt » avec assimilation de ce même article devant la consonne // du nom suivant comme en arabe classique. La structure syntaxique de la phrase est elle aussi tout à fait arabe avec une consécutive introduite par la conjonction *fa-* : <f-sm' l-hmw>, que l'on peut lire /fa-sami'at la-humūl/ « alors elle les exauça ».

Une autre inscription (présentée dans Robin, 1992, 115), également originaire de *Qaryat al-Fāw*, est la stèle funéraire d'un certain 'Iḡl Ibn Hofi'amm. Elle est datée de la fin du i^{er} siècle avant ec et est assez longue (10 lignes) pour autoriser une esquisse d'analyse linguistique. On y reconnaît l'article défini /'al/, assimilé devant les consonnes corales, mais aussi certaines omissions de la consonne ' qui laissent supposer qu'elle n'était guère articulée dans le dialecte arabe local. Christian Robin considère que l'inscription de 'Iḡl est le plus ancien texte en langue arabe connu à ce jour.

L'écriture sudarabique était plutôt bien adaptée à la notation de l'arabe puisqu'elle disposait d'un graphème distinct pour chacun des 28 phonèmes consonantiques de la langue (et même d'un graphème surnuméraire notant la 3^e sifflante inconnue de l'arabe). Ce n'est pourtant pas à partir de ces premières tentatives de fixation par écrit de la langue arabe que va se développer un système graphique proprement arabe. La raison essentielle en est que les derniers siècles qui précèdent l'apparition de l'islam et l'émergence des Arabes sur la scène de l'histoire ne sont guère favorables à la civilisation sudarabique, pourtant première grande civilisation de la péninsule. Impliqués, par leurs alliances, dans les conflits entre les deux grandes puissances de l'époque, l'Empire perse et l'Empire byzantin (et ses alliés éthiopiens), les royaumes

sudarabiques traversent des périodes de conflits sociaux et religieux graves et connaissent divers épisodes de guerre et d'occupation étrangère au cours du vi^e siècle. L'effondrement, vers la fin du même siècle, du barrage de *Ma'rib*, un grand ouvrage d'art qui faisait la fierté de l'Arabie du Sud, est souvent présenté comme marquant la décadence économique et politique des états sudarabiques. Leur influence sur les Arabes du Nord diminue et ceux-ci prennent même, petit à petit, le contrôle des grandes routes caravanières qui étaient une des sources de richesse de l'Arabie du Sud. La Mecque, ville arabe du Ḥiğāz, centre de pèlerinage et étape sur la route caravanière longeant la mer Rouge, prend une importance économique croissante, et la tribu des *Qurayš* qui la domine devient l'une des tribus les plus puissantes d'Arabie.

Dès le v^e siècle de nouveaux émirats arabes se développent dans les confins désertiques du nord de la péninsule. Au nord-ouest, les *Gassānides*, chrétiens monophysites, se mettent au service de Byzance dont ils garantissent, du moins en principe, la frontière sud avec l'ennemi héréditaire, les Perses. Ceux-ci leur opposent une autre tribu arabe, les *Lakmīdes*, convertis au christianisme nestorien, et qui font de Ḥīra, sur les bords de l'Euphrate, leur capitale. Il s'y développe une véritable vie de cour, et la poésie en langue arabe y est cultivée. D. Frolov (2000) considère même que Ḥīra a été le centre de l'une des deux grandes traditions poétiques préislamique, qu'il baptise « Ḥīran tradition », la seconde qu'il qualifie simplement de « bédouine » étant représentée un peu partout dans la péninsule, et en particulier dans le Ḥiğāz.

C'est vers les confins nord de la péninsule, dans les steppes syriennes, qu'il faut se tourner pour trouver les inscriptions arabes les plus tardives qui jalonnent l'évolution vers l'arabe historique.

L'une des plus anciennes, découverte en 1986 à 'En 'Avdat, dans le désert du Néguev, et étudiée par James Bellamy (1990), date du i^{er} siècle et présente, en écriture nabatéenne, un texte arabe inclus dans un texte en langue araméenne. Il s'agit d'une dédicace au dieu Obodas. Le texte arabe serait une sorte de louange adressée au dieu dans une langue assez élaborée. Malheureusement la lecture de ce texte est encore loin d'être éclaircie de façon convaincante, et l'on doit, en l'état actuel des connaissances, se contenter de dire qu'il s'agit bien d'arabe en raison de la présence de l'article défini préfixé <'l>, l'araméen ne connaissant qu'un article suffixé /-ā/. Ce texte contient deux noms communs <'lmwtw> « la mort » et <ğrḥw> « une blessure » qui exhibent le fameux suffixe w tant discuté pour les inscriptions nabatéennes, et qui est ici d'autant plus intrigant qu'il s'agit

de noms communs. Faut-il y voir un argument en faveur de la thèse de marques casuelles du nominatif, les deux noms étant, dans la lecture de Bellamy, des sujets de verbes ? Il est trop tôt pour le dire...

Quelques courtes inscriptions de la même aire géographique, également en écriture nabatéenne, mais plus tardives (milieu du iii^e siècle), notamment celles de *'Umm al-Ğimāl* et d'*al-Ḥiğr* appartiennent à la même étape. Elles présentent des occurrences de noms communs, précédées ou non de l'article défini, et exhibant le mystérieux suffixe *w* comme *<('l)qbrw>* = « (la) tombe ».

La plus célèbre de toutes les inscriptions en langue arabe et en écriture nabatéenne est celle de Al-Namāra, dite d'*Imru'l-Qays*, retrouvée à Al-Namāra, à 120 km au sud-est de Damas en 1901. Cette inscription fait, sur certains points, l'objet de lectures divergentes. Ainsi, après le syntagme *<mlk, 'l'rb>*, que tout le monde lit *malik(i) 'al'arab* = « le roi des Arabes », le mot graphique suivant est lu par certains comme le corroboratif : *kullihi* = « tous », alors que d'autres y voient le début d'une nouvelle proposition : *walaqabuhu* = « et son surnom [est] ». Cette importante divergence de lecture est due en partie au fait que l'écriture nabatéenne ne compte que 22 consonnes, alors que l'arabe en compte 28 de sorte que le même ductus consonantique peut être interprété de façons très diverses, comme cet exemple permet de le constater.

La dernière étape dans l'émergence de l'arabe tel que nous le connaissons est représentée par des inscriptions qui, par leur forme, préfigurent nettement l'écriture arabe, comme l'inscription de *Zabad*, près d'Alep (datée de 512) ou celle de *Ḥarrān*, dans le *Ḥawrān* (datée de 568). Ces inscriptions sont localisées dans la même région, *grosso modo* la steppe syrienne, et leur contenu n'est pas très différent de celui des inscriptions évoquées précédemment. Simplement elles sont globalement plus tardives (entre le milieu du iv^e et le début du vi^e siècle) et montrent diverses étapes de la formation de l'écriture arabe. Celle-ci a eu pour base l'écriture araméenne dont on vient de voir que l'arabe s'était un temps servi. Les spécialistes ont cependant discuté pour savoir si c'est à partir de la forme nabatéenne ou de la forme syriaque de l'écriture araméenne que le système graphique arabe s'est développé. On a évoqué en faveur du syriaque sa forme cursive, sachant que l'écriture arabe est massivement cursive. La découverte d'une forme cursive du nabatéen, distincte de la forme épigraphique, semble avoir fait abandonner l'hypothèse syriaque à la plupart des spécialistes. Ajoutons, dans cet ordre d'idées, que la cour des *Lakmīdes* de *Ḥīra* a été parfois citée comme ayant pu jouer un rôle dans l'élaboration de l'écriture arabe.

III. Le domaine linguistique arabe avant l'islam

Outre l'épigraphie, une autre source d'information nous permet de nous faire une idée du domaine linguistique arabe avant l'islam (Rabin, 1951). En effet, les grammairiens musulmans qui, à partir du ^{viii}^e siècle, entreprennent de faire une description systématique de la langue du Coran, ont laissé de nombreuses données sur la situation linguistique ancienne de la péninsule. Ces données, orientées vers l'élaboration d'une langue unifiée, ne visaient pas à reconstituer avec précision un tableau de la situation linguistique ancienne. Elles n'en restent pas moins une source précieuse d'information sur cette situation. Les indications des grammairiens arabes (désormais ga) permettent d'inférer que deux grandes aires dialectales existaient dans le domaine arabe : une aire occidentale, comprenant la région du Ḥiǧāz, et notamment La Mecque, et une aire orientale, couvrant le plateau du Naǧd. Les données des ga suggèrent que la zone occidentale, traversée par les routes caravanières et où existaient de nombreuses agglomérations sédentaires, points de rencontre de populations de toutes origines, était linguistiquement moins conservatrice que celle du Naǧd, surtout fréquentée par les grandes tribus nomades. Le plus souvent d'ailleurs, lorsque des divergences existent entre les deux aires dialectales, les ga ont tendance à préférer la variante du Naǧd, ne faisant en général une place à une variante ḥiǧāzienne que si celle-ci apparaît dans le texte coranique et ne s'écarte pas trop de la *koinè* poétique. C'est par exemple le cas de la « *mā ḥiǧāzienne* » consistant à mettre à l'accusatif le prédicat nominal d'un énoncé nié par *mā* « ne pas être ». Par contre la prononciation de l'occlusive glottale (*hamza*), pratiquement abandonnée dans le Ḥiǧāz, mais conservée dans le Naǧd et dans la *koinè* poétique, sera rétablie systématiquement dans le corpus coranique (cf. chap. IV).

Les ga donnent quelques indications sur les dialectes dits « yéménites », le plus souvent pour les stigmatiser. Selon certaines sources, quelques-uns de ces parlers ne faisaient pas subir d'inflexion casuelle aux démonstratifs, disant *Hāḍāni* « ces deux-là » en tout contexte, alors que les parlers du Naǧd fléchissent de telles formes. C'est là une des rares indications des ga relatives à une possible fluctuation de l'usage de la flexion casuelle.

On constate à cet égard que toutes les variétés d'arabe citées par les ga font

usage de la flexion casuelle, comme le montre bien, par exemple, le cas de la « *mā ḥiḡāzienne* » évoquée ci-dessus. Rien ne nous est dit des parlers arabes des steppes du nord, dont nous avons vu qu'ils avaient peut-être perdu l'usage de la flexion casuelle. Cela résulte d'un choix délibéré des ga, qui ont exclu de leurs recherches pour l'établissement d'une norme linguistique arabe les parlers considérés comme trop « corrompus » (*fāsid*).

Certains chercheurs, comme Jan Retsö [1] et Jonathan Owens [2], ont soutenu que l'aire arabophone était depuis toujours partagée en deux zones, l'une utilisant un arabe flexionnel, et l'autre un arabe sans flexions. Cette hypothèse, contredite par les données et les évolutions concordantes d'autres langues sémitiques comme l'akkadien et ses langues filles, l'est aussi par des postulats de base de la linguistique historique et comparée. Elle peut donc difficilement être prise au sérieux. Il est beaucoup plus plausible de penser que la variété non fléchie dérive d'une évolution historique, au demeurant parfaitement reconstituable, de la variété fléchie. Cette évolution a sans doute commencé très tôt à la périphérie de l'arabophonie, comme semblent en témoigner les données épigraphiques dont nous avons fait état, mais elle a fini par toucher l'ensemble du domaine, comme nous le verrons plus loin.

Notes

[1] J. Retsö , 'I'rāb in the Forebears of Modern Arabic Dialects, *AIDA-I* , 333-342, Paris, inalco, 1993.

[2] J. Owens, Case and proto-arabic, I & II «, *BSOAS*, 61-1, 51-73 et 61-2, 215-227, 1998.

Chapitre III

Aux sources de l'arabe classique

L'arabe classique est le résultat du travail de systématisation réalisé par les *grammairiens* à partir du *viii*^e siècle. Ce travail se fonde, entre autres, sur deux corpus textuels fondamentaux : celui de la poésie arabe préislamique, principal héritage littéraire de la langue arabe avant l'islam, et celui du Coran, texte fondateur de la nouvelle foi. Il convient donc de consacrer quelques pages à chacun de ces deux corpus de textes et à la variété de langue qui les a véhiculés.

I. La poésie arabe préislamique

La poésie arabe préislamique nous est connue par des recensions écrites dont les plus anciennes ne sont pas antérieures au début du *ix*^e siècle. Auparavant, ces poèmes se seraient transmis oralement, depuis leur composition par les poètes de l'anté-islam, grâce à des « récitants » (*ruwāt*) plus ou moins professionnels. Ces poèmes nous sont parvenus sous une forme fragmentaire, avec de multiples interpolations et variantes et des attributions parfois divergentes.

Le corpus de la poésie arabe préislamique consiste soit en courtes pièces (*qīṭa'*) ne dépassant pas quelques vers, soit en odes (*qaṣīda*) pouvant en comporter plusieurs dizaines, comme c'est le cas des dix « grandes odes » (*mu'allaqāt*), considérées par les arabophones cultivés comme des chefs-d'œuvre intemporels de la littérature arabe. Leur contenu reflète essentiellement la vie et les préoccupations des Bédouins : description de la faune et de la flore du désert, évocation du mode de vie des tribus nomades, liaisons amoureuses et séparations que la transhumance pouvait susciter, écho des conflits, voire des guerres entre clans avec leurs cortèges de jactance (*fakr*), de satire (*hiǧa'*), de panégyrique (*madīḥ*) ou de thrène (*riṭā'*).

Mais le plus remarquable, dans cette poésie, ce sont ses caractéristiques formelles. Il s'agit de poèmes monorimes, fondés sur une métrique quantitative

rigoureuse reposant sur le contraste des syllabes brèves et des syllabes longues. En fin de vers, à l'inverse des règles de pause de la langue ordinaire, les voyelles brèves finales ne sont jamais effacées, mais, au contraire, allongées. Quant au lexique, quoique rigoureusement soumis aux contraintes métriques, il est d'une foisonnante richesse et frappe par son archaïsme, son côté parfois « barbare », et un inimitable parfum de bédouinité.

Un fait qui a frappé les analystes modernes est la grande homogénéité linguistique de l'ensemble du corpus. Ce fait est *a priori* d'autant plus surprenant que les biographies des poètes, telles qu'elles sont rapportées par la tradition, révèlent qu'ils sont originaires de presque toutes les régions de la péninsule, même si le *Nağd* semble un peu plus prolifique. Cette étonnante homogénéité linguistique a suscité, au début du xx^e siècle, les doutes de certains spécialistes comme Ṭaha Ḥusayn (1927), qui ont considéré l'ensemble de cette poésie comme une pieuse forgerie des premiers grammairiens musulmans soucieux de donner un fondement philologique à leurs explications du texte coranique. Cette thèse a suscité des réactions très vives dans le monde arabe où on y a vu une attaque contre l'un des fondements de la culture littéraire arabe, voire de l'Islam. Mais elle a été également rejetée comme excessivement critique par des spécialistes plus modérés comme Régis Blachère (1952). Ce dernier a fait valoir que, même si les conditions dans lesquelles ces textes nous ont été transmis justifient bien des doutes sur le détail, même si l'analyse méthodique permet d'y reconnaître parfois la main du pasticheur, même enfin si la tradition elle-même nous apprend que des transmetteurs comme Ḥammād al-rāwīya (m. ~ 772) ou Ḳalaf al-'Aḥmar (m. ~ 796) sont connus pour avoir forgé de faux poèmes préislamiques, il est impossible de mettre en doute la représentativité de l'ensemble du corpus.

Mais c'est surtout avec le développement des recherches sur les littératures de tradition orale que l'on a pu mieux comprendre les caractéristiques du corpus poétique préislamique et reconnaître son authenticité au moins relative. Pour l'ancienne poésie arabe, ces recherches, inaugurées par James T. Monroe (1972), puis développées par Michael Zwettler (1978) et plus récemment par Bruno Paoli (2001), ont permis de mettre en évidence le caractère massivement « formulaire » de ce type de littérature. Les « formules » sont des modèles encapsulant à la fois des composants de la forme et du contenu d'un texte. Systématiquement employées, elles peuvent varier légèrement d'une occurrence à l'autre, et s'enchaînent de façon à constituer la trame d'une foule de poèmes, tous différents, mais présentant un « air de famille » indéniable. On a calculé que

les formules pouvaient représenter plus de 85 % d'un poème. Elles fournissent au poète les outils à la fois métriques et thématiques, lui permettant, grâce à un savoir-faire fondé sur une longue pratique, de produire des vers adaptés à chaque circonstance dans le mètre le plus approprié. Monroe montre, par exemple (1972, 29), que le poète dispose, selon le mètre employé, de divers synonymes pour parler des vestiges du campement de la bien-aimée, topos obligé de l'ouverture (*nasīb*) d'une *qaṣīda*. Certes la « liberté de création » s'en trouvait limitée, mais le public cherchait plus, dans la production des poètes, à retrouver des images et des idées familières, fondues dans une structure métrique attendue, que des innovations déroutantes. De là la grande homogénéité du corpus poétique préislamique. La langue de ce corpus, polie par des décennies d'usage « spécialisé », dépouillée de presque tous les particularismes dialectaux qui pouvaient en limiter la diffusion, et pratiquée à travers l'ensemble de l'Arabie, a été baptisée « *koinè* poétique » par les orientalistes et (*luḡa 'arabiyya*) *fuṣḥā*« (langue arabe) la plus claire » par les philologues musulmans. Comme l'a bien montré Dmitry Frolov (2000, chap. 2), les caractéristiques structurales de la *fuṣḥā* sont très fortement solidaires de son statut de langue de la poésie préislamique. En effet, les contraintes métriques très strictes de celle-ci, accordant à chaque syllabe, en fonction de son poids, un rôle précis dans l'élaboration du mètre, et interdisant l'effacement des voyelles brèves en finale de mot, fût-ce en fin de vers, impliquaient un fonctionnement optimal de la syntaxe désinentielle de la *fuṣḥā* dans ce qu'elle a de plus conservateur : maintien rigoureux des voyelles casuelles en toutes positions, liberté importante de l'ordre des mots, productivité maximale des schèmes de la morphologie. Il est même possible que l'usage poétique ait favorisé, pour des raisons de rythme et de rime, l'extension du domaine de la flexion casuelle à des catégories de mots qui à l'origine l'ignoraient (comme sans doute les démonstratifs, et peut-être, à une époque bien plus ancienne, les pluriels des noms). La propension, largement observée dans l'usage poétique, à traiter comme « triptotes » (c'est-à-dire dotés d'une flexion casuelle complète) des mots qui, dans l'usage normal de la langue, ne le sont pas, illustre parfaitement ce mécanisme. En somme, le corpus de la poésie préislamique se présente comme une parfaite « défense et illustration » de la *fuṣḥā*.

II. *Fuṣḥā* et autorité symbolique

On sait encore peu de chose sur l'histoire de l'élaboration de la poésie

préislamique (Frolov, 2000). La perfection formelle des poèmes préislamiques qui nous ont été transmis laisse supposer que l'on est en face d'une tradition extrêmement ancienne, et le fait que cette langue littéraire ait été adoptée dans l'ensemble de l'aire arabophone plaide dans le même sens. Il faut dire que la poésie, dans l'Arabie préislamique est loin d'être une simple « affaire de littérature ». Dans la société arabe traditionnelle, le poète est à la fois le héraut et le porte-drapeau de sa tribu : émanation du groupe, il le représente et parle en son nom dans tous les moments importants. Il est de tous les combats, de toutes les cérémonies, et l'on attend de lui, aux grandes heures de la vie de la tribu, qu'il trouve les mots magiques pour glorifier les combattants courageux, fustiger les lâches et les avares, prononcer invectives et malédictions contre les ennemis, lancer les appels à la vengeance ou immortaliser le souvenir des grands défunts. Sa parole, surinvestie par le groupe, est objet de vénération et de crainte. Cette parole d'ailleurs est considérée comme inspirée par quelque djinn (ar. *ġinn*), et à ce titre porteuse de redoutables pouvoirs. D'ailleurs, le mot arabe qui désigne le poète, *šā'ir*, signifie « celui qui sait par inspiration » (Blachère, 1952, II, 332).

