

# Méthodes et problématique de la classification des langues : cas des langues chamito-sémitiques

Mohand MAHRAZI\*

## Résumé :

Les langues changent dans le temps : de génération en génération apparaissent des différences, généralement légères et n'affectant que très modérément l'intercompréhension, mais qui, en s'accumulant, deviennent importantes. A force d'évoluer, certaines de ces langues finissent par mourir, et il arrive aussi qu'il en naisse ; chaque langue est un mélange tel qu'il est bien difficile de reconnaître après quelques générations, le substrat d'origine.

Depuis longtemps, la problématique de l'origine des langues a toujours suscité de nombreuses interrogations et hypothèses chez les linguistes et qui font appel à d'autres disciplines comme l'anthropologie, l'archéologie, la génétique, la préhistoire, etc.. Dès le Moyen Âge, l'idée d'une langue mère unique obsède les chercheurs. Pendant longtemps, certains ont cru que l'hébreu était la langue d'Adam et d'Eve, d'autres ont pensé au latin ou au grec... Cependant, la découverte du sanscrit au XIX<sup>e</sup> siècle bouleverse la situation.

De cette problématique, découlent plusieurs questions : les langues descendent-elles d'une langue unique ? Ou bien, les langues sont-elles apparues simultanément sous différentes formes en différents endroits ? Si c'est le premier cas, peut-on reconstruire cette langue originelle ? Si c'est le deuxième cas, existe-t-il des similitudes entre elles ? Les différences se situent à quels niveaux ? Peut-on les classer ? Si oui, comment et quels critères permettant de dire que telle langue est fille ou descend de telle autre, que telle langue appartient à tel groupe ? Par cet article, je voudrais donc présenter un aperçu sur les méthodes et les critères de classification, et enfin, terminer par une illustration en donnant un exemple sur les langues *chamito-sémitiques*.

## Abstract :

Languages change through time, differences show up from generation to generation, usually the differences are very slight and don't moderately affect the comprehension, but when they accumulate, they become important. Because of evolving so much some languages ends up dying, and sometimes new languages are born, every new language is a mixture even if it's difficult to admit through generalizations, the original substratum.

For a long time, the issue of language origins is always aroused by the countless interrogations and hypotheses by linguists that always call for other disciplines like anthropology, archeology, genetics prehistory...etc, ever since the middle ages, the idea of a unique mother language possessed researchers. For a long time, few believed that HEBREW was the language of Adam and Eve, others believed that it was Greek or Latin. However the discovery of the Sanskrit in the XIX<sup>th</sup> century changed the situation.

From this problematic, came several questions: are all languages descendents of the same and unique language? Or have all languages appeared simultaneously in different forms and different places? If it was the first case can we rebuild the original language? And if it was the second case, are there any similarities between all of them? How different are they? Can we classify them? yes, what are the criteria that would allow us to tell which language is descendent of which, and what category does it belong to? Through this article I would like to present a preview of the methods used and criteria of classification, and finally finish by an illustration, giving an sample of the chamito-Semitic languages.

\* M .C.A . faculte des Lettres et des Langues, Université Akli Mohand Oulhadj, Bouira.

## 1- Introduction :

Le nombre de langues dans le monde est estimé à plusieurs milliers. Ces langues ne constituent qu'une infime partie des langues parlées par l'Homme depuis son origine. Mais on ne saura jamais le nombre de langues éteintes puisque la plupart des langues disparues n'ont pas laissé de trace. Quant au nombre de locuteurs d'une langue, il varie d'une langue à une autre : il peut être très élevé (plusieurs dizaines de millions) ou très restreint (quelques centaines ou quelques dizaines, voire moins dans les situations de disparition prochaine).

Les langues changent dans le temps : de génération en génération apparaissent des différences, généralement légères et n'affectant que très légèrement l'intercompréhension, mais qui, en s'accumulant, deviennent importantes. Les langues changent aussi dans l'espace : si un groupe linguistique se sépare par des migrations ou la constitution de frontières, les changements apparus d'un côté et de l'autre peuvent donner d'abord des variantes dialectales (exemple: le français belge ou suisse par rapport au "français français" ou "standard") puis au bout de quelques siècles des formes sans intercompréhension, donc des langues différentes comme c'est le cas de plusieurs langues. Comme le disait Michel Malherbe (1983 : 13), "à force d'évoluer, certaines de ces langues finissent par mourir, et il arrive aussi qu'il en naisse, mais jamais du néant". Pour ce dernier, chaque langue est un mélange tel qu'il est bien difficile de reconnaître après quelques générations, le substrat d'origine. Ainsi le français n'est ni du latin ni du gaulois, et les deux ancêtres ont disparu dans l'usage, le second au point qu'on ne sait même plus à quoi il ressemblait exactement.

Les interactions entre langues, les emprunts lexicaux et syntaxiques, les changements propagés par vagues franchissant les frontières entre langues, les mélanges de populations, tout cela et d'autres facteurs encore rendent la situation donc bien plus complexe. Saussure, dans son *Cours de linguistique générale* (1994 : 371), écrit à ce propos :

« Demander à quel type de groupe de langues se rattache, c'est oublier que les langues évoluent ; c'est sous-entendre qu'il y aurait dans cette évolution un élément de stabilité. Au nom de quoi prétendait-on imposer des limites à une action qui n'en connaît aucune ? ».

