

qwertyuiopasdfghjklzxcvbnmqwerty  
uiopasdfghjklzxcvbnmqwertyuiopas  
dfghjklzxcvbnmqwertyuiopasdfghjkl  
zxcvbnmqwertyuiopasdfghjklzxcvbn  
mqwertyuiopasdfghjklzxcvbnmqwer  
tyuiopasdfghjklzxcvbnmqwertyuiop  
asdfghjklzxcvbnmqwertyuiopasdfgh  
jklzxcvbnmqwertyuiopasdfghjklzxcv  
bnmqwertyuiopasdfghjklzxcvbnmq  
wertyuiopasdfghjklzxcvbnmqwertyu  
iopasdfghjklzxcvbnmqwertyuiopasd  
fghjklzxcvbnmqwertyuiopasdfghjklz  
xcvbnmqwertyuiopasdfghjklzxcvbn  
mqwertyuiopasdfghjklzxcvbnmqwer  
tyuiopasdfghjklzxcvbnmqwertyuiopas  
ghjklzxcvbnmqwertyuiopasdfghjklz  
xcvbnmqwertyuiopasdfghjklzxcvbn  
qwertyuiopasdfghjklzxcvbnmqwerty

**Langues du monde**

**Ch.Bonnot**

2009-2010

Raphaël Grasset

## Sommaire

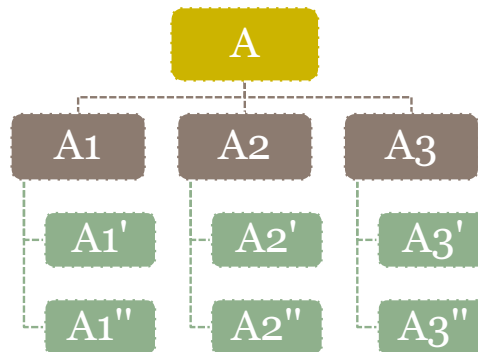
Les différentes classifications .....	3
La classification génétique .....	3
La classification aréale .....	3
La classification typologique .....	4
Classification génétique des langues .....	6
Arbres généalogiques des langues .....	6
Indo-European branches of the language tree .....	6
Arbre généalogique du proto-indo-européen .....	7
Indo-iranien <sup>(satem)</sup> .....	7
Arménien <sup>(satem)</sup> .....	8
Balto-slaves <sup>(satem)</sup> .....	8
Baltes .....	8
Slaves .....	8
Albanais <sup>(satem)</sup> .....	9
Grec <sup>(centum)</sup> .....	9
Germanique <sup>(centum)</sup> .....	9
Celtique <sup>(centum)</sup> .....	9
Italique <sup>(centum)</sup> .....	9
Tokharien .....	9
Hittite .....	10
Histoire de la classification des langues .....	10
Bases de l'analyse .....	11
Index des définitions .....	14

## Les différentes classifications

### La classification génétique

*Classification génétique* : établi et classe les langues en familles.

Les langues évoluent dans le temps. En évoluant elles tendent à se diversifier par un processus récursif :



*Protolangue* : langue non attestée par l'écriture.

A est une protolangue et A1, A2, A3 sont des langues filles. Le processus d'évolution des langues est continu.

Pour établir les familles de langues, on préfère se référer à la grammaire qu'au lexique qui peut être un apport.

*Cognats* : termes qui se ressemblent par la forme et le contenu (ex : *bone* en anglais, *bein* en allemand, *been* en hollandais, *ben* en suédois).

Les cognats permettent de construire des hypothèses sur la migration des langues. Ils sont donc d'intérêts pour les linguistes et les historiens.

### La classification aréale

*Classification aréale* : classe les langues en fonction des aires géographiques.

Par exemple, le Bulgare a pour famille génétique le slave du sud et comme classification aréale les langues balkaniques.

La connaissance des zones géographiques permet de comprendre l'évolution des langues et rends compte des convergences et ressemblances entre les langues parlées sur un même territoire et issues de branches différentes.

## La classification typologique

*Classification typologique* : classification non historique qui regroupe ensemble les langues qui se ressemblent (par exemple les langues à tons et les langues sans tons).