Mais l'usage poétique, surtout dans la forme complexe dans laquelle il se manifeste dans les vers élaborés (*qaṣīd*), n'est pas le seul que la *fuṣṣḥā* ait connu avant l'islam. La tradition nous apprend en effet que la parole d'un autre personnage était considérée avec crainte et respect dans la société arabe traditionnelle : celle du *kāhin*, terme que l'on peut rendre par « voyant » ou « devin ». Celui-ci fait partie du conseil de la tribu où il exerce des fonctions d'arbitre ou de juge. Or il utilise lui aussi la *fuṣṣḥā* dans ses vaticinations, faisant alors usage d'une forme de prose rythmée et rimée, connue sous le nom arabe de *saġ'*. Le *saġ'* n'a certes ni la régularité ni la complexité métrique de la poésie, et la rime y fonctionne de façon très différente : elle repose sur la forme que les mots ont à la pause, c'est-à-dire, notamment, après la perte de leur voyelle brève finale (fût-elle casuelle). Mais la parole du *kāhin* n'en est pas moins investie d'un pouvoir symbolique au moins aussi considérable que celle du poète. « Le *kāhin*, nous dit Blachère (1952, 191), vaticine donc dans un climat de crainte et de respect. Sa parole est toute chargée de solennité, de secrète efficence et de menace. » Des incantations en *saġ'* étaient aussi récitées par les pèlerins notamment au cours de circumambulations autour de lieux de culte comme la *Ka'ba* à La Mecque.

Un dernier personnage doit être évoqué dans cette présentation, trop sommaire, des détenteurs d'une « parole autorisée » dans la société arabe préislamique : c'est celui du *sayyid* ou chef de tribu. Cette fonction n'a jamais fait l'objet d'une

institutionnalisation formelle dans l'Arabie ancienne, et le *sayyid* ne se maintenait donc dans sa fonction que tant qu'il était reconnu comme tel par les membres de sa tribu. Pour cela, il devait constamment faire montre, au plus haut degré, des qualités que les Arabes aiment à se reconnaître : courage, générosité, longanimité et, bien sûr, éloquence. Dans les moments critiques, le *sayyid* devait, lui aussi, manifester sa capacité à utiliser la *fuṣḥā* et à délivrer à son auditoire, souvent en *sağ'*, un message propre à réveiller l'enthousiasme. Les philologues musulmans nous ont transmis un certain nombre de ces harangues dont l'authenticité est tout aussi sujette à caution, dans le détail, que les poésies préislamiques, mais qui n'en constituent pas moins des témoignages suggestifs des premiers embryons de prose oratoire (*ḵaṭāba*) en arabe.

Retenons de tout cela que la *fuṣḥā*, dans tous les usages de l'époque préislamique que nous pouvons reconstituer, apparaît comme fortement investie de puissance symbolique et associée, dans l'esprit des Arabes, à toutes les formes du pouvoir, humain ou supra humain. Ce fait ne doit pas être perdu de vue si l'on veut comprendre correctement l'étape suivante de l'histoire de la langue, celle de l'investissement de cette variété linguistique par le message coranique.

III. Le texte coranique

Le Coran (de l'arabe *qur'ān* « récitation ») est le texte sacré des musulmans, rassemblant l'ensemble des révélations reçues durant une vingtaine d'années, entre 610 et 632, par le Prophète de l'islam. Comme la poésie préislamique, le texte, tel que nous le connaissons aujourd'hui résulte d'un processus de fixation relativement lent. À l'origine, un certain nombre de « Compagnons » du Prophète (*ṣaḥāba*) en apprennent par cœur les versets, au fur et à mesure qu'ils sont récités par celui-ci. Ils vont constituer cette catégorie prestigieuse de musulmans dits « porteurs du Coran » (*ḥamalat al-qur'ān*). La tradition nous apprend que certains passages du texte avaient pu être notés, du vivant du Prophète, sur des supports hétéroclites, par ceux des Compagnons qui maîtrisaient l'écriture. Mais, outre que leur nombre était restreint, l'écriture arabe de l'époque était mal adaptée aux caractéristiques de la langue arabe et par suite incapable de donner de la forme orale de la révélation, une représentation précise et fiable. Pour donner une idée des insuffisances de ce système graphique, indiquons simplement qu'une seule forme de lettre, ou « *ductus* » (ar.

rasm), correspondait à cinq consonnes aussi différentes que *b*, *t*, *ṭ*, *n* et *y* en position initiale de mot, et que ni les interdentes ni les emphatiques ne disposaient d'un graphème spécifique ! Bref, cette écriture pouvait au mieux servir d'aide-mémoire aux récitants qui avaient déjà appris le texte par cœur. De cet état initial oral du corpus coranique, il est resté des traces significatives. Ainsi, lorsqu'est venu le temps (à partir du viii^e siècle) de fixer le texte dans une écriture réformée et enfin capable d'en donner une représentation précise, diverses variantes orales s'y manifestèrent : certaines de ces variantes, dites *'aḥruf*, c'est-à-dire « idiomes », portent sur des mots plus ou moins synonymes apparaissant dans un même contexte, en fonction des origines dialectales des transmetteurs. Ce premier type de variantes portait peu à conséquence, et la tradition rapporte que le Prophète, sans doute dans le souci de ne pas froisser de susceptibilités tribales, les considérait toutes comme « acceptables » car elles respectaient le sens. Elles ont tendu, avec le temps, à régresser, au fur et à mesure que l'écrit prenait plus d'importance et que les savants devenaient moins tolérants avec les variantes orales du texte sacré. D'autres, dites *qirā'āt*, « lectures », concerneront des lectures différentes du même *ductus* consonantique, lequel est affecté de points diacritiques différents selon les interprétations. Ces différences pouvaient être la source de divergences doctrinales plus sérieuses. Le *consensus doctorum* s'est efforcé d'en limiter le nombre à sept traditions dites « canoniques ». En pratique, il n'en circule plus que deux, celle de Ḥaḥṣ, majoritaire en Égypte, et par suite dans tout le monde musulman, et celle de Warṣ, encore largement suivie au Maghreb.

Selon la tradition, c'est peu après la mort du Prophète que se pose la question d'établir une recension écrite complète du texte coranique. Cette décision est attribuée au premier successeur du Prophète, le calife 'Abū Bakr (m. 634 ec). Le recueil ainsi obtenu n'aurait pas été le seul, et est resté, en tout état de cause, sans conséquences pratiques jusqu'à l'avènement du troisième calife, 'Uṭmān (m. 656 ec). Son califat est dominé par de sévères luttes politiques, portant entre autres sur la légitimité de son accession au pouvoir. Or le texte coranique, devenu pour tous les musulmans le fondement de toute autorité temporelle et spirituelle, fait l'objet de lectures divergentes où chacun des groupes en compétition prétend trouver la preuve de ses prétentions. 'Uṭmān fait donc établir, à partir du corpus d'Abū Bakr, une nouvelle recension officielle du texte coranique non sans susciter d'ailleurs de violentes protestations de la part de certains vieux Compagnons outrés d'avoir été écartés de la commission constituée à cet effet. Il fait aussi détruire tous les fragments anciens. Enfin, il fait faire des copies (sept ou plus selon les sources) qu'il fait distribuer dans tous

les grands centres urbains de l'empire musulman alors en pleine expansion. Le corpus 'uṭmānien est, en principe, la base du texte coranique tel qu'il nous est connu aujourd'hui. Il convient cependant de noter que, compte tenu du caractère encore très primitif du système graphique arabe de l'époque, l'initiative du calife 'Uṭmān ne résout pas définitivement le problème des multiplicités de lecture du Coran. La fixation du *ductus* consonantique ne suffisait pas à donner une représentation univoque du texte et il restait pratiquement nécessaire d'en connaître oralement une version pour pouvoir « lire » ce *ductus*. Ce problème ne reçoit de solutions techniques décisives que quelques décennies plus tard (nous y reviendrons), et ces solutions ne sont complètement adoptées que plus tard encore.

Le coran est composé de 114 chapitres nommés « sourates » (ar. *sūra*) de longueurs et de styles extrêmement variables, classés selon la tradition en sourates « mecquoises » ou « médinoises » selon que l'essentiel en a été révélé à La Mecque, la ville natale du Prophète, entre 610 et 622 ec, ou à Médine, la ville du Ḥiğāz qui l'accueille, lui et ses partisans, après leur émigration ou « hégire » (ar. *hiğra*) en 622, début de l'ère musulmane. Les sourates mecquoises sont plutôt courtes (parfois à peine trois versets), et formulées dans un *sağ'* au rythme et aux assonances très marqués. Leur contenu est dominé par les grands thèmes religieux : développements eschatologiques, appels à la purification et au repentir, tableaux apocalyptiques de la fin du monde et de la résurrection, description du séjour des élus au paradis et des damnés en enfer...

Au contraire, les sourates médinoises présentent des textes beaucoup plus longs (près de 300 versets pour la sourate 2, dite « la génisse »). Le texte en est généralement plus ample (certains versets sont de véritables paragraphes), et le *sağ'* s'y fait discret. Quant aux thèmes, ils sont essentiellement législatifs, mais sont souvent entrecoupés de récits qui donnent une version musulmane des grands récits de l'héritage monothéiste.

L'ensemble du texte est classé non pas chronologiquement mais par longueur décroissante des sourates, à l'exception de la sourate « liminaire » (ar. *fātiḥa*) qui ne compte que sept versets.

La langue du Coran est, sans conteste possible, de la *fuṣḥā*, malgré la présence de traits dialectaux du Ḥiğāz, et plus spécifiquement du parler de La Mecque. Le Coran souligne d'ailleurs à plusieurs reprises qu'il est délivré en une « langue arabe claire » (*bi-lisānin 'arabiyyin mubīn*) pour être parfaitement compris de

tous les Arabes (par exemple Cor. 12 : 2, 19 : 97, 26 : 192-199). Il paraît indubitable que le choix de la *fuṣṣḥā* pour délivrer le message coranique est tout à fait délibéré et qu'il obéit non seulement à la volonté d'en assurer la diffusion parmi tous les arabophones, mais aussi de l'inscrire dans la catégorie des productions linguistiques revendiquant l'autorité symbolique associée, comme on l'a vu, aux messages adoptant ce véhicule linguistique.

La langue du Coran fait, on l'a dit, un usage régulier de la prose rimée (*sağ'*), usage particulièrement sensible dans la période mecquoise. Les textes de cette période présentent d'autres caractéristiques linguistiques qui les rapprochent des productions, foncièrement orales, de l'ère préislamique, comme l'usage d'une syntaxe essentiellement paratactique, de serments ou d'imprécations, enfin d'un style visant plus à persuader qu'à démontrer. En outre, mais ceci vaut pour l'ensemble du texte, on y relève un très grand nombre de formules, dont certaines se répètent des dizaines de fois.

Significativement, les adversaires du Prophète musulman, pour dévaloriser son message, qui menaçait leurs intérêts matériels et leur autorité, ont voulu le réduire à n'être rien d'autre que de la poésie ou des vaticinations de devin, productions qui, on l'a vu, s'exprimaient dans une variété linguistique analogue. Le Coran proteste à plusieurs reprises (par exemple 21 : 5, 36 : 69, 52 : 29) contre de telles accusations, assurant que l'Apôtre de Dieu (ar. *rasūl 'Allāh*) n'est ni un poète ni un devin. Certains passages du Coran (par exemple 10 : 38, 11 : 13) mettent même les incrédules au défi de l'imiter s'ils en contestent l'origine divine. Ce défi semble, à l'origine, avoir concerné non pas la langue du Coran, présentée par le texte lui-même comme essentiellement familière aux Arabes, mais le contenu de son message, radicalement nouveau dans l'univers symbolique de l'Arabie polythéiste. C'est en tout cas cette nouveauté de contenu qui est la première évoquée lorsque se développe, surtout à partir du x^e siècle, l'un des grands thèmes de la littérature savante à propos du Coran, celui de son « inimitabilité » (ar. *'i'ğāz*).

Si le Coran rappelle, par de nombreux traits, les textes attribués par la tradition à la période antérieure, il est cependant incontestable qu'il a introduit dans la *fuṣṣḥā* des éléments nouveaux qui joueront un rôle fondamental dans le développement ultérieur de la langue arabe. Par sa forme, et notamment son ampleur, le Coran représente le premier véritable monument de la prose en langue arabe. Par son contenu, qui fait exploser les cadres mentaux traditionnels de la pensée arabe en y faisant entrer des thèmes métaphysiques, juridiques et

idéologiques radicalement neufs, il prépare la langue à assumer les tâches totalement inédites auxquelles vont l'appeler la fondation d'un ordre social et d'une civilisation nouvelle sans commune mesure avec ce que les Arabes ont connu jusque-là. Par sa variété stylistique (on y trouve discours et récits, descriptions et argumentations, exhortations et dispositions légales), et par le fait que tout lettré musulman l'apprend par cœur en tout ou en partie, il servira de modèle, conscient ou non, à tous les développements littéraires ultérieurs de la langue.

Surtout, le lien qui va se nouer entre la *fuṣḥā* et le message coranique va faire de ce qui était sans doute, à l'origine, un simple parler de tribu nomade des hauts plateaux de l'Arabie centrale et orientale, la langue liturgique d'une des religions majeures de l'humanité, et la langue officielle d'un empire qui pour un temps dominera la scène de l'histoire.

Chapitre IV

L'élaboration de l'arabe classique

Dans le siècle qui suit la mort du Prophète, les Arabes, unifiés par l'islam et profitant de l'état de mutuel épuisement dans lequel s'étaient mis Byzance et la Perse sassanide, vont conquérir un empire gigantesque, s'étendant de l'Espagne à la Transoxiane. Les populations de la plupart des régions conquises, écrasées sous leurs anciens maîtres, par les impôts, et parfois persécutées pour leurs opinions religieuses, ne leur opposent guère, le plus souvent, de sérieuse résistance.

I. Une situation inédite

Quoi qu'il en soit, les Arabes se retrouvent alors dans la situation de devoir gérer et administrer de vastes régions densément peuplées d'agriculteurs sédentaires, ainsi que certains des plus grands centres urbains du monde de l'époque, comme Damas, Ctésiphon ou Alexandrie. Les fils du désert, fort peu préparés à assumer de telles tâches, étaient en outre minoritaires dans leur nouveau domaine. Le risque était donc très grand que, noyés démographiquement et dépassés culturellement, ils ne s'acheminent vers une assimilation pure et simple par les peuples vaincus. Les décisions prises par leurs chefs, tant sur le plan de la gestion des populations arabes installées dans les territoires conquis que sur celui de la politique linguistique et culturelle du nouvel empire, suggèrent qu'ils avaient parfaitement perçu ce risque et qu'ils étaient décidés à le contrer.

Concernant l'occupation des territoires conquis, la politique adoptée presque partout fut de regrouper les conquérants et leurs familles dans des villes de garnison (*'amṣār*) spécialement construites à cet effet en divers points stratégiques : *Kūfa* et *Baṣra* en Iraq, *Fuṣṭāṭ* en Égypte, Kairouan en Tunisie. Ce fait est d'une importance capitale : en effet, les Arabes, quoique minoritaires dans l'ensemble de l'empire (excepté l'Arabie, bien sûr), ne l'ont jamais été sur les sites où ils résidaient. L'arabe a donc toujours été la langue de

communication dominante voire exclusive sur ces sites, et les non-arabophones qui viennent y vivre sont tout naturellement amenés à acquérir une forme, aussi rudimentaire soit-elle dans un premier temps, de la langue des conquérants. En outre, ils y sont exposés à la forme d'arabe normalement utilisée par les arabophones eux-mêmes et non à une sous-langue réduite à l'expression de quelques actes de communication propres aux relations entre maîtres et serviteurs. Les conditions sociolinguistiques d'apparition de pidgins ou de créoles n'ont donc jamais été réunies dans les *'amṣār*.

Seule la Syrie a dérogé à la politique de regroupement des conquérants dans des villes de garnison. Elle était déjà partiellement peuplée d'Arabes et restait menacée par une contre-offensive byzantine. On y autorisa donc des implantations de colonies rurales ou villageoises fonctionnant comme autant d'avant-postes de défense. Les grandes villes de Syrie (Damas, Homs, Alep), massivement désertées par la population d'origine grecque, furent occupées par les conquérants. Damas devint, à partir de 641 et pour plus d'un siècle, la capitale de l'État musulman. Durant toute cette période, elle fut soumise à une sévère restriction de l'immigration arabe pour préserver les avantages acquis par les contingents ayant participé à sa conquête, lesquels allaient ainsi fournir le premier noyau d'une véritable armée professionnelle (*muqātila*) au service du califat omeyyade.

II. L'arabe langue officielle

Pour la gestion administrative et économique des provinces conquises, les Arabes ont procédé avec un remarquable pragmatisme : ils ont, dans un premier temps, maintenu les anciens cadres administratifs locaux dans leurs fonctions, les laissant assurer les prélèvements fiscaux exigés des populations soumises. Pendant cette phase de transition qui dura plusieurs décennies, le grec en Syrie et en Égypte, le pahlavi dans les provinces orientales, et le latin au Maghreb et en Espagne continuèrent donc d'être les langues administratives.

C'est sous le règne de l'Omeyyade *'Abd-al-Malik Ibn Marwān* (685-705), que vont être prises les mesures décisives pour changer cet état de chose. Ce calife accède au pouvoir alors que l'autorité centrale est gravement contestée par des révoltes en Iraq et un début de sécession à La Mecque. Son règne va donc être marqué par la volonté d'écraser toute tentative séditeuse, de renforcer partout

l'autorité du pouvoir central et son contrôle sur les pouvoirs régionaux, et enfin de jeter les bases d'une unification matérielle et symbolique du nouvel État. Diverses décisions sont prises dans ce sens comme la construction de routes stratégiques, ou la création d'une monnaie unique. Sur le plan de la politique linguistique deux mesures fondamentales vont être adoptées : l'instauration de l'arabe comme seule langue administrative de l'État, et l'amélioration du système graphique de la langue pour permettre une lecture enfin sans ambiguïté des documents écrits.

III. La réforme de l'écriture arabe

Ce point revêtait une double importance : d'une part il était essentiel que les responsables politiques et administratifs des provinces puissent lire et comprendre sans ambiguïté les consignes écrites reçues des autorités centrales, et en retour leur communiquer des rapports écrits dont l'interprétation soit claire. Mais d'autre part il fallait mettre fin aux divergences sur le texte coranique. Nous avons vu au chapitre précédent que les mouvements de contestation du pouvoir faisaient tous référence au Coran considéré comme le fondement de toute légitimité politique dans l'État musulman. Or l'état extrêmement rudimentaire de l'écriture arabe de l'époque autorisait les lectures les plus divergentes du texte, et chacun pouvait ainsi y trouver des arguments légitimant ses prétentions. Il n'est donc pas surprenant que, parmi le train de mesures touchant à la réforme de la langue sous le règne de 'Abd-al-Malik, figure un effort de normalisation de l'orthographe coranique (Blachère, 1958, 90 sq.). Une commission fut formée pour proposer une réforme globale du système graphique et des normes orthographiques de l'arabe. Elle généralisa le principe consistant à différencier les consonnes ayant même *ductus* par des points diacritiques notés sur ou sous le corps de la lettre ainsi que la notation des voyelles longues par les consonnes dites « d'allongement » (*ḥurūf al-madd*) : désormais on dut noter *ī* comme <(i)y>, *ū* comme <(u)w> et *ā* comme <(a)'>. Seront aussi créés quelques signes graphiques visant à donner une image plus fidèle de la langue orale, comme le signe de gémination des consonnes (*šadda*). Un système de notation des voyelles brèves sous forme de petits signes accessoires suscrits ou souscrits est également adopté. Ce système, sans doute inspiré du syriaque, connut diverses variantes avant de se stabiliser. Ces signes, à la différence des points diacritiques, ne sont devenus d'un emploi courant que dans le texte coranique, que l'on jugea important de rendre avec la plus grande fidélité possible, et

accessoirement, quoique sans véritable systématique, dans quelques ouvrages savants ou didactiques comme des recueils de poésie ou des livres d'enseignement.