De cette problématique, découlent plusieurs questions : les langues descendent-elles d'une langue unique ? Ou bien, les langues sont-elles apparues simultanément sous différentes formes en différents endroits ? Si

c'est le premier cas, peut-on reconstruire cette langue originelle ? Si c'est le deuxième cas, existe-t-il des similitudes entre elles ? Les différences se situent à quels niveaux ? Peut-on les classer ? Si oui, comment et quels critères permettant de dire que telle langue est fille ou descend de telle autre, que telle langue appartient à tel groupe ? Par cet article, je voudrais donc présenter un aperçu sur les méthodes et les critères de classification, et enfin, terminer par une illustration en donnant un exemple sur les langues *chamito-sémitiques*.

Depuis longtemps, la problématique de l'origine des langues a toujours suscité de nombreuses interrogations et hypothèses chez les linguistes et qui font appel à d'autres disciplines comme l'anthropologie, l'archéologie, la génétique, la préhistoire, etc. Au début, dès le Moyen Âge, l'idée d'une langue originelle unique pour toute l'humanité hante les chercheurs. Pendant plusieurs siècles, certains chercheurs ont cru que l'hébreu était la langue d'Adam et d'Ève. D'autres, selon leurs croyances religieuses, ont fait référence au latin, au grec, à l'arabe... Enfin, la découverte du sanscrit par William JONES au dix-neuvième siècle bouleverse le mythe d'une langue unique.

Dire que deux idiomes appartiennent à la même langue ou constituent deux langues différentes, c'est pouvoir établir des liens ou des ressemblances entre eux de la même façon qu'il fallait justifier l'appartenance à une même famille linguistique de deux langues distinctes. Comme toutes espèces, selon des critères, les langues peuvent être hiérarchisées en catégories ou familles. Cette classification doit reposer sur l'utilisation de certains traits choisis selon les affinités, c'est l'un des rôles de la linguistique comparée.

La *linguistique comparée* (ou la *grammaire comparée* ou encore la *linguistique historique*) est apparue au XIX<sup>e</sup> siècle avec la découverte du sanscrit. La découverte de cette ancienne langue indienne et sa comparaison avec des langues européennes anciennes (grec, latin) et modernes met en évidence plusieurs correspondances entre cette langue et la plupart de ces langues d'Europe ; ce qui a donc amené les linguistes à établir une classification génétique des langues, qui avait pour objectif d'établir des relations de parenté entre les langues dans une perspective diachronique.

La principale méthode de travail repose sur la comparaison entre les différents états d'une même langue ou entre des langues différentes mais issues d'un même ancêtre. Elle permet, en relevant des concordances régulières phonétiques, syntaxiques, sémantiques, morphologiques, lexicales,... d'établir des parentés entre les langues. De là naît l'idée que des groupes de langues descendent d'une même source et pressentaient donc des filiations. C'est donc la linguistique comparée qui permet d'établir de manière scientifique l'existence des familles de langues .

## 2- Méthodes de classification

Etant donné la pluralité des méthodes de classification, je me contenterai ici, à présenter seulement les plus importantes ou du moins celles sur lesquelles la plupart des linguistes étaient d'accords.

Les linguistes distinguent d'emblée trois grands principes classificatoires des langues: le classement *généalogique* dit aussi *historique*, le classement *aréal* et le classement *typologique*.

### 2-1- Classement généalogique ou génétique

C'est la première classification proposée par les linguistes qui conduit à construire des arbres. Elle s'intéresse aux familles de langues descendant d'une langue présumée commune ou originelle. Cette génétique "provient d'une conception biologique de la langue qu'avait adoptée Franz Bopp au XIX<sup>e</sup> siècle"<sup>1</sup>. Ce dernier compare les langues à des êtres humains dont on pouvait suivre la naissance, la vie et la mort. Dans ce type de classement, les familles sont organisées hiérarchiquement en langues mères, langues filles, langues sœurs, etc. Avec la découverte du sanskrit et le début de la grammaire comparée, la méthode de classement se rationalise, mais sans abandonner entièrement l'idée de la monogenèse des langues. "La méthode est fondée sur la statistique lexicale portant sur un vocabulaire de base, et sur l'identification de correspondances phonétiques régulières, étant entendu que le changement des sons est régulier, ce qui permet de rendre les reconstructions proposées plus convaincantes" (Peyraube A. : 2000). Pour les comparatistes, c'est en définitive la loi phonétique qui constitue la preuve interne de la validité de la reconstruction. Ce mode de classification est toujours considéré comme étant supérieur aux autres (Greenberg 2001)<sup>2</sup>.