*Langues isolantes* : Langue ou le mot ne peut pas être découpé (suffixes, préfixes...) comme par exemple le chinois, le vietnamien.

*Langues flexionnelles* : Langue ou le mot peut être décomposé en racine plus éléments grammaticaux.

*Langues agglutinantes* : Langue ou le mot peut aussi être décomposé.

Il n'y a pas qu'une seule typologie. Il peut y avoir de nombreux critères différents comme la différence de structure :

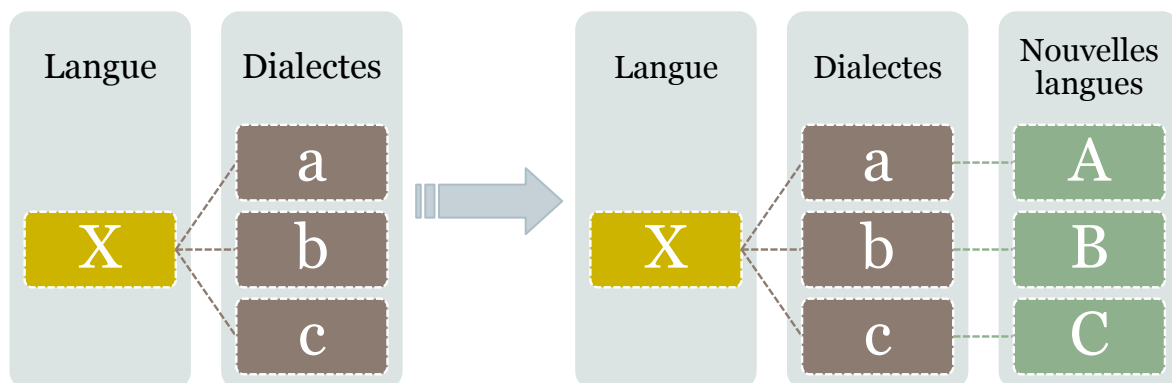
- Variation des tons : Asie, Afrique
- Par ressemblance phonique
- Par ressemblance grammaticale
- Langues segmentées et non segmentées (ex : ré-élec-tion)

L'intérêt de cette classification est de mieux comprendre les activités langagières. La linguistique permet d'expliquer la langue contrairement à la grammaire qui ne donne que les règles à appliquer.

### Combien de langues sont parlées dans le monde ?

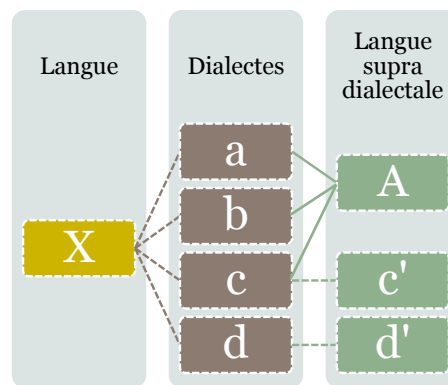
↳ Entre 5000 et 6528 langues et plus de 20000 si l'on compte les dialectes.

*Dialecte* : variation d'une langue soit sur une base régionale soit sociale (la langue des forgerons en Afrique par exemple).

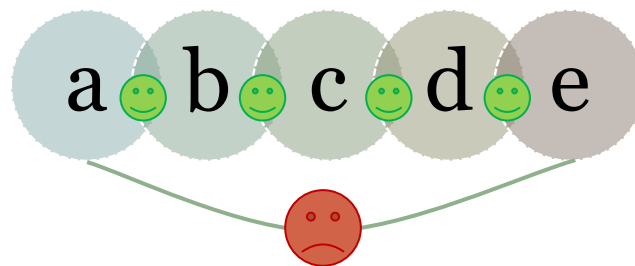


En cas de non interactions entre les dialectes, et pour des raisons sociales, régionales, d'isolement, de mode de vie...de nouvelles langues apparaissent.

