

Introduction: universaux linguistiques et typologie

Introduction à la typologie linguistique

David Blunier · Université de Poitiers L1 · Printemps 2025



Références

Sauf mention explicite, les contenus présentés ici sont tirés des références suivantes:

- Comrie, Bernard (1989). *Language universals and linguistic typology: syntax and morphology*. University of Chicago Press.
- Velupillai, Viveka (2012). *An introduction to linguistic typology*. John Benjamins.

Hors des données linguistiques et en l'absence de contraintes syntaxiques (accord), les formes du féminin et du masculin utilisées dans ce document sont distribuées au hasard.

Qu'est-ce que la typologie?

- La typologie est l'étude empirique des différentes propriétés des langues humaines.
- Elle se distingue à la fois de la **linguistique théorique** et de la **linguistique expérimentale (ou psycholinguistique)** par ses méthodes, mais son objectif est le même: **l'étude des universaux linguistiques**.

Qu'est-ce que la typologie?

- La typologie est essentiellement empirique et repose sur l'analyse systématique de différents systèmes linguistiques.
- Comme en biologie, la typologie repose sur la **classification** de ces systèmes en unités cohérentes.
- La base de ces classifications s'effectue sur ce qui constitue les propriétés d'un langage en tant que **structure linguistique**, indépendante des facteurs extra-linguistiques.

Genres et familles

- Comme en biologie, les langues humaines sont répertoriées en *genres (*genus, plur. genera*) et **familles**.
- Les **familles** représentent des groupes de langues possédant le même ancêtre commun.
- Les **genres** sont des sous-familles de langues, regroupées en fonction de propriétés communes directement observables sans passer par une étude diachronique (Dryer 1989).

La classification typologique complète du français est la suivante: *Indo-européen > italique > latino-falisque > latinique > romane > italo-occidentale > occidentale > gallo-ibérique > gallo-romane > français.*

Genres et familles

- Le français appartient donc à la famille de l'**indo-européen** et au genre **roman**.
- Certaines langues représentent à elles-seules leur famille ou leur genre, comme le coréen, le basque ou l'ainu (langue parlée au Japon, sur l'île d'Hokkaido); on dit de ces langues que ce sont des **isolats**.

Genres et familles

Table 2.1 The ten largest language families, ordered by size.

FAMILY	NO. OF LANGUAGES	APPROXIMATE AREA
Niger-Congo	over 1500	entire sub-Saharan Africa
Austronesian	over 1200	from Southeast Asia to Hawai'i; from Taiwan to New Zealand
Sino-Tibetan	ca 450	from the Himalayas to eastern China; from northern China to Thailand
Indo-European	close to 450	from Western Europe to Bangladesh; from Scandinavia to the Maldives
Trans-New Guinea	close to 400	Papua New Guinea and Indonesia
Afro-Asiatic	close to 400	Northern and Eastern Africa; the Middle East
Australian	over 250	Australia
Nilo-Saharan	ca 200	Central and Eastern Africa
Oto-Manguean	over 150	Mexico
Austro-Asiatic	over 150	from Eastern India to Vietnam; from Bhutan to Malaysia

Pidgins, créoles et autres "langues de contact"

- Un **pidgin** est un langage qui émerge lorsque deux groupes entrent en contact et doivent communiquer, mais ne possèdent pas de langage commun.
- Une de ces situations, typiquement, est la colonisation (triste histoire).
- Dans cette sorte de configuration, un groupe se met à emprunter des éléments du lexique d'un autre groupe (souvent les plus forts) et à l'incorporer dans sa propre langue.
- Le langage auquel on emprunte des éléments est appelé le **superstrat**, et la langue incorporant ces éléments le **substrat**.
- Un pidgin est donc une langue simplifiée qui n'est innée pour aucun des locuteurs.

Pidgins, créoles et autres "langues de contact"

- Un **créole** est un pidgin ayant évolué et s'étant imposé comme la langue native d'une communauté.
- La créolisation peut être rapide (une génération suffit).
- Contrairement au pidgin, le créole est une langue à part entière et est acquise par les enfants lors de la période d'acquisition (0-4 ans).

