

## INTRODUCTION

«Toute la terre avait une seule langue et les mêmes mots. Comme ils étaient partis de l'orient, ils trouvèrent une plaine au pays de Schinear, et ils y habitèrent. Ils se dirent l'un à l'autre: "Allons! Faisons des briques et cuissons-les au feu." Et la brique leur servit de pierre, et le bitume leur servit de ciment. Ils dirent encore: "Allons! Bâtissons-nous une ville et une tour dont le sommet touche au ciel, et faisons-nous un nom, afin que nous ne soyons pas dispersés sur la face de toute la terre." Le Seigneur descendit pour voir la ville et la tour que bâtissaient les fils des hommes. Et le Seigneur dit: "Voici, ils forment un seul peuple et ont tous une même langue, et c'est là ce qu'ils ont entrepris; maintenant rien ne les empêcherait de faire tout ce qu'ils auraient projeté. Allons! Descendons, et là confondons leur langage, afin qu'ils n'entendent plus la langue les uns des autres." Et le Seigneur les dispersa loin de là sur la face de toute la terre; et ils cessèrent de bâtir la ville. C'est pourquoi on l'appela du nom de Babel, car c'est là que le Seigneur confondit le langage de toute la terre, et c'est de là que le Seigneur les dispersa sur la face de toute la terre.» (Genèse, XI)

Si l'on en croit la Bible, il fut un temps où les êtres humains, parlant une seule et même langue, entreprirent de construire une tour qui pût monter jusqu'aux cieux. Mais, pour les punir de leur orgueil, Dieu sema la « confusion des langues », les hommes ne purent plus se comprendre, abandonnèrent la construction de la

tour et se dispersèrent. Ce mythe généra au xvi<sup>e</sup> siècle un débat suivi d'une sorte de guerre de succession<sup>1</sup>. Le débat reposait sur l'idée que, pour retrouver le Paradis perdu d'avant Babel, il fallait retrouver la langue prébabélique, la langue adamique, ou celle qui s'en rapprochait le plus. Dans cette compétition étaient en lice trois candidats, le latin, le grec et l'hébreu, parmi lesquels l'hébreu venait le plus souvent en tête, considéré comme la langue première. Puis apparut la guerre de succession : quelle était la langue vulgaire la plus proche de cette langue première ? Les Allemands, les Français et les Italiens voulaient chacun voir dans leur propre langue le vainqueur de cette recherche en paternité, et, si les arguments avancés étaient tout sauf scientifiques, on voyait aisément derrière ces débats un mélange de théologie et de nationalisme : retrouver la langue adamique, ou celle qui s'en rapprochait le plus, relevait de la théologie ; chercher à prouver que sa langue venait en ligne directe de cette langue première relevait d'un conflit nationaliste. Ces tentatives de retrouver la langue des origines ont été très vite abandonnées, mais le mythe de Babel, lui, est demeuré.

Comme une feuille découpée en de multiples petits morceaux de papier, réduite à l'état de confettis, la langue adamique aurait donc été pulvérisée en une multitude de « confettis linguistiques », les « confettis de Babel ». Bien sûr, l'épisode de la tour n'est qu'un mythe, mais qui s'est répandu dans le monde entier et a inspiré de nombreux poètes, romanciers et peintres. Or, il faut prendre les mythes au sérieux car ils façonnent les mentalités, les idéologies.

Ainsi, de ce mythe de Babel découlent deux idées qui ont fortement marqué et continuent d'imprégner notre vision des langues : d'une part, l'idée de la *malédiction*

---

1. Voir L.-J. Calvet, *Linguistique et Colonialisme*, Payot, 1974, p. 17-22.

*de Babel*, le plurilinguisme étant considéré comme une punition divine et la *babélisation* devenant un terme péjoratif, équivalant pour les langues à ce qu'est la *balcanisation* pour les territoires ; d'autre part, l'idée que, pour mener à bien une entreprise commune, il faut une langue commune. De la même façon que Dieu, en semant la confusion des langues, aurait rendu impossible la poursuite de la construction de la tour, il faudrait, pour mener à bien un quelconque projet, avoir la même langue.

