

**Champs linguistiques**

**Collection dirigée par**

Marc Wilmet (Université Libre de Bruxelles)  
et Ludo Melis (Katholieke Universiteit Leuven)

**NICOLE DELBECQUE (Éd.)**

# **Linguistique cognitive**

**Comprendre comment  
fonctionne le langage**

*Nouvelle édition augmentée, avec exercices et solutions*

**Préface de Jean-Rémi Lapaire**

C h a m p s   l i n g u i s t i q u e s



Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : [www.deboeck.com](http://www.deboeck.com)

© De Boeck & Larcier s.a., 2006  
De Boeck Université – Duculot  
Rue des Minimes 39, B-1000 Bruxelles

2<sup>e</sup> édition

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

Imprimé en Belgique

Dépôt légal :  
Bibliothèque nationale, Paris : septembre 2006  
Bibliothèque royale de Belgique : 2006/0035/009

ISSN 1374-089X  
ISBN 2-8011-1391-3  
ISBN13 : 978-2-8011-1391-2

## PRÉFACE

La rédaction d'un manuel de linguistique générale est à la fois un exercice périlleux et une épreuve de vérité. Il faut être capable de recenser les termes et notions incontournables de la linguistique, alors que les écoles sont nombreuses et les chercheurs divisés. Il faut ensuite fournir des résumés et inventer des exercices qui facilitent la mémorisation des concepts et la mise en œuvre des méthodes. Il faut enfin aborder avec fraîcheur et simplicité les questions complexes que se posent les spécialistes du langage : nature du signe, lien entre son et sens, relation entre langage et pensée, articulation entre langue, discours, individu et société. Et lorsque l'ouvrage se définit comme l'adaptation d'un livre original, composé en anglais, à plusieurs mains, *Cognitive Explorations of Language and Linguistics*, il faut de surcroît traduire et ajuster, en tenant compte des traditions descriptives de la francophonie. Tout cela Nicole Delbecq l'a non seulement osé, mais magnifiquement réussi, en s'appuyant sur une solide équipe de traducteurs et en veillant à ce que les diverses versions (française, espagnole, italienne, allemande, néerlandaise, grecque) restent suffisamment proches pour être utilisées en parallèle dans les départements de langues des universités européennes.

Le sommaire, à lui seul, établit la vaste zone de couverture du livre. Fidèle à Saussure, pour qui la linguistique était fondamentalement une branche de la sémiologie – cette « *science qui étudie la vie des signes au sein de la vie sociale* » – le livre aborde en premier la nature du signe linguistique et la fonction symbolique, ainsi que le rapport délicat entre catégorisation linguistique et catégorisation conceptuelle (Chapitre 1). La « base cognitive du langage » étant posée, la présentation des grands domaines de l'analyse linguistique peut débiter : **sémantique lexicale** (Chapitre 2), **morphologie** (Chapitre 3), **syntaxe** (Chapitre 4), **phonologie** (Chapitre 5), **sémantique générale** (Chapitre 6),



## Chapitre 10

### LA CLASSIFICATION ET L'ÉTUDE COMPARÉE DES LANGUES

Dans le chapitre 6 portant sur la sémantique interculturelle, nous avons étudié certaines similitudes et divergences dans les lexiques, grammaires et scripts culturels de diverses communautés linguistiques. Dans le présent chapitre, nous examinerons les classifications qui peuvent être faites des langues selon différents types de critères : le point de vue adopté peut être sociologique, génétique, typologique ou contrastif. Dans un premier temps, les langues peuvent être classées sur base de critères sociologiques externes à celles-ci, tels que leur statut et importance dans le monde. Ce genre d'étude comparée va de pair avec la question de savoir quelles sont l'origine et l'évolution des langues. D'autre part, il est également possible de les comparer à partir de caractéristiques structurelles internes aux langues. Bon nombre de langues sont génétiquement apparentées les unes aux autres ; elles forment des familles de langues. Néanmoins, des langues sans lien génétique peuvent également être regroupées à partir de certaines caractéristiques structurelles. La recherche typologique utilise comme critères les caractéristiques structurelles saillantes d'une langue – l'ordre des mots par exemple – dans le but de procéder à des regroupements typologiques entre les langues. À la question de savoir s'il existe des éléments universaux – tels les “primitifs sémantiques universaux” du chapitre 6 –, la réponse est positive, sur base de la théorie du prototype. Il est également possible de comparer de façon détaillée les caractéristiques structurelles ou “champs conceptuels” de deux langues. Ceci est le domaine de la linguistique comparée. Plus orientée vers la pratique, elle met l'analyse de la structure et du fonctionnement des langues au service de l'apprentissage des langues, de la traduction et de la confection de dictionnaires bilingues.

## 10.1 L'identification et le statut des langues

L'identification des langues se heurte à plusieurs difficultés (10.1.1). Du point de vue linguistique, le passage d'une langue à l'autre s'avère être graduel tant sur l'axe géographique que sur l'axe historique, et la délimitation de leur extension fluctue en fonction des critères utilisés (10.1.2). Du point de vue sociologique, le statut des langues se définit en termes politiques, institutionnels et démographiques (10.1.3).

### 10.1.1 Inventaire et délimitation des langues

À l'heure actuelle, il n'est toujours pas possible de fixer le nombre exact des langues parlées dans le monde. Certaines sources parlent de 5 000 langues, d'autres avancent le chiffre de 6 000. Les estimations variant à ce point, on est en droit de se demander pourquoi les linguistes ne parviennent pas à dresser l'inventaire des langues. Il semble qu'il y ait plusieurs raisons à cela.

Tout d'abord, nous manquons encore d'information quant aux langues parlées dans certaines régions du monde. En effet, il s'avère que certains territoires africains et australiens n'ont été que très peu explorés à cet effet. Remarquons entre parenthèses qu'une recherche sur le terrain requiert beaucoup de temps, de moyens financiers et de savoir-faire. De récents sondages font apparaître que dans certaines de ces régions, il reste un très grand nombre de langues à répertorier. Ainsi, Comrie (1987) annonce qu'à cet égard, et contre toute attente, la Nouvelle-Guinée s'avère être extrêmement digne d'intérêt : pas moins d'un cinquième de l'ensemble des langues parlées dans le monde y seraient localisées. Il se peut même que cette proportion soit à revoir à la hausse, dans la mesure où une partie des langues de cette île n'a pas encore pu être identifiée. Cette dernière observation vaut également pour bon nombre de langues parlées en Australie et en Afrique.

Une deuxième raison de l'imprécision est qu'il subsiste toujours un doute quant à la délimitation des langues : pour deux variétés linguistiques se trouvant dans le prolongement ou dans la proximité l'une de l'autre, il n'est souvent pas aisé de déterminer s'il s'agit de deux langues différentes ou seulement de dialectes appartenant à une seule langue. Même en Europe, où ce genre d'incertitude a pour ainsi dire disparu, le statut accordé à une variété linguistique – langue autonome ou dialecte – est autant le résultat de considérations politiques que de classifications linguistiques.

### 10.1.2 Critères linguistiques pour la reconnaissance des langues

Le critère le plus fréquemment utilisé pour reconnaître une langue est celui de **l'intelligibilité réciproque**. Lorsque deux interlocuteurs se comprennent mutuellement, on en conclura qu'ils parlent des dialectes de la même langue. Dans le cas contraire, s'ils ne se comprennent pas, on considérera qu'ils parlent chacun une langue différente. La réalité est cependant plus complexe et parfois paradoxale, même dans des régions d'Europe qui nous sont familières. La raison en est que la reconnaissance officielle d'une langue par l'État a pour corollaire l'établissement de frontières linguistiques officielles, fixées constitutionnellement. Dès lors, il est fait abstraction de certaines réalités du terrain. Dans des territoires étendus, tels l'Allemagne et l'Italie, les dialectes parlés par les habitants du nord sont pratiquement incompréhensibles pour les habitants du sud du pays, et vice versa. Il est bien connu que les Italiens de la région des Alpes ont besoin de sous-titres pour comprendre les dialogues de films mettant en scène la mafia sicilienne. L'intercompréhension n'est même pas garantie à l'intérieur d'une région comme la Flandre : les Ouest-Flamands ont beaucoup de mal à comprendre les Limbourgeois, et vice versa. En revanche, ces derniers comprennent mieux l'allemand que l'ostendais ou le brugeois. En effet, dans la zone frontalière entre l'Allemagne d'une part, et la Flandre et les Pays-Bas d'autre part, les habitants peuvent facilement se comprendre entre eux, bien qu'ils aient pour langue officielle respectivement l'allemand et le néerlandais. De leur côté, les langues scandinaves présentent elles aussi un indice non négligeable de compréhension mutuelle. Il semble donc que la compréhension mutuelle concerne généralement des dialectes très proches les uns des autres, qu'ils appartiennent ou non à la même langue officielle.

Il existe un autre problème concernant l'intelligibilité. La compréhension d'une langue ou d'un dialecte n'est pas une question de tout ou rien. Le degré de compréhension varie en fonction de facteurs tels la familiarité, la régularité du contact et la volonté de comprendre. Il arrive aussi que seul un des deux locuteurs comprenne la langue de l'autre.

Le problème de la compréhension mutuelle et celui des frontières linguistiques et dialectales se pose en termes de **continuum dialectal**. Ce concept permet de rendre compte de la compréhension mutuelle qui existe entre dialectes voisins ressortissant à des langues officielles différentes au sein de la même famille linguistique. D'autre part, le concept de continuum dialectal éclaire également le fait qu'il ne suffit pas que deux dialectes appartiennent à la même langue pour que la communication soit assurée. Plus la distance entre les dialectes est grande, plus la compréhension mutuelle sera problématique. Il n'en est pas

moins vrai que tous les dialectes d'un groupe ou famille de langues forment ce qu'on appelle un continuum. Comme le montre le tableau 1 ci-dessous, le continuum dialectal germanique va du Tyrol à la mer du Nord, du côté ouest, et à la mer Baltique, du côté nord.

Tableau 10.1 Quelques réalisations d'énoncés équivalents à *Comment allez-vous ?*

|                   | (sub)standard écrit            | réalisation phonétique |
|-------------------|--------------------------------|------------------------|
| bavarois          | <i>wia geht's da jetzat ?</i>  | via gerts da jetsat    |
| allemand standard | <i>wie geht's dir jetzt ?</i>  | vi: gerts die jetst    |
| bas-allemand      | <i>wo geit di dat nu ?</i>     | vo: gait di dat nu:    |
| néerlandais       | <i>hoe gaat het met u ?</i>    | hu: xɑ:t hət met y     |
| danois            | <i>hvordan har du det nu ?</i> | voɐdan ha: du de: nu:  |
| norvégien         | <i>hvordan har du det no ?</i> | vurdan har dy de: no:  |

À côté du continuum géographique, il existe également un **continuum historique**. À travers toutes les transformations que peut subir une langue, au point de donner naissance à d'autres langues, nombre de caractéristiques communes peuvent être préservées. Alors que certaines langues disparaissent de la carte – et deviennent des “langues mortes” –, de nouvelles langues surgissent à différents stades de l'histoire. Ce fut le cas du latin : les nombreuses variétés du latin (populaire) se propagèrent dans toute la zone de l'Europe allant de la Roumanie à l'Italie, et de l'Espagne jusqu'au nord de la France. On y assista à l'éclosion et au développement de langues indépendantes : le roumain, le sarde, l'italien et le rhéto-roman dans l'axe est-ouest, et l'espagnol, le portugais, le catalan, le provençal et le français, dans l'axe nord-sud.

Conventionnellement, on considère qu'une langue est morte lorsqu'elle n'est plus du tout parlée au quotidien. Toutefois, la mort d'une langue n'a pas lieu du jour au lendemain. Elle ne coïncide pas avec le décès du dernier locuteur de cette langue. La plupart du temps, une communauté linguistique connaît une période de transition au cours de laquelle ses membres abandonnent progressivement leur ancienne langue pour en utiliser une autre. Ainsi, il est possible que durant cette période transitoire l'ancienne langue reste encore vivante chez les aînés et qu'une partie de la communauté en garde la compétence (latente), alors même que la nouvelle langue est déjà communément employée.

Une autre question fondamentale se pose ici : comment peut-on constater la genèse d'une nouvelle langue si celle-ci ne se développe que progressivement à partir d'une variété d'une langue existante ? Là-dessus, les positions adoptées ne sont pas toujours consistantes. En effet, pour les langues romanes, nous avons l'habitude de dire qu'elles sont issues du latin. L'histoire retrace l'émergence de plusieurs langues différentes à partir de ce qu'il est convenu de considérer comme une seule langue. Or, pour le grec, qui a connu une tradition beaucoup plus longue que le latin, nous continuons à ne parler que d'une seule langue hellénique, nous contentant de distinguer entre le grec classique et le grec moderne, appelé aussi plus techniquement *dhemotikè* ‘langue du peuple’. Cette différence de traitement se justifie dans la mesure où, contrairement au latin, le grec n'a pas subi l'influence de différents substrats (voir chapitre 9.2) et que jusqu'à ce jour l'on hésite entre deux normes, la *katharévoussa*, “l'épurée”, plus proche de la forme classique, et la *démotique*, ou populaire, plus proche de la réalité parlée.