L'orthographe systématisée à l'occasion de cette réforme du système graphique fonctionne sur un principe simple : on note la forme phonétique du mot tel qu'il se réalise à la pause. Ce principe explique la plupart des (rares) conventions orthographiques de la langue, comme l'ajout d'une 'alif après l'accusatif indéfini, prononcé [ā] à la pause, ou la notation du suffixe féminin singulier *at*, prononcé [ah] à la pause, par un graphème spécial formé par un *h* surmonté des deux points diacritiques caractéristiques du *t* et baptisé *tā'marbūṭa*, « t liée ». Certaines conventions ont partiellement une source morphologique, comme la notation d'un ā long final résultant d'une consonne *y* sous-jacente sous forme d'un *y* non diacrité, dit « 'alif raccourcie » ('alif *maqṣūra*). D'autres enfin semblent purement graphiques, comme l'ajout d'une 'alif après les suffixes *-uw* ou *-aw* du pluriel des verbes, peut-être pour empêcher leur confusion avec la particule de coordination *wa-*. Les mots outils formés d'une simple syllabe brève, trop courts pour être des mots phonologiques, sont (avec quelques flottements pour la particule de coordination *wa-*) cliticisés au mot suivant. De même, les pronoms suffixes sont intégrés au mot précédent comme le veut la phonologie qui les inclut dans les règles d'accentuation de ce mot. Tout cela montre que la réforme orthographique se fondait, au moins implicitement, sur une analyse morpho-phonologique très cohérente de la langue.

Ces réformes ne sont pas immédiatement adoptées, tant s'en faut. Ainsi, l'étude de vieux manuscrits du Coran (Blachère, 1958, 92 sq.) révèle que les mesures de normalisation du texte coranique ne seront adoptées que progressivement : les signes diacritiques se sont généralisés d'abord, puis, plus lentement, les signes vocaliques. C'est seulement à partir du milieu du ix^e siècle que la *scriptio plena* s'impose définitivement dans la notation du Coran.

Pour des raisons différentes, la décision d'imposer l'arabe comme langue administrative n'a pas pu non plus passer instantanément dans les faits, surtout dans les campagnes reculées, par manque de personnel arabophone compétent pour prendre le relais des bureaucraties byzantines et sassanides. En Égypte, par exemple, il a fallu une génération après les décrets de 'Abd-al-Malik pour que le grec cesse d'être utilisé dans les documents officiels (cf. Björnesjö, 1996). Mais sur le long terme, cette décision, liée au développement considérable des administrations centrale et régionale, a provoqué la naissance d'une « caste » de

scribes (*kuttāb*) qui n'a cessé de grossir et de se spécialiser. Ces scribes constituèrent bientôt une composante extrêmement importante et influente de la couche sociale que l'on appella « l'élite » (*al-kāṣṣa*), et contribuèrent de façon décisive au développement de la langue en élaborant la prose administrative, en usage dans les chancelleries, notamment en adaptant à l'arabe la longue tradition persane en la matière. Plus généralement, l'adoption et le respect par les *kuttāb* des normes linguistiques de la *fuṣḥā* élaborées par la grammaire naissante (voir ci-dessous) constituent une forme déterminante d'intervention de l'État dans les pratiques linguistiques par les modèles de « bon usage » qu'ils contribuent à mettre en place et à diffuser. Une part importante de la production des grammairiens, des philologues et des lexicographes est d'ailleurs, explicitement ou implicitement, destinée à la caste des secrétaires de chancelleries.

IV. La grammatisation

L'effort constant des Omeyyades, durant leur siècle de règne, fut de maintenir une barrière aussi infranchissable que possible entre les conquérants arabes, confinés dans les villes de garnison, et les populations soumises. Ils allèrent même jusqu'à décourager les conversions des autochtones à l'islam, non seulement pour maintenir à un niveau constant les rentrées de la taxe de capitation perçue sur les non-musulmans, mais aussi et surtout pour empêcher les nouveaux convertis de venir grossir les rangs des « clients » (*mawālī*) des tribus arabes. Ces nouveaux convertis se mettaient en effet au service des conquérants, apprenaient l'arabe à leur contact, et pouvaient même contracter mariage au sein de familles arabes, ce qui avait pour effet de brouiller la frontière, que les Omeyyades voulaient étanche, entre vainqueurs et vaincus d'hier. Mais, à terme, la politique omeyyade était vouée à l'échec : l'afflux de richesses dans les villes de garnison y attirait de plus en plus de candidats à la vie citadine, qu'il s'agisse d'Arabes de la péninsule ou d'autochtones qui renonçaient à la dure vie des campagnes et cherchaient une promotion sociale en acquérant le statut de *mawālī*. Les anciens conquérants eux-mêmes, s'étant accoutumés au confort de la vie citadine, n'aspiraient plus qu'à profiter des avantages de leur situation de rentiers ou de négociants et abandonnaient la carrière militaire. Un courant de plus en plus important de l'opinion approuvait le mouvement d'islamisation des populations conquises et souhaitait que, conformément aux enseignements du Coran et de la Tradition du Prophète, tous les musulmans soient traités sur un pied d'égalité. Plusieurs révoltes anti-

omeyyades éclatèrent en divers points de l'empire, toutes sévèrement réprimées. Mais la dernière, provoquée au milieu du viii^e siècle par une vaste coalition partie de la province iranienne du *Ḳorāsān* aboutit, avec une surprenante rapidité, à la chute des Omeyyades et à leur remplacement par la dynastie 'Abbasīde qui se réclame d'un retour aux vraies valeurs de l'islam (Shaban, 1971). La nouvelle capitale, Bagdad, établie en 762 sur le territoire de l'ex-empire sassanide marque un déplacement du centre de gravité de l'islam vers l'est, ce qui favorisa incontestablement l'influence persane sur son histoire culturelle.

C'est dans ce contexte qu'il faut situer le vaste mouvement de grammatisation de la langue arabe qui a conduit à sa standardisation par l'élaboration des premières grammaires et des premiers dictionnaires. Ce mouvement est inauguré à *Baṣra* et *Ḳūfa*, les deux grandes villes d'Iraq, par de pieux « lecteurs » du Coran qui s'efforcent de donner du Texte Sacré des commentaires explicatifs (*tafsīr*) à l'intention de leurs coreligionnaires, notamment des *mawālī*, qui, bien qu'ils aient appris un arabe populaire pratiqué par tous dans les villes, ne maîtrisaient guère la langue du Coran, la *fuṣṣḥā*, qui à l'époque n'était plus usitée, selon toute vraisemblance, que dans l'aristocratie citadine et chez les Bédouins du désert. La tradition nous apprend que la « dégradation » (*fasād*) de la langue arabe parlée par la majorité des citadins était aussi un sujet de préoccupation car elle semblait devoir rendre les Arabes des villes eux-mêmes incapables de vraiment comprendre le Message Divin délivré pourtant en « arabe clair » (*Suyūṭī*, chap. 44).

Les premières analyses linguistiques de l'arabe apparaissent donc dans des commentaires explicatifs du Coran où elles sont présentées un peu au hasard, lorsqu'un verset soulève une difficulté lexicale ou grammaticale particulière. Progressivement elles deviennent plus substantielles et plus systématiques, et on y trouve des débats entre partisans d'explications divergentes. Les commentaires grammaticaux du Coran évoluent ainsi vers ce qui devient de véritables grammaires.

Dans leurs explications du Coran, les exégètes ont, tout naturellement, tendu à faire usage de leur connaissance de la poésie arabe préislamique où ils trouvaient des constructions ou des mots permettant d'éclaircir le Texte Sacré. C'est ainsi que commença le travail de recension systématique de la poésie ancienne, mais aussi des « informations » (*'aḵbār*) qui l'accompagnaient et aidaient à en expliquer les circonstances, notamment les récits des guerres tribales de l'Arabie

ancienne (*'ayyām*). Tout cet héritage de l'Arabie préislamique (*Ġāhiliyya*) était désormais perçu comme un moyen d'accès à la *fuṣḥā* et donc comme un auxiliaire indispensable de l'explication du texte coranique. On évalua que la période durant laquelle la langue des poètes était restée suffisamment pure pour que l'on pût se servir de leurs vers comme preuve (*ḥuġġa*) d'un fait de langue couvrait toute la période d'avant l'islam et celle du i^{er} siècle et demi de l'Hégire (donc jusqu'au milieu du viii^e siècle de l'ère courante environ). Au-delà de cette période, la majorité des grammairiens considérera que la langue des poètes devenait sujette à caution. On notera en passant que cette délimitation temporelle du champ des données intégrables dans la *fuṣḥā* exclu toute la floraison des poètes « modernes » (*muḥdatūn*) qui s'épanouiront dans la mythique Baghdad 'Abbasīde, comme *Baššār Ibn Burd* (m. 785) ou *'Abū Nuwās* (m. ~ 815) et enrichiront le patrimoine « classique » de la littérature arabe. Elle exclut aussi tous les grands prosateurs de langue arabe. C'est dire que de nombreux usages de la langue « classique » au sens historico-culturel de ce terme resteront totalement ignorés de la grammaire arabe traditionnelle.

Des parlers proches de la *fuṣḥā* sont restés en usage dans certaines tribus d'Arabie jusqu'au x^e siècle au moins (Fück, 1955). Les locuteurs de ces parlers seront utilisés comme informateurs directs dans la première période d'élaboration de la grammaire. On s'efforça donc de délimiter l'aire géographique dans laquelle des variétés représentatives de la *fuṣḥā* étaient supposées être encore parlées. En gros furent retenus comme informateurs valables les membres des tribus de l'Arabie centrale et orientale (cf. *Suyūṭī*, I, chap. 2). Les divers parlers arabes de la péninsule étaient à l'origine nommés *luġāt* « langues » : on disait par exemple *luġat Tamīm* « la langue des *Tamīm* ». La *fuṣḥā* apparaissait alors comme « la plus claire » de ces langues, au sens, sans doute, de « la moins marquée de particularismes ». Les *ga*, en en faisant la seule langue légitime, classèrent toutes les autres par rapport à elle et établirent, dans un espace où il n'existait que des différences, une hiérarchie où la *fuṣḥā* occupait le sommet. Graduellement, les autres « langues » sont devenues des *lahaġāt*, « dialectes » dévalorisés. Cette conséquence inéluctable de l'instauration d'une langue légitime, aggravée par l'écart qui se creusait entre la *fuṣḥā*, fixée par l'écrit, et les autres variétés, qui ont librement évolué dans l'usage oral, constitue un élément permanent du paysage sociolinguistique arabe.

Selon la tradition arabe (cf. par exemple *Suyūṭī*, chap. 44, qui compile des sources anciennes), la création de la grammaire arabe remonte à un certain *'Abū l-'aswad al-Du'alī* (m. ~ 688) à qui elle associe toute une postérité de

grammairiens dont elle nous dit même qu'ils s'étaient scindés entre deux écoles rivales, *Kūfa* et *Baṣra*. Le plus souvent ces grammairiens ne nous sont connus que de nom car la plupart de leurs travaux ne nous sont pas parvenus. Mais il n'y a pas lieu de douter de l'existence, depuis au moins le début du viii^e siècle, d'une intense activité de réflexion sur la langue, même si l'on a de plus en plus de raisons de douter de l'exactitude de l'image qu'en donne la tradition [\[1\]](#).

La plus ancienne grammaire de l'arabe qui nous soit parvenue a pour auteur un Persan, connu sous le (sur)nom de *Sībawayhi* (m. ~ 796). Cet ouvrage, simplement intitulé « Le Livre » (*al-Kitāb*) n'a pas été immédiatement reconnu, mais une fois qu'il s'est imposé, près de cent ans après la mort de son auteur, il a eu une influence considérable et a été étudié et cité par tous les grammairiens (même si, souvent, ils en ont déformé la pensée). Il continue d'être la matrice de tous les travaux ultérieurs sur la grammaire arabe (sur *Sībawayhi* et le *Kitāb*, cf. Carter, 2004). Il est intéressant de s'arrêter sur la hiérarchie des données utilisées par *Sībawayhi* dans la construction de sa grammaire, car elle permet de mieux cerner les bases véritables du processus de standardisation de la langue (Carter, 2004, 39 sq.) : en premier lieu vient le parler spontané des Bédouins pratiquant encore la *fuṣḥā*, ce que l'on désigne comme « le parler des Arabes » (*kalām al-'arab*), étant entendu que « Arabes » signifie dans ce contexte « Arabes bédouins », puis vient la langue artificielle de la poésie, suivie par la langue particulière du Coran, ensuite viennent les recueils de récits (*ḥadīth*) relatifs à la vie du Prophète, que l'on avait commencé à collecter vers la même époque que l'ancienne poésie, puis les proverbes et expressions idiomatiques et enfin les phrases et mots forgés *ad hoc*.

Cette hiérarchie est instructive à plusieurs égards. D'abord elle permet de relativiser l'idée selon laquelle la *fuṣḥā* était strictement identique à la langue de la poésie préislamique ou à celle du Coran : on voit qu'en fait ces deux corpus viennent seulement en deuxième et troisième position comme sources de la langue standard. Un examen approfondi des faits confirme ce premier jugement. Ainsi, *Sībawayhi*, dès le chapitre introductif du *Kitāb*, où il expose sa méthodologie, consacre une section aux particularités de la langue poétique pour montrer que nombre d'entre elles contreviennent à l'usage normal de la langue. De même, certaines formes coraniques, comme les formations verbales dérivées *yaḍḍakkaru* (Coran, 2 : 269) ou *lā tafarraqu* (Coran, 3 : 103), ne seront pas retenues dans la langue standard. En second lieu cette hiérarchie prouve formellement qu'au moins jusqu'à l'époque de *Sībawayhi* une forme d'arabe à flexions casuelles était encore d'usage courant chez les Bédouins. À ce propos,

Zwettler (1978, 134) a suggéré que la contribution des Bédouins à l'élaboration de la grammaire de l'arabe standard devait se réduire à des vers de poésie appris par cœur, ce qui ne suppose nullement qu'ils maîtrisaient encore une forme de langue quotidienne utilisant la flexion casuelle. Mais cette thèse ne résiste pas à un examen systématique des productions linguistiques qui leur sont attribuées par les ga. Pour ne donner qu'un seul exemple (emprunté à Carter, 2004, 57), *Sībawayhi* cite un comportement linguistique des Bédouins consistant à insérer après un nom entièrement fléchi comme *sayfun* « épée + nominatif + indéfini » une voyelle longue *ī* ayant pour fonction d'indiquer à l'interlocuteur que l'on n'a pas encore fini de parler, ce qui donne par exemple *hāḏā sayfunī...* « C'est une épée... » Un tel fait n'a aucune chance d'apparaître dans un vers de poésie, et son appartenance à la langue du discours quotidien est manifeste, ce qui invalide la théorie de Zwettler. Il paraît donc difficile de contester qu'à l'époque de *Sībawayhi* certaines tribus bédouines pratiquaient encore, comme langue conversationnelle, un arabe de type flexionnel, même si l'on ne doit pas identifier purement et simplement les parlers tribaux et la *koinè* poétique (cf. sur cette question Rabin, 1955). Pour ce qui est de l'arabe parlé dans les villes, on est bien moins renseigné. Il est cependant probable que le mélange de populations arabes d'origines diverses, notamment des « marges » de l'arabophonie, et dont la langue avait probablement évolué depuis longtemps vers un type non flexionnel, avait dû favoriser la généralisation dans l'usage populaire d'une forme d'arabe sans marques casuelles (cf. sur ce point Corriente, 1976, qui parle à ce propos d'une variété « nabatéenne » d'arabe). Les textes arabes sur papyrus étudiés par Hopkins (1984), et dont les plus anciens remontent au début du ix^e siècle, semblent conforter cette hypothèse car ils montrent qu'à côté de quelques tours qui rappellent encore la *fuṣḥā* le gros du corpus pointe incontestablement vers ce que nous voyons se développer plus tard comme du « moyen arabe » sans flexion. Cela n'implique pas cependant que l'arabe flexionnel avait, à l'époque, complètement disparu de l'usage parlé citadin, au moins dans certaines catégories sociales. Il paraît significatif à cet égard que certains récits rapportant les fautes de langue des *mawālī*, voire de certains Arabes, mettent dans leurs bouches des usages erronés de marques casuelles. Si ces marques avaient complètement disparu de tout usage courant on comprendrait mal que de telles fautes aient pu se produire. Les sources littéraires nous apprennent par ailleurs qu'à la cour des califes et dans la bonne société l'usage de la *fuṣḥā* est resté de rigueur au moins jusqu'au début du ix^e siècle. En outre, et jusqu'au x^e siècle, il était courant, dans la classe dirigeante, d'envoyer les enfants mâles faire des « séjours linguistiques » de longue durée dans des

tribus bédouines parlant encore un arabe « pur » (Fück, 1955, chap. 9) pour satisfaire à l'exigence de maîtrise orale de cette variété de langue dans la bonne société. Cela permet de penser que la *fuṣḥā*, outre sa conservation dans certaines tribus bédouines, était encore pratiquée dans l'aristocratie citadine aux premiers siècles de l'empire musulman, et expliquerait que les *mawālī* ou les Arabes citadins d'origine populaire aient pu vouloir, maladroitement, en imiter l'usage. Néanmoins, le fait que les enfants, même dans les milieux aristocratiques, en aient progressivement perdu la maîtrise, comme nous l'indiquent certains récits, et comme le suggère la nécessité des « stages linguistiques » dont nous avons fait état, montre que le recul de cette variété d'arabe devant la variété non flexionnelle était largement entamé dans les villes.