D'une certaine façon, l'histoire politique des langues est liée à ces deux idées, les uns mettant en avant le fait que le droit à la langue, à sa langue, si locale et minoritaire soit-elle, est imprescriptible, les autres invoquant un principe de réalité selon lequel les « grandes » langues sont incontournables. D'un côté, chacun défend son confetti ; de l'autre, on considère qu'il est des confettis plus importants ou plus utiles que d'autres. Ces positions, nous le verrons, sont toutes deux défendables.

Quoi qu'il en soit, dans le monde réel et plurilingue qui est le nôtre, l'histoire a lentement tendu à faire coïncider langues et frontières dans les États les plus anciens, en particulier les États européens. Un modèle centralisateur s'est ainsi répandu, avec l'idée qu'à un territoire politique donné doit correspondre une langue donnée, modèle dont on peut lire les effets dans le paradigme étymologique unique présidant au nom d'un pays, de son peuple et de sa langue. En France, il y a des Français qui parlent français ; en Allemagne, des Allemands qui parlent allemand ; en Chine, des Chinois qui parlent chinois ; au Japon, des Japonais qui parlent japonais, etc. C'est en quelque sorte le principe *loquor ergo sum*, « je parle donc je suis », ou plutôt « je dis qui je suis ».

Mais ce système de dérivation montre partout ses limites. D'une part, on ne parle pas que le français en France ou que le chinois en Chine. D'autre part, les

Suisses ne parlent pas suisse, les Sénégalais ne parlent pas sénégalais, les Indiens ne parlent pas indien... Chacun des « confettis linguistiques » de Babel n'est pas retombé sur un territoire délimité par ses frontières ou n'a pas créé ce territoire : ce sont les frontières tracées par l'Histoire et les politiques menées à l'intérieur de ces frontières qui ont agi sur les langues. Mais les langues transgressent les frontières. Certaines s'exportent, se répandent le long des voies de communication... Les pistes, les fleuves, les routes maritimes sont le vecteur de diffusion de langues : le malais dans les ports de l'archipel indonésien, le swahili le long des côtes est de l'Afrique, avant de traverser le continent vers l'ouest, le quichua le long de la cordillère des Andes, le lingala le long du fleuve Congo... D'autres langues sont valorisées par l'histoire politique, diplomatique, commerciale ou coloniale : le français en Europe aux XVIII<sup>e</sup> et XIX<sup>e</sup> siècles, l'espagnol et le portugais en Amérique du Sud, l'anglais aujourd'hui dans le monde... Ces langues prennent ainsi un *poïds* particulier, s'imposant aux autres, sans cependant réussir à oblitérer les spécificités locales.

Aujourd'hui, dans n'importe quel aéroport du monde, on lit des indications en deux ou trois langues, anglais, espagnol et basque à Bilbao, arabe, français et anglais à Casablanca, chinois et anglais à Hong Kong. Dans les avions, les informations sont généralement diffusées dans la langue du pays de la compagnie aérienne, dans celle du pays destinataire et en anglais. Mais les technologies de l'information et de la communication offrent aux « confettis de Babel » de nouvelles possibilités de diffusion. Internet, que certains avaient d'abord perçu comme devant être le lieu privilégié de la domination de l'anglais, est devenu au contraire un véritable forum plurilingue. Le téléphone, la télévision, Skype, les messages électroniques, Google, etc., sont désormais comme des autoroutes linguistiques le long

desquelles les langues se déplacent sans cesse, s'entrechoquent, se frottent.

