

## Avant-propos

Je propose dans cet ouvrage une série de chapitres<sup>1</sup>, conçus comme des essais, destinés à mettre en évidence la manière dont les langues fonctionnent, varient et se transforment, autrement dit la dynamique linguistique. Chaque chapitre peut être lu séparément, mais l'ordre n'est toutefois pas aléatoire.

L'ouvrage commence par une réflexion sur les relations entre langues et représentations du monde, sur l'arbitraire du signe, sur le relativisme et l'universalisme, avec en guise d'illustration et d'application une incursion chez des Amérindiens vivant dans le nord-est de l'Arizona, et chez les Eskimos de l'Alaska, du Canada et du Groënland, à propos de controverses sur l'expression et la notion du temps chez les premiers et sur le lexique de la neige chez les seconds. Avec ces premières explorations, on est nettement en dehors de la famille indo-européenne, ce qui donne une idée de la diversité linguistique.

Se présente ensuite une longue série qui va de l'indo-européen aux langues créoles – plus précisément le créole réunionnais – en passant par les langues de l'Antiquité que sont le grec classique et le latin et, bien entendu, également par le français, issu du latin oral courant, populaire, autrefois appelé latin « vulgaire » par les romanistes.

Les problèmes posés sont différents à chaque fois, mais un fil directeur les relie. De prime abord, il peut paraître étonnant au profane que le créole réunionnais, par exemple, provienne d'une proto-langue très ancienne et non attestée, dont l'existence est simplement hypothétique, mais pratiquement certaine, et qu'on a coutume d'appeler l'indo-européen.

En bref, l'histoire racontée ici par petites touches va de cette langue reconstituée par les méthodes de la linguistique comparative jusqu'à un ensemble de langues d'aujourd'hui. Il ne s'agit là que de l'une des nombreuses familles de langues

---

<sup>1</sup> La majorité de ces textes, mais non tous, sont des versions modifiées, améliorées, souvent considérablement enrichies et refondues, de contributions déjà publiées. Les références sont indiquées en note de bas de page au début de chaque chapitre.

parlées sur notre planète, et uniquement d'une bien modeste sélection au sein de cette famille. Mais le but est de montrer par des études de cas comment on a pu passer progressivement, imperceptiblement, d'une langue préhistorique très ancienne à des langues de notre époque, incluant les créoles.

Dans l'approche que je propose, les relations entre langue et discours constituent le socle sur lequel repose le cadre théorique. En effet, à mon sens, une langue est un ensemble systématique d'unités, de structures, de règles, principes et conditions qui permettent le discours. Le discours est à la fois un acte, l'énonciation, et le produit de cet acte, l'énoncé, celui-ci étant la seule manifestation observable par le linguiste, tout comme par les locuteurs eux-mêmes.

Encore faut-il préciser que seule une partie de l'énoncé est perceptible, à savoir sa composante phonétique, qui nécessite elle-même un encodage et un décodage complexes, par le locuteur et par le récepteur du discours, respectivement. Rappelons cette évidence : le son est le véhicule du sens, mais seul le son est perceptible. Le locuteur transforme son intention sémantique en sons, et le récepteur doit pratiquer la procédure inverse, consistant à reconstruire l'intention sémantique de son interlocuteur à partir de l'interprétation des sons que celui-ci a produits. Et dans ce cadre, gardons-nous d'oublier l'implicite !

Ainsi, il n'est possible de connaître la langue qu'indirectement, par l'observation et la description du discours, et si la langue est la condition de la mise en œuvre du discours, inversement c'est le discours effectif qui est la condition de l'appropriation de la langue par les nouvelles générations.

On comprend alors la difficulté du linguiste. Sa tâche consiste en effet à mettre au jour des structures et un fonctionnement qui lui demeurent en grande partie cachés. On est confronté à un paradoxe : les données dont on dispose ne sont pas celles que l'on aimerait observer. L'image des règles du jeu et de la partie que l'on joue est ici parlante : le linguiste est condamné à observer des parties et à en inférer les règles du jeu, qui ne lui sont pas fournies et ne sont pas accessibles. En l'occurrence, il s'agit d'un jeu dont les règles sont particulièrement complexes. Il s'agit même de *jeux*, au pluriel, chaque langue ayant ses propres unités et ses propres règles.