Une langue peut aussi devenir supra dialectale (comme le Gallo-romain) :



Pour décider de l'apparition d'une nouvelle langue, on doit fixer des critères de compréhension entre les formes utilisées. Cela doit prendre en compte les formes orales et écrites (chinois par exemple). Mais il est très dur scientifiquement de se défaire de la prise en compte de facteurs sociaux. Ainsi, l'Hindoustani comprend à la fois l'Hindi et l'Ourdou pour certains linguistes alors que pour d'autres, ce sont deux langues différentes. Exemple de compréhension entre zones géographiques proches avec changement dialectal progressif :



*Isolat* : famille avec une seule langue (ex : Kète, f. iénissénennes ; Basque, rattaché a aucune famille connue).

*Langue vernaculaire* : langue locale communément parlée au sein d'une communauté.

*Langue véhiculaire* : langue servant de moyen de communication entre populations de langues différentes.

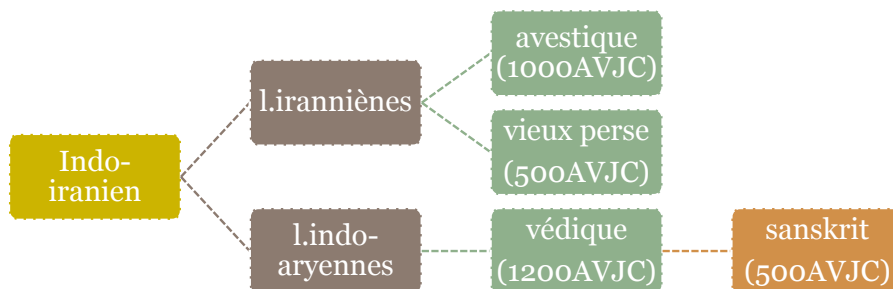


## Arbre généalogique du proto-indo-européen



Les différentes familles de langue reconnues à la fin du 19<sup>ème</sup> siècle :

### Indo-iranien (satem)



Indo-iranien provient des peuples parlant l'indo-européen partis vers l'est. Ce sont des langues documentées à date très anciennes :

- L'Avesta, texte religieux mazdéen de Zoroastre (1000AVJC)
- Les Védas, recueils de chants sacrés (1200AVJC)

Ces deux textes sont linguistiquement très proches et permettent à ces langues d'être classées de la même sorte.

En 500AVJC, le védique est codifié par le grammairien Panini et devient la « langue élaborée », le sanskrit. Le védique se fige et devient une grande langue culturelle en inde.

*Aryen* : signifie iranien en vieil iranien.

*Aria* : castes supérieures en vieil indien.

*Langues indo-aryennes* : langues européennes parlées en inde.

La langue romani n'a pas d'état ; elle appartient à un peuple ayant quitté l'inde au 11<sup>ème</sup> siècle qui s'est dirigé vers l'Europe en gardant sa langue.

### **Arménien (satem)**

L'arménien provient du Caucase. C'est une langue de diaspora qui a aboutit à deux normes :

- Arménien oriental en Arménie (ex : Tigran)
- Arménien occidental de la diaspora (ex : Dikran)

Les consonnes rondes et sonores sont inversées dans les deux langues. L'arménien a très longtemps été confondu avec les langues iraniennes.

### **Balto-slaves (satem)**

Les langues balto-slaves n'ont pas la même histoire que les langues indo-iraniennes. On n'a pas d'écrits balto-slaves. Vers 3000AVJC, les langues se sont séparées avant de se réunifier plus tardivement.

### **Baltes**

Les langues baltes comprennent :

- Le vieux prussien qui est une langue morte
- Le lituanien et le laiton qui, à l'écart des courants de migration, sont restées des langues très conservatrices, proches de l'indo-européen. Leurs premiers écrits datent du 16<sup>ème</sup> siècle.

### **Slaves**

Toujours conservatrices les langues slaves le sont tout de même moins que les langues baltes. Aujourd'hui leur utilisation est très répandue. Lors de l'expansion slave de 700APJC, il n'y avait encore qu'une seule protolangue slave. Les langues slaves comprennent :

- Slave oriental : russe, ukrainien et biélorusse
- Slave occidental : polonais, tchèque, slovaque (kachoube, sorabe, polabe)
- Slave méridional : serbe, croate, slovène, bulgare, macédonien

## **Albanais (satem)**

Parlé actuellement en Albanie, au Kosovo, en république de macédoine et au nord de la Grèce. On n'en sait pas grand-chose, les premiers textes datant du 16<sup>ème</sup> siècle. 90% du vocabulaire est emprunté à d'autres langues.