Comme nous pouvons le constater, il est très difficile – même pour les linguistes – d'identifier et de dénombrer avec exactitude les langues actuellement parlées dans le monde. Il reste à collecter des données supplémentaires et à affiner les critères pour arriver à un inventaire sur lequel tout le monde puisse être d'accord. Nous verrons dans la suite que la définition de ce qu'est une langue – et la classification des langues qui en résulte – est étroitement liée aux résultats et aux progrès de la sociolinguistique et des recherches en diachronie (concernant l'évolution des faits linguistiques dans le temps).

### 10.1.3 Le statut politique et international d'une langue

Au 16<sup>e</sup> siècle, une nouvelle conception de l'État se fit jour sous le règne de monarques puissants tels que Henri VIII, François I, Charles Quint ou encore Philippe II. Dans le façonnement de ce nouveau concept de l'État, la langue et la religion étaient appelées à jouer un rôle de tout premier ordre. Le slogan “un royaume, une langue, une religion” en reflète bien l'idée centrale.

Vu l'enjeu public et politique, il est clair que la décision de savoir quelle langue sera reconnue comme langue nationale et officielle dépend avant tout des autorités politiques, assistées ou non d'experts en linguistique. La définition linguistique d'une langue est une chose, la définition sociologique et politique en est généralement une autre. La situation du serbo-croate est révélatrice à cet égard. Bien que l'on parle cette langue tant en Serbie qu'en Croatie, l'alphabet usité est différent : en Serbie l'on se sert du cyrillique, en Croatie de l'alphabet latin. D'un point de vue linguistique il s'agit d'une seule langue, mais politiquement parlant, l'on en compte deux.



Un pays peut avoir intérêt à ne reconnaître qu'une seule langue officielle. Cela a été le fondement de la politique linguistique française, déjà du temps des colonies. D'autres pays que la France ont préféré officialiser plusieurs langues : ainsi, la Grande-Bretagne reconnaît l'anglais et le gallois, l'Espagne l'espagnol, le catalan, la galicien et le basque, la Belgique le français, le néerlandais et l'allemand, la Suisse le français, l'allemand, l'italien et le rhéto-roman. Il arrive même qu'un pays accorde le statut de langue officielle à un parler que les pays étrangers considèrent comme étant un dialecte. Dans le contexte européen, c'est le cas du luxembourgeois – langue parlée au Luxembourg – que certains linguistes définissent comme étant un dialecte de l'allemand. Néanmoins, contrairement aux dialectes allemands, le luxembourgeois possède une riche tradition littéraire. Qui plus est, ce parler est largement utilisé dans les médias tels les journaux, la radio et la télévision. Par conséquent, on ne peut pas assimiler son statut à celui des autres dialectes de l'allemand. Quoiqu'il en soit, promouvoir cette langue au rang de troisième langue officielle du Luxembourg – l'inscrivant ainsi parmi les langues officielles de l'Union Européenne – reste une décision politique.

En Malaisie, on utilise plusieurs variétés dialectales dérivées du malais bazaar pour communiquer dans la vie de tous les jours, mais le premier rang est réservé à la langue officielle qu'est le malais standard. En Indonésie, bien avant l'indépendance, les politiciens ont écarté l'idée d'adopter comme langue officielle une des langues nationales. Au lieu de favoriser le javanais, qui est utilisé par plus de 70 millions de personnes, ils optèrent pour une variété standard du malais, adaptée à l'Indonésie. Actuellement, elle est connue sous le nom d'indonésien bahasa ou malais indonésien.

Le statut de langue officielle n'est en aucun cas synonyme de prééminence typologique ou statistique. Or, comme l'octroi de ce statut en implique l'emploi systématique dans la fonction publique, il favorise à la longue l'instauration de règles et l'enrichissement lexical dans les domaines relevant de l'**usage institutionnalisé** tels l'administration, l'éducation, l'armée, la justice, etc. Dans de nombreux pays où des minorités jouissent d'une certaine autonomie, il existe des **lois linguistiques**. Celles-ci déterminent de façon précise quelle langue doit être utilisée en quelle occasion. Pareille spécification reflète souvent l'existence d'une gradation. Pensons notamment à la situation du gallois à côté de l'anglais au pays de Galles. Quand il existe plusieurs langues officielles, on assiste généralement à une sorte de répartition des rôles : au Luxembourg, le français reste la langue administrative (écrite), l'allemand la langue des médias, et le luxembourgeois s'est vu élever de langue véhiculaire ordinaire à langue parlée officielle.

L'importance d'une langue ne dépend pas seulement de son statut officiel, mais également du nombre de locuteurs. Pour déterminer quelles langues sont

actuellement les plus utilisées dans le monde, il existe plusieurs critères. Si l'on prend comme critère uniquement le nombre d'utilisateurs d'une même langue, les langues asiatiques arrivent en tête de classement, comme le montre le tableau 2 :

Tableau 10.2 Les langues les plus parlées au monde (en millions de locuteurs, natifs et autres), d'après Grimes 1996

|                      |     |              |     |
|----------------------|-----|--------------|-----|
| le chinois mandarin  | 885 | le portugais | 175 |
| l'anglais            | 450 | le russe     | 160 |
| l'espagnol           | 366 | l'arabe      | 139 |
| le hindi/ urdu       | 233 | le japonais  | 126 |
| le malais indonésien | 193 | le français  | 122 |
| le bengali / assam   | 181 | l'allemand   | 118 |

En revanche, si, en plus du nombre d'utilisateurs, on considère d'autres critères comme par exemple le nombre de pays où la langue en question jouit du statut de langue officielle, le nombre de continents où cette langue est utilisée, ou encore la puissance économique des utilisateurs, on obtient un tout autre résultat, comme l'indique le tableau 3 :

Tableau 10.3 Les langues les plus "internationales" du monde

| Langue    | locuteurs natifs <sup>1</sup> | pays où elle est officielle | continents où elle est parlée | PNB en millions de US \$ dans les pays clés <sup>2</sup> |             |
|-----------|-------------------------------|-----------------------------|-------------------------------|----------------------------------------------------------|-------------|
| anglais   | 300                           | 47                          | 5                             | 1,069                                                    | Royaume-Uni |
| français  | 68                            | 30                          | 3                             | 1,355                                                    | France      |
| arabe     | 139                           | 21                          | 2                             | 38                                                       | RAU         |
| espagnol  | 266                           | 20                          | 3                             | 525                                                      | Espagne     |
| portugais | 175                           | 7                           | 3                             | 92                                                       | Portugal    |
| allemand  | 118                           | 5                           | 1                             | 2,075                                                    | Allemagne   |
| malais    | 193                           | 4                           | 1                             | 167                                                      | Indonésie   |

(<sup>1</sup>Grimes 1996 ; <sup>2</sup>Fischer Weltalmanach 1997)

L'anglais arrive loin en tête tant par le nombre de locuteurs natifs que par le nombre de pays où il a le statut de langue officielle et par sa répartition sur l'ensemble des continents. Rien d'étonnant, dès lors, à ce que l'anglais soit devenu la langue "mondiale" par excellence.

Aussi importants que soient ces faits sociologiques, ils sont exclusivement basés sur des critères de comparaison externe. Dans la section suivante, nous abordons des critères d'ordre structurel interne.

## 10.2 La diffusion et la classification des langues

Comme nous l'avons vu dans la section précédente, les langues peuvent être classées à partir de critères sociologiques. Néanmoins, il est également possible de remonter aux sources pour en analyser la genèse et la diffusion (10.2.1). Ceci conduit à des regroupements sur base de la parenté génétique (10.2.2). Indirectement, cette approche nous renseigne également sur les flux migratoires qui ont marqué l'histoire.

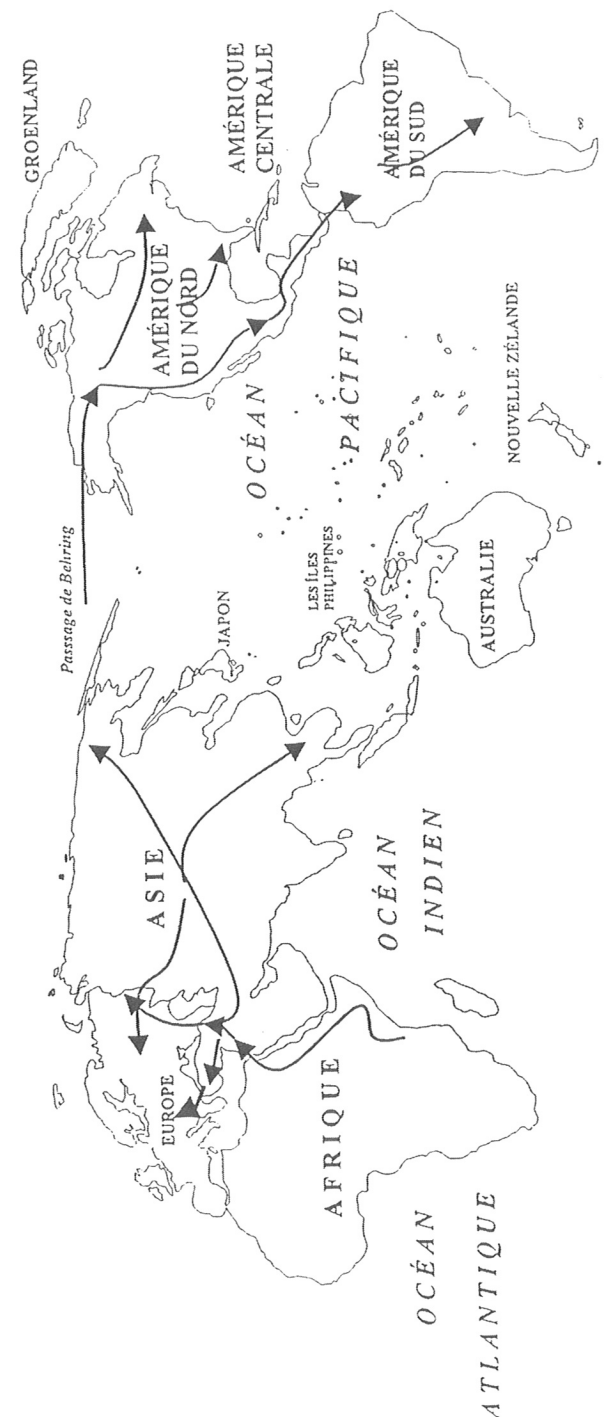
### 10.2.1 La genèse et la diffusion des langues

La comparaison des langues est un outil fondamental pour tenter de répondre à la question de savoir quelles sont l'origine, la nature et l'évolution d'une langue. Comme plusieurs domaines sont impliqués dans cette recherche, celle-ci implique la participation de diverses disciplines.

Selon Aitchison (1996), l'origine du langage coïncide avec celle de l'humanité en Afrique il y a 200 millénaires environ. Après avoir sommeillé durant quelque 150 millénaires, les langues se seraient mises à évoluer rapidement il y a à peine 50 000 ans, et ce selon des modalités très variées. On situe généralement la genèse de l'humanité à l'est des Grands Lacs, au Kenya. À partir de là, elle s'est non seulement propagée au sud et à l'ouest de l'Afrique, mais également vers le nord, et longeant la vallée du Nil, jusqu'au Proche-Orient. C'est ici que s'est effectuée la séparation entre les groupes se dirigeant vers l'Europe et vers l'Asie.

La branche asiatique de l'humanité s'est dispersée dans trois directions : (1) vers le sud-est de l'Asie et les archipels de l'équateur pour atteindre ensuite l'Australie et la Nouvelle-Zélande ; (2) vers l'Asie centrale (l'Inde actuelle) et à partir de là de nouveau vers l'Europe ; (3) vers le nord de l'Asie, et de là, traversant le détroit de Behring, jusqu'en Alaska et dans toute l'Amérique. La figure 1 (reproduite à la page suivante) retrace les principales migrations.

Figure 10.1 La diffusion des langues



Les migrations ont entraîné de multiples différenciations entre les langues, ce qui a donné naissance aux "familles de langues". De plus de en plus de peuples émigrèrent de l'Asie centrale vers l'ouest, donc vers l'Europe. C'est la raison pour laquelle l'indo-européen est aujourd'hui considéré comme étant l'une des principales familles de langues du monde.

Combien de temps a-t-il fallu pour que ces langues apparentées deviennent incompréhensibles les unes pour les autres ? Combien de langues ont pu disparaître ? Quelle est la durée requise pour qu'une nouvelle langue se développe : 20 ans ? Ou 200 ans ? Comment une langue se répand-elle ? Comment se fragmente-t-elle ?

Toutes ces questions requièrent une approche **pluridisciplinaire** : il est impossible de les traiter à fond sans avoir recours à l'archéologie, la physiologie, l'anatomie, l'anthropologie, la géographie humaine, l'éthologie (étude du comportement des espèces animales dans leur milieu naturel). Par la complémentarité des données provenant de ces différents domaines, il existe un espoir de répondre un jour à toutes ces questions.