Mais il demeure que, même si elles se déterritorialisent, les langues ont bien sûr un territoire d'origine, sur lequel elles continuent de fonctionner et de coexister localement avec d'autres langues. Entre un maelström mondial, un tourbillon qui s'apparente à un phénomène de gravitation, et la gestion, *in vivo* comme *in vitro*, dans les pratiques sociales comme dans les politiques linguistiques, du plurilinguisme dans les frontières d'un État, la diversité linguistique est à la fois locale et globale. Il est désormais impossible de penser une politique linguistique nationale en faisant abstraction de la situation mondiale des langues, et les règlements linguistiques des organisations internationales comme les politiques des grands ensembles linguistiques (anglophonie, francophonie, lusophonie...) ont des retombées sur les situations et sur les politiques locales.

Mais qu'est-ce que la *diversité linguistique*? Le mot *diversité*, issu du latin *diversitas*, a d'abord signifié «divergence», «contradiction», avant de prendre le sens courant aujourd'hui de «différence», «variété». Le mot est donc synonyme de *pluralité* et s'opposerait à l'uniformité, l'homogénéité ou la monotonie. On parle ainsi en France de gens «issus de la diversité», de «prise en compte» ou de «respect» de la diversité, de «lutte contre la discrimination et pour le respect de la diversité», étant sous-entendu que cette diversité doit se retrouver, en quelque sorte proportionnellement, dans les différentes instances de la société. *Le Monde* du 18 juin 2012 titrait par exemple «La diversité progresse dans l'Hémicycle», notant que, «essentiellement "blanche" jusqu'à présent, l'Assemblée nationale va refléter un peu plus les couleurs de la France après les élections législatives de dimanche 17 juin. Huit députés d'origine africaine, maghrébine, asiatique ou brésilienne vont rejoindre les

élus d'outre-mer ou originaires des départements sur les bancs du Palais-Bourbon».

À côté de cette diversité conçue comme pluralité anthropologique ou ethnique, on parle aussi de biodiversité, de diversité génétique, de diversité culturelle. Et c'est dans ce cadre qu'apparaît la notion de *diversité linguistique*. Elle est d'abord un fait. Nous verrons dans le premier chapitre de ce livre que l'on parle à la surface du globe terrestre près de 7 000 langues différentes, diverses dans leurs structures (les «familles» linguistiques, par exemple, présentent des caractéristiques communes à leurs membres et les différencient des membres d'autres «familles»), mais, surtout, qui assurent par leur nombre une diversité synonyme de pluralité. Entendue ainsi, la *diversité linguistique* signifie tout simplement qu'il existe «beaucoup» de langues et qu'elles sont «différentes». Les «confettis» de Babel sont nombreux et ne sont pas les dérivés immédiats ou lointains d'une supposée langue adamique.

Mais ce fait, ou cette situation, est en mutation. La *diversité linguistique* est considérée comme menacée, on parle de mort de langues, de langues en danger, comme on parle des menaces portant sur la biodiversité. La défense de la *diversité linguistique* est ainsi devenue une sorte de slogan, de discours politiquement correct qui tend à s'appuyer métaphoriquement sur une conception de l'écologie détournée de son sens premier. Si l'écologie est l'étude scientifique de l'environnement, de l'habitat, alors il peut exister une *écologie linguistique*, mais elle ne doit pas être confondue avec les discours militants – et, au demeurant, parfaitement recevables – sur la protection des espèces, par exemple. La protection de l'environnement inclut-elle la protection des langues en danger? Peut-on considérer que la diminution du nombre des langues peut mettre en danger la survie de l'espèce humaine, de la même façon que le réchauffement climatique ou une

croissance démographique exponentielle? Peut-on parler de *minorités audibles*, comme on parle de *minorités visibles*?