Le linguiste professionnel ne peut donc, dans ces conditions, que pratiquer une méthode hypothético-déductive<sup>2</sup>. Prôner une méthode inductive ou empirico-inductive serait naïf et contre-productif à plusieurs égards. Tout d'abord, ce serait oublier que le regard porté sur les données est déjà lourd d'hypothèses implicites ; en effet, de quelles données s'agit-il ? dans le cas de langues anciennes, nous ne disposons que de traces écrites qui sont des images imparfaites de l'oral, alors que les langues sont avant tout de nature vocale, l'écriture n'étant qu'un code secondaire, malgré son incontestable utilité.

Pour ce qui est des langues actuelles, à l'oral, qui est, répétons-le, prioritaire et primordial, les données sont avant tout des ondes acoustiques : or par le simple fait de les écouter, lorsqu'il s'agit d'une langue connue du descripteur, s'enclenche déjà chez celui-ci une construction d'hypothèses embryonnaires. Pour en être convaincu, il suffit d'écouter une langue dite étrangère, et l'on s'aperçoit alors que la réception des données, c'est-à-dire, j'y insiste, le décodage des ondes acoustiques, est un échec, ce qui revient à dire qu'aucune hypothèse sérieuse n'est possible sans l'aide d'un témoin locuteur de cette langue.

Lorsque le linguiste observe un discours produit dans une langue connue de lui, il n'est jamais dans une démarche inductiviste ou purement empiriste : il observe en effet en fonction d'un projet, qui oriente son écoute ou son regard.

En fait, dans tous les cas de figure, l'hypothèse ne peut jamais venir en fin d'un parcours d'observations : elle est nécessairement initiale, quelle que soit la langue étudiée, familière ou non. Le projet du linguiste, dans cet esprit, est la

---

<sup>2</sup> J'avoue être un fervent adepte de Karl Popper et j'approuve son concept épistémologique de réfutabilité : voir Karl R. Popper, *La logique de la découverte scientifique*, Nicole Thyssen-Rutten et Philippe Devaux (trad.), Paris, Payot, [1959] 1973 (*The Logic of Scientific Discovery*, Londres, [Hutchinson & Co, 1959], Londres/New York, Routledge, 1992 ; *Logik der Forschung*, Vienne (Autriche), Verlag von Julius Springer, 1935), et *Conjectures et réfutations. La croissance du savoir scientifique*, Michelle-Irène et Marc B. de Launey (trad.), Paris, [Payot, 1979], Payot & Rivages, 2006 (*Conjectures and Refutations. The Growth of Scientific Knowledge*, [Londres, Routledge & Kegan Paul, 1963], Londres/New York, Routledge, 2002).

construction d'un modèle explicatif, ce qui va au-delà de la simple description.

Encore quelques mots sur l'induction. Les partisans des méthodes dites empirico-inductives sont condamnés à se limiter au mieux à une description des prétendues données, mais jamais ils ne pourront atteindre la phase intéressante en science, qui est justement celle de l'explication. Or, dans la conception défendue dans les chapitres qui suivent, le but ultime est la phase explicative. Mais les données ne fourniront jamais la clé. Prenons l'exemple de la liaison en français, qui est l'objet de l'un des chapitres de ce volume<sup>3</sup> : seule une argumentation à partir de l'observation des données, avec une interaction dialectique entre les hypothèses et cette observation, permettra au linguiste de proposer un modèle explicatif après avoir compris les mécanismes profonds qui sont à l'œuvre au niveau de ce qu'il convient de considérer comme un véritable inconscient linguistique.

Rappelons aussi que les « vides », les « trous » syntaxiques sont légion dans le discours. Si un hispanophone dit *canta bien* ('il/elle chante bien'), où est le sujet de cette phrase ? Il n'est en tout cas pas dans les données, alors même que, syntaxiquement, il faut bien qu'il y ait un sujet<sup>4</sup> ! Et comment se fait-il que lorsqu'on demande à des francophones quel est le pluriel de *œil*, la majorité des personnes interrogées répondent spontanément *z'yeux* et non *yeux* ? Et d'où sort ce verbe *zieuter* ou *zyeuter*, enregistré par les lexicographes ? Pourquoi de nombreux locuteurs francophones ne font-ils pas la liaison dans *vingt euros, cent euros* ?<sup>5</sup> D'où vient cette innovation consistant à dire *le directeur a demandé à ce que je vienne*, énoncé pourtant plus complexe que la phrase standard *le directeur a demandé que je vienne* ? Comment expliquer cette extension curieuse des emplois de *à ce que* ? Là encore, la

<sup>3</sup> « La liaison en français : une évolution permanente ».