Nous verrons dans ce qui suit qu'il serait prématuré de considérer le paysage linguistique connu comme étant définitivement établi. Ceci vaut tant pour les informations dont nous disposons pour la situation actuelle, que pour la reconstitution – plus ou moins plausible – de phases antérieures dans l'évolution des langues.

### 10.2.2 La parenté entre les langues comme principe de classification

Traditionnellement, les classifications linguistiques ont pour dessein l'identification des **familles linguistiques**. Cette métaphore suggère l'existence d'un lien de parenté génétique entre certaines langues, lien semblable à celui qui existe entre les membres d'une famille. Selon cette hypothèse, les langues qui présentent un nombre important de ressemblances d'ordre phonologique, lexicologique, morphologique et syntaxique proviennent d'un même ancêtre linguistique. C'est ainsi qu'un certain nombre de langues telles que le hindi en Inde, les langues iraniennes, slaves, celtes, romanes et germaniques appartiennent toutes à une même famille linguistique, à savoir la famille indo-européenne.

L'identification des familles linguistiques requiert des recherches historiques très poussées. Elle implique la reconstitution de formes et de systèmes

linguistiques plus anciens appelés **proto-langue**, comme par exemple le proto-indo-européen. Nous avons pu observer dans le chapitre précédent (9.2) qu'il est possible de partir de mutations phoniques pour reconstituer les stades antérieurs à cette mutation et remonter ainsi au stade du proto-indo-européen. Dans le présent chapitre, nous nous concentrerons sur les conséquences occasionnées par les modifications historiques.

La typologie linguistique actuelle utilise une taxinomie assez complexe. En effet, à côté de la notion de famille linguistique, elle recourt notamment à deux autres métaphores provenant du domaine de la biologie, à savoir le **phylum** et la **lignée**. Au sommet de la taxinomie linguistique l'on retrouve le **phylum**. C'est l'équivalent de ce qui est considéré dans le monde végétal comme la souche primitive d'où est issue une série généalogique (p.ex. les fougères). En typologie linguistique, cette catégorie désigne un groupe de langues n'entretenant aucune relation avec d'autres (groupes de) langues. En descendant, la catégorie suivante est la **lignée**. Elle représente un groupe de langues appartenant à des familles linguistiques différentes, mais qui présentent des liens de parenté éloignés. C'est par exemple le cas en Afrique du nord et au Moyen-Orient, où la lignée afro-asiatique comprend plusieurs familles linguistiques dans les alentours du désert du Sahara. Il s'agit des familles cushitique, copte, sémitique, berbère et chadique. Ces cinq familles de la même lignée sont cependant plus éloignées entre elles que par exemple le hindi, le latin, le russe et l'anglais. La raison en est que ces quatre dernières langues appartiennent à une seule et même **famille linguistique**.

La notion de famille linguistique, qui se trouve au centre de cette classification, met en évidence les liens internes pouvant exister entre les différents membres d'une famille. Parfois – c'est notamment le cas pour l'indo-européen –, les différents niveaux, à savoir le phylum, la lignée et la famille coïncident entre eux. Par contre, dans d'autres cas ces trois niveaux sont bien distincts. C'est ce qui explique la complexité de la situation linguistique qui caractérise l'Afrique, l'Asie et les Amériques (cf. tableaux 4 et 5).

À un niveau inférieur, les familles linguistiques peuvent être subdivisées en différentes **branches**. Ainsi, l'embranchement occidental de la famille de l'indo-européen se ramifie en deux groupes : le groupe roman d'une part, et le groupe germanique de l'autre. Ces embranchements peuvent à leur tour se subdiviser en "sous-embranchements". Une famille est donc faite de branches, les branches de groupes, et les groupes de sous-groupes. Toutes ces notions sont illustrées dans le tableau 4 à l'aide de quelques langues d'Afrique. Le tableau 5 de son côté présente une classification des zones linguistiques majeures dans le monde. Le tableau 6 quant à lui offre une vue d'ensemble de la famille des langues indo-européennes.



Tableau 10.4 Niveaux taxinomiques de la classification des langues pour une partie de l'Afrique (d'après Moseley et Asher 1994 : 292)

| Phylum      | Lignée      | Famille  | Branche | Groupe | Sous-groupe |
|-------------|-------------|----------|---------|--------|-------------|
| niger-congo | benue-congo | bantoïde | bantou  | venda  |             |
|             | ubangi      |          |         | tsonga |             |
|             | gur         |          |         | sotho  |             |
|             | adamawan    |          |         | nguni  | zulu        |
|             | ijoidkwa    |          |         |        | xhosa       |
|             | kru         |          |         |        | swazi       |
|             | dogon       |          |         |        | ndebele     |
|             | mande       |          |         |        |             |
|             | atlontic    |          |         |        |             |
|             | etc.        |          |         |        |             |

Le phylum du niger-congo est constitué de plusieurs lignées : entre autres, le benue-congo, l'ubangi, le gur, etc. Le cas du benue-congo montre qu'une lignée n'est pas nécessairement composée de plusieurs familles. En effet, dans le cas des langues bantoues, les notions de lignée et de famille coïncident entre elles. Ceci signifie que, bien que le groupe linguistique bantou présente un lien de parenté éloigné avec les autres groupes linguistiques du congo-nigérien, la lignée bantoue ne forme qu'une seule et même famille. Celle-ci se compose d'une multitude d'autres embranchements qui ne sont pas tous représentés ci-dessus. Finalement, ces embranchements sont à leur tour composés d'une foule d'autres sous-groupes.

Le tableau 5 illustre toute l'importance de la notion de phylum. Le premier ensemble, celui des langues sub-sahariennes, se classe en trois phyla différents, à savoir le congo-nigérien, le khoisan et le nilo-saharien. Ceci implique qu'il n'existe aucun lien de parenté génétique entre ces trois phyla. Dès lors, on peut en déduire que les peuples utilisant ces langues migrèrent vers ces différents territoires d'Afrique bien avant l'apparition du langage humain, que Aitchison (1996) situe entre 150 000 et 50 000 ans avant J.-C.

La situation se présente fort différemment pour le second ensemble, englobant les langues parlées en Afrique du nord et au Moyen-Orient. En effet, elles appartiennent toutes à une même lignée : la lignée hamito-sémitique. Ceci signifie que toutes ces langues, c'est-à-dire le somalien, l'hébreu et l'arabe, les langues berbères et le hausa, sont toutes apparentées entre elles de loin ou de

près (Comrie 1987 : 155). Ce lien de parenté est exclusif de cette lignée : on ne le retrouve dans aucune des cinq autres zones distinguées. On peut donc en conclure qu'à l'exception de ce second ensemble, les cinq autres ensembles répertoriés dans le tableau 5 représentent plutôt des catégories géographiques que linguistiques.

Tableau 10.5 Les principales zones linguistiques du monde

- 1 les langues sub-sahariennes
  - 1.1 niger-congo, phylum comprenant e.a. la lignée / famille bantoue
  - 1.2 khoisan, phylum comprenant p.ex. les langues khoukhous (e.a. le nama)
  - 1.3 nilo-sahara, phylum comprenant e.a. la famille nilotique
- 2 la lignée afro-asiatique : Afrique du nord et Moyen-Orient
  - 2.1 famille cushitique, avec e.a. le somalien, plusieurs langues éthiopiennes
  - 2.2 l'ancien égyptien et le copte : langues éteintes
  - 2.3 famille sémitique, comprenant e.a. l'arabe, l'hébreu, l'aramien
  - 2.4 famille berbère
  - 2.5 famille chadique (e.a. le hausa)
- 3 les langues nostratiques : l'Inde, la Perse, l'Europe
  - 3.1 phylum indo-européen (voir le tableau 6)
  - 3.2 phylum kartvélique : le caucasien du sud, le géorgien
  - 3.3 phylum ouralique : le finnois, l'estonien, le lapon, le hongrois
  - 3.4 phylum altaïque : le groupe turc (e.a. le turc), le groupe mongole
- 4 langues austriques : Inde et Asie orientale
  - 4.1 phylum austrasien, comprenant e.a. le sous-groupe mon-khmer
  - 4.2 phylum dravidien, comprenant e.a. le tamil, le telugu
  - 4.3 phylum sino-tibétan, avec e.a. les familles chinoise et tibétano-burmienne
  - 4.4 le coréen, le japonais
- 5 langues australasiatiques et du Pacifique
  - 5.1 phylum austronésien (800 langues), lignées malaise, indonésienne, javanaise
  - 5.2 langues papoues (750 langues en Nouvelle-Guinée Papua)
  - 5.3 phylum australien (250 lg.), e.a. lignée pama nyungane (p.ex. le mbabaram)
  - 5.4 groupe polynésien
- 6 langues amérindiennes
  - 6.1 de l'Amérique du nord, comprenant e.a. les familles ou lignées suivantes :
    - 6.1.1 famille du groupe inuit
    - 6.1.2 famille athapaskane, comprenant e.a. le navaho
    - 6.1.3 famille wakashane, comprenant e.a. le kwakiutl, le nootka
    - 6.1.4 lignée uto-aztèque, comprenant e.a. le hopi
  - 6.2 langues de l'Amérique centrale, e.a. la famille maya
  - 6.3 langues de l'Amérique du sud

Le troisième ensemble correspond aux langues nostratiques. Ce regroupement remonte aux études du linguiste danois Pedersen (1924). Il est représentatif d'une approche typologique plus ancienne selon laquelle toutes les langues du monde seraient génétiquement apparentées entre elles. Les typologistes actuels sont plus prudents. Ils exigent toute une kyrielle de preuves empiriques avant d'admettre un quelconque lien de parenté entre les langues.

Le tableau 5 représente en quelque sorte un compromis entre ces deux approches. Il en ressort que l'on peut très bien se concentrer sur la description des ensembles géographiques, sans devoir pour autant se lancer dans l'étude des liens de parenté génétique pouvant exister entre les différents phyla d'un même ensemble géographique.

La famille des langues indo-européennes a été intégrée avec plusieurs autres phyla dans le groupe des langues nostratiques, sans qu'il y existe un quelconque lien de parenté entre elles (contrairement à ce que l'on affirmait il y a une bonne cinquantaine d'années).

Le tableau 6 donne un aperçu des langues indo-européennes. L'appellation des deux branches majeures, à savoir **Satem** et **Kentum** renvoie à l'ancien mot *cent*, respectivement en ancien iranien *satem*, et en latin *centum*. Nous avons déjà fait remarquer que la reconstitution des liens de parenté entre des ensembles et des familles de langues dépend des résultats de la recherche sur des phases antérieures de l'évolution linguistique. Ainsi, la forme commune de *cent* dans les langues **Satem** (à l'est) et des langues **Kentum** (à l'ouest) est la forme de l'indo-européen *\*k'mto*, le /k'/ étant une occlusive palatalisée. Dans les langues Satem de l'est, cette occlusive est devenue une fricative palatale, pour évoluer finalement jusqu'au /s/ que l'on retrouve dans *Satem* 'cent'. Dans les langues Kentum, le /k'/ s'est converti en l'occlusive vélaire /k/, ce qui explique pourquoi *cent* se dit *centum* avec [k] en latin et *hekatón* en grec.

Généralement, l'attribution de liens de parenté génétique aux groupes de langues indo-européennes est motivée par des raisons empiriques. Il existe, en effet, bon nombre de documents écrits qui permettent de retracer l'évolution des différents embranchements, groupes et sous-groupes linguistiques.

La plupart du temps, les ressemblances entre les langues peuvent être expliquées par des notions telles que 'lien de parenté génétique'. Néanmoins, il arrive parfois que les similarités soient le fruit d'un hasard ou d'un emprunt. L'on parle de **ressemblances fortuites** lorsqu'elles portent sur des mots isolés et concernent des langues géographiquement ou historiquement éloignées les unes des autres. Ainsi, Comrie (1987 : 8) fait état du vocable "dog" qui désigne

le concept "chien" en mbabaram, une langue australienne aborigène. L'hypothèse d'un emprunt effectué à l'anglais est exclue. Il serait cependant tout à fait erroné d'affirmer qu'il existe un lien de parenté génétique entre le mbabaram et l'anglais, car il existe une explication étymologique plausible selon laquelle ce mot trouverait son origine dans une forme proto-australienne. Dans le cas particulier de 'dog', il n'y a aucun doute possible. On dispose en effet d'une documentation abondante concernant la langue anglaise et la façon dont elle s'est développée. Or, quand il s'agit de langues moins connues, on ne dispose pas toujours de l'information nécessaire pour déterminer si l'existence de vocables similaires est oui ou non le fruit du hasard.