Langues des signes

- Langues à part entière, les LS sont encore aujourd'hui largement sous-étudiées par rapport aux langues orales (50 ans de recherche vs 300). Les données manquent pour établir des classifications stables.
- Cependant, Zeshan (2011) propose la classification suivante:

Langues des signes

BSL – Auslan – New Zealand Sign Language (NZSL: New Zealand)

Japanese Sign Language (NS: Japan) – Taiwanese Sign Language (TZS: Taiwan) –
South Korean Sign Language (SKSL: South Korea)

French Sign Language (LSF: France) – ASL – Russian Sign Language (RZY: Russia) –
Dutch Sign Language (NGT: the Netherlands) – Flemish Sign Language (VGT:
Belgium) – Quebec Sign Language (LSQ: Canada) – Irish Sign Language (TCE:
Ireland) – Brazilian Sign Language (LSB: Brazil)

DGS – (perhaps also other sign languages in Europe and the Middle East) – Israeli
Sign Language (ISL: Israel)

Swedish Sign Language (ST: Sweden) – Finnish Sign Language (SV: Finland)

ASL – Ugandan Sign Language (USL: Uganda) – Thai Sign Language (TSL: Thailand) –
Kenyan Sign Language (KSL: Kenya)

Italian Sign Language (LIS: Italy) – Argentine Sign Language (LSA: Argentina)

Chinese Sign Language (CSL: China) – Hong Kong Sign Language (HKSL: China)

Langues des signes

- Certaines de ces langues naissent d'un contact avec une langue pré-existante, souvent dans le cadre d'une institution (comme c'est le cas pour l'ASL);
- Souvent, les LS se "créolisent" en incorporant des éléments d'un superstrat sur un système de signes local pré-existant (*home signs systems).
- D'autres langages émergent spontanément, comme le Adamorobe Sign Language du Ghana ou la langue des signes du Nicaragua.

Il est important de noter que les langues des signes utilisées sur un territoire sont largement indépendantes des langues orales parlées sur ce même territoire. Par exemple, la langue des signes allemande (DGS, *Deutsche GebärdenSprache*) est une langue SOV (sujet-objet-verbe), alors que l'allemand est une langue SVO.

Langues des signes

- Pour aller plus loin, vous pouvez voir mon [cours d'introduction à la linguistique des langues des signes](#).

Universaux linguistiques

- Le but de la typologie est d'établir des **universaux linguistiques**, i.e. des propriétés communes à toutes les langues connues (ou à un groupe statistiquement significatif de celles-ci).
- Certains universaux sont **absolus**:
 - a. Toutes les langues (orales) ont des voyelles.
 - b. Toutes les langues peuvent référer à des entités.
 - c. Toutes les langues peuvent former des questions.

Universaux linguistiques

- D'autres sont **statistiques**, i.e. indiquent une probabilité élevée:
 - a. La plupart des langues utilisent la nasale /n/.
 - b. Dans la plupart des langues, le singulier est la forme de base et le pluriel est marqué.
 - c. La plupart des langues emploient une intonation ascendante pour former des questions oui-non.

Universaux linguistiques

- D'autres universaux sont **implicationnels**:

$X \rightarrow Y$

Si une propriété Y est observée dans une langue L , alors (très probablement) L possède également X .

- Exemples:

a. Si une langue possède le phonème /t/, alors elle possède également le phonème /k/.

b. Si une langue possède des réflexifs de première et seconde personne, alors elle possède aussi des réflexifs de troisième personne.

Universaux linguistiques

- Ces universaux peuvent être **bidirectionnels**:

$$X \leftrightarrow Y$$

Si une propriété Y est observée dans une langue L , alors (très probablement) L possède également X ; si X est observée, alors (très probablement) L possède également Y .

Si une langue L est d'ordre OV, alors L (très probablement) emploie des postpositions; si une langue emploie des pos-positions, alors L est (très probablement) OV.

Les méthodes de la typologie

- La linguistique contemporaine se fonde sur différents types de données, parmi lesquelles on peut trouver:
 - Des données de première main, provenant d'élicitations obtenues avec des locutrices natives;
 - Des grammaires descriptives, regroupant des données obtenues par des linguistes, le plus souvent en situation de terrain;
 - Des questionnaires;
 - Des textes;
 - Des résultats d'expériences;
 - Des bases de données.

Un exemple de base de données: **WALS** (*World Atlas of Language Structures*)

- Débuté en 1999, **WALS** est le résultat d'un effort collaboratif entre différents spécialistes en sciences du langage.
- WALS permet de consulter et de géolocaliser les propriétés de plus de 2678 langues.