

PARTIE I

Conférence inaugurale

1

Décrire les variétés du français : prolégomènes

A.REY

Directeur littéraire du Robert, Paris, France

Avant de réfléchir aux méthodes et aux procédures de description du français, il convient d'insister sur sa situation multiforme, situation normale pour toute langue de large diffusion. Celle-ci conduit à critiquer le caractère réducteur et abstrait de la linguistique pure, le caractère artificiel et volontairement approuvé de l'objet que vise cette science : une langue en tant que « système » unique, structuré. Ce système est construit autour de deux machines fonctionnelles, une phonologie et une morphosyntaxe. Il ne rend pas compte de caractéristiques pourtant fondamentales de toute « langue ». Cette notion même, et en l'espèce la notion de « français », correspond à un objet construit, à un modèle – non pas à un segment, encore moins à la totalité de l'objet observable et anthropologique qui nous concerne et nous importe entièrement, que ce soit dans l'intuition ou dans la pratique, dans la vie.

Ces banalités épistémologiques gagnent à être répétées inlassablement. Par exemple, il faut redire qu'une « langue » n'est rien si elle ne constitue pas un moyen social d'expression, de communication interhumaine, un élément de la pratique de groupe, et de la politique même. Historiquement, avant d'être étudiable en tant que système fonctionnel, le français est le produit de choix conscients et inconscients, d'attitudes, de réactions venant donc agir et réagir sur une situation historique. Le latin populaire parlé en Gaule, sur un substrat celtique et préceltique, avec des apports germaniques, n'a pas abouti spontanément à une langue, mais à une pluralité dialectale déjà simplifiée par la linguistique historique. Parmi ces produits spécifiques et proprement créolisés du latin, la volonté politique, au moment du partage de l'Empire carolingien, la volonté propagandiste chrétienne un peu plus tôt (concile de Tours) ont fini par imposer un usage, par l'opération essentielle de la notation écrite – alors interdite aux idiomes « vulgaires ». Cet usage, plus tard identifié à ceux de l'Ile-de-France et de Paris, peut s'appeler nationalement le « français ». Le mot étant issu,

ne l'oublions pas, d'un radical germanique qui reflète dans l'histoire la réalité d'un pouvoir politique et non pas l'origine, laquelle est latine et romane, bien qu'elle soit linguistiquement pertinente.

Le français donc, longtemps après le gotique de Wulfila et après d'autres langues germaniques, est le premier des « vulgaires » européens nés sur l'espace impérial romain, qui ont été à même de réaliser le programme posé au XVIII^e siècle par Dante : celui d'un « vulgaire illustre », c'est-à-dire d'un usage spontané et populaire se hissant au statut politique et social de « langue ».

Des idées assez nouvelles sur l'institutionnalisation des langues se sont récemment précisées, notamment celle de « colinguisme » qu'utilise Renée Balibar, surtout à propos des rapports entre latin et français jusqu'au XIX^e siècle. Cette notion rend mieux compte d'une situation sociolinguistique très générale que les termes de « bilinguisme » ou de « plurilinguisme », lesquels souffrent d'une grande ambiguïté, désignant à la fois des relations individuelles et collectives dans le maniement des idiomes. Aujourd'hui comme hier, le français, en tant que langue faite d'usages en situation historique, se définit par des colinguismes divers.

L'observation du français

Celle-ci ne peut se faire utilement que dans une situation pragmatique descriptible. Cette considération élémentaire et essentielle quant à la méthode suffit à mettre en question, au moins de ce point de vue, la notion générale de francophonie.

Les situations francophones sont profondément dissemblables. Cela, même si l'on écarte celle où le français, dans une société pratiquant une ou plusieurs autres langues, n'est qu'un élément du projet didactique de cette société (l'enseignement du français parmi d'autres langues étrangères). La typologie la plus simpliste, sans ambition scientifique, puisque les critères y sont mêlés et forcément pragmatiques, dégage au moins quatre situations elles-mêmes plurielles, que tout linguiste de terrain distingue parfaitement, mais que certains représentants du français européen ont tendance à amalgamer.