En étudiant des milliers de langues parlées dans le monde, les linguistes ont pu établir certains liens de parenté plus ou moins étroits entre des parlars dont plusieurs peuvent représenter des évolutions différentes d'un même prototype (Leclerc, 2014). Ainsi, ces langues sont réparties en familles qui sont à leur tour divisées en "sous-familles" ou "branches", comme c'est le cas par exemple de la branche romane, de la branche germanique, de la branche slave, etc.). Ces branches sont elles-mêmes, à leur tour, formées de certaines langues plus intimement apparentées entre elles qu'avec d'autres. Ainsi, les langues de la branche romane (français, espagnol, italien, espagnol, etc.) diffèrent de celles de la branche germanique (anglais,

<sup>1</sup> J. LECLERC, 2014 : « L'origine des langues », en ligne sur le site Internet :<http://www.axl.cefan.ulaval.ca/monde/familles>.

<sup>2</sup> Cité dans Alain Peyraube, 2002 : « La classification des langues en familles et macro-familles, avec une attention particulière portée aux langues d'Asie Orientale et d'Asie du Sud-Est ».

allemand, néerlandais, danois, etc.) et slave (russe, polonais, tchèque, slovène, etc.), mais elles appartiennent toutes à la même famille : la famille *indo-européenne* (Leclerc, 2014).

L'une des faiblesses de cette approche, qui se base sur la statistique lexicale, est la non conservation des langues de leur stock lexical et morphologique originaux. En effet, au fil du temps, les items lexicaux ou les affixes, même les plus basiques, disparaissent peu à peu et sont remplacés par des éléments nouveaux (Peyraube A. : 2000). Cependant, le seul élément dont on est certaine, c'est que, entre une multitude de langues diverses, un ensemble de traits communs existe et constitue une parenté indéniable. On ne connaît pas l'indo-européen primitif car, étant donné qu'aucun document écrit ne peut confirmer son authenticité ou qui laisse supposer son existence, mais en revanche, on connaît les langues qui en sont issues et ce qu'elles sont.

### 2-2- Classement aréal

Cette classification privilégie les considérations d'ordre géographique. Les traits communs existants au sein des différentes langues ne s'expliquent pas par une origine commune, mais simplement par les mélanges et les contacts séculaires entre elles. On définit plus précisément une aire linguistique par un ensemble de traits communs qui caractérisent les langues de cette aire et qui ont été introduits dans la langue par emprunt, après contact. Ces traits peuvent être à la fois aréaux et typologiques, mais ils ne peuvent être en même temps généalogiques et aréaux, car cela impliquerait deux scénarios historiques mutuellement exclusifs (Peyraube A. : 2000).

Si on examine une carte de la distribution géographique de traits linguistiques, on constate souvent que des langues proches géographiquement et qui partagent un trait linguistique majeur, en partagent beaucoup d'autres, même si elles appartiennent généalogiquement à des familles différentes. Il reste qu'il est souvent difficile de distinguer, dans les ressemblances entre les langues, celles qui sont le résultat d'emprunts par contact de celles qui pourraient être explicables par une même source. En l'absence de données historiques, comme c'est souvent le cas, origine commune et contagion sont deux hypothèses (Peyraube A. : 2000).

### 2-3- Classement typologique

August Schleicher et Wilhelm Von Humboldt étaient les premiers à avoir proposé ce type de classification. "La classification typologique des langues a pour but leur description et leur regroupement en fonction de certaines

caractéristiques communes de leurs structures, sans rechercher nécessairement l'établissement de généalogies ou de familles de langues"<sup>1</sup>.

Dans ce mode de répartition, les langues sont rangées en types en se fondant sur les similarités des structures linguistiques, critères phonétiques ou phonologiques, morphologiques ou syntaxiques.

### 2-3- 1- Les critères phonétiques ou phonologiques

En fonction de leur système vocalique, les langues se distinguent : "les langues à trois voyelles ([i], [u], [a]), les langues à double articulation antérieure (voyelles non arrondies et arrondies), les langues à double articulation postérieures (voyelles non arrondies et arrondies), les langues à double durée vocalique, etc."<sup>2</sup>

De même, en fonction du système consonantique, "certains linguistes ont tenté de classer les langues en fonction des modes d'articulation: les langues à consonnes occlusives limitées (une seule), à consonnes fricatives limitées (seulement le [t]), les langues à consonnes pré-nasalisées ou post-nasalisées, les langues à clics, les langues à deux modes articulatoires, etc. D'autres distinguent les langues «à tons» (comme le chinois, le vietnamien, le birman), les langues «à accent tonique fixe» (tchèque, finnois, hongrois), les langues à accent tonique «à valeur phonologique» (russe), etc."<sup>3</sup>

### 2-3- 2- Les critères morphologiques

Les critères d'ordre morphologique font distinguer quatre principaux types : le type *flexionnel*, le type *agglutinant*, le type *isolant* et enfin le type *polysynthétique* ou *incorporant*.