## **Grec (centum)**

C'est une langue attestée depuis très longtemps (1400AVJC). A la base, il existait différents dialectes puis, l'unification politique d'Athènes unifie les dialectes qui deviennent la langue supra dialectale Koinè (langue commune). Koinè est même devenue un nom commun.

*Koinèisation* : formation d'une langue commune à partir de dialectes apparentés.

*Standardisation* : vient d'en haut.

## **Germanique (centum)**

Le germanique ne comprend pas que l'allemand puisque l'anglais en fait aussi parti. Il recouvre les langues des peuples qui habitaient le sud de la Scandinavie et les côtes baltes et ont émigrés vers le sud et l'est en remplaçant les celtes (wisigoth, ostrogoth...). Les langues germaniques comprennent :

- Germanique oriental : le gothique est aujourd'hui une langue morte mais on la connaît grâce à la traduction d'une bible au 4<sup>ème</sup> siècle. C'était la langue des germains de l'est près de la mer noire.
- Germanique nordique (l.scandinaves) : suédois, danois, islandais (irlandais conservateur par isolation)
- Germanique occidental : Anglais (500APJC), Allemand haut (montagnes du sud), Allemand bas (plaines du nord)

## **Celtique (centum)**

Les langues celtiques sont les seules langues indo-européennes dont tous les locuteurs sont bilingues.

## **Italique (centum)**

Ce sont des langues peu connues à cause de l'extension de Rome qui a imposé le latin. Après sa chute, les langues se diversifièrent (français, portugais, espagnol, italien, roumain).

## **Tokharien**

C'était une langue parlée dans le Turkestan chinois vers 800APJC. On a retrouvé des textes bilingues tokharien-sanskrit qui ont permis de la comprendre. Elle était à l'est alors que c'était une langue centum ressemblant au celtique.

## Hittite

Langue parlée en Anatolie. On a retrouvé des tablettes en écriture cunéiforme traduites en acadien. Le hittite s'est détaché du tronc indo-européen très rapidement.

## Histoire de la classification des langues

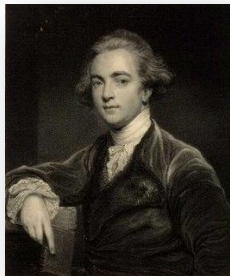
C'est à la fin du 18<sup>ème</sup> siècle que l'on commence prendre conscience que les langues sont différentes et c'est peu après, début 19<sup>ème</sup>, que l'on commence à les classer. Mais l'intérêt pour les langues est bien plus ancien : au 14<sup>ème</sup>, on s'interrogeait sur les langues indiennes (Amérique), pendant la réforme on s'intéresse à la théorie de l'hébreu langue originelle et la normalisation de l'imprimerie relance les débats...

**Etymologie populaire** : activité que tous les sujets parlant ont.

ex : *Insula britannica* viendrait du mot hébreu *baratanac*.

« Avant de chercher l'origine d'un mot dans une langue étrangère, il faut l'avoir décomposé, l'avoir dépouillé de toutes ses inflexions grammaticales, & réduit à ses éléments les plus simples. »

**Turgot**, article sur l'étymologie dans l'*Encyclopédie*



Sir **William Jones** (28 septembre 1746 - 27 avril 1794) est un linguiste et chercheur sur l'Inde ancienne, célèbre en particulier pour sa redécouverte de la famille des langues indo-européennes (le néerlandais Marcus Zuerius van Boxhorn avait déjà fait la découverte que l'avestique, l'ancien perse, et le sanskrit étaient apparentés aux langues européennes dans un groupe qu'il appelait « scythique », et cela dès le XVI<sup>e</sup> siècle, mais son travail n'a pas connu de postérité).