La *diversité linguistique* est enfin un projet, ou un programme d'action. De nombreuses organisations internationales en ont fait leur cheval de bataille, non qu'elles veuillent multiplier encore le nombre des langues, mais au contraire parce que, craignant la disparition de beaucoup d'entre elles, elles militent pour leur préservation. On peut citer ici la déclaration de l'Unesco sur la diversité culturelle, qui ne fait qu'implicitement référence aux langues, ou la Charte européenne des langues régionales et minoritaires, qui appelle explicitement à leur protection. La diversité linguistique est donc d'abord un fait, une situation. Elle est ensuite une situation menacée, qui a mené à des prises de position (et, parfois, à des discours idéologiques) portant sur la nécessité de la défendre. Enfin, la diversité linguistique est l'objet de résolutions, de programmes, d'actions, dont nous présenterons une analyse critique.

Après avoir présenté la situation linguistique du monde (chapitre 1), le poids des différentes langues dans le monde (chapitre 2) et l'utilité de cette approche pour la détermination de politiques linguistiques (chapitre 3), nous analyserons le problème des langues en danger (chapitre 4), verrons ce qu'est la diversité linguistique (chapitre 5), puis nous interrogerons sur la possibilité d'élaborer des politiques pour le maintien de la diversité (chapitre 6) et sur l'avenir linguistique du monde (chapitre 7).





## LA SITUATION LINGUISTIQUE DU MONDE

On compte aujourd'hui dans le monde environ 7000 langues. Ethnologue.com, le site du SIL (Summer Institute of Linguistics<sup>1</sup>), en dénombrait 7105 en 2013, précisant qu'il s'agissait de «*known living languages*», de «langues vivantes connues». Or, en 1996, il annonçait 6703 «langues vivantes connues». Faut-il en conclure que leur nombre aurait augmenté en une quinzaine d'années? Il est certes possible que de nouvelles langues apparaissent. Il arrive également que l'on découvre dans des régions reculées des langues que l'on n'avait pas encore répertoriées. Mais la différence entre les chiffres de 2012 et ceux de 1996 repose sur tout autre chose...

D'une part, la façon de recenser les langues varie selon des critères qui peuvent être scientifiques, politiques ou idéologiques. Ainsi, les linguistes spécialisés dans les études chinoises considèrent qu'il y a sept langues chinoises différentes, alors que les discours officiels chinois stipulent qu'il n'y en a qu'une, le mandarin (ou putonghua, 普通話), et que les six autres (le hakka, le wu, le yué, le min, le xiang et le gan) sont des dialectes. Le site Ethnologue.com, pour sa part, recense «treize langues chinoises». La différence de classification (une langue et six dialectes dans un cas, sept langues dans l'autre) tient ici à une opposition entre un point de vue

---

1. [www.ethnologue.com](http://www.ethnologue.com)

scientifique et une position idéologique. Mais il peut aussi se produire que des formes un temps considérées comme des dialectes d'une même langue soient ensuite considérées comme des langues différentes, ou que des formes autrefois considérées comme une seule langue soient subdivisées en plusieurs langues. Il y a alors une « augmentation » du nombre de langues, produite par le regard que les linguistes ou les locuteurs portent sur elles. Examinons rapidement trois exemples.

### UNE LANGUE OU DES LANGUES ?

Le premier exemple concerne la France, où l'on distingue traditionnellement entre langue d'oc et langue d'oïl. Mais faut-il mettre ces deux appellations au singulier ou au pluriel ? Faut-il, par exemple, parler d'une langue occitane (qui aurait des variantes régionales) ou de plusieurs langues qui seraient le languedocien, le gascon, le limousin, l'auvergnat, le provençal, le provençal alpin et le niçard ? Ce débat donne parfois lieu à des contradictions étranges. Ainsi, le 17 octobre 2003, le conseil de la région Paca adoptait un texte disant que « la langue provençale et la langue niçoise sont les langues régionales de la région Provence-Alpes-Côte d'Azur ». Il s'agit donc, pour les conseillers, ou du moins pour la majorité d'entre eux, de deux langues. Mais, le 5 décembre de la même année, le même conseil adopte un autre texte précisant que « la langue occitane ou langue d'oc est la langue régionale de la région Provence-Alpes-Côte d'Azur : le provençal rhodanien, le provençal maritime, le niçard et l'alpin sont les formes régionales de la langue occitane ou langue d'oc ». Le provençal et le niçois ne sont donc plus des langues, mais des formes différentes d'une même langue, l'occitan.