Le français est la langue maternelle de la grande majorité du groupe social (la majorité des locuteurs est « héréditaire ») ; il a le statut de langue nationale, officielle, institutionnelle et il l'a seul. En cas de bilinguisme, il est en situation dominante. Normalisé, enseigné, il se transmet sans effort politique appuyé (même si une politique de la langue est affichée), ni prise de conscience particulière, ou du moins par un réglage où le contact avec d'autres langues – qui peut d'ailleurs être conflictuel – ne le menace pas fondamentalement. Cette situation n'élimine pas les difficultés dans certains secteurs, et permet le sentiment d'une crise quant à la qualité de la langue ou à son efficacité, sentiment d'ailleurs quasi universel et non pas propre au français.

Telle est la situation du français en Europe occidentale, France, Belgique wallonne, Suisse romande, avec des cas particuliers dans les zones de contact (Alsace, Luxembourg, Catalogne française, Corse...). Cependant, même dans le cas dominant, des plurilinguismes se développent et s'absorbent au gré des immigrations. En outre, un colinguisme s'est développé avec la forme internationale la plus neutralisée de l'anglais, qui joue un rôle explicite et important (affaires, sciences), sans parler d'effets plus cachés, mais assez sérieux, dont les systèmes de communication massive sont les principaux véhicules.

Le français, langue maternelle de la majorité, acquiert des caractères spécifiques par rapport à ses origines ; il a dû s'affranchir d'un bilinguisme officiel et peut hésiter entre plusieurs normes possibles. La compétence passive d'une partie des locuteurs est alors distincte de la compétence active, par acceptation partielle et par influence subie d'une norme européenne du français, différente de la norme ou des normes effectives. On a reconnu la situation québécoise.

Dans les autres espaces francophones d'Amérique du Nord (Acadie, Louisiane, Ontario, etc.) le français, linguistiquement proche en tant qu'usage, est dans une autre situation sociolinguistique, plus précaire, plus menacée, avec des interférences plus fortes du fait du contact des langues.

Cette situation est caractérisée par un colinguisme avec l'anglais beaucoup plus actif et menaçant que celui qui caractérise le français européen.

Le français peut être l'élément officiel et normalisé d'un bilinguisme ou d'un trilinguisme où la langue maternelle est, au moins pour la majorité des francophones, un créole. Cette situation, commune aux Caraïbes et à l'océan Indien, concerne à la fois des « départements ou territoires » dépendant politiquement de la France ou bien des États indépendants comme Maurice ou Haïti, dans lesquels le créole tend à acquérir un statut différent, plus officiel et plus accepté culturellement.

Les créoles concernés étant génétiquement et lexicalement liés au français, un continuum français-créole peut parfois être évoqué, la créolisation du français régional ayant pour pendant une francisation plus ou moins forte des créoles, au moins lexicalement. Cette situation a fait l'objet d'études très poussées.

Enfin, le français, langue de colonisation relativement récente par rapport aux situations précédentes (XIX^e - première moitié du XX^e siècle), joue un rôle institutionnel dans les zones du monde où de nombreuses langues entièrement différentes sont pratiquées en tant que langues maternelles. Des pays d'Asie ont fait partie de cet ensemble, aujourd'hui représenté essentiellement par l'Afrique. Mais cette situation, vous la connaissez mieux que quiconque, et vous savez combien ses modulations sont considérables. D'abord selon le statut de la langue ou des langues maternelles. Malgré le plurilinguisme arabe-berbère au Maghreb, ce sont alors les relations entre les usages de l'arabe dialectal - lesquels sont comme adossés à la référence historique et religieuse de grand prestige qu'est l'arabe dit « classique » - et les usages du français qui sont en cause. L'analyse des situations scolaire, médiatique et socioculturelle spontanées du français, plus encore que sa position institutionnelle et idéologique - très contestée, comme on sait -, révèle alors plusieurs cas distincts. Cependant, le français en Afrique maghrébine présente des traits communs notables par rapport aux situations sociolinguistiques de l'Afrique subsaharienne.

Alors, le nombre et la variété des langues maternelles, la diversité de leurs statuts, leur degré de fonctionnement écrit, de normalisation, de description systématique en vue d'un enseignement, définissent en creux l'espace dit « francophone ».