- **Le type flexionnel** : c'est un mode où "les radicaux sont pourvus d'affixes grammaticaux variables et exprimant plus ou moins à la fois, par exemple, le genre, le nombre et le cas, ou la personne, le temps, le mode, la voix, etc."<sup>4</sup>Ce type se caractérise par trois traits<sup>5</sup> :

<sup>1</sup> J. LECLERC, 2014 : « L'origine des langues », en ligne sur le site Internet : <http://www.axl.cefan.ulaval.ca/monde/familles.htm>

<sup>2</sup> Idem

<sup>3</sup> Idem

<sup>4</sup> J. LECLERC, 2014 : « L'origine des langues », en ligne sur le site Internet : <http://www.axl.cefan.ulaval.ca/monde/familles.htm>

<sup>5</sup> J. POITOU, 2009 : "Problématique de la typologie des langues" en ligne sur le site Internet : <http://j.poitou.free.fr/pro/html/gen/typologie.html>

- a- Une même unité lexicale peut avoir plusieurs formes selon les valeurs grammaticales qui lui sont associées. Ces variations de forme se font à l'aide de deux procédés :
  - l'affixation (suffixation, préfixation, infixation)
  - les alternances internes, vocaliques ou consonantiques.
- b- Les signifiants correspondant aux différentes valeurs grammaticales ne sont pas autonomes les uns par rapport aux autres, ils sont fusionnés (ou amalgamés), voire même aussi fusionnés avec le signifiant de l'unité lexicale.
- c- Les segments représentant les valeurs grammaticales correspondent souvent à plusieurs valeurs distinctes.

La plupart des langues européennes sont des langues flexionnelles. En latin par exemple, la désinence *-arum* peut exprimer le génitif, le féminin et le pluriel du mot *rosarum* ; de plus, elle assigne le mot auquel elle est accolée à la classe du nom, de l'adjectif, ou du pronom, mais pas à celle du verbe. En français, on retrouve la flexion de l'article (le, la, les), la flexion de certains substantifs et adjectifs (cheval-chevaux, original-originaux), la flexion verbale...

- **Le type agglutinant** : ce type se caractérise par le fait que toutes les relations grammaticales se marquent par l'adjonction des affixes (suffixes, infixes, préfixes) aux bases. A chaque affixe correspond une relation grammaticale (genre, nombre, cas, ...). Saussure (1994 : 283) définit l'agglutination comme suit : « L'agglutination consiste en ce que deux ou plusieurs segments originellement distincts, mais qui se rencontraient fréquemment en syntagme au sein de la phrase, se soudent en une unité absolue ou difficilement analysable ». En voici quelques exemples d'agglutination en français et en turc<sup>1</sup> :

<sup>1</sup> Exemple extrait de (Sternemann et Gutschmidt 1989 : 81) cité dans Jacques Poitou 2009.

<sup>9</sup> Exemple extrait de J. LECLERC, 2014 : « L'origine des langues », en ligne sur le site Internet : <http://www.axl.cefan.ulaval.ca/monde/familles>.

Français :**ceci** de deux mots *ce* et *ci***toujours** de deux mots *tous* et *jours***aujourd'hui** de quatre mots *au*, *jour*, *d'* et *hui*Turc :**ev** = maison**evim** = ma maison**evlerim** = mes maisons**evimden** = de ma maison**evlerimden** = de mes maisons

- **Le type isolant** : ce type se caractérise par le fait que chaque mot a une seule forme et ne peut être modifié que ce soit par flexion ou par dérivation. Les valeurs grammaticales, les catégories comme le pluriel, sont marquées soit par l'adjonction d'un mot supplémentaire, soit par l'ordre des mots. Parmi les langues isolantes les plus connues, on peut citer le chinois, le cantonais, le vietnamien, le laotien et le cambodgien, etc. Voici un exemple en chinois<sup>1</sup> :

ta	chi	fan	le	Ta	Chi	Le	Fan
<i>il</i>	<i>mang</i> <i>e</i>	<i>repas</i>	<i>passé</i>	<i>Il</i>	<i>Man</i> <i>ge</i>	<i>passé</i>	<i>Repas</i>
= <i>Il prit son repas.</i>							

- **Le type polysynthétique ou incorporant** : dans ce type de langues, les relations grammaticales peuvent être exprimées en ajoutant des éléments, variés qui n'ont pas nécessairement d'existence autonome par ailleurs, à un radical unique ou en lui faisant subir des transformations, sans qu'aucune partie de la phrase ne puisse changer de place. Le linguiste danois Hjelmslev donne l'exemple du groenlandais, dans une phrase comme *kavfiliorniarumagaluarpunga*, qui en un seul mot, signifie *Je voudrais faire du café*.

**2-3- 3 - Les critères syntaxiques**

Le linguiste américain Joseph H. Greenberg considère que l'ordre des constituants dans la phrase ou le critère syntaxique est aussi un critère de classification des langues. Par exemple, en français, l'ordre syntaxique le plus courant est : sujet + verbe + complément (S-V-C). Comme on peut



---

aussi trouver d'autres ordres : S-O-V (sujet-objet-verbe) pour le latin, V-S-O (verbe-sujet-objet) pour l'espagnol et le portugais américain. Aujourd'hui,

l'ordre des constituants dans la phrase, est sans doute un des critères les plus utilisés d'une classification typologique des langues car l'ordre des mots d'une langue reste un trait structural stable, imperméable à l'emprunt externe.