Tableau 10.6 Langues indo-européennes (phylum=lignée=famille)

| Branches                                                          | Groupes                                         | (Sous-groupes)                                                                                      | Langues individuelles |                                                                                                                                                                        |
|-------------------------------------------------------------------|-------------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| I<br>N<br>D<br>O<br>-<br><br>E<br>U<br>R<br>O<br>P<br>É<br>E<br>N | S<br>A<br>T<br>E<br>M                           | 3.1.1.1. indo-iranien :                                                                             | indo-aryen<br>iranien | hindi, nepali, bengali, sanskrit (éteint)<br>persan, kurde                                                                                                             |
|                                                                   |                                                 | 3.1.1.2. arménien                                                                                   |                       |                                                                                                                                                                        |
|                                                                   |                                                 | 3.1.1.3. anatolien :                                                                                | hittite               |                                                                                                                                                                        |
|                                                                   |                                                 | 3.1.1.4. tokharien                                                                                  |                       |                                                                                                                                                                        |
|                                                                   |                                                 | 3.1.1.5. balto-slave :                                                                              | balte<br>slave :      | lituanien, letton<br>oriental : russe, biélorusse, ukrainien<br>occidental : polonais, tchèque, slovaque<br>méridional : bulgare, serbo-croate, slovène,<br>macédonien |
|                                                                   |                                                 | 3.1.1.6. albanais :                                                                                 | albanais              | albanais                                                                                                                                                               |
|                                                                   | K<br>E<br>N<br>T<br>U<br>M                      | 3.1.2.1. hellénique                                                                                 |                       | grec (moderne)                                                                                                                                                         |
|                                                                   |                                                 | 3.1.2.2. latin                                                                                      |                       |                                                                                                                                                                        |
|                                                                   |                                                 | 3.1.2.2a. roman                                                                                     |                       |                                                                                                                                                                        |
|                                                                   |                                                 | roman oriental :                                                                                    |                       | roumain, italien, sarde                                                                                                                                                |
|                                                                   |                                                 | ibéro-roman :                                                                                       |                       | espagnol, catalan, portugais                                                                                                                                           |
|                                                                   |                                                 | gallo-roman :                                                                                       |                       | français, provençal                                                                                                                                                    |
|                                                                   |                                                 | rhéto-roman                                                                                         |                       |                                                                                                                                                                        |
|                                                                   |                                                 | 3.1.2.3. celtique<br>goidélique<br>cymrique                                                         |                       | irlandais, écossais, gaélique<br>gallois, breton, cornique (éteint)                                                                                                    |
|                                                                   |                                                 | 3.1.2.4. germanique :                                                                               |                       |                                                                                                                                                                        |
|                                                                   | germanique oriental<br>germanique septentrional | gotique (éteint)<br>suédois, norvégien, danois,<br>islandais, faroëen                               |                       |                                                                                                                                                                        |
|                                                                   | germanique occidentale                          | anglais, frison, néerlandais,<br>afrikaans, haut-allemand,<br>bas-allemand, luxembourgeois, yiddish |                       |                                                                                                                                                                        |



### 10.3 La typologie et les universaux linguistiques

La discipline linguistique appelée **typologie** s'occupe de répertorier les caractéristiques communes à diverses langues, qu'il y ait ou non des liens de parenté entre elles. Il existe plusieurs approches typologiques. Dans ce qui suit nous nous intéressons plus particulièrement à l'une des questions centrales en typologie, à savoir la recherche d'universaux linguistiques. Nous y avons déjà fait allusion dans les chapitres 6 et 7. Un certain nombre de concepts universels – également appelés primitifs sémantiques – ont été résumés dans le tableau 2 du chapitre 6. La discussion dont ils faisaient l'objet concernait "l'alphabet de la pensée humaine". Elle indiquait que l'appareil conceptuel humain est basé sur un ensemble restreint composé d'une soixantaine de concepts fondamentaux. Il s'agit des premiers concepts à être exprimés dans toutes les langues du monde. Dans ce contexte, la classe de mot à laquelle ils appartiennent – ou le type de morphème – n'a aucune importance. Ainsi, la forme des concepts *je* et *tu* peut varier : ils peuvent être exprimés par un pronom comme en français, ou par un suffixe comme en latin (*veni* 'je suis venu(e)'; *venis* 'tu es venu(e)'). En *nootka*, une langue amérindienne qui n'utilise pas de noms, la référence au locuteur (*je*) ou à l'allocutaire (*tu*) est exprimée à l'aide d'affixes accolés aux verbes.

On peut généraliser et résumer ces faits en disant que les concepts universaux sont exprimés soit sous forme de lexèmes (morphèmes libres), soit sous forme de morphèmes liés ; on parle alors d'**universaux idéationnels**. Or, la fonction première du langage étant la communication, on peut aussi admettre l'existence d'**universaux interpersonnels**, qui règlent la façon dont nous communiquons les uns avec les autres. Les maximes qui ont trait au caractère qualitatif, quantitatif et pertinent de la communication (cf. 7.3.1) peuvent être intégrées dans l'ensemble des **universaux interpersonnels** ou **pragmatiques**.

Bien que ces types d'universaux linguistiques ne soient pas dépourvus d'intérêt, ils ne présentent pas la moindre complexité interne. Il s'agit de simples inventaires dont les éléments peuvent aisément être vérifiés – et donc confirmés ou infirmés – pour peu qu'il soit possible de faire le tour des cinq ou six mille langues du monde. Or, le type d'universaux qui suscite l'intérêt de la typologie linguistique – qui est le successeur de la linguistique comparée –, est d'ordre plus complexe. La typologie prend appui sur l'analyse de tous les éléments du répertoire des sons, des morphèmes, des unités lexicales, des structures syntaxiques, etc., pour établir les possibilités combinatoires et découvrir les contraintes qui pèsent notamment sur l'ordre linéaire.

Pour une bonne illustration de cette approche dans le domaine lexical, nous nous tournons vers la nomenclature des couleurs dans les différentes langues.

Avant, cette terminologie était considérée comme étant spécifique à chaque langue. L'on admettait simplement que chaque langue segmentait à sa façon le spectre des couleurs, pour donner ensuite un nom à chacun des segments distingués arbitrairement. Grâce à l'étude comparative d'une centaine de langues, les chercheurs Berlin et Kay (1969) ont découvert qu'il y a en réalité une impressionnante régularité dans la distribution de la terminologie des couleurs (voir aussi Heider 1972). Dans les langues où apparaissent de tels termes, ils sont au moins deux, à savoir "clair" et "foncé" (ou "blanc" et "noir"). S'il y a trois termes, le troisième sera "rouge", et s'il y en a quatre ou cinq, les termes suivants seront "jaune" et/ou "vert". Le sixième terme est toujours "bleu", et le septième toujours "brun". Les quatre suivants sont – en ordre variable – "pourpre/violet", "rose/mauve", "orange" et "gris". Nous résumons ceci dans le tableau 7 :

Tableau 10.7 Distribution universelle des termes de couleurs

| stade 1 | stade 2 | stade 3 | stade 4 | stade 5 | stade 6 |
|---------|---------|---------|---------|---------|---------|
| blanc   |         | jaune   |         |         | gris    |
|         | < rouge | < ou    | < bleu  | < brun  | < rose  |
| noir    |         | vert    |         |         | orange  |
|         |         |         |         |         | pourpre |

Cet agencement montre que le contraste entre deux extrêmes comme blanc et noir constitue un bon point de départ pour construire des concepts. Dans les langues où l'on ne retrouve que deux termes de couleurs, à savoir clair et foncé (ou blanc et noir), cette opposition est plus saillante que celle entre les couleurs particulières, celles-ci pouvant être catégorisées comme étant *claires* (rouge, jaune, bleu, etc.), ou *foncées* (rouge, jaune, bleu, etc.).

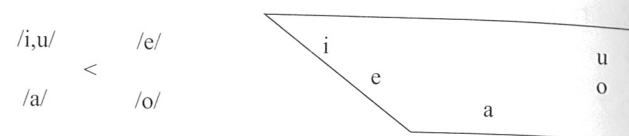
Le détail des mécanismes en jeu n'est cependant pas encore bien connu. En effet, des recherches plus récentes (Kay et al. 1991) ont démontré que la situation n'est pas aussi nette : il arrive qu'une ou plusieurs des couleurs du dernier stade apparaissent plus tôt. Il n'empêche que dans les grandes lignes on peut maintenir qu'il existe un haut degré de systématisme dans le développement des noms de couleurs.

Ce même principe pourrait être appliqué à l'usage qui est fait du système vocalique dans les différentes langues. Le fait que toute langue possède au moins deux voyelles constitue un trait phonologique universel. Nous pourrions aller plus loin en affirmant que si une langue ne compte que trois voyelles, celles-ci doivent occuper des positions cardinales diamétralement opposées, à savoir, /a/, /i/ et /u/ ; c'est la situation que Kelz (1976) a relevé notamment pour le guarani. Dans la mesure où une langue possède quatre voyelles, celle qui se rajoute sera

soit le /o/, soit le /e/. Si la langue en compte cinq, il s'agira de ces cinq voyelles cardinales-ci ; c'est le cas par exemple en espagnol. (Rappelons que l'italien possède 9 voyelles, l'allemand 11, l'anglais 14, et le français 16 (cf. chapitre 5.4.3)).

Par analogie avec la distribution des termes de couleurs, nous pouvons émettre l'hypothèse que la configuration du système vocalique d'une langue obéit à un ordre naturel de complexité dont l'élaboration de base est représentée dans le tableau 8 :

Tableau 10.8 Distribution universelle des voyelles de base



En morphologie, ce principe de base peut également être appliqué, par exemple au système des affixes. De la classification des affixes il ressort que les *suffixes* sont utilisés le plus fréquemment. Ensuite viennent les *préfixes*, suivis des *infixes* et éventuellement des *circonfixes* (également appelés "morphèmes discontinus", tel *ge- ...-t* dans le participe passé du néerlandais (*gewerkt* 'travaillé')). Le tableau 9 reflète cet ordre de préférence.

Tableau 10.9 Distribution universelle des affixes

suffixe < préfixe < infixes < circonfixe

Dans le passé, la typologie linguistique focalisait son attention sur les universaux syntaxiques ou grammaticaux. Greenberg (1963, 1966<sup>2</sup>) a résumé les idées qui prévalaient déjà à l'époque. Il passe en revue les possibilités dont témoigne l'ordre des mots, tel SVO (sujet-verbe-objet), VSO, etc. La façon dont il présente ses conclusions est en bien des points semblable aux résultats auxquels aboutirait une approche en termes de prototypes. Ce qui vaut pour les catégories en général, vaut également pour les universaux : à l'image des membres centraux et des membres périphériques d'une catégorie, l'on pourra dégager d'une part des configurations linéaires centrales (ou prototypiques) et d'autre part des ordres non préférentiels ou marginaux. Il n'y a donc pas lieu de s'inquiéter pour les soi-disant exceptions. Elles s'intègrent dans un classement hiérarchique où les différents ordres possibles trouvent leur place en fonction des tendances observées. Le tableau 10 offre une vue d'ensemble de l'enquête menée par Greenberg (1966 : 107) pour 30 langues différentes :

Tableau 10.10 Classement préférentiel des types d'ordre syntaxique

|      |   |      |   |     |   |     |   |     |
|------|---|------|---|-----|---|-----|---|-----|
| SVO  | > | SOV  | > | VSO | > | VOS | > | OVS |
| (13) |   | (11) |   | (6) |   |     |   |     |

(Les chiffres renvoient à 30 langues sélectionnées parmi toutes les lignées possibles)

Sur base de cette recherche ciblée, l'on peut conclure, d'une part, que SVO et SOV représentent les ordres prototypiques ou centraux, et d'autre part, que VSO est, bien que moins prototypique, encore fréquemment utilisé. Ces trois ordres linéaires ont un point commun, à savoir que le sujet précède l'objet. Cette tendance formelle peut être mise en relation avec un ordre de priorité conceptuelle, dont le principe a été avancé au chapitre 4 : un flux d'énergie est transmis de l'Agent au Patient, ou – en termes plus généraux – une forme de contrôle est exercée par un Agent, un Possesseur ou un Expérimentateur (la personne faisant l'expérience de quelque chose) sur un Patient. (Pour ces termes, nous renvoyons au chapitre 4 (4.2.4 et 4.2.5)). Ces priorités conceptuelles se reflètent iconiquement dans les trois principaux types d'ordre linéaire, dans la plupart des langues du monde. L'ordre inverse, c'est-à-dire celui où l'objet précède le sujet, est marginal (ou périphérique). L'ordre VOS se retrouve dans la langue amérindienne appelée *cœur d'alène* ; VOS et OVS quant à eux, sont utilisés dans deux autres langues amérindiennes, à savoir le *siuslaw* et le *coos* (Greenberg 1966 : 110).