<sup>4</sup> Certains diront que dans *canta bien*, le sujet est le *-a* de *canta* : mais alors, quand je dis *Juanita canta bien* ('Juanita chante bien') si le *-a* est sujet, quelle est la fonction de *Juanita* ? Ce simple exemple montre qu'il est plus satisfaisant de postuler un sujet « zéro » dans *canta bien*. Ce problème est abordé à propos du latin dans le chapitre « Nécessité de l'abstraction : l'exemple de la proposition infinitive en latin ».

<sup>5</sup> Voir le chapitre « La liaison en français : une évolution permanente ».

réponse n'est pas dans les données, et se contenter de constater les faits ne nous avance à rien sur le plan scientifique : ce qui est attendu du linguiste, c'est une explication.

La linguistique telle qu'elle est conçue dans cet ouvrage est une science avant tout argumentative. En outre, il faut comprendre et accepter que les arguments ne sont pas « dans » les données : ils sont le produit du travail intellectuel et rationnel du linguiste.

Le paradoxe de la linguistique est double : c'est une science plutôt « dure », mais dont l'objet est « mou » ; et cet objet est inobservable, parce qu'il est de nature cognitive. On ne trouvera en effet jamais les langues ailleurs que dans les cerveaux ou, si l'on veut, dans les esprits. Le problème est que ceux-ci sont insondables, ce qui rend les langues littéralement inaccessibles. D'où le rôle de l'hypothèse.

Il ne faudrait pas déduire de ce qui précède que le « terrain » est négligé, bien au contraire. On lit trop d'ouvrages ou d'articles prenant pour argent comptant ce qui a déjà été dit, sans vérification et sans enquête. La preuve de cet intérêt pour la vérification se trouve d'ailleurs dans le chapitre sur le relativisme linguistique<sup>6</sup>. À sa lecture, on pourra constater que les mythes linguistiques, inlassablement répétés, y sont dénoncés après une recherche minutieuse, longue et patiente. L'idée sous-jacente associée à cette vérification est la séduction de l'erreur : les mythes sont tenaces parce qu'ils séduisent, et il est difficile de résister à la tentation.

Il va de soi que toute science humaine a nécessairement une base empirique. La vraie question est : comment concevoir et traiter cette base ? Dans cette logique, ce qu'on appelle empirisme consiste en général en *excès* d'empirisme, avec pour conséquence que l'on se prive de la dimension argumentative et explicative, et que l'on tourne le dos à la composante cognitive, pourtant centrale et cruciale quand on s'intéresse au langage. Le défaut majeur de l'empiriste est qu'il n'a pas compris que le discours n'est pas la langue : il n'en est que l'activité en partie seulement observable que celle-ci rend possible. Il est bien de décrire la communication interpersonnelle, mais ce n'est pas

---

<sup>6</sup> « Relativisme, séduction de l'erreur, mêmes et contre-mêmes ».

suffisant : le but ultime est la compréhension et l'explication des structures cognitives qui en sont la condition.

Toutes les réflexions qui précèdent sont en fait liées à une particularité bien précise : la langue n'est pas une « chose », elle n'est pas un objet indépendant de l'esprit, qui aurait une existence autonome et qui serait transmis de génération en génération comme un héritage à préserver. On ne transmet pas une langue à un jeune enfant, c'est lui-même qui la construit.

La faculté de langage, don génétique de l'espèce, permet à l'enfant, lors de son appropriation de la langue, de recréer les structures linguistiques mentales à partir de l'observation et de l'assimilation inconscientes du discours des autres, de son entourage. Cependant, n'ayant d'existence que psychique, la langue n'est jamais recréée à l'identique et l'appropriation produit nécessairement un état global distinct de celui qui a présidé à la production des données discursives observées. C'est en partie de ce fait que naît la variation, dans l'espace et dans le temps, et c'est ce qui explique que ce terme de variation figure dans le titre de l'ouvrage. La diachronie est ainsi au cœur de la réflexion, même si ce n'est pas toujours explicite.

Insistons : l'enfant émet inconsciemment des hypothèses sur les structures de la langue grâce à sa faculté innée. Or un tel mode d'appropriation implique nécessairement des changements linguistiques, de génération en génération, et c'est ainsi que latin, français et créole réunionnais, pour ne prendre que ces exemples, sont les produits de changements incessants et de fragmentations linguistiques depuis des temps très reculés.