Alors, plusieurs langues ou une seule langue ? Linguistes et groupes militants s'opposent sur ce point, et

il est évident que ces débats ont des retombées sur le nombre officiel de langues parlées dans le sud de la France. Le linguiste Jean-Marie Klinkenberg décrit une seule langue divisée en trois grands groupes dialectaux (gascon, nord-occitan, moyen occitan<sup>1</sup>). Un autre linguiste, Bernard Pottier, divise également ce qu'il appelle les «parlers d'oc» en trois zones (gascon, oc du Nord, oc du Sud), mais il parle à propos de cet ensemble d'un «certain nombre de dialectes romans présentant des traits linguistiques communs<sup>2</sup>» – sans jamais utiliser le mot «langue». Bien entendu, on peut se demander s'il y a intercompréhension entre ces différentes formes, mais ce critère n'est pas toujours suivi. Ethnologue.com, par exemple, ne liste que l'occitan et précise qu'il comprend des dialectes – l'auvergnat, le gascon, le languedocien, le limousin et le provençal –, mais ajoute que «la situation dialectale est très fragmentée, avec une intelligibilité limitée entre quelques variétés». Et le linguiste Philippe Blanchet écrit, pour sa part, que «norvégien et suédois, hindi et ourdou, valencien et catalan, corse et italien, provençal et occitan fonctionnent comme langues distinctes malgré une proximité typologique et une certaine intercompréhension<sup>3</sup>».

On peut aussi rechercher dans l'Histoire des traces anciennes de cette (ou de ces) langue(s), mais rien n'est très probant. L'ordre des chevaliers de Saint-Jean, devenu ordre de Malte, par exemple, fut organisé en 1319 d'une façon originale : les chevaliers étaient regroupés selon leur origine non pas géographique ou politique mais linguistique, en «langues», l'idée étant de les

---

1. J.-M. Klinkenberg, *Des langues romanes*, Louvain-la-Neuve, Duculot, 1994, p. 223-228.

2. B. Pottier, «La situation linguistique en France», in *Le Langage*, Bibliothèque de la Pléiade, 1968, p. 1154-1158.

3. P. Blanchet, «Pour une (socio)linguistique de la complexité», *Cahiers de sociolinguistique*, n° 8, Presses universitaires de Rennes, 2003, p. 285.

regrouper de sorte qu'ils puissent communiquer entre eux. Or, ces langues, au nombre de huit, étaient les langues de Provence, d'Auvergne, de France, d'Italie, d'Aragon, d'Angleterre, d'Allemagne et de Castille. Auvergnat et provençal étaient donc considérés à l'époque comme deux langues différentes, et il faudrait admettre qu'elles se seraient ensuite fondues en une seule, la langue d'oc. Une langue ou des langues? Il n'est décidément pas facile de répondre à cette question.