### Remarques :

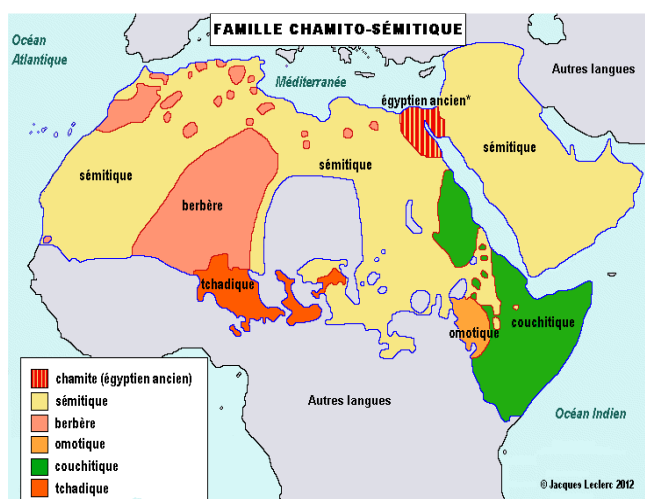
- 1- Aucun de ces types n'existe sous une forme pure dans la réalité ; certaines langues les mêlent, le français par exemple, procède par plus d'un type : type flexionnel (*cheval / chevaux*), type isolant (*je suis / tu es*), type agglutinant (*épais / épaisse*). Un syntagme comme *porte-manteau* est isolant, alors qu'une opposition du genre *pomme / pommier* est flexionnelle.
- 2- Le *basque*, l'*eskimo-aléoute* et le *nigéro-congolais* comptent parmi les langues qui ont connu un isolat linguistique, c'est à dire qu'on leur a pas trouvé de parenté certaine ; ainsi, chacun d'eux peut donc constituer une famille à part

### 3- Les langues *chamito-sémitiques*<sup>1</sup>

Parmi les principales familles de langues distinguées par les spécialistes, on peut citer les langues *indo-européennes*, les langues *chamito-sémitiques*, les langues *nigéro-congolaises*, les langues *amérindiennes*, les langues *australésiennes*, les langues *papoues* et *australiennes*, les langues *bantoues*, les langues *ouralo-altaïques*, les langues *sino-thaïes* et les langues *dravidiennes*. Dans notre étude, nous nous limiterons seulement aux langues *chamito-sémitiques*.

<sup>1</sup> Ce nom est tout à fait arbitraire, il est une pure invention des linguistes de la fin du XVIIIe siècle. "Sous l'influence de la Genèse (Bible judéo-chrétienne), les linguistes européens présentèrent les Hébreux, les Araméens, les anciens Égyptiens et les Arabes comme les descendants de Sem (d'où *sémitique*) et de Cham (d'où *chamite*), les fils du patriarche Noé" (LECLERC 2014). Pour Lionel Galand (1999), la composition double du terme *chamito-sémitique* ne doit pas induire à l'erreur courante qui laisse supposer qu'il existe un "chamitique" à côté d'un "sémitique. En effet, *chamito-sémitique* ne peut être entendu comme le nom d'un ensemble où entrent sur un pied d'égalité le *sémitique* et d'autres groupes apparentés, distincts du sémitique, mais non moins distincts l'un de l'autre... Pour éviter cet amalgame, les Anglo-saxons préfèrent parler de la famille *afro-asiatique*, ce terme paraît aussi inapproprié car comme le précise David Cohen (2006 : 10), "ces langues ne sont pas africaines, ne couvrent pas l'Afrique –ni même la majeure partie de l'Afrique –, naturellement, et ne sont pas non plus véritablement asiatiques –et le sont même encore moins dans ce sens".

Il faut attendre le début du vingt-et-unième siècle pour que les linguistes rassemblent toutes les langues de l'Asie de l'Ouest et de l'Afrique présentant des similitudes entre elles en une grande famille linguistique appelée *chamito-sémitique*. Ces langues font partie des familles les plus importantes du monde, tant par leur histoire – elles ont été dans le passé des langues de grandes civilisations : l'égyptien, le babylonien, le sumérien, le phénicien, le cananéen, l'hébreu, etc. – que par leur nombre –elles sont parlées par près de 300 millions de locuteurs – et leur distribution géographique de leurs locuteurs qui s'étend du nord de l'Afrique (du Maghreb jusqu'au Nigeria et une partie du Cameroun, en passant par l'Éthiopie, l'Erythrée et la Somalie) et de l'Île de Malte, ainsi que dans tout le proche Orient, pour s'arrêter aux frontières de l'Iran<sup>1</sup>. Actuellement, parmi les langues *chamito-sémitiques* les plus parlées est l'arabe, avec ses variétés dialectales, avec environ 200 millions de locuteurs.



<http://www.axl.cfan.ulaval.ca/monde/famarabe.htm>

<sup>1</sup> J. LECLERC, 2014 : « L'origine des langues », en ligne sur le site Internet : <http://www.axl.cfan.ulaval.ca/monde/familles>.