Un aspect essentiel de ces différentes situations nationales et régionales est la place qu'occupe le français dans l'évaluation des locuteurs, d'où procèdent leurs attitudes vis-à-vis des langues constituant la sphère de leur colinguisme.

La sociolinguistique peut décrire, à condition de disposer d'informations contrôlées (enquêtes, etc.), des configurations variées dans les jugements de valeur, depuis le français placé « en haut » (France, parfois Afrique) jusqu'au français « en bas » (Acadie, Louisiane, Québec avant la révolution tranquille), en passant par un français « égal » (Suisse) et un français « à côté » (Maghreb - par rapport à l'arabe) qui n'exclut ni évaluation supérieure ni infériorisation. Mais il faut distinguer ici les attitudes subjectives et les comportements

objectifs. Un Maghrébin pourra fort bien considérer le français comme une langue étrangère imposée par la colonisation et vouloir l'inférioriser par rapport à l'arabe, tout en le valorisant dans ses pratiques (écoute de la radiotélévision, études scientifiques, etc.). Les paradoxes sont fréquents dans toute société plurilingue et pluriculturelle.

La francophonie, en tant qu'espace sociogéographique, est complexe et problématique à divers titres. Certes, il faut décrire les variétés et les spécificités des usages du français dans les divers pays francophones d'Afrique, mais on ne peut les décrire, ces usages de la langue, que si on les aborde à la fois par divers angles d'approche : linguistiques *stricto sensu*, avec une méthodologie variationniste, mais aussi politiques, idéologiques, historiques, socio-économiques, didactiques. Je n'aurai que le temps de mentionner l'importance de la phonétique (en Afrique, effets de substrats, nature de l'usage français colonial : de France ou de Belgique, en l'espèce, et implications sur la morphosyntaxe), et celle, plus cruciale encore, de la syntaxe. À ce propos, il faut rappeler la dramatique inégalité du niveau de connaissance entre la syntaxe du français écrit normalisé et unifié, où les variables relèvent surtout des rhétoriques et des stylistiques, et celle des usages oraux spontanés qui révèlent des structures toutes différentes. Si l'on étudie les variétés de français spontané pratiquées en France, en elles-mêmes et en éloignant la référence au français écrit de l'école, on s'aperçoit que les tendances à la pidginisation connues et étudiées en milieu urbain africain ont des corrélats, sinon des analogies, en milieu banlieusard français, du côté des enfants d'immigrés des grands ensembles.

Ici, je sors du sujet : une étude contrastive des usages parlés spontanés du français, beaucoup y songent ; mais les moyens requis – tant financiers qu'humains – seraient immenses, et les intérêts politiques et institutionnels ont beaucoup de raisons de ne pas les fournir.

Cela dit, reste un point essentiel et plus aisément étudiable, le lexique.

Le lexique

Je pense, par conviction et par intérêt professionnel, que, si la syntaxe est au centre de l'étude de la langue en soi, le lexique est au centre de celle du langage et des langues en tant qu'objets sociaux. D'abord, les variantes locales et sociales du lexique reflètent un ensemble pertinent de conditions d'emploi, de situations concrètes. Ensuite, le lexique forme avec les terminologies le point d'articulation entre langage, visions du monde et appréhension du réel. Enfin, c'est le lexique, ensemble peu systématique par rapport à la grammaire, qui fait l'objet principal de descriptions didactiques indispensables, les dictionnaires, ces descriptions abordant tous les aspects de la langue, grammaire comprise.

Les vocabulaires d'une langue, on le sait, sont le lieu le plus visible de la variation. À tel point que l'optique variationniste et différentielle, exceptionnelle en syntaxe, y est habituelle. Décrire les traits lexicaux différentiels d'un usage est une démarche naturelle, plus économique que la description fonctionnelle de cet usage, et souvent efficace. Un dictionnaire contrastif d'un usage du français a en effet l'avantage, en soulignant les différences, de manifester des caractères propres à son sujet.