Cette façon de penser en termes de prototypicité est également présente dans la liste de 45 universaux établie par Greenberg. L'on y retrouve des formulations restrictives telles que *presque toujours*, *avec une fréquence qui n'est pas le reflet d'une distribution aléatoire* ou encore *avec une fréquence trop élevée pour être le fruit du hasard*. Le tableau 11 reprend quatre des cinq premiers universaux syntaxiques formulés par Greenberg (1966 : 110) :

Tableau 10.11 Universaux syntaxiques

1. Dans les phrases déclaratives à sujet nominal et à objet nominal, l'ordre prédominant est généralement un ordre où le sujet précède l'objet.
2. Dépassé, invalidé depuis.
3. Les langues dans lesquelles prédomine l'ordre VSO sont toujours prépositionnelles.
4. Les langues où l'ordre normal est SOV tendent à être postpositionnelles.
5. Dans les langues où l'ordre prédominant est SVO et où le génitif suit le nom qui le régit, il est habituel que l'adjectif suive également le nom (par exemple : *l'amour de Dieu* > *l'amour divin*).

Le genre d'universaux énoncés sous les points 3, 4 et 5 du tableau 11 ci-dessus portent le nom d'**universaux implicationnels**. On indique par là que le fait que S, V et O se présentent dans un certain ordre, implique aussi un ordre fixe entre d'autres éléments de la phrase. On en dérive par exemple la position de la préposition, qui en théorie peut se placer devant le nom (prépositionnel), ou derrière le nom (postpositionnel). Les énoncés 3 et 4 signalent que si le verbe précède l'objet (V + O : *il monte les bagages* / *il monte l'escalier*), la préposition précède le nom (*il monte dans sa chambre* / *il grimpe sur l'arbre*) ; par contre, si l'objet précède le verbe (O + V) – comme en hongrois (*"Jean les bagages monte"*) –, alors la préposition suit le nom, elle se met "en postposition" (*"il grimpe l'arbre sur"*). Dans la phrase hongroise *Zoltán a fa alatt fut 'Zoltán l'arbre sous court'* diffère de la phrase française correspondante *Jean court sous l'arbre* par l'antéposition du nom à la préposition et par la position finale du verbe.

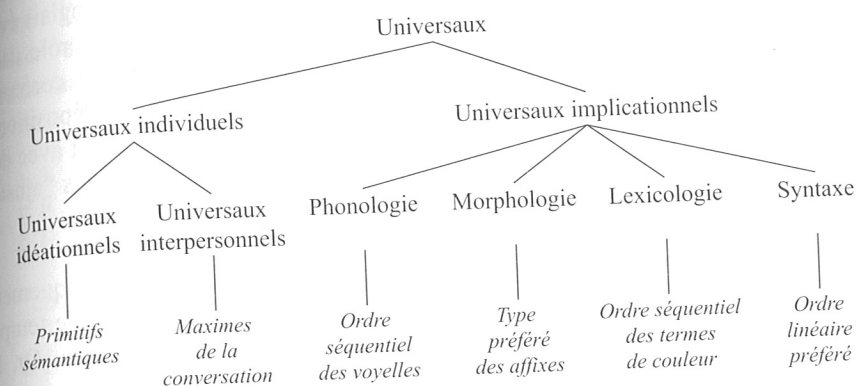
Les langues qui présentent les mêmes universaux implicationnels relèvent du même type d'ordre linéaire. Mais il n'existe pas nécessairement de lien de parenté entre elles. L'emploi de postpositions est non seulement un point commun à de nombreuses langues du phylum ouralique – auquel appartient notamment le hongrois (cf. tableau 5, section 3.3) –, mais on le retrouve également parmi celles du phylum altaïque – comprenant notamment les groupes de langues turque et mongole (tableau 5, section 3.4) – et même dans celles parlées au Japon et en Corée. Sur base de cette caractéristique commune qu'est la postposition, il avait été proposé antérieurement de regrouper toutes ces langues dans la grande famille ouralo-altaïque. Or, l'hypothèse du lien de parenté génétique n'a pas pu être maintenue faute d'indices supplémentaires pour l'étayer. Notons que dans le tableau 5, ces langues sont classées sous trois phyla différents.

Par souci de clarté, nous reprenons dans le tableau 12 ci-contre les différents termes concernant les universaux.

Tous les types d'universaux que nous avons abordés jusqu'ici se sont vus attribuer le nom d'**universaux substantiels**. Ceci signifie qu'il s'agit d'éléments issus d'une liste restreinte d'universaux parmi lesquels une langue peut effectuer des choix. Plus tard, Chomsky les rebaptisa **universaux paramétriques**, car selon lui, il n'existe aucune propriété *p* commune à toutes, ni même à la plupart des langues. En effet, chaque langue sélectionne une variante grammaticale appartenant à un paramètre *P*, défini comme un ensemble universel et immuable de propriétés. Auparavant, Chomsky avait déjà imaginé un second type majeur d'universaux qu'il avait appelé **universaux formels**. La dénomination "formel" tient au fait que ces universaux, qui sont indissociables

de l'acquisition des langues, représentent les contraintes "formelles" internes à la grammaire de ces langues.

Tableau 10.12 Types d'universaux



Tout enfant ne souffrant pas d'un handicap psychique ou physiologique important est capable d'apprendre au moins une langue, et de la maîtriser au bout de quelques années. Ceci vaut indépendamment des différences d'environnements social, culturel ou national. Ainsi, moyennant un contact suffisant avec des usagers d'une langue donnée, n'importe quel enfant sera en mesure d'en acquérir les structures et les règles propres avec une efficacité stupéfiante, et ce même si ses performances intellectuelles dans d'autres domaines ne sont pas aussi bonnes. Ceci laisse à supposer qu'il doit y avoir dans le cerveau un **dispositif langagier particulier** – probablement inné – qui est activé chaque fois qu'une personne est confrontée à de nouvelles données linguistiques. Ce dispositif (connu sous le nom de LAD '*language acquisition device*') contiendrait tous les universaux formels et orienterait l'enfant dans sa découverte des structures d'une langue et du langage en général. Selon cette approche, l'apprenant s'adapterait à n'importe "input" et serait à même d'en dériver des règles.

De nos jours, la notion de LAD comme capacité "aveugle" semble quelque peu dépassée. Actuellement, l'on insiste beaucoup plus sur l'intense activité – tant visible qu'invisible – que l'enfant déploie dans le processus d'acquisition d'une langue. Dès avant la naissance il se familiarise avec le système phonologique de sa langue maternelle par la voix de sa mère. Ses premiers balbutiements – encore prélinguistiques – sont un exercice de production des phonèmes et des intonations intériorisés. L'interaction avec les éducateurs (parents, parents d'accueil) et avec les autres enfants facilite et accélère le processus. On assiste alors à l'apparition de constructions de son propre cru, qui n'appartiennent pas à la langue maternelle et qui ne pourraient donc en principe



pas être admis par le LAD. Ces formes et constructions impossibles finissent par disparaître grâce au contact permanent. Tant il est vrai que l'enfant s'adapte complètement à son milieu. Il ne met donc pas en œuvre des concepts préétablis mais seulement ceux qu'il entend manier par les personnes de son entourage, comme l'illustre bien l'exemple des enfants coréens et anglais évoqué dans le chapitre 6 (tableau 1). La nationalité ne joue aucun rôle dans l'apprentissage d'une langue : il suffit que les enfants d'origine coréenne soient exposés à l'anglais parlé pour en assimiler le système des prépositions. De même, il suffit aux enfants d'origine anglaise d'entrer en contact avec des enfants d'origine coréenne pour s'approprier le système des relations "ajusté-non ajusté".

En d'autres termes, il semble que l'être humain est bel et bien génétiquement programmé pour apprendre au moins une langue, et ce en très peu de temps. Cette disposition innée est une capacité propre à l'espèce humaine. Or, la nature exacte du phénomène et les rouages du mécanisme ne sont pas encore entièrement connus.

## 10.4 La linguistique comparée

Les hypothèses issues des recherches sur les universaux ne manquent pas d'intérêt pour la linguistique théorique et pour les approches interdisciplinaires. La liste des universaux linguistiques et l'inventaire des groupes et familles linguistiques constituent le résultat d'enquêtes impliquant des dizaines voire des centaines de langues. L'ampleur de ces recherches et la priorité qu'elles accordent aux ressemblances entre les langues visent à dégager des vues d'ensemble. La comparaison de langues individuelles y tient peu de place. En revanche, la **linguistique comparée** ne se concentre que sur deux langues ou quelques-unes tout au plus. La comparaison systématique de deux langues permet de mesurer avec précision le degré de convergence entre elles et de vérifier en profondeur dans quelle mesure le comportement d'un élément déterminé est divergent dans les deux langues. Bien souvent, ce genre de comparaison permet de déceler des correspondances multidimensionnelles et ouvre de nouvelles perspectives cognitives. L'on peut ainsi être amené à réviser les définitions et les catégories en vigueur même dans des domaines assez solidement établis, tels la systématique des temps et aspects du verbe. Cependant, le rôle de la linguistique contrastive ne s'arrête pas à cette dimension théorique. En effet, elle a également des visées plus pratiques, servant notamment de support à l'apprentissage des langues, à la traduction et à la rédaction de dictionnaires bilingues.

### 10.4.1 Les équivalents de la forme "progressive" de l'anglais

Comme nous l'avons vu dans le chapitre 4 (4.4.5), la forme "progressive" du verbe constitue un élément d'ancrage qui permet au locuteur de placer en point de mire l'aspect interne d'un événement. D'autres langues que l'anglais connaissent des formes morphologiquement équivalentes, mais celles-ci s'avèrent être nettement moins usitées. (Les formes progressives sont mises en italiques. Le signe "=" signale qu'il s'agit d'une traduction mot à mot.)

- a. A : What are you doing ? I am writing a card.
- b. F : Qu'est-ce que tu es en train de faire ? J'écris une carte.
- c. I : Cosa stai facendo ? Sto scrivendo una cartolina.  
= 'Que restes-tu faisant ? Je reste écrivant une carte.'
- ou : Cosa fai ? Scrivo una cartolina.  
= 'Que fais-tu ? J'écris une carte.'
- d. N : Wat ben je aan het doen ? Ik schrijf een kaart.  
= 'Qu'es-tu à le faire ? J'écris une carte.'
- e. Al : Was machst du gerade ? Ich schreibe eine Karte.  
= 'Que fais-tu en ce moment ? J'écris une carte.'

Tous ces exemples donnent à voir comme ancrage le déroulement de l'événement. Or, seule la forme anglaise impose cette perspective interne. En néerlandais et en italien, les formes progressives et non progressives existent côte à côte. Là où la forme progressive accentue la vision interne, la vision exprimée par la forme non progressive est neutre. L'allemand ne connaît pas de forme progressive ; c'est l'adverbe *gerade* 'maintenant, en ce moment' qui suggère l'aspect de l'action.

En poussant l'analyse un peu plus loin – en vue d'une possible généralisation –, l'on s'aperçoit que la forme progressive anglaise peut également exprimer la notion de 'durée'. Il suffit pour cela qu'elle se combine avec le "perfectif" : *have been X-ing*. Dans les autres langues, cette notion de 'durée' est exprimée à l'aide de locutions adverbiales.

- a. A : He has been crying for an hour.
- b. F : Il pleure depuis une heure.
- c. I : Sta piangendo da un'ora  
= 'Il reste pleurant depuis une heure.'
- ou : Piange da un'ora.  
= 'Il pleure depuis une heure.'
- d. N : Hij weent al een uur.  
= 'Il pleure déjà une heure.'
- ou : Hij is al een uur aan het wenen.  
= 'Il est déjà une heure pleurant.'



- ou : Hij heeft al een uur geweend.  
= 'Il a déjà une heure pleuré.'  
ou : Hij is al een uur aan het wenen geweest.  
= 'Il a déjà une heure pleurant été.'

Nous remarquons que, tout comme la forme progressive anglaise, la forme progressive néerlandaise peut se combiner avec une forme perfective ; ceci n'est pas le cas pour l'italien. Dans ces deux langues, les formes non-perfectives permettent d'ailleurs également d'exprimer l'aspect 'durée', mais celle-ci doit alors être explicitement indiquée par "al een uur" et "da un'ora".

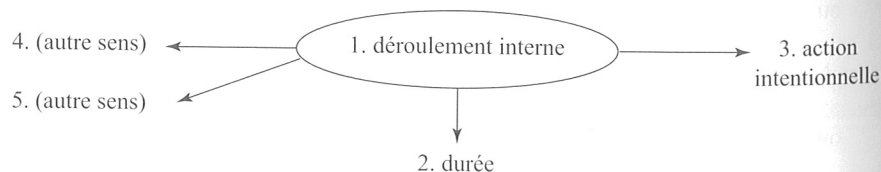
Le fait qu'en néerlandais, italien et allemand la notion de durée peut toujours être exprimée sous forme perfective non-progressive, démontre clairement à quel point cette forme est neutre dans ces trois langues. En anglais, en revanche, l'ancrage de la 'perspective interne' s'est transformé en "patron habituel" – au sens Whorfien du terme (cf. chapitre 6.3) –.

Inversement, on peut aussi constater que là où la forme progressive anglaise peut encore avoir d'autres significations telles "intention ou plan dans le futur proche", ce n'est pas le cas pour les formes équivalentes dans les autres langues.

