

1. Qu'est-ce qu'une famille de langues ?

Une famille de langues est un regroupement de plusieurs langues linguistiquement reliées, descendantes d'une langue-ancêtre commune (appelée **protolangue**).

La plupart des langues du monde appartient à une famille définie et les langues qu'on ne peut regrouper avec d'autres sont généralement appelées **isolats**.

Seules les **langues créoles** ne sont ni des isolats, ni des membres d'une famille linguistique, mais constituent un modèle à part.

2. Des liens génétiques

Quand on compare le français, l'espagnol, le portugais, l'italien, le roumain, par exemple on est frappé d'un ensemble de ressemblances qui leur donnent un « air de famille ». Cet « air de famille » ne se retrouve pas en comparant le français avec l'allemand. Mais si l'on compare l'allemand avec l'anglais, le néerlandais, le suédois ou le danois, on retrouve entre ces langues un autre « air de famille ».

L'idée la plus naturelle est que si ces langues se ressemblent, c'est parce qu'elles sont des évolutions divergentes, des « descendantes » d'une même langue antérieure (ou « protolangue »), aujourd'hui disparue. Pour les cinq premières langues citées ci-dessus, appelées **langues romanes** nous connaissons l'ancêtre commun (le **latin**), pour les quatre suivantes, appelées **langues germaniques**, on n'a pas de documents écrits dans la langue ancêtre, mais on peut très raisonnablement penser qu'elle a existé. C'est en comparant les langues entre elles, en recherchant des règles constantes sur leurs ressemblances (et leurs divergences), que les linguistes arrivent à établir ces classifications génétiques. Cette méthode s'appelle la **linguistique comparative**. La classification des langues en familles de langues s'appelle *classification génétique* : deux langues appartenant à la même famille sont *apparentées génétiquement*.

3. Des ressemblances parfois trompeuses

Attention cependant. Des ressemblances entre deux ou plusieurs langues peuvent être dues à leur parenté génétique (les formes ressemblantes étant issues d'une même forme commune ancienne), mais elles peuvent avoir d'autres origines :

- Les **emprunts** : Ainsi le fait qu'on ait en français *tomate* et en aztèque *tomatl* ne prouve pas que ces deux langues soient apparentées, mais simplement qu'il y a eu un contact entre ces deux langues. Une plante jusque-là inconnue en Europe est arrivée avec le nom qu'on lui donnait dans son pays d'origine. Le français a donc « emprunté » un mot à une autre langue et l'a adapté à son vocabulaire.

- Le **hasard** : Les langues ont un système de sons limité pour exprimer des milliers de notions complexes. Si on prend deux langues au hasard parlées à des grandes distances l'une de l'autre, il y a fort à parier qu'on trouve toujours 3 ou 4 mots qui se ressemblent par la forme et par le sens. Ainsi si « dieu » en grec se dit *theos* et en aztèque, *teotl*, c'est probablement un hasard. C'est aussi très certainement un hasard si pour dire « chaud », le hmong, langue asiatique, dise *sov* (à prononcer comme le français « chaud »).

On ne peut donc parler de parenté génétique que si l'on trouve, non une ressemblance frappante mais isolée, mais un ensemble convergent de ressemblances, même partielles.

4. Nombreuses familles, familles nombreuses

De la même façon qu'il y a des langues « sœurs », issues d'une langue commune parlée 1000 ou 2000 ans auparavant, on peut imaginer des langues « cousines », issues de langues parlées plus anciennement. C'est en mettant en évidence des ressemblances systématiques et convergentes que certains linguistes ont au XIXe siècle posé l'existence d'une grande famille indo-européenne, première famille linguistique à être identifiée, qui comprend les langues romanes, les langues germaniques, les langues slaves, le grec et d'autres. Et si l'apparement du français et du russe vous paraît surprenant, essayez-donc de comparer le français au kurde ! Ce sont pourtant des **langues indo-européennes**. Appartenir à une même famille n'est pas un gage de ressemblance apparente, ni **d'intercompréhension** entre les locuteurs de ces langues. Certaines familles de langues peuvent avoir un grand nombre de divisions internes. Ces divisions sont parfois elles-mêmes appelées « familles » ou « sous-familles », ce qui peut engendrer une certaine confusion. Plusieurs termes peuvent être utilisés pour parler des divisions internes, des sous-divisions, des sous-sous-divisions, etc. Il n'y a pas pour le moment de réel consensus quant à leur dénomination, on parle ainsi de « **groupe**, de **branches**, de **sous-groupes**, etc ». Quand une famille regroupe un grand nombre de langues et de divisions internes on parle parfois de « superfamille » ou « macro famille ». C'est le cas de la famille **Niger-Congo**, qui regrouperait entre 1300 et 1500 langues (les chiffres divergent selon les sources), représentant entre le cinquième et le quart de l'ensemble des langues de la planète. On peut regrouper dans une même famille des langues géographiquement très distantes, parfois même parlées sur des continents différents. Ainsi la famille **eskimo-aléoute** regroupe des langues parlées en Sibérie orientale et de l'autre côté de l'Océan pacifique en Alaska, séparées par plusieurs milliers de kilomètres d'océan. On trouve d'ailleurs des langues eskimo-aléoutes dans tout l'extrême nord de l'Amérique de la côte pacifique à la côte atlantique et même au

Groenland. Dans le même ordre d'idée, les langues de la famille **austronésienne** sont parlées dans tout le Pacifique Sud, en Asie du sud-est et on trouve même des langues austronésiennes à Madagascar, au large du continent africain !

A l'inverse certaines régions du monde présentent une très grande diversité et une grande complexité dans les regroupements génétiques des langues. Ainsi, si on compte trois familles de langues en Europe, le **continent américain** dans son ensemble possède environ la **moitié** des familles de langues autochtones de la planète, alors que ces 400 langues totalisent environ 25 millions de locuteurs natifs. De nombreuses familles de langues amérindiennes regroupent moins de 15 langues. Il faut noter que de telles diversités sont partiellement dues à « l'isolement » de certaines populations et aussi aux manques de données sur ces langues, rendant difficile leur classification.

La classification des langues est une source de débat permanente entre les linguistes, et cela fait que le nombre même des familles, leurs dénominations et leurs compositions peuvent varier beaucoup d'une source à l'autre.