Ces considérations rendent d'autant plus significatif le fait que les ga aient choisi comme norme linguistique la variété flexionnelle de la langue, variété dont il est très vraisemblable qu'elle n'était plus couramment en usage chez la majorité des arabophones citadins. Ce choix n'est évidemment pas le fruit du hasard. Il résulte de la structure du champ culturel de la société musulmane d'alors, champ dominé par deux courants symboliques majeurs, celui de l'arabité et celui de l'islam. Le premier reflétait les intérêts de la caste conquérante, attachée à ses privilèges et peu soucieuse de les partager avec les autochtones, même convertis. Cette caste avait donc tout intérêt à encourager la promotion comme langue officielle de la variété d'arabe la moins aisément accessible aux non-Arabes, celle dont la maîtrise était alors l'apanage des conquérants. Le second courant, celui de l'islam, était en principe favorable à une fusion de tous les musulmans dans la nouvelle société, proclamant « qu'un Arabe ne peut être supérieur à un non-Arabe que par la foi ». Mais le Coran ayant été délivré dans cette même *fuṣḥā*, c'était une bonne raison pour vouloir en faire la langue légitime de la nouvelle communauté. L'adoption de la *fuṣḥā* comme langue officielle était donc un dénominateur commun des deux courants dominants de la nouvelle société et par suite l'instrument linguistique tout désigné du dispositif de contrôle social (Carter, 1983). Dès lors la *fuṣḥā* sera un élément irréfragable de l'édifice culturel arabo-musulman, et toute recherche qui portera sur la langue lui donnera une place essentielle, tout particulièrement au phénomène de la flexion qui en est l'élément le plus manifeste. Il est très significatif, à cet égard, que celle-ci ait reçu, dans la terminologie grammaticale, le nom de *'i'rāb* qui signifie étymologiquement « arabiser, rendre arabe, donner un caractère arabe », sachant que dans ce contexte « arabe » renvoie aux Arabes bédouins, mythiques détenteurs de la *fuṣḥā*. Le *'i'rāb*, justement parce qu'il n'était plus d'un usage courant que parmi les grandes tribus arabes bédouines, et peut-être l'aristocratie

arabe des villes, était devenu un trait linguistique inégalement distribué parmi les arabophones et par là même fortement discriminant du point de vue sociolinguistique. Il remplissait ainsi toutes les conditions pour fournir à ceux qui en auraient conservé (ou acquis) l'usage un « profit de distinction », selon l'expression de Pierre Bourdieu (2001, 85). Dans l'histoire de la culture arabe, quiconque maîtrisera le *'i'rāb* bénéficiera de ce profit de distinction et pourra le convertir non seulement en capital symbolique (prestige, statut de locuteur « autorisé »), mais aussi en profits matériels de toutes sortes (diplômes, distinctions, postes). La maîtrise du *'i'rāb* constitue donc l'objectif de toute grammaire de l'arabe (Kouloughli, 2007).

La lecture des premiers traités de grammaire arabe déroute le lecteur moderne, les sujets y étant traités selon une démarche assez arbitraire. À partir du x^e siècle, sous l'influence de la logique grecque, diffusée par un vaste mouvement de traduction (cf. chap. suivant), s'impose l'idée qu'un exposé scientifique doit s'organiser selon une démarche systématique. Le premier grammairien à mettre en pratique cette idée est *Ibn al-Sarrāġ* (m. 928), qui avait étudié la logique avec le philosophe *al-Fārābī* (m. 950). Son *Kitāb al-'uṣūl fī l-naḥw* (Le livre des fondements de la grammaire) présente l'ensemble de la matière à partir d'un exposé méthodique fondé sur les propriétés grammaticales des parties du discours et les mécanismes d'assignation du *'i'rāb*. Progressivement cette démarche est adoptée par tous les grammairiens.

Les recherches lexicographiques, contemporaines de celles sur la grammaire, ont eu la même motivation : recueillir toutes les données aidant à l'explicitation du Coran. Ici encore on a eu recours à des enquêtes linguistiques auprès des tribus bédouines ayant gardé une langue « pure ». Un effet de ces enquêtes « tous azimuts » est que le lexique général de l'arabe s'est démesurément gonflé de doublets dialectaux ou de formes plus ou moins synonymes et reflète, d'une manière hypertrophiée, l'univers sémantique des Bédouins.

Les premiers travaux lexicographiques sont de courtes épîtres recensant tous les mots relatifs à un champ lexical donné : les parties du corps, le cheval, les phénomènes atmosphériques, *etc.* Ces épîtres se développent en véritables traités en accumulant de plus en plus de données, pas toujours selon une logique bien claire. Ces ouvrages semblent avoir été principalement destinés aux secrétaires des chancelleries (*kuttāb*) soucieux d'enrichir leur vocabulaire. Le premier dictionnaire organisé sur la base purement formelle des consonnes radicales (*ġuḍūr* ou *'uṣūl*) sous-jacentes à tout lexème arabe, est attribué au génial al-

Ḳalīl Ibn 'Aḥmad (m. ~ 791 ?), principal maître de Sībawayhi, et inventeur, selon la tradition, de la première formalisation de la métrique arabe. Son dictionnaire est connu sous le nom de *Kitāb al-'ayn*, du nom de la consonne fricative laryngale sonore de l'arabe, qui lui a servi comme point de départ du classement phonétique des racines. On se demande encore si al-Ḳalīl est vraiment l'auteur de cet ouvrage car l'idée d'un recensement du lexique à partir des racines suppose une vision abstraite et totalisante de la langue qui paraît assez stupéfiante à une date aussi précocce. Elle est en tout cas une parfaite illustration de l'idée qu'un dictionnaire vise à rassembler de façon exhaustive et en juxtaposant tous les usages possibles, la totalité des ressources lexicales d'une langue, réalisant par là une sorte de matérialisation de l'idée abstraite de langue élaborée par la grammatisation. Cette approche formelle du lexique ne sera cependant généralisée qu'à partir du x^e siècle par des auteurs comme Ibn Durayd (m. 933), dans sa *Ġamhara*, ou Ibn Fāris (m. 1004) dans son *Kitāb Maqāyīs al-Luġa*. Ils renoncent d'ailleurs au classement par points d'articulations des racines pour adopter un ordre alphabétique. Al-Ġawharī (m. 1002) dans son *Ṣiḥāḥ* adopte un classement à partir de la dernière radicale des mots, proposant ainsi une sorte de dictionnaire de rimes répondant aux attentes du public, la prose rimée (*sağ'*) étant devenue le style d'écriture à la mode dans les chancelleries et les assemblées littéraires (*mağālis*). Le développement ultérieur de la lexicographie arabe se poursuit dans le sens du recensement exhaustif du lexique de la *fuṣṣḥā*, notamment de ses mots rares et peu connus (*ġarīb*). Elle culmine dans le gigantesque dictionnaire d'Ibn Manẓūr (m. 1311) intitulé significativement *Lisān al-'Arab* (La langue des Arabes). Par contre, l'énorme lexique issu du développement des sciences tant religieuses que profanes reste superbement ignoré des grands lexicographes. C'est seulement à une période plus tardive que des ouvrages comme le *Kitāb al-Ta'rīfāt* (Livre des définitions) de 'Alī al-Ġurġanī (m. 1413) comblent cette lacune.

Notes

[1] Cf. là-dessus R. Talmon (2003), *Eighth-Century Iraqi Grammar : A Critical Exploration of pre-Kalilian Arabic Linguistics*, Winona Lake, Harvard Semitic Studies Series 53, et G. Bohas, J.-P. Guillaume et D. E. Kouloughli (2006).

Chapitre V

Constitution du domaine arabe

La période qui va du ix^e au milieu du x^e siècle, voit l'apogée de la puissance politique, économique et culturelle de l'empire arabo-musulman, en même temps qu'une sorte d'âge d'or de la langue arabe. Désormais dotée d'un outillage linguistique qui n'a plus rien à envier à celui des vieilles langues de civilisation dont elle prend la place, et portée par la vitalité économique et culturelle de ce qui était alors devenu la société la plus dynamique et la plus développée du monde, la *fuṣḥā* voit s'épanouir toutes ses potentialités. Parallèlement, se met en place une nette bipolarisation entre langue savante et langue populaire, en même temps que s'étend considérablement l'aire géographique arabophone.

I. L'essor de l'arabe classique

On peut considérer que trois grandes sphères d'activité intellectuelle vont, par leur développement, contribuer, chacune à sa manière, à l'essor de la *fuṣḥā* : celle des « sciences islamiques » (*al-ʿulūm al-ʾislāmiyya*), celles des « sciences profanes » (*al-ʿulūm al-daḡīla*), essentiellement issues de la traduction des œuvres du patrimoine philosophique et scientifique des vieilles cultures du Moyen-Orient, enfin celle de la littérature (*ʿadab*).

Sous le nom de « sciences islamiques » on inclut la totalité des activités intellectuelles ayant le Coran pour source ou pour but. Ainsi la grammaire et la lexicologie, dont nous avons déjà parlé, font partie de ce domaine, puisque leur objet premier est l'explicitation du texte coranique. Mais la recension et l'étude de la poésie préislamique et des textes relatifs à l'Arabie d'avant l'islam sont aussi, sans paradoxes, intégrées dans ce domaine en raison du rôle que la connaissance de ces textes peut jouer dans l'exégèse coranique [\[1\]](#). Les traditions relatives à la vie du Prophète et les récits des conquêtes (*futūḥat*) font aussi, tout naturellement, partie de ce domaine. Mais le Coran contient également les

fondements du droit musulman tant civil que pénal et même commercial. Il s'ensuit que le droit (*fiqh*) et toutes les disciplines qui s'y rattachent relèvent du champ des sciences islamiques. Il faut admettre d'ailleurs que ces diverses disciplines, tout en explorant des champs de savoir bien distincts, présentent un indéniable « air de famille ». Cela est dû en grande partie à une méthodologie foncièrement identique, et qui n'est pas sans rappeler la démarche scolastique en usage dans les activités théoriques analogues dans l'Occident médiéval, avec leur cortège de définitions, de thèses, d'objections et de réponses. En outre, de nombreux spécialistes travaillent dans plusieurs de ces domaines à la fois et tous sont formés à une sorte de « tronc commun » de plusieurs de ces matières.

La formation à ces disciplines est assurée dans des institutions financées par le mécénat public ou privé, à travers la constitution de fondations religieuses (*waqf*). Elle comporte en gros deux niveaux : celui de l'école élémentaire (*maktab* ou *kuttāb*) où l'on acquiert les bases de la grammaire et où l'on apprend par cœur le Coran, puis celui de la formation intermédiaire et, pour les plus motivés, supérieure, assurées dans des collèges (*madrassa*) comme la *Niẓāmiyya* à Bagdad, ou dans des grandes mosquées (*ḡāmi'*). Certains de ces établissements tels *al-'Azhar* au Caire, *al-Zaytūna* à Tunis ou *al-Qarawiyyīn* à Fez auront un véritable statut d'université et formeront les cadres juridiques et religieux dont la société musulmane a besoin.

Sur le plan linguistique, l'ensemble des sciences islamiques donne naissance à une énorme terminologie technique extrêmement spécialisée, mais aussi et surtout à un nouveau type de prose, que l'on peut qualifier de technique et qui se caractérise, malgré une certaine lourdeur, par une indéniable précision analytique. En particulier, les articulations logiques de l'argumentation y sont rendues explicites par l'usage de particules qui, à l'origine essentiellement coordinatives, conformément au génie de la vieille *fuṣṣḥā* orale, acquièrent de nouvelles valeurs de subordination permettant d'exprimer toutes les nuances du raisonnement.

Quant à l'expression « sciences profanes », elle désigne l'ensemble des disciplines intellectuelles constituées dans les cultures que les Arabes ont trouvées dans les territoires conquis et auxquelles ils accèdent par la traduction. Elles incluent, outre la philosophie (*falsafa*), les disciplines scientifiques (mathématiques, sciences naturelles, médecine, astronomie) et parascientifiques (astrologie, alchimie) que les cultures grecque, syriaque, perse, indienne, voire chinoise ont élaborées et qui circulent sous diverses formes dans le Moyen-

Orient de l'époque. Il faut ici signaler que les Perses sassanides avaient fondé, dans la ville de *Ġundīšāpūr*, une école de médecine qui était devenue une véritable académie des sciences en accueillant philosophes, scientifiques et médecins chassés de l'empire byzantin par les persécutions religieuses ou en faisant venir des savants de l'Inde voisine. Une importante activité de traduction d'ouvrages scientifiques du grec en syriaque semble y avoir eu lieu. Le calife al-Ma'mūn (m. 833) en fondant à Baghdad la célèbre « Maison de la sagesse » (*Bayt al-Ḥikma*) poursuit cette politique éclairée au profit de la langue arabe. Des générations de brillants traducteurs, tels Ḥunayn Ibn 'Ishāq (m. 877), traducteur de Galien et d'Hippocrate, ou Qusṭā Ibn Lūqā (m. ~ 912), traducteur d'ouvrages de mathématiques et d'astronomie, ont en un temps remarquablement court enrichi la langue arabe de tout l'héritage philosophique et scientifique alors en circulation au Moyen-Orient. La découverte et l'étude de ce précieux patrimoine suscitent un engouement certain malgré l'hostilité de musulmans plus conservateurs comme Ibn Hanbal (m. 855) qui condamne l'étude de ces disciplines qu'il soupçonne de nourrir l'impiété.

Le transfert en arabe du corpus philosophique et scientifique grec avait été en partie préparé par sa traduction préalable en syriaque, langue sémitique cousine de l'arabe, et l'on peut penser que ce travail a facilité la traduction ultérieure vers l'arabe. Il n'en reste pas moins que ce transfert a nécessité un important travail d'adaptation de la *fuṣṣḥā*.

Au niveau du lexique, il a fallu la doter d'un nouveau vocabulaire conceptuel, et notamment rendre de façon précise et cohérente les paradigmes de notions que le grec exprime au moyen de sa morphologie dérivationnelle, procédé pratiquement inconnu en arabe. Les traducteurs ont parfois essayé le calque structurel, comme dans *lā-wuġūd* (inexistence), ou le détournement de suffixes dérivationnels de la langue, comme *-iyya*, suffixe féminin marquant l'origine, et devenu un moyen très productif de former des notions abstraites comme *kayf-iyya* (qualité) formé sur l'interrogatif *kayf* (comment ?). Le plus souvent, ils ont eu recours à la prodigieuse richesse du lexique arabe pour trouver une racine susceptible d'exprimer « en bloc » la nouvelle notion. Mais ils n'ont pas hésité à recourir à l'emprunt pur et simple. Contrairement à une opinion répandue, l'arabe n'a jamais répugné aux emprunts comme l'attestent la poésie préislamique et le Coran. Mais indiscutablement, c'est dans la période initiale de son développement scientifique et technique qu'il en a fait le plus, notamment au grec, au pahlavi et au syriaque, lorsque la forme du mot emprunté permettait son intégration dans la structure morphologique de la langue. Si au contraire sa

forme restait perçue comme trop « étrangère », il arrivait souvent qu'il soit finalement remplacé par un mot senti comme plus « arabe ».

Au niveau de la syntaxe, la tendance à développer l'usage des connecteurs logiques et de la subordination, tendance déjà signalée à propos des sciences coraniques, joue ici aussi pleinement. Signalons en passant un trait qui semble caractériser les traductions arabes de textes philosophiques, à savoir la tendance à introduire systématiquement entre le thème (*mubtada'*) et son propos (*ḵabar*) dans la phrase « nominale » un pronom de séparation (*ḍamīr faṣl*) pour rendre la copule du grec. Cette construction n'est en principe courante en arabe que si les deux termes de la phrase sont également définis. Son usage quel que soit le statut définitoire de ces deux termes semble bien une sorte de « calque » structurel de la phrase grecque.

Les œuvres scientifiques directement rédigées en arabe suivent assez vite les traductions, voire les accompagnent. Directement conçues par des arabophones, elles perdent le caractère un peu raide et artificiel des traductions, mais elles gardent, inévitablement, les traits d'une sous-langue spécialisée et exigent un lectorat savant spécifiquement formé à leur utilisation.

Le troisième courant qui concourt à l'élaboration de la culture arabe classique est celui de la littérature dans l'acception très particulière donnée alors au terme arabe de *'adab* et qui désigne non pas simplement la prose littéraire mais l'éthique et l'art de vivre qu'elle a pour mission de véhiculer dans la nouvelle culture.

La prose littéraire en langue arabe n'est pas née *ex nihilo*, et l'on peut lui assigner trois sources au moins : les deux premières, qui remontent à la période préislamique, sont d'une part celle des récits relatifs aux guerres tribales (*'ayyām al-'arab*), et d'autre part la harangue publique (*ḵuṭba*) que les chefs de tribus délivraient dans les grandes occasions ; la troisième, est celle des récits relatifs à la vie du Prophète (*sīra*) et des conquêtes islamiques (*futūḥāt*) auxquels on peut ajouter les récits et sermons délivrés par les conteurs populaires (*quṣṣās*) même si, probablement, la langue qu'ils utilisaient a dû, assez tôt, diverger des normes de la *fuṣḥa*.

Avec la prose administrative « de qualité », à laquelle on associe les noms de 'Abd al-Hamīd al-Kātib (m. 750) et de Ibn al-Muqaffa' (m. 757), on entre dans le *'adab* proprement dit, avec sa double dimension éthique et esthétique.

Concernant la première, les deux auteurs mettent constamment en avant, dans ce qui nous est parvenu de leurs écrits, l'importance pour l'homme de culture de respecter une sorte de code d'honneur où la fidélité envers ses amis, la sincérité à l'égard des supérieurs et la bienveillance pour les subordonnés sont des règles intangibles. Quant à la dimension esthétique, leurs écrits manifestent par la recherche d'une élégance sans affectation et d'une précision sans lourdeur qu'ils avaient une haute conception de l'écriture. La traduction-adaptation du persan à l'arabe du célèbre *Kalīla wa-Dimna*, par Ibn al-Muqaffa' reste l'un des premiers monuments de la prose littéraire arabe.

Mais c'est avec l'œuvre immense et protéiforme de Ġāḥiẓ (m. 869) que le '*adab*' prend toute sa dimension. Cette œuvre naît dans un contexte de bouillonnement culturel caractérisé notamment par l'émergence d'un courant idéologique « anti-arabe » porté par la catégorie sociale des *mawālī* d'origine persane et contestant les privilèges sociaux et l'arrogance des descendants des conquérants Arabes. Ce courant, connu sous le nom de *šu'ūbiyya*, ne conteste pas l'hégémonie linguistique de la langue arabe, dans laquelle il exprime ses revendications, mais exige une stricte égalité entre tous les musulmans, et l'intégration, dans l'édifice culturel de la nouvelle société, des apports des vieilles cultures dont sont issus ses représentants. Dans certains de ses écrits Ġāḥiẓ semble se poser, face à la *šu'ūbiyya*, en défenseur d'un arabisme pur et dur. Mais *mawlā* lui-même, universaliste convaincu, et en outre sectateur du rationalisme religieux *mu'tazilite*, il défend et illustre dans l'ensemble de son œuvre une exceptionnelle ouverture à tous les apports culturels, qu'ils soient d'origine grecque, persane, indienne... ou arabe. Ce syncrétisme universaliste restera une caractéristique constante du '*adab*' de bon aloi. L'autre héritage durable de Ġāḥiẓ, qui reste aussi une caractéristique du meilleur '*adab*', est l'usage d'une langue à la fois souple et précise, faisant souvent usage de constructions parallèles, de citations du Coran et de la poésie et n'ayant qu'un recours léger à la prose rimée (*sağ'*) dont l'abus deviendra, par contraste, une plaie de la prose littéraire décadente (à partir du xii^e siècle).

Pour conclure ce bref survol des champs d'activité intellectuelle qui ont influé sur le devenir de l'arabe, insistons sur l'effet convergent qu'ils ont eu sur la constitution d'une véritable langue écrite, propre à exprimer précisément des idées abstraites et à marquer l'enchaînement des raisonnements et des argumentations. Les sources de la *fuṣṣḥā*, poésie préislamique ou Coran, étaient originellement de l'oral, fixé secondairement par l'écriture. La langue qu'elles utilisaient restait profondément marquée par cette origine, utilisant un lexique

essentiellement concret et faisant un usage quasi exclusif de la juxtaposition (parataxe) et de la coordination pour suggérer les relations entre les idées. Il était souvent indispensable de connaître le contexte extralinguistique pour bien comprendre les références, comme le montre le rôle fondamental des commentaires pour une compréhension correcte du Coran. Au contraire, la langue élaborée par les spécialistes des sciences religieuses ou profanes, et, dans une mesure non négligeable celle des scribes (*kuttāb*) et des écrivains (*'udabā'*), est de l'écrit authentique, destiné à être lu et compris de façon « anonyme » et universelle. Pour ce faire, cette langue a développé tout un jeu de marques formelles de cohérence et de cohésion textuelle permettant de reconstituer avec le moins d'ambiguïté possible le réseau des références et l'enchaînement des idées. Par suite, cette langue, même si elle continue à faire usage du *'i'rāb*, n'est plus l'antique *fuṣḥā*, compacte et synthétique. C'est une nouvelle *fuṣḥā*, dont la syntaxe est de plus en plus analytique, et où le *'i'rāb* n'est plus, au vrai, qu'un ornement de moins en moins fonctionnel, remplacé qu'il est par un ordre des mots de plus en plus strict (cf. Kālīl, 1978 : chap. 2 qui propose d'appeler « arabe néo-classique » cette nouvelle forme de la *fuṣḥā*).