On aura compris que dans la conception défendue ici, on ne saurait distinguer une linguistique de la « langue » et une autre de la « parole » : il existe pour l'auteur de ces lignes une seule linguistique, dont l'objet est à la fois la quête de cette langue invisible, et l'ensemble des relations entre langue et discours, ou encore, la manière dont le discours est produit à partir de la langue, vue comme compétence cognitive, ou enfin, la manière dont l'enfant construit sa langue à partir de l'observation inconsciente du discours des autres.

En raison de ces relations entre langue et discours, toute langue varie à tout moment de son histoire. Aucune n'est uniforme : variation et changement sont inhérents à toute langue,

ce qui impose une vision de la synchronie non seulement dynamique, mais également variationniste. Dans cet esprit, la variation est au cœur de la problématique linguistique.

On lit parfois que la langue serait un code. C'est une erreur. En fait, la langue envisagée comme un ensemble de structures mentales ne saurait constituer un code. Les notions de code et de codification s'appliquent uniquement au domaine normatif et prescriptif, et non au domaine cognitif, et il faut se garder de confondre les règles de ces codes sociaux avec les règles linguistiques mentales postulées par le linguiste pour rendre compte des processus cognitifs dans la dialectique des relations entre langue et discours. La « codification » est une régulation secondaire qui se surimpose aux règles purement linguistiques, lesquelles relèvent pour partie des profondeurs de la réalité mentale. Dans le présent ouvrage, cette dimension normative ne sera guère prise en compte, l'objet étant ici le fonctionnement proprement linguistique. Cela ne revient évidemment pas à ignorer le rôle des normes et de la notion de standardisation. Nous verrons leur importance notamment dans le cas de l'histoire complexe du latin.

En tout cas, l'accent sera mis essentiellement sur le fait que les langues se modifient sans cesse et se fragmentent. L'exemple classique est la transformation de variétés non standard du latin en diverses langues romanes : français, occitan, portugais, castillan, catalan, italien, roumain, etc.

Dans ce cadre, la créolisation n'est qu'un cas particulier de processus qui mènent à la création d'une nouvelle langue par un ensemble de locuteurs hétérogènes, à partir de données discursives qui sont initialement étrangères à une partie importante d'entre eux. Les conditions historiques propres à la créolisation ont accéléré et accentué la dérive par rapport aux variétés initiales. Les nouvelles variétés résultant du contact des populations se sont affranchies du français, tout en gardant un « air de famille » avec celui-ci.

Il arrive que le changement linguistique provoque à la longue ce qu'il convient d'appeler un glissement typologique. Ainsi, le français et les créoles à base française sont-ils sortis de la galaxie typologique du latin, dont ils sont issus directement ou indirectement. Ce glissement typologique ne peut se consta-

ter que si l'on confronte des périodes relativement distantes de l'histoire, car le glissement est progressif.

Le latin possédait une flexion nominale casuelle et une flexion verbale riches, un ordre des mots et des groupes relativement libre dans la phrase, pas d'article dans le syntagme nominal, etc. Tout cela a disparu. Le français a pratiquement perdu tout le système casuel, n'en gardant qu'un faible résidu pour les pronoms personnels (*il, le, lui...*) ; le français oral a perdu la flexion nominale, à quelques exceptions près (*cheval, chevaux*, par exemple) ; sa flexion verbale s'est considérablement réduite, avec comme conséquence la diminution du nombre de marqueurs d'accord avec le sujet, ce qui a entraîné l'impossibilité de sujet non exprimé<sup>7</sup> (sauf à l'impératif). La perte de la flexion casuelle a eu pour conséquence un ordre des mots plus contraignant, qui est désormais utilisé comme procédé d'indication des relations syntaxiques. Il est clair que l'on a assisté à un glissement typologique majeur par rapport au latin de l'époque classique.

Enfin, dans les essais qui suivent, on pourra constater qu'une place de choix a été réservée au grec classique et au latin et, d'une manière plus générale, à ce qu'on appelle l'Antiquité. Il y a à cela trois raisons principales : la première est la passion de l'auteur de ces lignes pour les langues anciennes ; la seconde est qu'elles révèlent un univers d'une richesse culturelle exceptionnelle ; la troisième, qui n'est pas la moindre, est que leur enseignement est en voie de disparition.

---

<sup>7</sup> Cf. l'exemple espagnol vu *supra*.