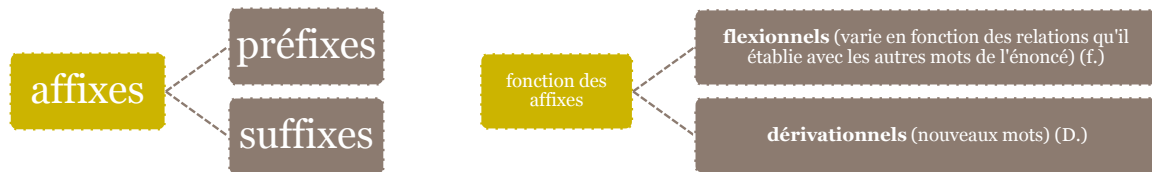
En 1786, il prononce un discours sur la langue sanskrite qui met en évidence l'origine commune des langues indo-européennes. Deux ans avant, il avait écrit un essai sur les polythéismes grecs, italiens et indien et en avait conclu qu'il existait un lien entre eux.

Par la suite les études s'enchaînent sur des thèmes différents en **grammaire comparative**:

- 1816 : **Franz Bopp** écrit son « Système de **conjugaison** de la langue sanskrite comparé à celui des langues grecques, latines, persanes et germaniques ».
- 1818 : **Rasmus Rash** publie ses « Recherche sur les **origines** de l'ancienne langue noroise et islandaise ».
- 1819 : **Jakob Grim** publie une « **Grammaire** allemande » qui étudie l'évolution des différents dialectes allemands.

## Bases de l'analyse

La structure nominale des langues indo-européennes est très riche. Ainsi pour comparer les langues, on a utilisé les **affixes flexionnels** car c'est un système clos de 6 éléments seulement (les personnes) qui est difficilement empruntable à une autre langue (contrairement aux racines). Ceci est à la base une idée de Panini.



(ré/organ/is/er/ons)

(D./radical/D./f./f.)  
(.../futur/1ère pers. Plur.)

**Morphème** : Plus petite unité porteuse de sens (ou monème, ici entre parenthèse (̄))

C'est ce qu'à étudié Bopp dans son livre sur le sanskrit :

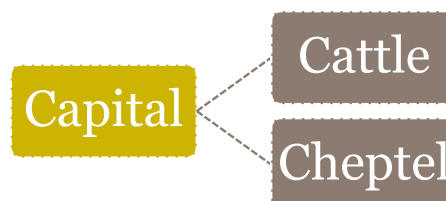
- 1<sup>ère</sup> pers. : *ami* → *mâm* (moi)
- 2<sup>ème</sup> pers. : *ati* → *tam* (toi)

De son côté, Grimm a montré que les consonnes germaniques différaient toutes de la même façon des langues indo-européennes et qu'il s'agissait donc de la même famille.

Par la suite, on a commencé à se demander si on pouvait recréer la forme mère. Pour cela on étudie chaque consonne :

- Père
- \*Pater → forme hypothétique du début
- \*PHter → forme hypothétique actuelle  
(\* signifie que le mot est hypothétique)

Ainsi, les mots changent de forme mais aussi de sens. A l'époque, le père était plutôt le protecteur du village que le père génétique. Un autre exemple :



En 1861, **Auguste Schleicher** (19 février 1821 – 6 décembre 1868) publie un compendium de la grammaire comparée des langues Indo-Européennes. Botaniste de formation, il est très influencé par les idées de Darwin. Pour lui les langues fonctionnent comme les espèces : une mutation est héritée par toute la descendance. C'est la « Stammbaum theory ».

Une autre idée issue du Darwinisme est celle des étapes d'évolution des langues :

- **Isolantes** : mots pas analysables (on ne peut pas découper les mots)
- **Agglutinantes** : radical auquel on ajoute des affixes sans règle particulière
- **Flexionnelles** : langues à paradigme (perfection de l'esprit selon lui)

**Paradigme** : ensemble des formes d'un mot en fonction du rôle qu'il a dans la phrase (déclinaisons, conjugaisons).

Le problème de la reconstruction d'une langue mère c'est que les différentes reconstitutions ne sont pas datées. Les formes elles-mêmes ne sont pas forcément homogènes et tous les sons d'un mot reconstruit n'appartiennent pas forcément à la même époque :



**Sonnantes** : consonnes intermédiaires entre des consonnes et des voyelles.

**Saussure** a défini trois coefficients sonantiques en rapport avec l'évolution des langues (H<sub>1</sub>, H<sub>2</sub>, H<sub>3</sub>).