Signalons pour finir qu'à partir du xi^e siècle le papier, invention chinoise, a commencé à être d'un usage courant dans le monde musulman, ce qui a favorisé une plus large circulation de l'écrit.

II. La bipolarisation linguistique

La période qui voit le plein épanouissement de la *fuṣḥā* voit aussi se mettre en place la structure linguistique bipolaire qui va désormais caractériser le monde arabophone. À l'un des pôles de cette structure on trouve précisément la *fuṣḥā*, qui devient de plus en plus une langue exclusivement écrite, alors qu'à l'autre pôle se développe l'arabe parlé, diversifié selon l'origine tribale des locuteurs, leur mode de vie (citadins, villageois, paysan ou encore plus ou moins nomade), leur implantation géographique (villes de garnison, petites agglomérations urbaines, colonies villageoises).

Les spécialistes sont loin d'être d'accord sur la genèse de cette situation et d'abord sur l'origine de l'arabe parlé. Le débat s'est noué notamment autour de la caractérisation de la différence entre les deux variétés d'arabe. Pour certains auteurs, comme Blau (1977), Versteegh (1984), ou Ferguson (1959 et 1989) les

parlers arabes d'avant les conquêtes islamiques étaient foncièrement semblables entre eux (en tenant compte des différences dialectales dont nous avons fait état précédemment) et ne différaient de la *koinè* poétique qu'autant que la variété littéraire d'une langue diffère de ses variétés non littéraires. Pour eux, le dénominateur commun entre ces diverses variétés d'arabe est qu'ils appartiennent à l'« ancien arabe » (*Old Arabic*), de type « synthétique », par opposition au « néo-arabe » (*New Arabic*), de type « analytique ». La différence entre les deux types serait essentiellement que le premier tend à marquer les fonctions morphosyntaxiques par des procédés flexionnels (morphèmes liés, flexion interne) alors que le second le fait plutôt par des morphèmes libres (particules, prépositions). Le passage de l'ancien arabe au néo-arabe serait, dans cette optique, un changement net et radical de type linguistique (Blau, 1972). Ce point de vue ne fait cependant pas l'unanimité. Ainsi, pour Federico Corriente (1971), la perte du système casuel est l'aboutissement d'un processus qui était déjà en germe depuis longtemps et qui est dû au fait que ce système n'avait dans le fonctionnement effectif de la langue qu'un rendement fonctionnel négligeable : la structure syntagmatique de l'arabe et l'ordre des mots rendaient les marques casuelles redondantes, ce qui permettait de s'en passer sans aucun dommage pour la communication. D'un point de vue différent, mais qui plaide dans le même sens, Retsö (1993) fait valoir que l'expression du passif par affixation comme le font les parlers arabes actuels n'est pas moins synthétique que celle que l'arabe classique réalise par apophonie. Cette seconde façon de voir a pour conséquence de minimiser les différences entre arabe ancien et néo-arabe et de suggérer que le passage de l'un à l'autre est une question de degré et non de nature.

Quoi qu'il en soit l'accord entre Blau, Versteegh et Ferguson ne va guère au-delà de l'idée que l'écart entre les deux variétés d'arabe est postérieur aux conquêtes islamiques. Sur les conditions d'apparition de cet écart, les trois auteurs divergent profondément.

Pour Blau, ce sont les dialectes arabes sédentaires qui auraient donné naissance au type néo-arabe, et les changements considérables qui ont affecté les populations arabes à la suite des conquêtes en ont favorisé la diffusion. Ainsi s'expliqueraient à la fois les ressemblances observées entre dialectes (dues à la communauté de type linguistique) et leurs différences (reflétant les variations dialectales qui existaient déjà en Arabie). Les deux critiques que l'on peut opposer à cette théorie (Corriente, 1971 et 1976) concernent d'une part la thèse d'une apparition brusque et « catastrophique » du néo-arabe après l'islam et

d'autre part l'accentuation artificielle de la différence entre les deux variétés d'arabe. Sur le premier point en effet tout porte à croire que le type néo-arabe avait commencé d'exister bien avant l'islam. Sur le second, l'examen attentif de la *fuṣḥā*, notamment à son époque « néo-classique », de même que l'étude sans préjugés simplificateurs des dialectes montrent que la thèse des deux types linguistiques radicalement distincts est une vue de l'esprit.

Selon Versteegh les formes d'arabe spontanément produites par les *mawālī* ont inévitablement dû être des « pidgins », ces formes de langue rudimentaires qui se développent dans les situations où l'urgence de communiquer l'emporte sur toute autre considération. Les enfants des *mawālī* auraient ensuite transformé ces pidgins en créoles, moyens de communication plus étoffés et plus stabilisés. Ce sont ces formes créolisées d'arabe qui seraient à l'origine des parlers arabes actuels dont la diversité s'expliquerait largement par le fait qu'ils sont nés sur des substrats linguistiques divers et en l'absence de toute forme de régulation. Versteegh postule qu'ensuite une phase de « décréolisation » se serait produite du fait que « la dissémination de l'arabe classique comme langue prestigieuse de culture et de religion introduisit un modèle qui affecta la situation linguistique à un degré tel qu'entre parlers dialectaux et langue standard apparut un continuum linguistique analogue à la diglossie du monde arabophone actuel » (Versteegh, 1997, 109). Cette théorie a été diversement critiquée (Fergusson, 1989, et Holes, 1995). On a fait valoir qu'aucun document historique ne vient lui donner une ombre de soutien, alors que le caractère drastique des bouleversements qu'elle postule n'aurait pu passer inaperçu des grammairiens ou des chroniqueurs qui nous ont laissé de nombreuses observations sur les pratiques linguistiques du petit peuple des villes. On a aussi souligné que les témoignages les plus anciens des productions linguistiques en arabe non standard, en particulier le corpus de papyrus antérieurs au x^e siècle étudié par Simon Hopkins (1984), révèlent que la langue reflétée dans ces écrits éphémères et sans aucune prétention à la correction linguistique présente de nombreux traits néo-arabes, mais rien qui ressemble à la simplification drastique de la grammaire que supposent pidginisation ou créolisation. On a aussi fait observer que la forte ressemblance entre les parlers des diverses régions du Monde arabe, dont les langues substrat (araméen, copte, berbère) sont très différentes les unes des autres, est difficilement compatible avec l'hypothèse de Versteegh, car dans tous les cas documentés la langue maternelle marque fortement de son empreinte le créole résultant. Il faut noter en passant que la conception que se fait Versteegh de la « décréolisation » comme processus capable de réintroduire de la complexité morphologique dans un créole ne semble avoir aucun précédent attesté dans la

pourtant vaste littérature sur ce sujet.

Ferguson, pour sa part, pense que le fait que des populations arabes d'origines tribales diverses se sont trouvées mêlées dans les villes de garnison a entraîné un mélange des anciens dialectes tribaux et l'apparition de ce qu'il appelle une « *koinè* militaire » laquelle aurait été répandue aux quatre coins de l'empire par les mouvements de troupes. À l'appui de sa théorie, il cite 14 traits linguistiques que l'on trouverait, selon lui, dans tous les dialectes arabes actuels, mais qui seraient inconnus de la *fuṣḥā*. Il attribue ces traits et leur diffusion pan dialectale à la *koinè* militaire. Cette théorie a elle aussi été critiquée (notamment par Cohen, 1970, et Miller, 1986). Les critiques ont porté sur le fait que les ressemblances entre dialectes peuvent certes résulter d'une origine commune, mais aussi de développements convergents, plausibles pour des parlers apparentés, ou encore de tendances évolutives profondes partagées par tous ces parlers (par exemple la perte du duel a affecté de nombreuses langues sémitiques sur des périodes très longues). À cela il faut ajouter la possibilité d'une influence unificatrice *a posteriori* de certains parlers localement promus à une position de prestige, sans oublier, dans cette fonction, la *fuṣḥā* dont l'influence normalisatrice est toujours à l'œuvre. Ces critiques n'invalident pas complètement l'hypothèse de Ferguson (David Cohen lui-même admet que la langue composite parlée par les troupes arabes a bien été un input de l'arabisation). Elles contestent plutôt la thèse « monogénétique » qui la sous-tend, et font valoir qu'il n'est pas réaliste de supposer que tous les parlers arabes actuels descendent nécessairement d'une seule souche linguistique, quelle qu'elle soit. La typologie des parlers arabes actuels, que certains traits rattachent clairement aux dialectes arabes préislamiques reste en tout état de cause l'argument fondamental contre la thèse monogénétique.

Pour conclure cette discussion, nous dirons que les données disponibles laissent penser que la variation linguistique a caractérisé le domaine arabe depuis des temps très reculés, et que, selon toute vraisemblance, une variété d'arabe sans flexion était déjà en usage dans les marges de l'arabophonie avant l'islam. Cette variété d'arabe s'est largement répandue parmi les arabophones, notamment le petit peuple, dans les villes de garnison. C'est probablement ce type d'arabe que les *mawālī* ont appris spontanément au contact des arabophones citadins, sans doute de façon approximative pour la première génération, mais de façon quasi native pour les générations suivantes, souvent issues de mariages mixtes (*muwalladūn*). En effet, il n'existait pas de politique de ségrégation des *mawālī* susceptible de bloquer leur apprentissage de l'arabe à un stade de pidgin ou de

créole. En outre l'arabe auquel ils étaient régulièrement exposés, s'il n'était pas, tant s'en faut, conforme aux normes des grammairiens, n'avait rien non plus d'une langue mutilée, réduite aux seules exigences d'une interaction pragmatique éphémère. Enfin, pour des raisons évidentes de promotion sociale, les *mawālī*, ayant adopté l'islam, avaient tout intérêt à s'arabiser car la participation à de nombreuses activités économiques et sociales exigeait la maîtrise de cette langue. Quant à la *fuṣḥā*, langue liturgique de l'islam et langue officielle de l'État musulman, elle a été préservée un temps sous sa forme orale dans les tribus les moins en contact avec les milieux sédentaires, et peut-être aussi dans les milieux aristocratiques des villes comme signe identitaire et de différenciation d'avec le *vulgum pecus*. Codifiée et fixée par les *fuṣḥā*, elle se maintiendra, au moins en apparence, sous une forme intangible par la suite. En fait, l'examen de son usage écrit ultérieur montre qu'elle n'a pas complètement échappé aux tendances évolutives à l'œuvre dans toutes les variétés d'arabe. De même, l'évolution des parlers bédouins proches de la *fuṣḥā*, bien que retardée, a finalement été la même que celle des autres parlers arabes. Cela montre qu'il s'agit bien d'un processus d'évolution linguistique interne de la langue et non de quelque accident provoqué par des événements historiques contingents, même si de tels événements ont pu accélérer cette évolution.

III. L'expansion de l'arabe

À la veille des conquêtes islamiques, la situation des pays limitrophes de l'Arabie est à peu près la suivante : Le Yémen parle et écrit une forme de sudarabique « ancien », mais dans les montagnes omanaises, le Dhofar et l'île de Socotra, on pratique le sudarabique « moderne », une langue sémitique distincte, qui a survécu jusqu'à nos jours dans quelques régions du sud-ouest de la péninsule. L'ensemble de cette région est en contact permanent avec des tribus arabophones depuis des siècles. Les populations de l'ensemble du Croissant Fertile (Palestine, Syrie, Iraq) parlent une forme ou une autre d'araméen, même si les langues officielles sont le grec dans les régions contrôlées par les byzantins et le pahlavi dans l'empire sassanide. Ces langues officielles sont pratiquées par une fraction de la population des grandes villes, fraction étroite en Syrie, plus importante en Iraq. Diverses langues religieuses écrites sont également en usage, comme l'hébreu ou le syriaque (variété savante d'araméen utilisée par les églises chrétiennes dissidentes de Byzance) ; le *hinterland* perse parle et écrit le pahlavi. Les populations autochtones sont en contact, depuis des siècles, avec des tribus

arabes dont certaines sont sédentarisées alors que les autres continuent de nomadiser dans la steppe. Des échanges de toutes sortes existent entre ces tribus et les paysans autochtones, et diverses formes de bilinguisme arabe-araméen existent. En Égypte, le copte est la langue parlée par le peuple et écrite par le clergé national. Le grec, langue administrative, est pratiqué par la bourgeoisie hellénisée à Alexandrie et dans les centres urbains du Delta. Il semble d'autre part qu'il existait déjà, bien avant les conquêtes arabes, des implantations de nomades Arabes non seulement dans le désert du Sinaï, mais aussi entre la mer Rouge et le Nil. En Afrique du Nord, le berbère est la langue majoritaire mais sa forme écrite n'est pratiquement pas usitée. C'est le latin qui sert de langue administrative. Il est probable qu'une part de la population urbaine pratique une forme de proto-roman. Il n'est pas non plus impossible qu'ait subsisté dans la région de Carthage des poches de locuteurs du punique, langue sémitique des anciens carthaginois. Il n'y a eu aucun contact avec l'arabe avant les invasions musulmanes. Enfin en Espagne et dans les îles de la Méditerranée, la langue parlée par les autochtones est une forme de proto-roman et la langue administrative le latin. Ici il n'y a pas eu de contact avec l'arabe avant la conquête musulmane.

En relativement peu de temps, l'arabe va devenir la langue parlée de tout le Croissant fertile, de l'Égypte et d'une large partie de l'Afrique du Nord. Il va aussi, pendant plusieurs siècles être parlé par une fraction de la population de la partie de l'Espagne sous domination musulmane, ainsi qu'en Sicile et à Malte. La *fuṣḥā*, connaîtra dans ces régions un brillant développement. Les invasions normandes des îles de la Méditerranée (x^e et xi^e siècles) puis la *Reconquista* chrétienne en Espagne (xi^e-xv^e siècle) chassent l'arabe de la plupart de ces territoires. Le cas de Malte est tout à fait singulier : conquise par les Arabes en 870, la petite île est complètement arabisée, si bien que malgré sa conquête par les Normands et le fait qu'elle ait été, depuis, coupée du hinterland arabophone, on continue à y pratiquer une forme tout à fait caractéristique d'arabe de type maghrébin.

À l'autre extrémité de l'Empire musulman, en Iran, l'arabe écrit a connu également une période de développement brillant, mais n'a jamais réussi à devenir la langue parlée d'une fraction significative de la population. En tant que langue littéraire il est, à partir du milieu du x^e siècle, concurrencé par la renaissance du persan. Il n'y subsiste, par la suite, que comme langue religieuse et juridique.

L'arabisation des régions encore aujourd'hui arabophones ne s'est cependant pas faite de façon uniforme. Le monde araméophone semble bien avoir entièrement adopté l'arabe, dès le ix^e siècle, peut-être en raison de la proximité linguistique des deux langues, mais aussi grâce à l'existence préalable de couches importantes de la population du Croissant fertile qui étaient plus ou moins bilingues. Un troisième facteur, peut-être le plus décisif, a été l'afflux massif et continu de tribus arabes de la péninsule. Ce phénomène a d'emblée été important en Iraq, ce pays servant en quelque sorte de relais pour les troupes arabes en route pour l'Iran et les provinces plus orientales de l'empire pendant la période des conquêtes. En Syrie, l'immigration arabe a été bloquée pendant toute la période omeyyade, mais a repris dès la chute de la dynastie de Damas. En Égypte, le processus a été beaucoup plus lent. Concernant la langue écrite, il fut relativement rapide puisqu'il semble que le premier papyrus bilingue grec arabe date de la fin du vii^e siècle (sous le règne de 'Abd al-Malik Ibn Marwān) et le dernier d'à peine un demi-siècle plus tard, le premier document monolingue arabe se situant à mi-chemin, au début du viii^e siècle. Mais ces données ne nous disent rien sur la progression de l'arabe comme langue orale. Sur ce point il faut chercher des témoignages indirects. Par exemple, lors de son voyage en Égypte en 832 le Calife *al-Ma'mūn* a besoin d'interprètes. Mais dès le x^e siècle, on sait que le clergé copte utilisait l'arabe, le copte n'étant plus guère compris. Il a survécu néanmoins longtemps encore comme langue liturgique et savante. Au Maghreb, l'arabisation a été encore plus lente. La distribution des dialectes arabes maghrébins montre qu'il faut y distinguer des parlers sédentaires (reconnaissables notamment à la prononciation sourde du *q* et à la transformation des interdentes en occlusives) et des parlers bédouins. La première catégorie résulte de la « première arabisation » qui date du viii^e siècle et ne touche que les villes et leurs alentours immédiats d'une part et quelques régions rurales d'autre part. Les parlers bédouins résultent de la « seconde arabisation », provoquée par l'arrivée à partir du milieu du xi^e siècle de tribus bédouines originaires de l'Arabie occidentale : les *Banī Hilāl* qui occupent l'Algérie actuelle, les *Banī Sulaym* qui s'installent en Tunisie et les *Banī Ma'qīl* qui ont été littéralement « déportés » au milieu du xii^e siècle vers le Maroc par les souverains almohades pour grossir le rang de leurs troupes. C'est cette seconde vague qui plante l'arabe dans l'intérieur du Maghreb, assurant ainsi le succès définitif de l'arabisation de la région. L'Espagne et les îles de la Méditerranée semblent n'avoir connu que l'équivalent de la « première arabisation ». C'est peut-être ce qui explique la fragilité de leur arabisation.

Il n'y a eu aucune politique d'arabisation forcée. L'État musulman s'est contenté d'imposer l'usage de l'arabe dans les écrits administratifs et juridiques en acceptant, tant que c'était nécessaire, l'usage de la traduction et de l'interprétariat. C'est donc essentiellement le désir d'intégration sociale qui a favorisé l'arabisation. Il convient de noter aussi que les processus d'arabisation et d'islamisation sont foncièrement distincts : ainsi, les peuples d'Iran, et une partie des Kurdes et des Berbères se sont islamisés sans s'arabiser, alors que les chrétiens d'Orient, les Coptes et les Juifs se sont arabisés sans changer de religion, et en gardant éventuellement une langue liturgique distincte de l'arabe. Cependant, le processus d'islamisation en donnant naissance à la catégorie sociale des *mawālī* et des *muwalladūn* a certainement été un facteur important dans l'unification sociale et, progressivement linguistique de nombreuses régions conquises. Une condition nécessaire du succès final de l'arabisation semble cependant avoir été, pratiquement partout (à l'exception de Malte), un apport important et renouvelé de populations arabophones.

Notes

[1] C'est ce qui explique l'émoi suscité dans les milieux musulmans traditionalistes lorsque Ṭaha Ḥusayn (1927) conteste l'authenticité de ces textes.