Deuxième exemple avec ce que l'on appelle aujourd'hui l'ex-Yougoslavie. Il s'agissait d'un pays fondé en 1918 sous le nom de Royaume des Serbes, Croates et Slovènes, puis rebaptisé en 1929 Royaume de Yougoslavie, devenu République fédérative populaire de Yougoslavie après la Seconde Guerre mondiale, puis, en 1963, République fédérative socialiste de Yougoslavie, avant de connaître en 1992 la sécession de quatre républiques fédérées, la Bosnie-Herzégovine, la Croatie, la Macédoine et la Slovénie. Dans le même temps fut créée une République fédérale de Yougoslavie, regroupant la Serbie et le Monténégro, qui prit en 2003 le nom de Communauté d'États Serbie-et-Monténégro et fut dissoute en 2006, à l'indépendance du Monténégro. Il existe une contrepartie linguistique à cette histoire politique tortueuse : de 1945 à 1992, la langue officielle de la République étaient le serbo-croate (parfois appelé croato-serbe), le macédonien, l'albanais et le slovène étant également langues officielles, tandis que le hongrois et le roumain étaient enseignés dans certaines régions seulement. Mais «serbo-croate» était l'appellation la plus commune, et c'est sous ce nom qu'on l'enseignait dans les universités étrangères, en particulier en France, à l'Inalco (Langues'O). Les linguistes distinguaient dans cet ensemble «serbo-croate» entre différents dialectes, selon la façon de dire *quoi*: le kajkavien (parlé dans la région de Zagreb, où *quoi* se dit *kaj*), le tchakavien (parlé sur la côte Adriatique, où *quoi*

se dit *tcha*) et le chtokavien (parlé dans une partie de la Croatie, en Serbie, au Monténégro et en Bosnie-Herzégovine, où *quoi* se dit *chto*). Mais ces différences ne correspondaient à aucune division politique ou nationale; il s'agissait simplement d'une variation géographique. Ajoutons à cela que les Croates, catholiques, écrivaient le «serbo-croate» en caractères latins, tandis que les Serbes et les Monténégrins, orthodoxes, l'écrivaient en caractères cyrilliques. Enfin, entre la Serbie et la Croatie, il y avait quelques différences lexicales. Mais tout cela n'empêchait pas les gens de se comprendre! Après les différentes sécessions et indépendances, chacun des nouveaux pays appelle la langue d'un nom différent: le serbe, le bosniaque, le croate, le monténégrin... Nous pouvons donc avancer que, dans cette partie du monde, le nombre des langues a «augmenté», cette augmentation n'étant pas due au fait que le «serbo-croate» a subitement éclaté en de nouvelles formes, mais due à des événements politiques et à des sentiments nationalistes.



**Billet de banque yougoslave émis en 1978  
et rédigé dans les langues officielles**

Notre troisième exemple concerne une langue que l'on appelait en Inde, au début du xx<sup>e</sup> siècle, l'hindoustani.

Cette langue était parlée dans le nord de l'Inde par des populations hindouistes et musulmanes. Les premières l'écrivaient en caractères devanagari (l'alphabet du sanskrit) et les secondes, en caractères arabes. Au moment de l'indépendance de l'Inde, en 1947, le pays fut divisé en deux. Le Pakistan a pris l'hindoustani comme langue officielle, sous le nom d'ourdou, et l'Inde l'a pris comme l'une de ses langues officielles, sous le nom de hindi. L'ourdou emprunte systématiquement des mots à l'arabe ou au persan, le hindi emprunte pour sa part au sanskrit, et les deux formes divergent donc chaque jour un peu plus. Là aussi, le nombre des langues a donc « augmenté », pour des raisons religieuses et politiques. Du fait de l'Histoire, des oppositions religieuses, des haines et des sentiments identitaires, ce qui était considéré à une certaine époque comme une langue, l'hindoustani, fut considéré ultérieurement comme deux langues : hindi et ourdou.

Le parallélisme entre cet exemple et le précédent, celui du serbo-croate, saute aux yeux. Il est d'ailleurs intéressant de noter que, dans la citation donnée plus haut à propos de l'occitan et du provençal, Philippe Blanchet ne citait pas que ces deux langues, mais « norvégien et suédois, hindi et ourdou, valencien et catalan, corse et italien, provençal et occitan ».