Selon les spécialistes, les langues *chamito-sémitiques* présentent une unité de structure profonde, sur le plan grammatical et syntaxique. Les langues "chamito-sémitiques" sont des langues à flexion dans lesquelles chaque mot présente une racine constituée d'une suite de consonnes (le plus souvent trois<sup>1</sup>, parfois deux ou quatre).

David COHEN indique 57 traits répandus dans l'ensemble des groupes chamito-sémitiques parmi lesquels on peut citer :

- les fortes restrictions sur les groupes de consonnes dans la syllabe ;
- le nombre restreint de timbres vocaliques ;
- la mise en relation possible (quoique non systématique) entre un lexème et une racine (définie comme un ensemble de consonnes). Par exemple NZL (notion de descente), qui est en arabe à la base de *nazala* (NZL + accompli, "il est descendu"), *yanzilu* (NZL + inaccompli, "il descend"), *manzil* (NZL + schème locatif, "auberge : lieu où l'on descend"), etc. les formants des schèmes ainsi que les préfixes et suffixes dérivationnels, qui sont fondés sur une série limitée de phonèmes, commune à toutes les branches du chamito-sémitique (ʾ, h, š, s, t, n, m, y, w, a, i, u) ;
- la distinction des phrases à prédicat nominal et à prédicat verbal ;
- une opposition de genre (masculin-féminin), et de nombre (singulier, pluriel, et dans certaines familles duel) ;
- pour le nom, la distinction entre un état construit et un état indépendant, dans de nombreux dialectes appartenant à toutes les branches ;
- pour le verbe, l'adjonction obligatoire au thème verbal d'une marque de personne représentant le terme sujet ;
- une organisation du paradigme verbal fondée sur l'opposition d'un thème simple à des dérivés, qui sont formés soit par augmentation du thème de base, soit par affixation d'un morphème de dérivation ;
- une prédominance de l'aspect (avec une distinction originelle entre un statif et un processif) sur le temps ;
- le moyen comme voix marquée dans certains états

<sup>1</sup> Ce triconsonantisme des racines est un trait fondamental des langues chamito-sémitiques. Cette structure trilitère discontinue est fondamentale ; les éléments radicaux exprimant l'idée lexicale restent apparents dans tous les mots formés sur une même racine et reliés étroitement à l'expression d'un concept. Pour Marcel Cohen (1947 : 58-59), les bilitères et monolitères, qui sont fréquentes en berbère et en couchitique, ne sont en fait que des formes réduites.

-l'existence de deux types de conjugaison, l'une suffixée stative et verbominale d'origine et l'autre préfixée processive, plus proprement verbale, dans presque toutes les branches ;

- développement d'une forme modale, souvent plus simple ;
  - des formants de dérivation nominale communs à plusieurs branches : suffixes en *-n*, en *-y* ;
  - un préfixe déverbatif *m-* répandu en chamito-sémitique ;
  - des interrogatifs en *m-* et d'autres pouvant être reconstruits comme *'ayy* ;
  - des séries pronominales (autonomes / clitiques) relativement proches dans les diverses branches. Actuellement, le *chamito-sémitique* compte cinq branches principales : le *sémitique*, l'*égyptien*, le *couchitique* (Afrique de l'Est), le *berbère* et le *tchadique*<sup>1</sup>.
- **Les langues sémitiques** : cette famille tire son nom du fils de *Noé*<sup>2</sup>, "*Sem*". Certaines langues sont attestées depuis 2000 ans avant notre ère (akkadien, ougaritique, éblaïte, phénicien, etc.). Cette famille est la plus importante du groupe chamito-sémitique : elle regroupe l'arabe (200 millions de locuteurs), l'hébreu (4,6 millions) et les langues akkadiennes anciennes (le babylonien et l'assyrien) vivantes en Irak. D'après Karl-G. PRASSE, ces langues sémitiques sont intimement liées si bien que nous pouvons presque toujours trouver les correspondances d'un mot d'une langue dans les autres ; mais en revanche, lorsque nous abordons les autres branches, la situation change. La majorité de ces langues sont, aujourd'hui, éteintes comme : l'akkadien, l'éblaïte, le babylonien, l'ougaritique, le cananéen, le moabite, le phénicien, le samaritain.

<sup>1</sup> L'hypothèse d'appartenance du *tchadique* à la famille *chamito-sémitique* ne fait pas encore véritablement l'unanimité des spécialistes ; elle a été incorporée plus récemment par les linguistes, même si l'hypothèse est ancienne. Ils estiment que l'on a pas encore apporté la preuve irréfutable que les ressemblances entre eux sont bien dues à une origine génétique commune et non à des emprunts résultant des contacts prolongés entre *chamito-sémitique*.

<sup>2</sup> Ce nom est aussi arbitraire selon David Cohen (2006: 13), car pour lui, *sémitique* n'a rien à voir avec *Noé* et son fils *Sem*. [...] Ce dont il s'agit, ce n'est pas de situer les langues par rapport à *Noé*, mais de donner, arbitrairement, un nom.