Chapitre VI

Apogée et déclin

Dès l'accession au pouvoir des 'Abbāsīdes, des forces centrifuges avaient commencé à détacher du pouvoir central diverses régions de l'Empire : à l'ouest une principauté omeyyade avait été fondée en Espagne dès la fin du viii^e siècle, et divers pouvoirs indépendants s'étaient constitués au Maghreb. À partir du x^e siècle ce processus de fractionnement s'accélère partout. Partis du Maghreb, les Fâtimides fondent en Égypte un califat chiite. La lignée omeyyade d'Espagne proclame un califat autonome. Les provinces orientales deviennent elles aussi pratiquement autonomes, même si leurs gouverneurs, devenus héréditaires depuis les Bouyides, reconnaissent formellement l'autorité 'Abbāsīde. D'ailleurs les califes deviennent les otages des chefs de leur armée formée de Turcs. Au milieu du xi^e siècle une dynastie turque, les Seldjoukides, contrôle Bagdad où le calife n'est plus qu'une figure symbolique. La prise de Bagdad par les Mongols en 1258 met fin au califat 'Abbāsīde. Des principautés seldjoukides survivent encore jusqu'au début du xiv^e siècle pour être ensuite remplacées par les Ottomans qui vont peu à peu constituer un empire qui perdure jusqu'au xx^e siècle. Une évolution analogue a lieu en Égypte : les esclaves soldats mamelouks y exercent le pouvoir effectif du xiii^e au xvi^e siècle, avant de subir la domination ottomane qui s'étend jusqu'au Maghreb central. L'émiettement du pouvoir encourage, dans les marges du monde musulman, des entreprises militaires étrangères : en Orient les Mongols et les Turcs, en Méditerranée orientale les croisés, en Méditerranée occidentale les Normands, les Espagnols et les Portugais, vont avec plus ou moins de succès s'efforcer de prendre le contrôle de territoires jusque-là sous autorité musulmane. Les envahisseurs orientaux se convertissent finalement à l'islam et favorisent même son introduction en Inde. Les croisés sont finalement repoussés au xiii^e siècle. Par contre la Reconquista mettra fin à la présence arabe en Espagne au xv^e siècle, en même temps que les Turcs ottomans portent le coup de grâce à Byzance et instaurent un califat turc sunnite qui n'est aboli qu'après la Première Guerre mondiale au xx^e siècle. Un relatif *statu quo* s'établit ensuite entre le xvi^e et le xix^e siècle quant aux grands

équilibres politiques et militaires dans le monde arabo-musulman.

I. L'évolution de l'arabe classique

Jusqu'au milieu du xi^e siècle, on peut dire que la *fuṣḥā* n'a guère souffert des vicissitudes de la situation politique et même, paradoxalement, qu'elle en a profité : en effet, la multiplication des cours califales ou simplement princières signifiait celle des riches mécènes qui mettaient tous un point d'honneur à ce que leurs assemblées (*maǧālis*) de lettrés et de savants n'aient rien à envier à celles des principautés voisines, voire à celle des califes. On assiste donc, jusqu'à la fin du xi^e siècle et même encore plus tard à une véritable floraison d'œuvres littéraires et scientifiques en langue arabe, non seulement dans l'orbite des trois cours califales de Bagdad, Le Caire et Cordoue, mais aussi dans des émirats comme Alep (Syrie) ou Rayy (Iran).

La production dans le domaine des sciences, connaît son apogée, stimulée par la création d'institutions savantes (grandes bibliothèques, observatoires) et par la relative liberté d'expression qu'un islam encore ouvert rendait possible. C'est l'époque qui voit naître les œuvres des grands philosophes de langue arabe comme Ibn Sīnā (m. 1037) en Orient, ou Ibn Bāǧǧa (m. 1139), Ibn Ṭufayl (m. 1186), Ibn Rušd (m. 1198) ou Ibn Maymūn (m. 1204) en Espagne. C'est aussi l'âge d'or de la production scientifique en langue arabe (médecine, mathématiques, astronomie, sciences naturelles).

Les sciences islamiques connaissent alors leur période de maturité et de systématisation. De cette époque datent les grandes synthèses théologiques et juridiques qui ont donné à l'islam les cadres référentiels de ses grandes orientations doctrinales (notamment le sunnisme et le chiisme institutionnels).

Les genres littéraires traditionnels poursuivent leur développement, qu'il s'agisse de la poésie, avec al-Mutanabbī (m. 965), 'Abū Firās al-Ḥamdānī (m. 968) ou al-Šarīf al-Raḍī (m. 1016), le 'adab avec 'Abū Ḥayyān al-Tawḥīdī (m. 1023) ou 'Abū l-'Alā al-Ma'arrī (m. 1058). Mais de nouveaux genres littéraires apparaissent comme les « *séances* » (*maqāmāt*) véritables exercices de virtuosité linguistique, utilisant un vocabulaire rare et savant pour raconter en prose rimée (*saǧ'*) les tribulations de héros picaresques. Le premier grand maître en sera al-Hamaḍanī (m. 1008).

À partir du xii^e siècle, la production en langue arabe classique donne des signes d'essoufflement. Certes la période va avoir ses très grands noms, comme al-Ḥarīrī (m. 1122) pour la *maqāma* ou Zamaḡšarī (m. 1144) pour la grammaire et l'exégèse coranique. Mais la poésie décline, malgré les tentatives d'un Ibn al-Farīdī (m. 1235) d'en détourner les thèmes vers la mystique ou celles d'un Bahā' al-Dīn Zuhayr (m. 1253) d'en rapprocher la langue de l'usage quotidien. Surtout, la prose littéraire commence à souffrir d'un mal redoutable qui va peu à peu la ronger : le *sağ'* s'y répand de plus en plus imposant aux écrivains des contorsions d'expression et des recherches de vocabulaire rare (rime oblige !) qui vont caractériser l'écriture arabe décadente. Ces signes de déclin vont s'accroître aux siècles suivants, lorsque le pouvoir politique va, dans de nombreuses régions, échapper totalement aux Arabes. Dans les provinces orientales, c'est le persan qui devient la langue de la littérature, voire de l'administration, ne laissant à l'arabe que les domaines religieux et juridique. Dans les territoires sous contrôle mongol, turc ou même berbère, les activités littéraires, philosophiques ou scientifiques ne suscitent guère d'intérêt et l'usage de l'arabe est de plus en plus cantonné au domaine des sciences islamiques. Le fond est atteint lorsque le turc remplace l'arabe comme langue de l'administration au cœur même du monde arabophone. Il devient alors de plus en plus difficile de trouver des écrits arabes qui soient simplement corrects du point de vue de la langue, sans même parler de l'aspect esthétique. C'est que la *fuṣṣḥā*, devenue une langue purement écrite, dépend crucialement pour sa reproduction de l'existence d'institutions d'État qui la protègent et prennent en charge sa transmission par l'enseignement. Il est frappant qu'entre le xiii^e et le xv^e siècle la plupart des œuvres en langue arabe classique se veulent des synthèses comme le dictionnaire encyclopédique de l'arabe d'Ibn Manẓūr (m. 1312), l'œuvre grammaticale d'Ibn Hišām (m. 1360) ou, un peu plus tard, l'immense compilation de Suyūṭī (m. 1505), comme si ces auteurs ressentaient le besoin, en ces temps d'incertitude, de rassembler tout le savoir accumulé par les siècles fastes de peur qu'il ne soit définitivement perdu. Même l'œuvre d'un Ibn Ḳaldūn (m. 1406), pourtant radicalement nouvelle, a un aspect de testament intellectuel.

Du xvi^e au xix^e siècle la langue arabe classique entre dans une longue léthargie éclairée seulement par les faibles lueurs d'œuvres mineures qui ne sont le plus souvent que de pâles imitations des grands modèles du passé.

II. Le moyen arabe

À partir du xi^e siècle, on voit apparaître des écrits qui, tout en ayant manifestement un statut littéraire, ont recours, à des degrés divers, à des usages linguistiques qui rompent avec les règles de la *fuṣḥā*. Le cas le mieux étudié est celui du *muwašṣaḥ*, poésie strophique qui apparaît vers la fin xi^e siècle en 'Andalus et qui non seulement rompt avec les règles de la métrique arabe traditionnelle, mais va même jusqu'à introduire des passages entiers en arabe dialectal, voire en hispano-roman ! Il existe d'autres exemples, moins manifestes, d'intrusion de l'arabe dialectal dans la littérature. Évoquons en passant les nombreux cas que relève Johann Fück (1955, chap. 10 et 11), d'usage de formes « néo-arabes » dans des textes qui, comme celui du géographe Yāqūt (m. 1229), se veulent en *fuṣḥā*. Mais signalons surtout le cas de 'Usāma Ibn Munqid (m. 1188) qui écrit son autobiographie, le *Kitāb al-i'tibār*, dans une langue délibérément familière et où il n'est pas difficile de reconnaître, sous-jacent, son parler maternel syrien. Et que dire de l'émergence du genre, très populaire, des grandes gestes (*siyar*) qui racontent, originellement dans la langue des conteurs populaires, les exploits de héros comme le poète préislamique 'Antar(a) ou des *Banī Hilāl* conquérants du Maghreb ? Que dire aussi du récit universellement connu des *Mille et une nuits* ('*alf layla wa-layla*) ? Les versions arabes qui circulent de ces textes ont été plus ou moins réécrites en arabe standard, mais là encore il est aisé de reconnaître les tournures dialectales derrière l'ajout de marques de 'i'rāb. Dans tous ces textes, l'intention littéraire est manifeste et l'on ne répugne nullement à y mélanger les registres de langue, notamment par la citation de nombreux vers tirés du patrimoine littéraire classique. Tout cela montre que la *fuṣḥā* « pure et dure » est devenue une langue figée, vouée par la polarisation linguistique et culturelle à n'être plus qu'une langue artificielle et cérémonieuse, et que, dans les usages vivants et créatifs, elle est de plus en plus vouée au métissage avec les parlers vivants.

Cela nous amène à la question du « moyen arabe » (*Middle Arabic*). Si l'on suit Versteegh (1997, 114) qui écrit que « dans les études modernes sur l'arabe le nom collectif donné à tous les textes qui présentent des déviations par rapport à la grammaire arabe classique est celui de moyen arabe » alors, sans aucun doute, tous les types de textes dont nous venons de faire état sont du moyen arabe. Versteegh cependant met en garde contre les risques d'ambiguïté liés à l'usage de ce terme, et notamment celui de croire (sur le modèle des expressions « moyen français » ou « moyen anglais » par exemple) qu'il s'agit de désigner un état de langue intermédiaire entre « l'ancien arabe » (*Old Arabic*) et l'arabe moderne [1], ce que laissait pourtant entendre Joshua Blau (1966) dans l'introduction à sa grammaire de l'arabe chrétien. Cet auteur s'est cependant

rétracté (Blau, 1988) et a redéfini le moyen arabe comme « un terme général [pour désigner] les textes dans lesquels arabe classique, néo-arabe et éléments pseudo-corrects alternent ». Versteegh adopte aussi cette définition en y ajoutant une caractéristique, à savoir que le moyen arabe est typiquement de l'arabe « fautif », les écarts constatés entre la norme grammaticale et l'usage effectif étant dus à l'insuffisante compétence en arabe classique du scripteur. Les deux autres causes possibles d'écarts, selon lui, sont le désir de rapprocher les textes, quand il s'agit de dialogues, de la langue parlée, et dans le cas spécifique des chrétiens et des juifs, le fait qu'ils ne sont pas soumis, comme les musulmans, aux normes de l'arabe coranique.

Il convient de remarquer, d'abord, que la notion de moyen arabe, définie comme un fourre-tout où l'on met les textes de toutes les époques contenant des traits non standards, ne présente aucun intérêt théorique précisément en raison de son caractère hétérogène. En outre, la caractérisation des écarts en termes de fautes est normative et peu scientifique, surtout appliquée à des textes au statut très différent. Il nous paraît beaucoup plus éclairant d'approcher au moins une partie de ces textes en termes « d'écriture populaire », selon la suggestion de M. Doss (1996) qui souligne que cette caractérisation ne doit pas être assimilée à l'écriture « dialectale », dont elle situe l'apparition seulement au ^{xix}^e siècle.

III. La dialectalisation

Nous avons vu que la différenciation dialectale existe dans le domaine arabe aussi loin que remonte notre documentation. L'existence de la *koinè* poétique n'a pas aboli ces différences (ce n'était nullement son objet) et même le texte coranique leur a donné, dans une certaine mesure, droit de cité. Quant à la thèse d'une « *koinè* militaire » qui aurait servi de point de départ aux parlers arabes actuels, trop d'objections peuvent lui être opposées, et en particulier le fait qu'une partie de la typologie de ces parlers ne s'explique qu'en référence à des distinctions dialectales qui remontent à l'Arabie préislamique. L'une de ces distinctions, qui traverse l'ensemble du domaine arabe indépendamment des grandes zones dialectales dont nous ferons état dans un instant, est celle qui oppose « parlers sédentaires » (désormais ps) et « parlers bédouins » (désormais pb). Les pb sont plus conservateurs que les ps : conservation des interdentes, passées à des occlusives dans les ps, conservation des diphtongues *aw* et *ay* qui passent souvent à [ō] et [ē] ou à [ū] et [ī] dans les ps, prononciation affriquée du

ğ que les ps réalisent [j] ([g] en basse-Égypte et certaines régions du Yémen) et prononciation sonore [g] du phonème *q* de l'arabe classique, que les ps réalisent toujours sourd [q] ou [']. Il faut souligner que cette classification n'a plus rien à voir avec le mode de vie des intéressés : la plupart des locuteurs d'un pb sont aujourd'hui des sédentaires. En outre des phénomènes de recouvrement partiel de ps par des pb se sont produits au cours de l'histoire donnant des parlers mixtes où les traits ps et pb se mélangent. Signalons aussi que dans les ps, on distingue, en fonction de divers critères, les parlers citadins des parlers villageois.

Outre ce critère d'héritage dialectal ancien, les trois paramètres qui concourent à spécifier les différences dialectales sont le substrat linguistique, la localisation régionale et l'appartenance socioculturelle.

Concernant le substrat, il semble raisonnable de postuler que les grands domaines linguistiques auxquels appartenaient les ancêtres des locuteurs actuels de l'arabe (araméen, copte, berbère) ont exercé une certaine influence sur les caractéristiques du parler arabe résultant. Il existe de nombreuses études tendant à établir ce lien, mais la plupart manquent de rigueur scientifique, et la question reste donc ouverte à l'investigation.

Pour ce qui est de la localisation, on peut classer l'ensemble des parlers arabes actuels en cinq grands groupes : Les parlers de la péninsule Arabique, qui présentent le plus de traits archaïques (par exemple un duel encore productif) ; les parlers mésopotamiens ; les parlers du *Šām* (Syrie, Liban, Palestine, Jordanie) ; les parlers de la vallée du Nil, qui s'étendent du Delta au Soudan, et parmi lesquels il faut distinguer celui du Caire qui, grâce au cinéma et aux médias, jouit d'une diffusion et d'un prestige uniques dans le Monde arabe ; les parlers maghrébins qui s'étendent du centre de la Libye à la Mauritanie et qui se caractérisent par une réfection du système des préfixes de l'inaccompli. Ces cinq groupes, il convient de le remarquer, correspondent aux cinq grandes zones politico-économiques en lesquelles le monde arabophone s'est fragmenté à partir du xvi^e siècle.

Pour ce qui est du facteur socioculturel, la religion peut être liée à certaines différences linguistiques : ainsi les parlers juifs présentent des traits phonétiques spécifiques. D'autre part, il a été observé que les parlers des femmes présentaient des traits qui leur sont propres.

Notes

[1] Idéalement, il faudrait en effet garder l'expression « moyen arabe » pour désigner un tel état de langue et parler « d'arabe moyen » pour la langue mélangée dont il est question ici. Mais l'usage a ses raisons...

Chapitre VII

Le renouveau contemporain

Du xvi^e au xix^e siècle, l'ensemble du monde arabophone, à l'exception de l'ouest du Maghreb et du sud de la péninsule Arabique, est sous domination ottomane. Le turc y est la langue officielle des administrations régionales, et l'arabe n'y est plus que la langue religieuse et, partiellement, juridique, même si, dans certaines provinces, l'usage administratif de l'arabe se maintient faute de personnel maîtrisant le turc. Il n'y a pas eu, jusqu'au début du xx^e siècle, de politique de turcisation des régions arabophones, mais l'arabe était délibérément marginalisé, les autorités ottomanes exigeant l'usage exclusif du turc dans les fonctions officielles.

Dans le courant du xix^e siècle, l'Empire ottoman traverse des difficultés de tous ordres qui lui valent dans les chancelleries le titre « d'homme malade de l'Europe ». Il perd à l'ouest l'Algérie puis la Tunisie, passées sous domination française, et à l'est des pans entiers de ses possessions européennes. Le besoin de réformes se fait sentir et débouche sur le mouvement des « tanzimat » (1839-1876) visant à moderniser la vie politique dans l'Empire.

I. La renaissance arabe (nahḍa)

C'est dans ce contexte que va se développer le vaste mouvement politique et culturel qui sort peu à peu le Monde arabe de sa léthargie séculaire et qui est désigné sous le nom arabe de *nahḍa* (réveil, renaissance). Ce mouvement a deux sources : l'une en Syrie-Liban, l'autre en Égypte (Hourani, 1993, chap. 18).

C'est dans la région syro-libanaise que naissent les premiers thèmes du nationalisme arabe, en réaction à l'oppression turque mais aussi comme une réponse laïque aux tensions qui opposaient les diverses communautés confessionnelles et avaient parfois pris un tour dramatique. L'arabisme, qui rejetait les approches en termes de religion, était tout naturellement porté à

thématiser l'unité culturelle et linguistique des Arabes et à réclamer une restauration de l'arabe dans ses prérogatives de langue administrative et culturelle. En même temps, il prônait la modernisation de la culture arabe, voire sa laïcisation par l'ouverture à la culture occidentale. Celle-ci commençait à être mieux connue grâce à la création par des missions catholiques ou protestantes d'institutions d'enseignement supérieur comme le Collège protestant syrien (1866) ou l'Université Saint-Joseph (1875). Dans ces établissements, l'arabe était utilisé pour enseigner des matières modernes comme la médecine, les sciences naturelles ou la sociologie. Toute la région connaît alors une grande effervescence intellectuelle : on traduit en arabe des œuvres du patrimoine littéraire occidental comme l'*Iliade*, la Bible, le théâtre français et anglais. On compile de nouveaux dictionnaires et même, sous la direction de Buṭrus al-Bustānī (m. 1883), une encyclopédie visant à initier le lecteur arabophone aux sciences et aux techniques modernes. Ces ouvrages connaissent une large diffusion grâce à la généralisation de l'imprimerie. Divers journaux en langue arabe sont fondés ainsi que des périodiques à caractère politique, littéraire, ou même scientifique comme le célèbre *Muqataṭaf* qui finira par émigrer en Égypte pour cause de promotion du darwinisme.

Les effets de toutes ces entreprises éditoriales sur la langue arabe furent considérables : si tous les auteurs mettaient un point d'honneur à livrer au lecteur une prose parfaite quant à la grammaire et au style, ils s'efforçaient en même temps de libérer la langue des lourdeurs et des artifices dont les siècles antérieurs l'avaient chargée, et d'utiliser une terminologie technique et scientifique à la fois rigoureuse et fidèle au génie de l'arabe. De ces efforts va peu à peu émerger l'arabe standard moderne (asm), forme modernisée de la *fuṣḥā*.