Ces trois exemples, rapidement évoqués, présentent donc des points communs. Mais que nous apprennent-ils ? Qu'il n'est pas facile de distinguer entre langues et dialectes ? Que ces classifications peuvent être arbitraires, ou idéologiques ? Que la promotion d'un parler au statut de langue peut être politique ? Que les locuteurs peuvent décider de ce statut ? Tout cela à la fois, certes. Mais il faut bien parvenir à nous entendre, à savoir de quoi nous parlons. Aussi, pour éviter de retomber sans cesse dans le même débat, avons-nous choisi de prendre en compte les langues listées dans le code ISO 639. Il s'agit d'un code « alpha-3 », qui répertorie à ce jour 7 480 langues du monde – parmi lesquelles

certaines sont disparues et d'autres sont des «macro-langues» – et leur attribue un code de trois lettres, par exemple *fra* pour le français, *urd* pour l'ourdou, *brv* pour le croate, *oci* pour l'occitan, etc.

Nous ne sommes pas nécessairement toujours d'accord avec les décisions de cette classification – nous y reviendrons notamment à propos de l'arabe –, mais du moins a-t-elle le mérite d'exister et de constituer une liste de référence.

### **DERRIÈRE LES LANGUES, DES LOCUTEURS**

Une chose est de savoir combien de langues sont parlées à la surface du globe ; une autre est de savoir combien d'individus les parlent. Et se pose ici une première question : qu'est-ce que veut dire «parler une langue»? Quels rapports y a-t-il entre un francophone natif, un Anglais ou un Allemand qui aurait étudié le français au cours de sa scolarité, et un commerçant turc qui, dans le bazar d'Istanbul, peut interpeller en français des clients potentiels, plaisanter avec eux, annoncer ses prix et marchander? Pouvons-nous considérer qu'ils parlent la même langue, de la même façon? Le simple bon sens nous indique que les uns ont le français pour langue maternelle (ou langue «première»), que les autres l'ont appris comme langue étrangère, qu'ils le pratiquent plus ou moins bien, l'utilisent pour des raisons culturelles, professionnelles ou encore touristiques. Mais il faut ajouter à ces catégories celle des locuteurs qui, sans avoir le français pour langue première, l'utilisent tous les jours parce qu'ils vivent dans un pays dont il est la langue officielle. C'est à la fois le cas des Alsaciens qui parlent alsacien chez eux mais vivent le reste de leur vie en français, ou des Sénégalais qui ont le français pour langue officielle, l'apprennent à l'école, l'utilisent au bureau, mais ont pour langue première le joola, le

sereer ou le peul, et peuvent en outre parler le wolof, langue dominante dans leur pays. Compter le nombre de locuteurs d'une langue est donc une activité malaisée, et il nous faut au moins distinguer entre ceux qui l'ont pour langue première et ceux qui l'ont pour langue seconde ou troisième, le rapport entre ces deux groupes permettant d'ailleurs, comme nous le verrons plus loin, de calculer l'éventuelle *véhicularité* de la langue.

Commençons par le premier groupe, celui des locuteurs de langue première, et considérons quelques-unes des langues déjà citées : l'occitan, le serbe, le macédonien, le mandarin, le français, l'ourdou et le hindi. Combien ces langues ont-elles de locuteurs ? Le tableau ci-dessous répond à cette question, étant entendu que nous ne prenons en compte que les locuteurs ayant ces langues pour langue première et que nos données sont empruntées à Ethnologue.com, qui a tendance à sous-estimer le nombre des locuteurs : certains militants occitans, par exemple, parlent de 7 à 8 millions de locuteurs, certains défenseurs de la francophonie parlent de 120 millions de locuteurs... Mais nous considérerons que, statistiquement, les erreurs et les sous-estimations se répartissent également, et que par conséquent elles ne changent pas grand-chose à la hiérarchie qui apparaît dans ce tableau.

**Quelques langues et leur nombre de locuteurs**

LANGUES	NOMBRE DE LOCUTEURS
Mandarin	845 millions
Hindi	181 millions
Français	67 millions
Ourdou	60 millions
Serbe	7 millions
Macédonien	2 millions
Occitan	2 millions