- **Les langues égyptiennes anciennes** : leurs évolutions commencent avec l'ancien égyptien dont la plus ancienne forme remonterait à près de 3000 avant notre ère. Cette famille englobe l'ancien, le moyen et le nouvel égyptien et le démotique. Aujourd'hui, excepté le copte, qui est encore utilisé comme langue liturgique de l'Eglise copte orthodoxe, toutes ces langues sont éteintes.
- **Les langues couchitiques** : du nom de *Coush*, fils de *Cham*, elles sont attestées depuis le 17<sup>e</sup> siècle et comptent environ 50 langues. Ces langues se répartissent de la frontière de l'Egypte jusqu'au nord de la Tanzanie. On peut subdiviser cette famille en cinq groupes, le couchitique du nord (bédja), le couchitique central (agaw), le couchitique de l'est (somali (8,3 millions de locuteurs dans le monde), saho, afar, oromo, ...), le couchitique du centre ouest (sidamo, darasa, kambata, ...) et le sud-couchitique (iraqw, ...).
- **Les langues tchadiques** : cette branche rassemble environ 140 langues et dialectes, situés aux alentours du lac Tchad. En dehors du haoussa, parlé par environ 22 millions de locuteurs et le seul à être reconnu par les Etats (Bénin, Burkina Faso, Cameroun, Ghana, Niger, Nigeria, Soudan, Togo, Tchad), elles sont peu parlées et se rattachent aux langues nigéro-africaines. Les plus importantes sont : le haoussa, le mandara, le ngala, le bana, etc.
- **Les langues berbères ou amazighes** : ces langues sont parlées en Afrique du Nord dans une dizaine de pays : Algérie, Maroc, Tunisie, Libye, Niger, Mali... Selon les spécialistes, cette famille est très homogène. Elle est subdivisée en quatre groupes : les parlers septentrionaux (kabyle, rifain...), méridionaux (touareg...), les dialectes orientaux (Siwa, Audjila...) et le zenaga (parlé en Mauritanie).

#### 4- Conclusion

Tout au long du dix-neufième et du vingtième siècle, les linguistes tentent de regrouper les langues du monde en familles linguistiques selon divers critères. En général, ils tentent de reconstruire des *proto-langues* pour les familles de langues dont ils sont spécialistes. trois types de classement ont été utilisés : le classement *généalogique* qui conduit à construire des arbres ; le classement *typologique* qui cherche les similarités et les ressemblances des structures linguistiques en mettant aussi au jour les divergences les plus importantes, enfin le classement *aréal* qui privilégie les

contacts géographiques et s'appuie sur des termes d'extension et de recouvrements d'aires.

La première classification supposait que toutes les langues descendaient d'une seule ou d'un même ancêtre, elle s'appelle aussi la classification *historique*. Une telle classification opère avec des niveaux hiérarchiques, c'est-à-dire, les familles sont organisées hiérarchiquement en langues mères, langues filles et langues sœurs, etc. Le français, par exemple, appartient au groupe des langues romanes, lui-même dépendant de la famille indo-européenne, que les "unificateurs" peuvent aussi rattacher à la macro-famille eurasiatique.

La deuxième classification, tente de prendre ses distances par rapport à l'étude de la filiation des langues. Elle se donne pour objet de décrire l'ensemble des phénomènes qui constituent une langue et de comparer ensuite les langues entre elles. Dans cette classification, les langues sont rangées en types, en se basant sur les similarités des structures linguistiques : morphologiques, sémantiques, syntaxiques et très rarement phonétiques. Selon les critères morphologique, on dénombre quatre types de langues : les langues flexionnelles (latin, russe,...), les langues agglutinantes (le turc), les langues isolantes (le chinois) et les langues polysynthétiques (le groenlandais). Quant aux critères syntaxiques, ils classent les langues suivant l'ordre des constituants principaux de la phrase assertive : Sujet, Verbe, Objet.

La troisième classification privilégie les considérations d'ordre géographique. Cette démarche considère que les faits homologues observés au sein de la diversité des langues ne s'expliquent pas par une origine commune, mais ils sont dus simplement aux contacts et les mélanges des sociétés. Cette conception a entraîné de grandes révisions de certaines reconstructions comme par exemple, l'unité du groupe ouralo-altaïque (langues ouraliennes et langues altaïques) qui n'est plus reconnue. Il reste qu'il est souvent difficile de distinguer, dans les ressemblances entre les langues, celles qui sont le résultat d'emprunts par contact de celles qui pourraient être explicables par une même origine. En l'absence de données historiques, comme c'est souvent le cas, origine commune et contagion sont deux hypothèses envisageables.