- a. anglais : *I am not taking* the train today.  
b. français : Je ne vais pas prendre le train aujourd'hui.  
c. italien : Non intendo prendere il treno oggi.  
= '(Je) ne pense prendre le train aujourd'hui.'  
d. néerlandais : Ik ga vandaag niet de trein nemen.  
= 'Je vais aujourd'hui pas le train prendre.'  
a. allemand : Heute nehme ich den Zug nicht.  
= 'Aujourd'hui prends je le train pas.'

Les diverses significations de l'aspect progressif anglais peuvent, dès lors, être considérées comme membres d'une catégorie radiale, telle qu'elle est résumée dans le tableau 13 ci-dessous.

Tableau 10.13 Réseau radial de quelques significations de la forme progressive anglaise



Notre brève analyse contrastive a pris comme point de départ un patron morphologique (la forme progressive). Dans la comparaison il a fallu prendre en compte trois éléments d'ancrage, à savoir l'*aspect interne*, la *durée* et l'*intention dans un futur proche*. En même temps, cette analyse montre à quel point les correspondances entre deux langues peuvent être partielles et trompeuses. Ainsi, contrairement au temps présent en anglais, les temps présents en italien et en néerlandais peuvent exprimer à la fois l'aspect interne et l'aspect duratif.

Nous avons pris en considération les correspondances et les divergences existant entre les langues. Actuellement, la recherche s'intéresse surtout aux divergences entre les langues. Dans la mesure où l'approche contrastive permet de jeter une lumière nouvelle sur des différences de sens qui s'avèrent pertinentes dans bien des domaines, on peut affirmer que la linguistique comparée est une branche importante de la linguistique descriptive.

#### 10.4.2 Aspects méthodologiques de la linguistique comparée

La linguistique comparée fit son apparition dans le climat du behaviorisme des années 50. Selon ce courant de pensée, la science porte sur le comportement observable mais n'a pas accès à la "boîte noire" contenue dans l'esprit. La théorie behavioriste a aussi été appliquée au processus d'apprentissage d'une langue étrangère. On parlait de l'idée que les caractéristiques structurelles par lesquelles la **langue-cible** (la langue étrangère à apprendre) diffère substantiellement de la **langue-source** (la langue maternelle de l'apprenant) représentent l'obstacle majeur dans l'apprentissage de la langue, et méritent par conséquent une attention et un effort soutenus. On considérait que le risque d'erreur était automatiquement plus grand dans les domaines où la langue-cible diffère de la langue-source. Cette affirmation a dû être réajustée à mesure que l'on a mieux compris la relation entre la structure des langues d'une part et les difficultés liées à l'apprentissage d'autre part. En effet, au fil du temps on a pris conscience de plusieurs choses. Premièrement, on s'est rendu compte qu'il n'existait aucun lien entre le degré de dissemblance linguistique et l'effort mental requis. Deuxièmement, on a observé que la compétence de l'apprenant pouvait être entravée autant par des divergences mineures que par des divergences majeures (contrairement à ce qu'affirmait le mouvement behavioriste). Dès lors, la description des langues ne suffit pas pour pouvoir retracer les processus mentaux mis en œuvre par l'apprenant. Ces découvertes ont eu pour effet de discréditer partiellement la linguistique contrastive telle qu'elle était pratiquée dans les années '70.

Il est néanmoins démontré – et de nouvelles vérifications empiriques le confirment – que l'apprentissage d'une seconde langue ne pourra être fructueux que si l'apprenant arrive à faire la part des choses et distingue entre les structures et catégories linguistiques acquises précédemment et celles qui sont propres à la langue étrangère. Ceci signifie que face à de nouvelles données linguistiques, l'apprenant se verra dans l'obligation de réviser les anciennes catégories, schémas et prototypes, et ce à tous les niveaux de l'organisation linguistique. Une telle mise à jour implique que l'ancien dispositif mental soit adapté aux données spécifiques de la langue étrangère concernée. Dans ce contexte, plusieurs besoins se font sentir : celui d'analyses contrastives, et plus particulièrement, celui d'outils destinés à clarifier la portée des dissemblances que l'on peut rencontrer dans telle situation. Il faut que les dissemblances soient rendues intelligibles aux yeux des apprenants et des traducteurs.

Il y a plusieurs manières de se lancer dans des comparaisons linguistiques. En effet, on peut partir de catégories telles qu'elles sont décrites dans la grammaire traditionnelle, se baser sur l'ensemble du lexique ou prendre comme point de départ des listes de mots ou d'expressions, ou encore partir d'un ensemble de textes. La façon dont on présentera et explicitera les correspondances et divergences existant entre les langues dépendra d'une part de la langue choisie comme langue-cible, et d'autre part du cadre théorique servant de support.

Dans les précédents exemples, nous avons pris pour point de départ la forme progressive de l'anglais. Notre but était de découvrir les catégories conceptuelles véhiculées par cette construction verbale. Parallèlement, nous nous sommes demandé à partir de quelles formes les francophones, italophones, néerlandophones et germanophones abordent la forme anglaise. Nous pourrions également adopter un point de vue plus théorique et partir des catégories conceptuelles 'aspect interne' et 'aspect duratif' pour en analyser les diverses formes d'expression en anglais, français, italien, néerlandais, allemand. Le tableau 14 en donne l'aperçu.

Pour les finalités de l'enseignement des langues et de la traduction il est également indiqué de développer des instruments de travail basés sur des éléments linguistiques qui sont familiers à l'apprenant, tels que des mots, des expressions, des champs sémantiques particuliers, certains types de textes. L'information que l'on transmet concernant le vocabulaire d'une langue devrait provenir de l'examen approfondi de l'usage qui est fait des mots et des expressions dans des contextes authentiques. Les recherches sur le lexique font apparaître la nécessité de recourir à de vastes ensembles de données avant de pouvoir établir avec précision l'entrée qu'il convient de donner à un lemme – c'est-à-dire les définitions recouvrant les significations que peut prendre un vocable – et d'en fournir les traductions possibles. À cet égard, les lexiques

existants ne suffisent pas car la compétence linguistique d'un ou même de plusieurs lexicographes (auteurs de dictionnaires) n'est pas suffisamment fiable. Il convient de compléter les sources disponibles en combinant l'information qu'elles contiennent avec celle que l'on peut tirer d'un **corpus** de textes qui aura été soigneusement composé et qui sera constamment remis à jour. L'on entend par corpus (ou ensemble de corpus) une vaste collection de textes écrits ou oraux qui reprennent le plus grand nombre de variétés possibles.

Tableau 10.14 L'expression de l'aspect interne et de l'aspect duratif

|                | Aspect interne                                         | Aspect duratif                                             |
|----------------|--------------------------------------------------------|------------------------------------------------------------|
| a. anglais     | forme progressive                                      | parfait progressif<br>(+ adjectif duratif)                 |
| b. italien     | gérondif / présent<br>(+ adjectif temporel ponctuel)   | (gérondif) parfait + adjectif duratif                      |
| c. néerlandais | progressif / présent<br>(+ adjectif temporel ponctuel) | prés. / progr. / parf. / parf. prog.<br>+ adjectif duratif |
| d. allemand    | présent + adjectif temporel<br>ponctuel                | prés. / parfait + adjectif duratif                         |

La linguistique comparée quant à elle s'appuie sur des **corpus parallèles**, c'est-à-dire des (ensembles de) textes dont on peut raisonnablement supposer qu'ils rendent le même contenu dans des langues différentes (cf. Paulussen 1995).

La première étape consiste à étudier les variantes qui apparaissent pour un mot ou une expression à l'intérieur d'une seule langue. Cet inventaire permet de mettre en lumière, soit la disparition, soit l'émergence de certains mots ou de certaines expressions idiomatiques. À ce stade, l'analyse montrera aussi que certaines expressions sont tellement peu fréquentes, qu'on peut les ignorer, alors que d'autres ont un sens tellement précis qu'elles ne peuvent être utilisées que dans des contextes bien particuliers. La combinaison récurrente de certains mots peut caractériser certains types de textes. Elle peut également nous mettre sur la piste de contextes particuliers où certains mots ou expressions évoquent des **connotations** de négativité, de mépris, d'enthousiasme, de distance, etc. Prenons l'exemple du choix du verbe *mettre* ou *passer* en association avec un intervalle temporel (*un certain temps*) suivi de *à + infinitif*. Les deux signifient 'employer, utiliser X temps pour faire quelque chose'. Or, seule l'expression avec *mettre* (1a) se laisse paraphraser par 'faire quelque chose en X temps' (1b) ; le temps utilisé peut être mesuré en fonction du travail accompli ; sans autre indication contextuelle, on aura tendance à penser que le

locuteur trouve ce temps long, vu qu'il dispose du tour négatif pour marquer qu'il le trouve plutôt court (1c). Avec le verbe *passer*, en revanche, il ne peut pas être question de performance (2a) : d'une part, l'expression laisse planer le doute sur l'accomplissement (2b), et d'autre part, le tour négatif ne peut être compris que sur le mode ironique – c'est-à-dire signifiant exactement le contraire de ce qui est dit –, ce qui est révélateur de la connotation de dépit.

- (1) a. Il a *mis* tout l'après-midi à réparer la gouttière.  
b. Il a eu besoin de tout l'après-midi mais il y est arrivé.  
c. Il n'a mis qu'un après-midi à réparer la gouttière.
- (2) a. Il a *passé* tout l'après-midi à réparer la gouttière.  
b. Au bout de quatre heures {il avait enfin terminé / il a laissé tomber et a appelé un plombier pour le faire}  
c. \* ? Il n'a passé qu'un après-midi à réparer la gouttière.
- (3) a. \*Il met {sa vie / son temps} à mettre de l'ordre  
b. Il passe {sa vie / son temps} à mettre de l'ordre

Le fait que seul *passer* se combine avec des intervalles clairement non délimités tels *sa vie*, (*tout*) *son temps* corrobore l'idée qu'il s'agit d'une dépense de temps inutile ou exagérée (3).

La deuxième étape est **contrastive**. À ce niveau, l'on a déjà déterminé quel sera le point de départ de l'analyse destinée à recueillir les éléments équivalents dans l'autre langue. Dès lors, on essaiera de réunir la même quantité de données afin de repérer l'ensemble des variantes et des **collocations** ou contextes bien spécifiques qui s'avèrent être typiques du mot ou de l'expression que l'on étudie. En parcourant des corpus parallèles ou analogues à ceux que l'on avait analysés dans la langue-source, on s'arrêtera sur les items correspondants. Il suffit qu'il y ait un item commun ou tant soit peu identifiable comme équivalent pour que la comparaison entre les deux langues puisse commencer. La confrontation des résultats obtenus dans les deux langues pourra se faire à partir du moment où suffisamment d'information aura été récolté de part et d'autre. C'est en l'occurrence ici que commence la véritable comparaison entre les langues en question.

### 10.4.3 Expressions verbales : étude comparée de l'anglais et du français

Les verbes constituent une catégorie lexicale privilégiée. Dans le chapitre 4, nous avons vu que les expressions verbales sont décisives pour la classification syntaxique et sémantique des grilles événementielles et des schémas constructionnels. Les constructions phrastiques de base peuvent être considérées comme l'éventail des variantes possibles d'expressions à noyau verbal. Ainsi,

l'inventaire des combinaisons possibles entre un verbe et les noms compatibles avec ce verbe – notamment en position sujet et objet – consiste à énumérer les schémas constructionnels dans lesquels ce même verbe peut se présenter. En effectuant le classement parallèle pour des verbes appartenant à deux langues différentes, on obtient un instrument de travail intéressant. Étant donné que le nombre de verbes existant dans les langues européennes n'est pas illimité, une telle classification est tout à fait réalisable.

En guise d'illustration, nous allons comparer le verbe anglais '*to count*' et ses équivalents en français à partir de quelques exemples simplifiés. S'agissant d'une activité de base, l'on peut dire que *compter* se trouve au centre d'un domaine conceptuel important. Dans les civilisations dites avancées, les sciences empiriques et formalisées ont fourni pour l'ensemble des opérations mentales liées à cette activité des définitions bien précises, communément acceptées dans la société. Par conséquent, l'on peut s'attendre à ce que l'organisation du champ lexical correspondant soit dans ses grandes lignes la même pour les principales langues européennes. En principe, ceci doit faciliter le passage d'une langue à l'autre, dans la mesure où l'on peut projeter la signification des noms et des verbes sur leurs équivalents dans l'autre langue. Voyons si cette prédiction se vérifie pour la paire '*to count*' / '*compter*'.