En Égypte, le « réveil arabe » résulte de la politique de modernisation de l'État entreprise sous le règne du Khédive Muḥammad 'Alī (m. 1849). Ce général de l'armée ottomane, originaire d'Albanie, prend le pouvoir au Caire dans la foulée de l'expédition napoléonienne d'Égypte (1798-1802). Impressionné par la supériorité militaire, technique et scientifique des Européens, il décide de dégager l'Égypte de l'emprise ottomane et de la lancer dans un ambitieux programme de réformes. Il modernise l'agriculture, lance une industrie d'État, refond l'administration, crée des écoles modernes, envoie de jeunes officiers se former en Europe. Surtout, il encourage le développement des sciences et des techniques, notamment en créant en 1835 une école de langues et en 1841 un bureau de traduction en vue d'ouvrir le pays à la connaissance de la culture occidentale. On y traduit, à partir du français surtout, des ouvrages touchant à

tous les domaines : mathématiques, géographie, sciences naturelles, histoire, philosophie, arts de la guerre. Dans son entreprise de modernisation, le Khédive trouve un allié actif et efficace en Rifā'a al-Ṭaḥṭāwī (m. 1873), un jeune azharite conquis, après un séjour d'étude en France, par les idées de progrès. C'est lui qui prendra la direction du bureau de traduction, mais aussi celle du périodique *al-waqāi' al-miṣriyya* dont il fait le premier journal politique d'Égypte. Par la presse et par le livre, qui devient plus accessible grâce à la fondation d'une imprimerie nationale à Būlāq en 1822, un public de plus en plus large va pouvoir lire des textes variés et non plus simplement religieux. Un des effets de ce mouvement est que l'arabe cesse d'être ce qu'il était devenu depuis des siècles : une langue figée, monopole des clercs.

À la fin du xix^e siècle, un durcissement de la politique turque à l'égard des nationalistes et libéraux arabes au Levant pousse nombre d'entre eux à l'action politique clandestine ou à l'exil en Égypte. Mais le développement de l'asm se poursuivra et s'étendra peu à peu à toutes les parties du Monde arabe.

II. La situation actuelle

Le démembrement de l'Empire ottoman après la Première Guerre mondiale ouvre la voie aux revendications nationalistes dans les pays arabes. Ces revendications ne cesseront de s'amplifier dans l'entre-deux-guerres pour aboutir, dans la seconde moitié du xx^e siècle, à l'indépendance de l'ensemble de ces pays.

Aujourd'hui, l'arabe est seule langue officielle dans 18 pays : Algérie, Arabie Saoudite, Bahreïn, Égypte, Émirats arabes unis, Irak, Jordanie, Koweït, Liban, Libye, Maroc, Oman, Palestine (territoires autonomes), Qatar, Soudan, Syrie, Tunisie, Yémen. Il a le statut de langue co-officielle dans six pays : avec le français en Mauritanie, à Djibouti, au Tchad et dans l'archipel des Comores, avec l'hébreu en Israël, et avec le Tigrina en Érythrée.

Dans la majorité de ces pays, une politique active de promotion de la langue officielle a été menée : l'arabe y est imposé dans l'enseignement, l'administration, la justice, les médias et plus globalement dans l'ensemble de l'environnement social (signalisation, publicité). Des « Académies de langue arabe » ont été créées notamment en Syrie (1918), Égypte (1932), Iraq (1947) et

Jordanie (1976). Ces institutions se proposent de veiller à ce que les besoins en terminologie moderne soient satisfaits dans le respect des règles établies par la tradition grammaticale (Kālīl, 1979). Sous leur égide, ou celles d'organismes spécialisés de la Ligue arabe, des dictionnaires généraux ou spécialisés ont été publiés. Mais leur influence réelle sur l'évolution de la langue reste négligeable, en partie en raison de leur manque de coordination, mais aussi du fait que la couche sociale qui vit de la maîtrise de la *fuṣḥā* (religieux, juristes, professeurs) s'oppose à tout projet de réforme de la langue, comme le montre l'exemple significatif de l'échec du projet de réforme de l'écriture arabe tenté par l'Académie du Caire en 1938 (Meynet, 1971).

En pratique, donc, c'est surtout l'usage, et notamment celui de la presse et des médias, qui détermine l'évolution de la langue. Toutes les études sur les processus d'élaboration de l'asm (par exemple Monteil, 1960, Blau, 1981, ou Holes, 2006) montrent que les calques lexicaux et phraséologiques de l'anglais et du français y ont joué un rôle déterminant notamment à travers la traduction hâtive des dépêches d'agences de presse par des journalistes plus pressés de communiquer une information que de vérifier la conformité de leur traduction aux canons de la *fuṣḥā*.

Les écrivains ont eu, eux aussi, une influence importante, quoique moins voyante, sur l'élaboration de l'asm. En effet, si quelques-uns ont essayé de revivifier des genres classiques, comme al-Muwayliḥī (m. 1930) pour la *maqāma*, ou Ṣawqī (m. 1932) pour la poésie traditionnelle, la plupart, fascinés par la littérature occidentale, se sont efforcés d'adapter à la langue arabe des genres littéraires qui y étaient inconnus, comme le roman, la nouvelle ou le théâtre. Ces efforts impliquaient l'adaptation de la langue à des exigences expressives nouvelles, et de nombreux débats témoignent des difficultés rencontrées pour rendre la langue littéraire capable d'exprimer de façon crédible des dialogues entre personnages de roman et plus encore de théâtre. Des œuvres comme celle de Tawfīq al-Ḥakīm (m. 1987), père du théâtre arabe, ou de Naḡīb Maḥfūẓ (m. 2006), prix Nobel de littérature en 1988, témoignent du succès de ces entreprises.

Dans les pays du Maghreb où, durant l'époque coloniale, l'arabe avait été à des degrés divers exclu des sphères officielles, les gouvernements issus des indépendances ont fait de sa restauration dans ses prérogatives de langue d'État une composante essentielle de leur politique culturelle. Cette politique dite « d'arabisation » y a fonctionné comme un élément central du dispositif de

légitimation politique (Grandguillaume, 1983) et a bien souvent signifié, en Algérie et au Maroc, le refus de donner droit aux revendications identitaires et culturelles des citoyens berbérophones, qui représentent pourtant un pourcentage non négligeable de la population. Cette situation semble cependant en voie de se normaliser. Des problèmes analogues se sont posés en Iraq pour le kurde.

Au niveau de l'usage oral quotidien, les indépendances ont eu pour effet de promouvoir un des parlers, généralement celui de la capitale, au statut de parler véhiculaire national. Ce parler, dont le centralisme étatique favorise involontairement la diffusion, est utilisé dans les échanges avec des compatriotes usant de parlers différents. Le cas du parler du Caire est à cet égard le plus spectaculaire.

Dans la plupart des pays arabes, notamment ceux qui ont connu des régimes populistes socialisants ou que les revenus du pétrole ont rendu riches, une politique active de scolarisation générale des enfants et parfois d'alphabétisation des adultes a contribué à élargir considérablement le nombre des citoyens ayant une connaissance au moins élémentaire de l'asm. La presse a connu partout un important développement. Mais surtout la diffusion très large de la radio et de la télévision a donné à la langue officielle une propagation sans équivalent à aucune autre époque de l'histoire de l'arabe. La connaissance, au moins passive (compréhension) de cette variété d'arabe a donc considérablement progressé même dans les campagnes et les milieux modestes des villes. Cela a contribué à diminuer la distance qui séparait depuis des siècles la langue standard des divers parlers : le lexique de la langue orale s'est partout chargé de mots directement importés de l'asm, *via* les médias, et touchant à tous les domaines de la vie sociale : politique, économie, sports, activités artistiques, santé, travail. Ces emprunts sont souvent reconnaissables car leur structure phonologique contredit souvent des tendances profondes des « dialectes » : présence de consonnes interdentes (souvent réalisées comme des sifflantes en Égypte et en Orient alors qu'elles étaient transformées en occlusives dans la phonologie dialectale), présence de voyelles brèves en syllabes ouvertes alors que de telles voyelles étaient effacées dans les dialectes, notamment maghrébins. Même la construction des syntagmes et des énoncés peut emprunter à l'asm : on entend ainsi, dans certaines variétés d'arabe parlées par les locuteurs instruits des expressions comme « *law mā 'uhdiratš* » (si elle n'avait pas été gaspillée) avec un passif à inflexion vocalique, typiquement « classique » encadré par une négation à morphème discontinu « *mā...š* » typiquement « dialectale ». On est alors dans le registre que certains ont proposé d'appeler « arabe médian » et qui

apparaît comme un niveau intermédiaire entre l'asmet l'arabe « dialectal ».

C'est dire que la situation linguistique actuelle du monde arabe est plus complexe que le modèle à deux niveaux bien distincts que proposait Charles A. Fergusson (1959) en parlant de « diglossie » ne le laisse imaginer. Si ce modèle avait quelque réalité dans le Monde arabe de la première moitié du xx^e siècle, il ne correspond plus à rien aujourd'hui. Les modèles actuels de la réalité sociolinguistique arabe (Kouloughli, 1996), qu'ils soient « multistratifiés », comme chez Sa'īd Badawī (1973), triglossiques comme chez Terry Mitchell (1986) ou en termes de « continuum » linguistique comme chez Shahir El-Hassan (1977) insistent tous sur l'inadéquation de l'approche diglossique.

Pour donner une idée concrète du comportement linguistique d'un arabophone d'aujourd'hui (Holes, 2006), il vaut mieux partir du postulat que les deux variétés linguistiques clairement distinctes et opposées que sont « le classique » et « le dialectal » ne sont que deux abstractions. Tout acte de parole réel se situe quelque part entre ces deux pôles abstraits et résulte d'un jeu complexe de paramètres sociolinguistiques (situation de communication, rapports entre les interlocuteurs, nature des informations échangées, enjeux de l'acte de parole, etc.) qui contribuent à définir où l'on va se situer dans l'échelle des registres linguistiques disponibles. Cette évaluation est du reste « dynamique » : au cours de l'échange elle peut se réajuster, provoquant un déplacement vers plus ou moins de formalité.

L'opération fondamentale qui commande le mécanisme d'ajustement est le nivellement, lequel a pour effet d'inhiber les formes perçues comme trop locales pour les remplacer par des formes plus communes. Quand l'échange linguistique s'établit entre locuteurs du même pays, c'est souvent vers le « parler véhiculaire », celui de la capitale, que s'effectue le nivellement. Le recours aux ressources de la langue standard s'impose néanmoins si la discussion s'oriente vers des sujets à caractère conceptuel marqué : thèmes idéologiques, techniques, scientifiques, bref ceux pour lesquels les parlers locaux ne disposent pas toujours d'un lexique et d'une phraséologie adaptés. On ramène le curseur vers une variété plus locale quand les thèmes se font plus familiers : humour, questions domestiques ou privées. Quand l'échange s'établit entre locuteurs de pays différents, en particulier de zones dialectales différentes, alors le nivellement s'oriente vers l'asm.

La variété d'arabe que l'on nomme « arabe parlé cultivé » (*Educated Spoken*

Arabic) ou encore « arabe médian » est une variété en voie de stabilisation qui utilise les ressources lexicales de l'asm et une partie significative de sa morphologie et de sa syntaxe, avec diverses simplifications et notamment un abandon presque total du *ʿirāb*, et qui tolère un recours assez large à des formes parlées senties comme suffisamment communes et donc légitimes à un niveau régional ou panarabe. C'est dans cette variété que l'on rencontre les exemples les plus spectaculaires de mélange de systèmes linguistiques. L'arabe médian s'utilise notamment dans des situations de communication interarabes et sur les médias lorsqu'un expert répond de façon spontanée à des questions ou expose sans texte écrit un point de vue.

L'écrit est en principe moins sujet à variation, sa production et sa diffusion étant, jusqu'à une date récente (avant l'avènement des « blogs » et des sms), soumis à un certain contrôle social. Pourtant, il serait faux de croire que la chose écrite en arabe l'est uniquement en asm. Déjà les écrivains avaient ressenti la nécessité de s'éloigner de la langue standard lorsqu'ils voulaient donner un peu de naturel et de spontanéité à leurs dialogues. *A fortiori* les dramaturges ont éprouvé le besoin de disposer d'une langue qui puisse sans artifice être jouée sur une scène et ont développé à cet effet le concept de « langue tierce » (*luġa t̡āliṭa*) que l'on peut voir comme une forme d'arabe médian. Aujourd'hui la bande dessinée, les scénarios, la poésie, la correspondance privée ont introduit toutes sortes de transitions entre l'asm et l'« écrit dialectal ». Il est trop tôt pour dire comment l'écrit électronique, qui ne manque pas d'adeptes dans le Monde arabe, viendra s'insérer dans l'ensemble de ces pratiques et infléchir l'évolution d'ensemble dont on peut dire d'ores et déjà qu'elle a arraché l'arabe au monopole des spécialistes de la langue.

Annexe – L’arabe à vol d’oiseau

I. Phonétique – Phonologie

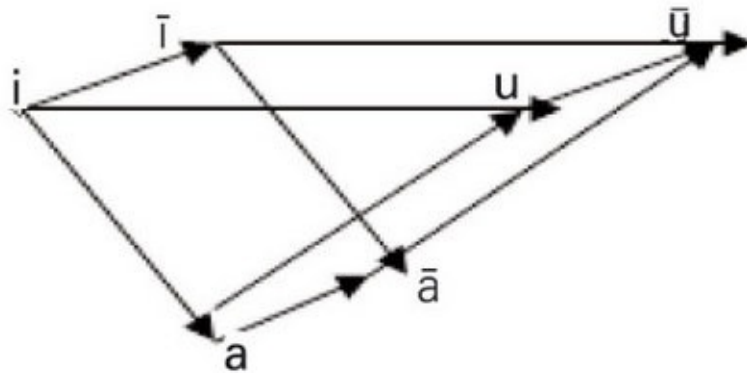
Le système phonologique de l’arabe est caractérisé par l’abondance des consonnes (28) qui contraste avec la pauvreté des voyelles (3). Le système consonantique est défini par 9 modes d’articulation (repris dans les 9 lignes du tableau ci-dessous) et les 9 points d’articulation suivants (repris dans les 9 colonnes du tableau) : 1 *labiale* ; 2 *labiodentale* ; 3 *interdentale* ; 4 *dentale* ; 5 *palatale* ; 6 *vélaire* ; 7 *uvulaire* ; 8 *pharyngale* ; 9 / *glottale*.

Les 4 paires séparées dans le tableau par une barre oblique s’opposent par le trait d’ « emphase » : les « emphatiques » se distinguent des « non-emphatiques » par une articulation secondaire caractérisée par un exhaussement du bloc pharyngolaryngien qui augmente le volume arrière de la cavité buccale provoquant une « bémolisation » de la voyelle. Cette bémolisation peut s’étendre par anticipation ou par inertie aux syllabes adjacentes.

Il faut noter les nombreuses consonnes « gutturales » (colonnes 6 à 9) caractéristiques du conservatisme articulatoire de l’arabe.

Le consonantisme présente peu de variantes allophoniques. Les principales caractérisent plutôt des variantes régionales et/ou sociales.

Le système vocalique se résume au « triangle vocalique » enrichi d’un contraste brèves/longues. Il est résumé dans le tableau ci-dessous :



Ce système présente une grande variabilité liée à sa pauvreté phonologique : en particulier, sous l'influence des emphatiques ou des pharyngales, le spectre des voyelles subit un abaissement du second formant qui leur donne un timbre « bémolisé » caractéristique.

L'arabe connaît trois types syllabiques de base : CV (syllabe « légère »), CVV, CVC (syllabes « lourdes »). Les types CVCC et CVVC (syllabes « surlourdes ») ne peuvent apparaître qu'en fin de mot (suite à la chute pausale d'une voyelle finale), ou dans des contextes morphologiques spéciaux. Ils sont interdits en métrique arabe classique.

L'arabe interdit donc les groupes de deux consonnes en initiale de syllabe et ceux de trois à l'intérieur du mot. Si de telles séquences sont générées par la morphologie ou la syntaxe, elles sont normalisées par des processus d'épenthèse ou d'anaptyxe.

Tout mot phonologique autonome porte un accent de mot qui combine intensité et hauteur et dont le placement obéit à une règle simple : on accentue la syllabe lourde la plus proche de la fin du mot, à l'exclusion de la dernière et sans remonter au-delà de l'antépénultième. Si le mot ne présente pas une telle syllabe, on accentue la dernière si elle est surlourde, ou alors l'antépénultième (la pénultième si le mot n'a que 2 syllabes) quel que soit son poids.

À la pause, l'arabe efface les voyelles brèves finales ainsi que le suffixe *-n* de l'indéfinition : *kitābun/* > [kitab], *katabtu* > [katabt].

II. Morphologie

La base morphologique d'un mot lexical (nom ou verbe) s'analyse comme l'entrelacement d'une racine et d'un schème (cf. ci-dessous) sur lesquels peuvent s'agglutiner affixes et clitiques grammaticaux : *al-kitāb-a* = le-livre-acc. ; *wa-katab-a-hā* = et-écrivit-il-elle = et il l'écrivit ; *ma'a-hā* = avec elle. Les mots grammaticaux (particules) sont inanalysables. Tout morphème de taille CV ou inférieure est cliticisé : *bi-hā* = avec elle. Cette contrainte est reflétée dans la graphie.

La racine est en général un triplet ordonné de consonnes radicales notées R_1 , R_2 et R_3 . Les racines bi, quadri ou pentaconsonantiques existent mais sont plus rares et peu productives. Le schème est constitué de voyelles et éventuellement de consonnes qui peuvent s'insérer en toute position autour des radicales. Les mots de même racine exhibent une certaine parenté sémantique, les mots de même schème une parenté fonctionnelle. Ainsi : *kataba* « il a écrit », *kātib* « écrivain », *maktaba* « bibliothèque » et *kitāb* « livre », tous de racine /*k-t-b*/ partagent l'idée d'écriture. De même : *maktab* « bureau », *maḥall'ab* « stade », *maṭ'am* « restaurant » tous de schème /*ma-R₁R₂-a-R₃*/ partagent la notion de « nom de lieu ».

L'arabe a trois cas, marqués, en général, par la suffixation au nom des trois voyelles : *u* = nominatif *a* = accusatif et *i* = génitif. Le masculin et le féminin sont distingués. Cette distinction ne fonctionne pleinement que pour les humains : au pluriel les non-humains sont morphologiquement des féminins. L'arabe distingue le singulier, le duel et le pluriel.

Il y a quatre catégories de noms : le substantif obéit à une organisation notionnelle spécifique (cf. ci-dessous) ; l'adjectif supposant un support substantif dont il exprime une propriété, s'accorde en genre, nombre et cas avec le support. Il accepte en général gradation et comparaison ; le nom verbal exprime un processus verbal nominalisé et peut régir un premier actant au génitif et un ou deux autres à l'accusatif (dans la syntaxe ancienne, il pouvait régir, à l'état indéfini, un actant au nominatif) ; enfin l'adverbe, incident au verbe à l'adjectif ou à un autre adverbe, était traité à l'origine comme une fonction complément du verbe, mais tend à se constituer en catégorie propre en arabe moderne.

Déictiques et pronoms peuvent remplir des fonctions syntaxiques de noms, mais ils s'en distinguent par leur caractère inanalysable. Les pronoms se répartissent en libres et clitiques.

Les principaux processus de dérivation qui peuvent affecter les noms sont : l'adjonction du suffixe *-at* qui, pour les adjectifs marque le féminin, et pour le substantif, selon leur catégorisation notionnelle, soit le féminin, soit la singulation ; celle de */-āt/* pluralisation de la précédente ; *-iyy* qui marque la relation ou l'origine ; */-ū(na)* (nom.)/*-ī(na)* (acc. ou gen.)/ marque la pluralisation et est réservé aux substantifs masculins humains ; */- ā(ni)* (nom.)/*-ay(ni)* (acc. ou gen.)/ marque partout le duel. Le pluriel n'est que rarement marqué par une dérivation affixale : en général il est associé à un schème distinct de celui du singulier (on parle alors de pluriel « interne ou brisé ») : *kitāb/kutub* = livre(s). Bien que certains schèmes de singulier soient régulièrement associés à certains schèmes de pluriel, cette situation est loin d'être générale. Le diminutif est marqué par des schèmes spécifiques.