**Références bibliographiques**

- ACHAB K., 2005 : "Le système de genre et son origine en berbère et en chamito-sémitique", in « FAITS DE LANGUES ». *Les langues chamito-sémitiques (afro-asiatiques)*. Volume 1, revue de linguistique N° 26. Edition OPHRYS, Paris. pp. 97-128.
- Benveniste E., 1966 : *Problèmes de linguistique générale*. Paris, Gallimard.
- BRETON R., 1998 : *Géographie des langues*. Ed. Casbah, Alger.
- BRUGANTELLI V., 2006 : "La négation berbère dans le contexte chamito-sémitique", in « FAITS DE LANGUES ». *Les langues chamito-sémitiques (afro-asiatiques)*. Volume 2, revue de linguistique N° 27. Edition OPHRYS, Paris. pp. 65-72.
- CENTRE INTERNATIONAL D'ETUDES PEDAGOGIQUES : "Le groupe des langues indo-européennes : le triomphe de la grammaire comparée [COMPARAISON 18]" article en ligne consulté le 22 mars 2007 sur le site Internet : <http://www.ciep.fr/publications/genetique/genetique21.htm>
- CHAKER S., 1998 : *Manuel de linguistique berbère II*. ENAG Editions, Alger
- COHEN D., 2006 : "Entretien avec David COHEN", réalisé par Antoine Lonnet et Amina Mettouchi, in « FAITS DE LANGUES ». *Les langues chamito-sémitiques (afro-asiatiques)*. Volume 2, revue de linguistique N° 27. Edition OPHRYS, Paris. pp. 9-26.
- COHEN D., 1968 : "Les langues chamito-sémitiques", in Martinet A. (éd.), *Le Langage*, Paris, p. 1288-1329.
- COHEN M., 1947 : *Essai comparatif sur le vocabulaire et la phonétique chamito-sémitique*. Paris, Champion.
- GALAND L., 1999 : "Langue(s) Berbère(s)", Article paru dans L'écho du Parc, 22, décembre 1999, pp. 12-19 (Association des anciens personnels du Lycée du Parc, Lyon).
- HAGEGE C., 1982 : *La structure des langues*. Paris, PUF.
- JAKOBSON R., 1938 : "Sur la théorie des affinités phonologiques entre les langues", *Actes du 4ème congrès international des linguistes*. Copenhague. Repris dans Troubetzkoy N., *Principes de phonologie*. Paris : Klincksieck, 1947, pp. 351-365.



- LAKS B., 2002 : "Le comparatisme : de la généalogie à la génétique", *Langages* 146, pp. 19-45.
- LECLERC J., 1986 : *Langue et société*. Mondia Editeurs.
- LECLERC J., 2001 : *Aménagement linguistique dans le monde*, en ligne sur le site Internet : <http://www.tlfq.ulaval.ca/axl/>
- LECLERC J., 2014 : "La famille chamito-sémitique ou afro-asiatique", article en ligne sur le site Internet : <http://www.tlfq.ulaval.ca/AXL/monde/famarabe.htm>
- LECLERC J., 2014 : "L'origine des langues", article en ligne sur le site Internet : <http://www.tlfq.ulaval.ca/axl/monde/origine-langues.htm>
- LONNET A. & METTOUCHI A. 2006 : *Les langues chamito-sémitiques (afro-asiatiques)* (volume 1), en ligne sur le site Internet : [http://lettres.univ-lemans.fr/fdl/numeros%20parus/numeros%20fdl/n26/presentation\\_26.htm](http://lettres.univ-lemans.fr/fdl/numeros%20parus/numeros%20fdl/n26/presentation_26.htm)
- MALHERBE M., 1983 : *Les langages de l'humanité : une encyclopédie de 3000 langues parlées dans le monde*, SEGHERS, Paris.
- MEILLET, A., 1928 : *La Méthode comparative en Linguistique historique*, Oslo Réimpr. Paris, Champion, 1966.
- MEILLET A., 1964 : "Introduction à l'étude comparative des langues indo-européennes", Paris, 1937 [reprint : Alabama,], pp. 51-81.
- PEYRAUBE A., 2002 : "Le problème de la classification des langues". En ligne sur le site Internet : [http://www.reseau-asie.com/rf/ed\\_fr\\_0\\_7.html](http://www.reseau-asie.com/rf/ed_fr_0_7.html)
- PEYRAUBE A., 2002 : "La classification des langues en familles et macro-familles, avec une attention particulière portée aux langues d'Asie Orientale et d'Asie du Sud-Est", CRLAO (CNRS et EHESS), article en ligne sur le site Internet : <http://htl.linguist.jussieu.fr/speyra.pdf>
- POITOU Jacques 2009 : "Problématique de la typologie des langues" en ligne sur le site Internet : <http://j.poitou.free.fr/pro/html/gen/typologie.html>
- RUHLEN M., 1997 : *L'origine des langues*. Ed. Belin, Paris

- SAUSSURE F., 1994 : *Cours de linguistique générale*. Ed ENAG. Algérie.
- WIKIPEDIA : "Typologie des langues" article en ligne sur le site Internet : [http://fr.wikipedia.org/wiki/Typologie\\_des\\_langues](http://fr.wikipedia.org/wiki/Typologie_des_langues)
- WIKIPEDIA : "Langues par famille", article en ligne sur le site Internet : [http://fr.wikipedia.org/wiki/Langues\\_par\\_famille](http://fr.wikipedia.org/wiki/Langues_par_famille)