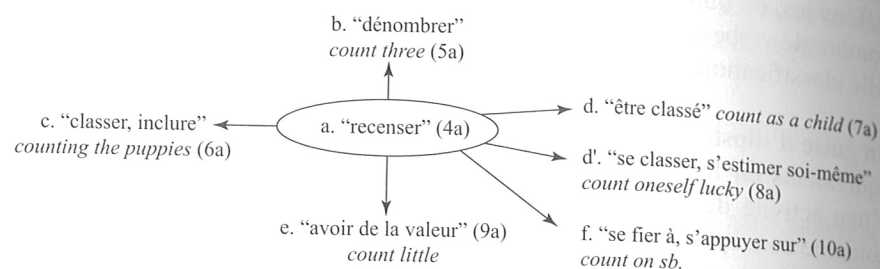
- (4) a. *The porter counted our bags.*  
b. Le porteur compta nos sacs.
- (5) a. *I count three before screaming.*  
b. Je compte jusqu'à trois avant de crier.
- (6) a. *Fifty dogs, counting the puppies.*  
b. Cinquante chiens, en comptant les chiots.
- (7) a. *For the ticket, he still counts as a child.*  
b. Pour le ticket, on le compte encore comme un enfant.  
b'. Pour le ticket, il est encore compté comme un enfant.
- (8) a. *He can count himself lucky.*  
b. Il peut s'estimer veinard.
- (9) a. *Your feelings count little with him.*  
b. Vos sentiments ne comptent guère pour lui.
- (10) a. *Do not count on me.*  
b. Ne compte pas sur moi.

Comme l'indiquent les exemples ci-dessus, le verbe anglais '*to count*' est le plus souvent traduit en français par le même verbe, à savoir '*compter*'. L'on observera cependant que le nombre d'éléments morphologiques supplémentaires mis en œuvre – prépositions, pronoms clitiques, particules – est plus grand en français qu'en anglais. Essayons maintenant d'établir le réseau radial du verbe '*to count*' à partir de ces sept occurrences. Ceci nous permettra d'une part de voir comment les différents sens sont reliés les uns aux autres et,



d'autre part, de dégager à la fois les similitudes et les divergences entre les deux langues.

Tableau 10.15 Réseau radial de 'to count'



Pour l'analyse contrastive nous partons du principe qu'il y a lieu de parler de correspondance sémantique directe entre deux (ou plusieurs) verbes dans les deux langues à comparer dans la mesure où les participants nucléaires – le sujet et l'objet (ou le complément) de la voix active – endossent les mêmes rôles sémantiques. Dans l'exemple ci-dessus du verbe anglais 'to count' et de ses équivalents français, cette condition est parfaitement remplie. En effet, les deux premières significations du verbe anglais 'count' sont des significations littérales qui présentent, tout comme leurs équivalents français, la configuration suivante :

1. 'count objects' = compter des objets = Agent-Patient (exemple 4)
2. 'count three' = compter jusqu'à trois = Agent-But (exemple 5)  
(en anglais, 'three' n'est pas Patient mais But, ce qui devient plus clair lorsque la trajectoire est exprimée – *count (up) to 10* – ou lorsque la source est exprimée, comme c'est le cas dans l'expression 'count from 1 to 10')

Les deux significations suivantes du verbe anglais 'to count' sont également littérales, mais elles ont un sens plus étendu. En effet, 'counting' est en l'occurrence utilisé dans le sens de 'en comptant (= y compris)' et de 'classant, classifiant'. La différence entre les deux se reflète au niveau des rôles sémantiques : la configuration de base, Agent-Patient, fait place à la combinaison Agent-Essif ou Patient-Essif. Notons cependant qu'en français, il y a une distinction entre les deux selon qu'il y a un agent ou non. Lorsqu'il n'y a pas d'Agent exprimé, le français recourt à la voix passive.

3. 'counting the puppies' = en comptant les chiots = Agent – Patient (exemple 6)
4. 'count as a child' = être compté comme un enfant = Patient – Essif (exemple 7)
5. 'count oneself lucky' = s'estimer heureux = Agent – Essif (exemple 8)

Dans les deux derniers exemples ci-dessous, le verbe anglais 'to count' est utilisé au sens figuré, tant en français qu'en anglais.

6. 'count little' = ne guère compter = Patient (exemple 9)
7. 'count on somebody' = compter sur quelqu'un = Agent-But (exemple 10)

Dans certaines expressions telles que 'gold counts as a special metal' et 'I do not count it against him', on utilisera plutôt d'autres verbes en français. On traduira ces phrases plutôt par : "l'or est un métal précieux" et "je ne lui en veux pas".

Bien que ces analyses soient très utiles, notamment pour déterminer la structure sémantique d'un verbe à partir des rôles attribués aux participants, et pour en retracer systématiquement les équivalences conceptuelles avec le pendant dans une autre langue, elles ne sont cependant d'aucun recours lorsqu'il s'agit d'expressions toutes faites, qui sont particulières à une langue. Un exemple typique est l'expression française 's'attendre à'. Cette expression peut très bien servir d'équivalent à l'anglais 'to count on' dans des énoncés tels que 'count on a heritage' 's'attendre à (recevoir) un héritage'. Par ailleurs, elle peut également exprimer des connotations négatives comme dans les expressions suivantes : 's'attendre au pire' ou 'Il faut s'attendre à des averses abondantes'. Lorsqu'on emploie 's'attendre à' avec une connotation négative, on est en plein dans le domaine de ce qui est **idiomatique**. Ceci signifie qu'au lieu d'être représentatives de constructions syntaxiques courantes, ces expressions sont propres à la langue, et présentent dès lors des caractéristiques uniques à bien des égards. Sauf ironie, l'anglais 'to count on' n'admet pas pareille extension de sens ; il faudra donc traduire par un autre verbe, tel 'to expect'.

À présent nous allons partir du verbe français 'compter', et voir comment il est traduit en anglais. Pour cela, il faut d'abord faire la distinction entre 'compter' au sens propre et 'compter' au sens figuré.

Dans un premier temps, nous présentons les significations littérales de ce verbe :

- A. *Compter* dans le sens de 'calculer', 'recenser', 'dénombrer'.

Exemple :

- a) Le porteur compta nos sacs.
- b) The porter counted our bags.

Ou encore :

- a) Je compte jusqu'à trois avant de crier.
- b) I count three before screaming.

Nous constatons que le verbe français *compter* se traduit dans les exemples ci-dessus par 'to count'.

- B. *Compter* dans le sens de 'inclure'.  
Exemple :  
a) Cinquante chiens, en comptant les chiots.  
b) Fifty dogs, counting the puppies.

Ici aussi le verbe français *compter* est traduit en anglais par le verbe 'to count'.

- C. *Compter* dans le sens de 'prévoir'.  
Exemple :  
a) Il faut bien compter 10 jours pour finir le travail.  
b) You must reckon 10 days to finish the work.

Nous remarquons qu'ici le verbe *compter* est traduit non pas par le verbe anglais 'to count', mais bien par 'to reckon'.

- D. *Compter* dans le sens de 'faire payer'.  
Exemple :  
a) On ne lui a pas compté le café.  
b) He wasn't charged with the coffee.

Ici encore, le verbe français *compter* n'est pas traduit pas le verbe anglais 'to count', mais bien par 'to charge with', qui est un verbe prépositionnel fixe.

Ainsi, l'on constate qu'à chaque fois que le verbe *compter* est utilisé dans un sens littéral, il est transitif, c'est-à-dire qu'il est suivi d'un objet direct. Cette observation vaut également pour sa traduction anglaise, sauf pour le verbe 'to charge with', qui prend une préposition.

Nous passons à présent à l'analyse du verbe *compter* lorsqu'il est employé au sens figuré. Force est de constater qu'à tous les coups il devient intransitif :

- A. *Compter* dans le sens de 'valoir', 'avoir de l'importance'.  
Exemple :  
a) Vos sentiments ne comptent guère pour lui.  
b) Your feelings count little with him.

Ici le verbe *compter* est traduit en anglais par 'to count'.

- B. *Compter* dans l'expression 'compter pour du beurre'.  
Exemple :  
a) Cela compte pour du beurre.  
b) That doesn't matter.

Dans cette expression, l'on traduira le verbe 'compter' plutôt par 'to matter' en anglais.

- C. *Compter* dans le sens de 'figurer parmi un groupe', 'être au rang des ...'.  
Exemple :  
a) Il compte parmi mes amis.  
b) He is one of my friends.

Pour traduire cette phrase, l'anglais utilisera plutôt le verbe 'to be' suivi d'un attribut.

- D. *Compter* dans le sens de 'se considérer comme faisant partie d'un groupe'.  
Exemple :  
a) Il peut se compter parmi les gens heureux.  
b) He can consider himself a happy man.

*Compter* est ici utilisé dans un sens plus abstrait ; il désigne métaphoriquement l'état d'une personne. L'anglais utilisera dans le cas présent plutôt un autre verbe tel que 'to consider', construit réflexivement ('himself') et avec attribut.

- E. *Compter* dans le sens de 'tenir compte de'.  
Exemple :  
a) Il faut compter avec l'opinion.  
b) You've got to reckon with public opinion.  
You have to take account of public opinion.  
You have to take public opinion into account.

Nous voyons que l'anglais utilisera, tout comme le français, un verbe suivi d'une préposition.

- F. *Compter* dans le sens de 'compter sur quelqu'un'.  
Exemple :  
a) Ne compte pas sur moi.  
b) Do not count on me.  
Do not rely on me.

L'anglais donnera la préférence soit au verbe 'to count' suivi d'une préposition fixe, soit au verbe prépositionnel fixe 'to rely on'.

Tous les exemples que nous venons de donner illustrent bien que le verbe français 'compter' peut être traduit en anglais par le même lexème, c'est-à-dire par le verbe 'to count'. Néanmoins, il est également possible de l'exprimer à l'aide de lexèmes différents tels 'to reckon (with)', 'to charge (somebody) with', 'to matter', 'to rely on', 'to consider oneself as', 'to be one of'...

Toutes les observations que nous avons faites jusqu'à maintenant ne représentent qu'une infime partie de ce qu'une analyse contrastive détaillée pourrait nous faire entrevoir à propos de deux verbes anglais et français. Même dans un champ sémantique aussi fondamental que celui de *compter*, nous observons que d'autres verbes prennent parfois le relais de, respectivement, 'compter' et 'to count'. Toutes les correspondances peuvent être analysées en termes de configurations de rôles sémantiques. Cependant, la valeur heuristique de ce genre d'analyse ne s'arrête pas là. Dans la mesure où il en émerge également un nombre d'expressions figées, ceci conduit à une meilleure compréhension de l'impact de l'**idiomaticité**. Ainsi, dans chaque langue il existe un très grand nombre de tournures qui se caractérisent par leur caractère unique. Elles échappent, par conséquent, à l'analyse en termes de configurations de rôles sémantiques et de grilles événementielles telles que nous les avons utilisées ici et décrites dans le chapitre 4. Les expressions idiomatiques présentent une configuration **idiosyncrasique**, ce qui veut dire qu'elle est *ad hoc*, non productive et souvent imprévisible. Le recours aux réseaux radiaux, quant à lui, constitue un instrument privilégié pour faire ressortir les relations existant entre les différentes significations, que celles-ci soit productives ou idiosyncrasiques.

## 10.5 Résumé

L'étude comparée des langues relève de plusieurs disciplines et présente des enjeux théoriques très variés. D'un point de vue sociologique il importe notamment de déterminer combien de langues sont actuellement parlées dans le monde. Ici se pose la question de savoir quels critères il convient d'utiliser pour les répertorier et comment on distingue une langue d'un dialecte. À cet égard, nous avons vu que la **compréhension mutuelle** entre interlocuteurs n'est pas un critère tout à fait fiable. La raison en est que deux dialectes voisins dans le **continuum dialectal** peuvent répondre à ce critère tout en appartenant à deux langues différentes. De même, il est possible que deux dialectes éloignés issus de la même langue ne satisfassent pas à ce critère. Qui plus est, il existe également un **continuum historique**. Les écarts, même légers, que produit l'évolution des variantes linguistiques peuvent donner lieu à la naissance de nouvelles langues. Le développement des langues romanes à partir du latin en est un bon exemple. Avec l'apparition du concept d'État moderne, unitaire et centralisé, apparaît aussi celui de **langue officielle** : la décision politique d'attribuer un statut privilégié à un dialecte régional – généralement celui de la zone la plus influente – est lourde de conséquences tant linguistiques que sociologiques. L'**institutionnalisation** en favorise d'autant plus la diffusion qu'elle se voit souvent renforcée par l'imposition de **lois linguistiques**. Pour mesurer l'importance d'une langue, plusieurs paramètres sont dès lors à

prendre en considération : au nombre de locuteurs et de pays s'ajoute le statut politique de la langue et le profil socio-économique des locuteurs natifs.

Une deuxième approche est motivée par le souci interdisciplinaire de mettre en lumière l'origine, l'évolution et la propagation de l'espèce humaine à travers l'étude du développement des différentes langues. Cette approche regroupe les différentes langues à partir de liens de parenté. Établir les **liens de parenté génétiques** entre les langues revient à constituer les familles linguistiques et à reconstituer des **proto-langues** telles que par exemple les langues **indo-européennes**. La reconstitution s'appuie sur l'étude diachronique des **mutations phoniques** et des autres **changements structurels** qui sont à la base de la différenciation entre les langues. À côté des familles linguistiques, d'autres notions ont fait leur apparition, constituant ainsi une taxonomie. Le plus haut niveau de cette taxonomie est appelé **phylum** linguistique, ce qui signifie "ensemble de langues géographiquement ou historiquement défini, ne présentant aucun lien de parenté avec les autres groupes linguistiques". Ces phyla sont constitués de **lignées** dont les membres sont apparentés entre eux de manière éloignée. Ces lignées peuvent, à leur tour, être composées de plusieurs **familles**, comme par exemple la famille des langues indo-européennes. De telles familles peuvent elles aussi être constituées de différents embranchements, comme par exemple la branche des langues Kentum (à l'ouest) par opposition aux langues Satem (à l'est). Les branches peuvent à leur tour se subdiviser en sous-embranchements, comme par exemple le groupe des langues romanes, celui des langues germaniques, celui des langues celtes, etc.