La composition est un procédé pratiquement inconnu de l'arabe. Les tentatives modernes pour l'utiliser dans la constitution d'un vocabulaire scientifique restent marginales.

Quant à la morphologie verbale, elle distingue deux paradigmes : l'un, dit « accompli », se forme en ajoutant des suffixes de personne à la base verbale idoine. L'autre, dit « inaccompli », se forme en ajoutant à la base verbale idoine des préfixes de personne et des suffixes de genre et/ou de mode (sur la valeur de ces deux paradigmes, voir III, ci-dessous). Pour les formes les plus simples du verbe, on a une alternance $R_1VR_2VR_3/R_1R_2VR_3$ entre bases verbales d'accompli et d'inaccompli. Pour les autres formes, la base verbale présente normalement une structure unique pour les deux paradigmes. Dans tous les cas, les voyelles de la base exhibent des alternances propres au thème vocalique du verbe.

Les bases verbales opposent systématiquement la diathèse subjective (active) à la diathèse objective (passive) par alternance vocalique du schème : *kataba* = il a écrit ; *kutiba* = il a été écrit. La diathèse objective exclut l'expression de l'agent et attribue le nominatif au patient ou à un circonstant.

Plusieurs types de verbes peuvent être formés sur une même racine. On distingue traditionnellement le verbe « nu » (forme I des arabisants) formé sur un schème qui ne comporte que des consonnes radicales et des voyelles brèves thématiques

et le verbe « augmenté » (formes II à XV des arabisants) formé sur des schèmes plus complexes pouvant comporter des voyelles longues, la gémiation de R_2 ou de R_3 , et/ou divers affixes (préfixes et/ou infixes). Par rapport au verbe nu qui exprime un procès élémentaire, les verbes augmentés expriment diverses modalités : factitivité, transitivisation, réciprocité, réflexivité, etc. Ainsi, pour la racine <'l-m> on a : 'alima « savoir » ; 'a'lama « informer » ; 'allama « enseigner » ; ta'allama « apprendre » ; 'ista'lama « se renseigner ». Les grammaires recensent 15 formes augmentées théoriquement possibles. En fait toutes ne sont pas attestées pour une même racine.

L'accompli des verbes « nus » se répartit selon trois thèmes vocaliques fondamentaux : $R_1aR_2aR_3$ caractérise les verbes ayant pour source le sujet, $R_1aR_2iR_3$ les verbes ayant une source extérieure au sujet, et $R_1aR_2uR_3$ les verbes ayant pour siège le sujet (verbes d'état). Les voyelles des schèmes inaccomplis présentent des alternances vocaliques corrélatives.

Il existe des noms « déverbaux » exprimant des notions reliées au verbe et dont certains obéissent à des règles de formation systématique : le nom verbal associé au verbe nu présente une très grande variété de schèmes (plusieurs dizaines). Il peut y en avoir plusieurs pour un même verbe. Pour les verbes augmentés, les schèmes sont moins nombreux et plus prédictibles ; les participes, actif et passif, obéissent à des schèmes fixes de formation ; de même les noms de lieu, d'instrument et quelques autres.

III. Syntaxe et sémantique

L'organisation syntagmatique obéit à un principe général simple : les éléments déterminatifs tendant à précéder la tête du syntagme, et les éléments qualificatifs à la suivre. Pour le syntagme nominal, les démonstratifs, déterminants, particules modales, classificateurs, quantificateurs, précèdent la tête, alors que les adjectifs, compléments de nom, appositions, relatives, la suivent (à l'exception du déterminant indéfini -n, post-posé au nom). De même, pour les syntagmes verbaux, les opérateurs de modalité, d'aspect, d'auxiliation précèdent le verbe-tête, alors que les spécifiants de procès (objets, circonstants, adverbes...) le suivent. En discours, il peut être dérogé à l'ordre des arguments du verbe en fonction de paramètres comme la différence de détermination et/ou d'empathie et avec des effets de sens précis.

En ce qui concerne les verbes, on a déjà parlé de l'opposition entre formes à suffixes et à préfixes. Cette opposition est interprétée par les arabisants comme une opposition purement aspectuelle : la forme à suffixes de personnes serait un « accompli » et celle à préfixes un « inaccompli ». Une prise en compte plus complète du fonctionnement textuel des deux paradigmes suggère que la valeur la plus profonde de l'opposition pourrait être plutôt modale qu'aspectuelle : la forme à suffixes, exprimant le « certain », aurait certes une interprétation aspectuelle d'accompli, mais aussi des interprétations temporelles renvoyant au passé et au présent résultatif, et des interprétations modales expliquant son utilisation dans les énoncés performatifs, et l'expression des souhaits, hypothèses et conditions. La forme à préfixes, exprimant le « possible », aurait une interprétation aspectuelle de type inaccompli, mais aussi des interprétations temporelles renvoyant au présent actuel, au futur, à l'habituel, au gnomique, ainsi que des interprétations modales comme l'éventuel.

L'arabe connaît trois grands types de phrase simple : la phrase thématique (dite « nominale »), caractérisée par un thème défini (nom, pronom ou déictique) suivi d'un prédicat nominal (défini ou indéfini) ou verbal qui s'accorde en genre et en nombre avec lui. Elle exprime, en fonction de la nature et de la définition de ses arguments, diverses relations : attribution, identité, appartenance, *etc.* La phrase verbale, ayant pour ordre canonique un verbe suivi de son sujet et de ses éventuels compléments, est caractérisée par le fait que le sujet ne peut s'accorder qu'en genre mais non en nombre avec le verbe : =*ǧā'a (a)l-walad-u* = est-venu l'enfant = l'enfant est venu, =*ǧā'a (a)l-'awlād-u* = est-venu les enfants = les enfants sont venus. Sémantiquement, elle présente un processus ou un état qu'elle spécifie ensuite par les arguments du verbe. La diathèse objective ne diffère que par l'interprétation du sujet grammatical comme objet du procès : *ḍuriba walad-u-n* = a été frappé un enfant (nom.) = un enfant a été frappé ; la phrase locative, caractérisée par la donnée d'un localisateur défini (généralement un syntagme prépositionnel) suivi d'un localisé : *fī l-bayt-i walad-u-n* = dans la maison [est] un enfant = il y a un enfant dans la maison. Elle exprime la localisation, l'existence, la possession. Elle peut aussi avoir une valeur modale, le localisateur exprimant la modalité et le localisé la phrase modalisée : *min al-ǧarīb-i 'an taqūl-a ḍālika* = de l'étrange que tu dises cela = il est étrange que tu dises cela.

Les modalités sont souvent exprimées par des particules spécialisées dont la position normale est le début de la phrase (ou du syntagme modifié) : on peut renforcer l'assertion thématique par la particule *'inna* ; la négation des phrases

thématiques ou locatives se fait au moyen du verbe défectif d'état *laysa* « ne pas être » : *laysa (a)l-walad-u marīḏ-a-n* = n'est-pas l'enfant (nom.) malade (acc.) ; *laysa fī (a)l-bayti 'ahad-u-n* = n'est-pas dans la maison quelqu'un (nom.) = il n'y a personne dans la maison ; celle des phrases verbales se fait par diverses particules spécialisées : *lā*, *mā*, *lam* (parfois aussi *laysa* avec la forme à préfixes) : *lā yafhamu* = non il-comprend = il ne comprend pas : l'interrogation totale est exprimée par la particule *hal* : *hal ra'ayta (a)l-walad-a* = est-ce que tu as vu l'enfant ? ; *hal huwa marīḏ-u-n* ? = est-ce que lui [est] malade ? Elle peut l'être aussi par la particule *'a* (cette dernière clitique en raison de sa structure), mais celle-ci sert plutôt à sélectionner le syntagme interrogé : *'a-gadan ta'tī* = est-ce demain [que] tu viens ? ; cliticisée à *laysa* elle exprime l'interro-négation : *'a-laysa l-walad-u marīḏ -a-n* ? l'enfant n'est-il pas malade ?

La relative est construite par simple juxtaposition si l'antécédent est indéfini : *ḡā'a walad-u-n ta'rifu-hu* = est-venu un enfant tu connais-lui = que tu connais. Si l'antécédent est défini elle est introduite par l'article de phrase *'allaḏī* (qui s'accorde avec l'antécédent en genre, nombre et, pour le duel, en cas) : *ḡā'a (a)l-walad-u 'allaḏī ta'rifu-hu* = est-venu l'enfant que tu connais-lui. La fonction de la relative est marquée par un pronom de reprise « sur site ».

Les complétives sont introduites par deux complétiviseurs : *'an* et *'anna* (ce dernier prenant la forme *'inna* si le verbe introducteur est une forme quelconque du verbe *qāla* = « dire »). *'an* introduit les complétives qui ont une structure de phrase verbale (le verbe est au subjonctif s'il s'agit d'un verbe à préfixe) : *'urīdu 'an yakruḡa* = « je veux qu'il sorte ». *'anna* introduit les autres complétives : *ra'aytu 'anna(a)l-walad-a marīḏ-u-n* = j'ai vu que l'enfant [était] malade. *'an* et *'anna* sont corrélés respectivement aux procès considérés comme simplement envisagés ou comme effectifs.

Les énoncés conditionnels et hypothétiques sont constitués de deux phrases interdépendantes : cette interdépendance est souvent marquée par le choix du même paradigme verbal pour les deux phrases, et une particule spécialisée introduit le couple. Dans la syntaxe classique, la particule *'in* introduit les conditionnelles et la particule *law* les hypothétiques. Dans la langue moderne *'iḏā* (quand) tend à se substituer à *'in*.

Le système casuel de l'arabe obéit à des tendances générales simples, malgré quelques idiosyncrasies : le nominatif marque le thème et le propos des phrases thématiques, le localisé des phrases locatives, et le sujet des phrases verbales. Il

marque aussi le patient dans la diathèse objective (car c'est le sujet grammatical dans ce cas). L'accusatif marque les objets, premier et second du verbe ainsi que tous les compléments de phrase. Il marque aussi, ce qui est compréhensible, le sujet et le prédicat des verbes de perception et d'évaluation (qui sont en fait des objets phrastiques) ; mais il marque aussi, ce qui est moins évident, le prédicat des verbes d'état. On a donc à côté de *al-walad-u marīḍ-u-n* = l'enfant (nom.) [est] malade (nom.), *ḥasibtu al-walad-a marīḍ-a-n* = j'ai cru l'enfant(acc.) malade (acc.), mais aussi *kāna al-walad-u marīḍ-a-n* = était l'enfant (nom.) malade (acc.). On trouve aussi l'accusatif après un certain nombre de particules de sur-thématisation ou de modalisation du nom : *'inna (a)l-walada marīḍ-u-n* = oui, l'enfant [est] malade ; *layta l-šabāb-a ya'ūdu* = si la jeunesse (acc.) revenait ! Le génitif marque le complément de la préposition, *fil-bayti* = dans la maison (gen.), et celui du nom, *kitābu l-walad-i* = [le] livre [de] l'enfant (gen.). La génitivation neutralise, avec les constructions à nom verbal, l'opposition sujet/objet : *ḍarbu (a)l-waladi* = le frappement de l'enfant, soit que l'enfant frappe, soit qu'il soit frappé.

IV. Le lexique

Le système « racine-schème » s'est montré d'une grande puissance génératrice surtout pour les noms et verbes associés à des racines triconsonantiques. En outre la régularisation analogique, accentuée par la grammatisation, a joué un rôle très important notamment dans la constitution du lexique savant. Il est encore sollicité, à l'époque moderne, pour combler les lacunes du vocabulaire scientifique et technique de l'arabe.

La phonologie et la morphologie contraignent fortement les emprunts : la langue a tendance à n'intégrer complètement que les mots que leur structure formelle permet d'intégrer dans le système racine-schème : ainsi *film* identifié à un mot de type $R_1iR_2R_3$ donne le pluriel arabe *'aflām* = « films ».

Mais dans la majorité des cas, on a plutôt recours à l'immense stock des racines traditionnelles (les grands dictionnaires de la langue en recensent plus de 9 000) pour répondre aux nouveaux besoins lexicaux de la langue. L'arabe reste donc une langue à lexique largement endogène.

Bibliographie

- Badawī M. , (1973), *Mustawayāt al-‘arabiyya l-mu‘āšira fī miṣr*, Le Caire, Dar al-Ma‘ārif.
- Bellamy J. , (1990), « Arabic verses from the first/second century : The inscription of 'En 'Avdat », *Journal of Semitic Studies*, n° 2573-79.
- Björnesjö S. , (1996), L'arabisation de l'Égypte : le témoignage papyrologique, *Égypte / Monde arabe*, n° 27-2893-106.
- Blachère R. , (1952), *Histoire de la littérature arabe des origines à la fin du xv^e siècle de J.-C.*, Paris, Adrien-Maisonneuve.
- Blachère R. , (1958), *Introduction au Coran*, Paris, Maisonneuve & Larose.
- Blau J. , (1966), *A Grammar of Christian Arabic*, (3 vol.), Louvain, Imprimerie orientaliste.
- Blau J. , (1972), On the problem of the synthetic character of Classical Arabic as against Judaeo-Arabic (Middle Arabic), *Jewish Quarterly Review*, n° 63260-269.
- Blau J. , (1977), The beginnings of the arabic diglossia : A study of the origins of neo-arabic, *Afro-Asiatic Linguistics*, n° 4175-202.
- Blau J. , (1988), *Studies in Middle Arabic and its Judaeo-Arabic Variety*, Jerusalem, The Magnes Press.
- Blau J. , (1981), *The Renaissance of Modern Hebrew and Modern Standard Arabic*, Los Angeles, University of California Press.
- Bohas G. , Guillaume J.-P. et Kouloughli D. E , (2006), *The Arabic Linguistic Tradition*, Georgetown, Georgetown University Press.
- Bourdieu P. , (2001), *Langage et pouvoir symbolique*, Paris, Le Seuil.
- Carter M. G. , (1983), *Language Control as People Control in Medieval Islam*, al-'Abḥāt, n° 3165-84.
- Carter M. G. , (2004), *Sībawayhi*, Oxford, Tauris.
- Cohen D. , (1970), Koinè, langues communes et dialectes arabes, *Études de linguistique sémitique et arabe*, 105-125, La Haye, Mouton.
- Corriente F. , (1971), On the functional yield of some synthetic devices in arabic and semitic morphology, *Jewish Quarterly Review*, n° 6220-50.
- Corriente F. , (1976), From old arabic to classical arabic the pre-islamic, Koinè, Some notes on the native grammarians' sources, attitudes and goals, *Journal of Semitic Studies*, n° 2162-98.

- Diakonoff I. M. , (1994), *Afro-Asiatic Languages* Londres, Encyclopaedia Britannica. , dvd-rom version Diem W. , (1973), Die nabatäische Inschriften und die Frage des Kasusflexion im Altarabischen, *Zeitschrift der Deutschen Morgen ländischen Gesellschaft*, n° 123227-237.
- Doss M. , (1996), Réflexions sur les débuts de l'écriture dialectale en Égypte, *Égypte / Monde arabe*, n° 27-28119-145.
- El-Hassan S. , (1977), Educated spoken arabic in Egypt and the Levant : A critical review of diglossia and related concepts, *Archivum Linguisticum*, n° 8-2112-132.
- Ferguson C. A. , (1959 a), Diglossia, *Word*, n° 15325-340.
- Ferguson C. A. , (1959 b), The Arabic Koine, *Language*, n° 25616-630.
- Ferguson C. A. , (1989), Grammatical agreement in classical arabic and the modern dialects : A response to Versteegh's pidginization hypothesis, *al-'Arabiyya*, n° 225-17.
- Ferrando I. , (2001), *Introducción a la historia de la lengua árabe : nuevas perspectivas*, Zaragoza, Pórtico Librerías.
- Frolov D. , (2000), *Classical Arabic Verse : History and Theory of 'Arūd*, Leiden, Brill.
- Fûck J. , (1955), *'Arabīya : Recherches sur l'histoire de la langue et du style arabe*, Paris, Didier.
- Garbini G. , (1984), *Le lingue semitiche : studi di storia linguistica*, Naples, Instituto Universitario Orientale.
- Grandguillaume G. , (1983), *Arabisation et politique linguistique au Maghreb*, Paris, Maisonneuve & Larose..
- Hetzron R. , (1976), Two principles of genetic classification, *Lingua*, n° 3889-108.
- Holes C. , (2006), *Modern Arabic : Structures, Functions and Varieties*, Georgetown, Georgetown University Press.
- Hopkins S. , (1984), *Studies in the Grammar of Early Arabic*, Oxford, Oxford University Press.
- Hourani A. , (1993), *Histoire des peuples arabes*, Paris, Le Seuil.
- Ḥusayn T. , (1927/1973), *Fī l-'adab al-ġāhiliyy*, *Œuvres complètes*, vol. V, Beyrouth, Dār al-Kitāb al-Lubnānī.
- Ḳalīl H. , (1978, et 1979) *al-Muwallad (I & II) : dirāsa fī numuwwi wa-taṭawwuri l-luġa l-'arabiyya ba'da l-'islām* Le Caire, et Alexandrie, al-Hay'a l-Miṣriyya l-'amma li-l-Kitāb.
- Kouloughli D. E. , (1994), *Grammaire de l'arabe d'aujourd'hui*, Paris, Pocket - Langues pour Tous.
- Kouloughli D. E. , (1996), Sur quelques approches de la réalité

sociolinguistique arabe, *Égypte / Monde arabe*, n° 27-28 287-299.

Kouloughli D. E. , (2007), *Le résumé de la grammaire arabe par Zamaḵṣarī*, Lyon, ens-Éditions.

Meynet R. , (1971), *L'écriture arabe en question*, Beyrouth, Dar El-Machreq.

Miller A. M. , (1986), The origin of the modern arabic sedentary dialects : An evaluation of several theories, *al-'Arabiyya*, n° 19 47-74.

Mitchell T. F. , (1986), What is educated spoken arabic ?, *International Journal of the Sociology of Language*, n° 6 17-32.

Monroe J. T. , (1972), Oral composition in pre-islamic poetry, *Journal of Arabic Literature*, III, 1-54.

Monteil V. , (1960), *L'arabe moderne*, Paris, Klincksieck.

Paoli B. , (2001), Meters and formulas : The case of ancient arabic poetry, *Linguistic Approaches to Poetry, Belgian Journal of Linguistics*, n° 15 113-136.

Rabin C. , (1955), The beginnings of classical arabic, *Islamica*, 4, 19-37.

Robin Ch. , (1992), *L'Arabie antique de Karib'il à Mahomet*, Aix-en-Provence, Edisud.

Shaban M. A. , (1971), *Islamic History : A New Interpretation*, Cambridge, Cambridge University Press.

Suyūṭī Ġalāl al-Din (m. 1505) , al-Muzhir, Le Caire, Dār 'Iḥyā' al-Kutub al-'Arabiyya, 1958.

Versteegh K. , (1984), *Pidginization and Creolization : The Case of Arabic*, Amsterdam, John Benjamins.

Versteegh K. , (1997), *The Arabic Language*, Edinburgh, Edinburgh University Press.

Zwettler M. , (1978), *The Oral Tradition of Classical Arabic Poetry : Its Character and Implications*, Columbus, Ohio State University Press.