A la fin du 19<sup>ème</sup> siècle, les Néogrammairiens (« Jung grammatisher ») considèrent qu'il ne doit pas y avoir d'exceptions et que tout doit être expliqué. Les plus connus sont :

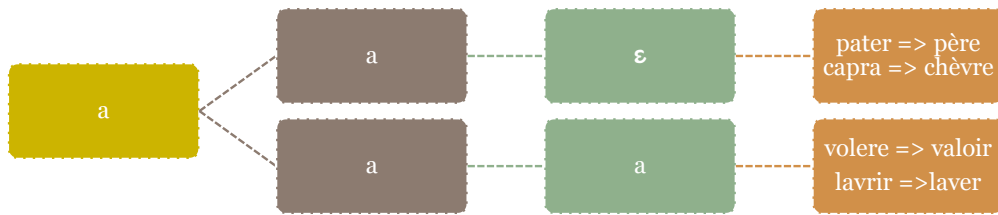
- Brugman
- Paul
- Leskien
- Verner

Ils utilisent deux types de reconstruction :

- Interne : raisonnement à l'intérieur même de la langue
- Externe : comparaison avec d'autres langues apparentées

Ils analysent les exceptions de la loi phonétique absolue comme des interférences avec d'autres lois. Cela donne lieu à deux explications :

- Par le contexte : loi de Verner sur l'accent tonique par exemple
- Par analogie : lorsque l'évolution phonétique rend les formes adhérentes pour le système, le locuteur récrée les racines pour unifier la forme.



*Isoglosse (une)* : ligne imaginaire séparant deux aires dialectales.

L'isoglosse des langues anciennes (ici Taokas et Favorlang) entoure beaucoup de nouvelles formes alors que les dernières formes créées (ici amis, puyuma et paiwan) sont entourées par la dernière et plus petite isoglosse.

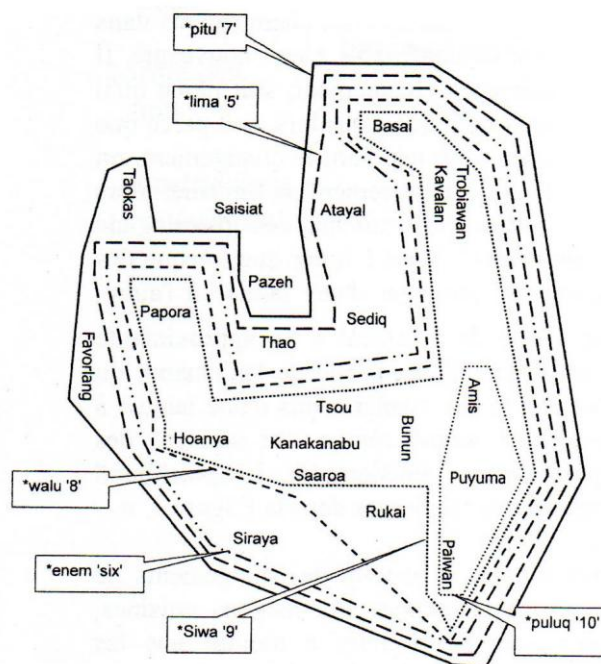


Figure 1: isoglosses des nombres de 5 à 10 dans les langues austronésiennes de Taiwan

## Index des définitions

<b>A</b>		<b>K</b>	
Aria.....	8	Koinèisation.....	9
Aryen.....	8		
<b>C</b>		<b>L</b>	
Classification aréale.....	3	Langue véhiculaire.....	5
Classification génétique.....	3	Langue vernaculaire.....	5
Classification typologique.....	4	Langues agglutinantes.....	4, 12
Cognats.....	3	Langues flexionnelles.....	4, 12
		Langues indo-aryennes.....	8
		Langues isolantes.....	4, 12
<b>D</b>		<b>M</b>	
Dialecte.....	4	Morphème.....	11
<b>E</b>		<b>P</b>	
Etymologie populaire.....	10	Paradigme.....	12
		Protolangue.....	3
<b>I</b>		<b>S</b>	
Isoglosse.....	13	Sonnantes.....	12
Isolat.....	5	Standardisation.....	9