L'étude des critères linguistiques structurels et internes est du ressort de la **typologie linguistique**. Cette troisième approche s'intéresse tout particulièrement aux **universaux linguistiques**. Ceux-ci sont décelables à tous les niveaux linguistiques : au niveau lexical, phonologique, morphologique et syntaxique l'on parle d'**universaux idéationnels**, alors qu'au niveau pragmatique il s'agit d'**universaux interpersonnels**. Ces différentes séries, qui contiennent entre autres des universaux de nature **implicationnelle**, constituent l'ensemble des **universaux substantiels**. En plus de cet ensemble, rebaptisé **universaux paramétriques** par Chomsky, celui-ci lança la notion d'**universaux formels**, représentant l'ensemble des exigences formelles auxquelles une langue doit satisfaire. L'aptitude de l'enfant à apprendre assez rapidement n'importe quelle langue s'explique, selon lui, en vertu d'un **dispositif inné**, contenant notamment l'ensemble des universaux formels. De nos jours, cette théorie a fait place à une vision plus complexe, qui attribue à l'enfant un part bien plus active dans le processus d'acquisition.



Enfin, on peut aussi envisager les approches contrastives. La **linguistique comparée**, qui est spécialisée dans la comparaison de deux ou de plusieurs langues, a des visées plus concrètes. Elle entend faciliter la traduction et l'apprentissage des langues étrangères ainsi que la rédaction de dictionnaires bilingues. L'analyse contrastive peut prendre comme point de départ une forme morphologique dans une **langue-source** – telle que la forme progressive des verbes anglais – et commencer par en inventorier les significations. Ensuite, elle pourra examiner comment ces différentes significations peuvent être exprimées dans la **langue-cible**. Cette méthode permet de décrire les convergences et les divergences entre la langue-source et la langue-cible. Néanmoins, elle ne permet pas de prévoir quelles difficultés surviendront lors de l'apprentissage. Quand on apprend une langue étrangère, on doit être prêt à abandonner ses anciens schémas et prototypes, et à les réviser constamment pour y intégrer de nouveaux faits linguistiques. Pour étudier à fond les variantes possibles d'un mot, d'un groupe de mots ou d'une expression, la linguistique comparée ne peut se passer de grands **corpus parallèles**, tant oraux qu'écrits. Ceux-ci permettent d'en répertorier les différents contextes possibles et de faire apparaître quelles sont les **collocations** typiques en fonction du type de discours ou du genre de textes étudiés. Le dépouillement de ces données permet de distinguer les emplois centraux de ceux qui ne le sont pas, et de mettre au point l'instrument de travail précieux que constituent les réseaux radiaux correspondants pour les langues à étudier. Pour l'analyse comparée des verbes, l'on se tournera par ailleurs vers l'analyse en termes de **configurations de rôles** sémantiques : s'il y a équivalence entre les rôles sémantiques des participants, il y aura également équivalence conceptuelle. Dans la mesure où ce genre d'analyse cesse d'être opérationnel quand on a affaire à une tournure idiomatique – vu que ce qui est idiomatique s'écarte par définition des schémas constructionnels productifs –, l'on peut dire qu'en procédant de la sorte on voit clairement où commence le domaine de l'idiomaticité.

## 10.6 Lectures conseillées

Les comparaisons linguistiques présentées dans Comrie (1987) et Coulmas (1995) mettent en lumière le statut sociologique des langues. Aitchison (1996) et Beakin (1996) sont de bonnes introductions à la recherche interdisciplinaire concernant l'origine, l'évolution et la diffusion des langues. Robins (1973) et Hoenigswald (1973) informent sur l'histoire de la classification des langues. Un excellent atlas des langues du monde est celui de Moseley et Asher (1994). Ramat (1987) propose une initiation à la typologie linguistique, Shopen (1985) en donne une vision d'ensemble. Les premiers ouvrages classiques sur les universaux linguistiques concernant l'ordre des mots sont ceux de Greenberg

(1963, 1966<sup>2</sup>, 1973). Comrie (1981) et Goddard & Wierzbicka (éd. 1994) ont également étudié les universaux sémantiques. Weinreich (1953) constitue un ouvrage de référence sur le phénomène du contact entre les langues. Une bonne introduction à l'analyse contrastive est celle de James (1980). Krzeszowski (1990) offre une étude plus théorique sur la linguistique comparée. Paulussen (1995) propose la description d'un corpus parallèle trilingue.

## 10.7 Applications

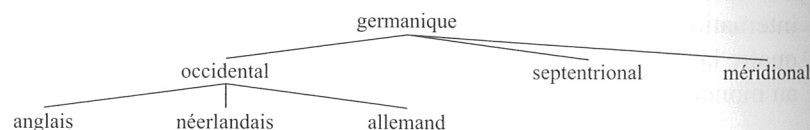
1. Nous avons vu qu'à l'intérieur d'une langue on trouve un continuum dialectal. Pensons à la Grande-Bretagne, l'Italie... Les aires dialectales y sont relativement bien démarquées, ce que vous avez certainement pu remarquer lors de vos voyages.  
Avez-vous fait les mêmes observations en parcourant la France ? Selon vous, la distance géographique y joue-t-elle autant que dans d'autres pays ? Avez-vous observé que le continuum dialectal est modulé par d'autres facteurs ? par l'âge ? la profession ? la formation ?...  
(Exemple : Il y a moins de différences entre un médecin lillois et un médecin marseillais qu'entre un agriculteur du Vaucluse et de Lorraine.)
2. Situez sur la carte de la figure 1 les différents phyla linguistiques du tableau 5.
3. Expliquez à partir des faits repris dans le tableau 3 pourquoi l'anglais et le français sont les deux langues les plus internationales du monde. Qu'est-ce qui les différencie entre elles ? Qu'est-ce qui les différencie de l'arabe et de l'espagnol ? Seriez-vous tenté de dire que l'espagnol est "plus international" ? Convient-il de relativiser quelque peu l'importance de ces quatre langues ? Qu'est-ce qui empêche le chinois, la langue la plus parlée au monde, de devenir la première langue internationale ?
4. Les expressions concernant la langue ou son évolution contiennent beaucoup de métaphores qui bien souvent renforcent la métaphore conceptuelle « la langue est un organisme vivant ». Nous pensons par exemple à des expressions telles « la mort d'une langue », « la naissance d'une nouvelle langue », « le maintien de la langue », etc. Essayez de trouver dans des ouvrages sur l'évolution de la langue comme ceux de Aitchison (1991, 1996, 1997), Beakin (1996) ou même de Darwin (1859) d'autres métaphores de ce type. Voyez si d'autres métaphores conceptuelles font également leur apparition dans ce domaine.

Voici un fragment qui peut servir de point de départ :

Il reste une préoccupation à ajouter, celle de la disparition des langues. Quarante-vingt-dix pour cent des langues du monde sont peut-être en danger. Près de 6 000 langues sont couramment parlées dans le monde. La moitié d'entre elles sont moribondes, étant donné qu'elles ne sont plus apprises par la nouvelle génération de locuteurs. De 2 500 autres, l'on peut dire qu'elles se trouvent en zone de danger, dans la mesure où elles sont parlées par moins de cent mille personnes. Il en reste une soixantaine, à peine un dixième de l'ensemble actuel, qui feront figure de survivants dans un siècle d'ici. Certaines langues finiront inévitablement par se scinder, à l'image du latin qui a fini par se répartir sur les différentes langues romanes. De nouvelles langues pourront ainsi émerger. Mais la diversité sera beaucoup plus réduite. De la magnifique gerbe composée de l'éventail actuel des langues, il ne restera qu'un petit bouquet de quelques différentes fleurs tout au plus. (traduction de Aitchison 1997 : 95)

5. Quelles sont les langues européennes de la famille indo-européenne (voir tableau 6) qui ont le statut de langues officielles, et lesquelles ne l'ont pas ? Soulignez les langues non officielles. Consultez éventuellement Ammon (réd. 1994) ou Haarmann (1993).
6. Prenons (a) le grec, (b) le finnois, (c) le gallois. À quelles familles et à quelles branches appartiennent ces trois langues européennes ? Consultez les tableaux 5 et 6. Voyez-vous quelles sont leurs langues "sœurs" ? Disposez-vous d'assez d'information pour constituer l'arbre généalogique ?

Voici, à titre d'exemple, l'arbre généalogique qui correspond à l'anglais et aux langues "sœurs" que sont le néerlandais et l'allemand :



7. Voici quelques contextes où apparaît le verbe *compter*. Ils sont suivis de leur traduction anglaise.

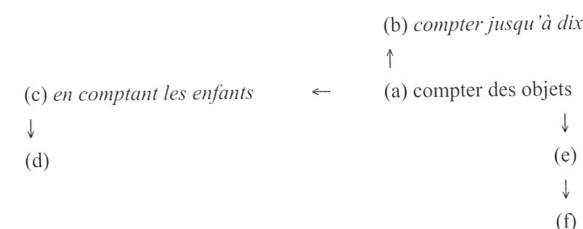
(i) français

- a. Nous devons encore compter les enfants.
- b. Mon petit-fils sait déjà compter jusqu'à 10.
- c. Il y avait 50 personnes, en comptant les enfants.
- d. Ma famille compte 150 personnes.
- e. Pour lui il n'y a que l'argent qui compte.
- f. Ce goal ne compte pas.

(ii) anglais

- a. We must still count the children.
- b. My grandson can already count till 10.
- c. There were 50 people, counting the children.
- d. That family comprises 150 members.
- e. Money is the only thing that matters to him.
- f. That goal is not valid.

- (a) Complétez le réseau radial de *compter* donné ci-dessous à partir des exemples de (i). Comparez avec le réseau radial du verbe anglais *to count* (tableau 15).



- (b) La configuration des rôles est-elle la même pour les phrases (d), (e), et (f) dans les deux langues ?
- (c) Voyez-vous une raison qui explique pourquoi les significations qui correspondent à (d), (e), et (f) ne sont pas rendues par *to count* ?
- (d) Voici quelques emplois figurés de *compter*. Que nous apprend la traduction anglaise ? En quoi les réseaux radiaux de *compter* et de *to count* se distinguent-ils ?
  - (i) a. Il faut bien compter dix jours pour faire cela.
  - b. On ne compte plus ses gaffes.
  - c. C'est un succès qui compte.
  - d. Ils nous apportèrent leurs connaissances, sans compter leur bonne volonté.
  - e. Ils n'ont pas compté le café.
  - f. La ville compte quelques très beaux monuments.
  - g. On compte ce roman parmi les meilleurs de l'année.
  - h. Ses jours sont comptés.
  - i. C'est le résultat qui compte.
  - j. On avait compté sans la grève.
- (ii) a. You must {allow / reckon on it taking} a good ten days.
- b. We've lost count of his blunders.

- c. It's an important success
- d. They gave us their knowledge, {not to mention / to say nothing of} their goodwill.
- e. They didn't charge for the coffee.
- f. The town has some very beautiful monuments.
- g. This novel {is considered to be / ranks} among the best of the year.
- h. His days are numbered.
- i. It's the result that {matters / counts}.
- j. We hadn't {reckoned on there being a strike / allowed for the strike}.

## SOLUTIONNAIRE

### Chapitre 1. La base cognitive du langage

#### 1.6.1.

- (a) Signe symbolique : le triangle inversé signale par convention qu'il faut céder la priorité.
- (b) Signe iconique : image indiquant le danger potentiel de chutes de pierres.
- (c) Symbolique : le sens attribué aux signes de l'alphabet morse est purement conventionnel.
- (d) Indexical : le givre est un indice de gel.
- (e) Symbolique : la correspondance entre l'aiguille et les chiffres indique conventionnellement la vitesse de la voiture.
- (f) Indexical : le bruit de l'alarme qui se déclenche révèle la présence de quelqu'un.
- (g) Indexical : les pleurs signalent que le bébé a besoin de quelque chose.
- (h) Indexical : en agitant la queue le chien montre son contentement.
- (i) Symbolique : (le port de) l'alliance est une convention.
- (j) *Indexical* : le poing levé est expression de colère et/ou de menace, modulable par des expressions faciales ou corporelles. Le poing serré a également acquis une valeur *symbolique* : il est signe d'appartenance à un mouvement social, et plus spécifiquement au parti communiste ou socialiste. De façon générale, il est pris conventionnellement comme signe de désaccord.
- (k) Symbolique : celui qui porte un piercing entend exprimer une certaine attitude.