Une anecdote soulignera cet aspect quasi politique des choses. Alors que l'espagnol d'Amérique a suscité de nombreux dictionnaires contrastifs, le Mexique a jugé indispensable d'élaborer un dictionnaire fonctionnel intégral de son usage, ne faisant aucun recours au castillan en tant que source. Il en est résulté un petit dictionnaire, dû à l'excellent spécialiste qu'est Fernando Lará. Or ce dictionnaire, étant élémentaire, décrit surtout, d'après

un corpus exclusivement mexicain, des unités de haute fréquence. Il en résulte que les différences apparentes avec un dictionnaire espagnol d'Europe y sont faibles, puisque les mexicanismes typiques, indianismes ou autres qui ne sont pas des unités de haute fréquence, y sont peu nombreux. Des réactions nationalistes négatives se sont produites à ce sujet car les nombreuses originalités de l'ouvrage portaient sur des nuances sémantiques, des connotations et de la phraséologie, éléments mal perçus. En outre, les mexicanismes les plus usuels, diffusés dans le monde hispanique tout entier, figurent déjà dans les dictionnaires espagnols produits en Europe.

Ainsi, les différences entre usages ne se laissent décrire dans un dictionnaire ou un lexique que si le corpus utilisé les manifeste ; or, la prise en compte habituelle et majoritaire de discours écrits institutionnalisés (imprimés, le plus souvent), y compris les journaux, peut donner l'illusion d'une relative homogénéité avec la référence implicite (intuitive ou identifiée trompeusement avec un dictionnaire de référence « standard »). Par exemple, le discours didactique, administratif et même journalistique en français présente des traits stylistiques communs et respecte en général des règles assez homogènes, alors même qu'il reflète la variation locale. Ce reflet est alors discret ou même normalisé (re-francisé) par rapport aux productions spontanées, normées de manière « endogène ».

Les divergences lexicales les plus visibles portent sur des domaines du vocabulaire bien délimités : univers référentiel naturel (animaux, plantes, détails géographiques), culturel (habitat, vêture, nourriture, coutumes, religions, fêtes...), institutions, etc. D'autres différences plus profondes entre les usages portent sur la syntaxe, la phraséologie, les effets de fréquence et les connotations, les stratégies de discours et risquent alors d'échapper. On peut aussi plaquer sur un dictionnaire d'un seul usage (en l'espèce un dictionnaire français de France) un inventaire différentiel, mais il s'agit alors d'un hybride, bien différent d'un dictionnaire fonctionnel d'un autre usage.

Le dictionnaire différentiel ou contrastif se justifie comme mesure provisoire et préparatoire. Sa principale qualité est économique. En effet, dans l'histoire, c'est lui qui apparaît chaque fois que l'on veut souligner l'existence d'un usage par rapport à une norme historiquement première. Celle-ci n'est pas forcément dominante sur le plan quantitatif : le cas du portugais au Brésil le montre bien. Seules des circonstances socio-économiques expliquent que les dictionnaires anglo-américains soient fonctionnels (mais il existe des recueils d'américanismes aux États-Unis), alors que les variétés de l'espagnol en Amérique (à l'exception du mexicain, on l'a vu) sont décrites par des dictionnaires contrastifs.

Pour le français, les recueils, inventaires et même « trésors » élaborés sur le modèle contrastif restent indispensables et précieux. Mais ils présentent, nous le savons tous, divers inconvénients. La question, sachant que seul ce type d'ouvrages est réalisable dans des circonstances effectives données, et définies par l'économie, les moyens, les possibilités documentaires et rédactionnelles..., est d'atténuer ces inconvénients, en attendant d'avoir les moyens d'une description fonctionnelle intégrale d'un usage, quel qu'il soit.

Il existe de nombreuses interprétations possibles de la description différentielle. Ajouter des entrées nouvelles au dictionnaire ne suffit pas. Mais, si l'on tient compte de la phraséologie, des nuances de sens, de la fréquence, de l'importance idéologique de certains termes, le modèle devient très riche. Encore faut-il prendre en compte la pragmatique : stratégies discursives reflétant des attitudes sociales, niveaux d'usage, oral ou écrit, actif ou passif, de nature géographique, sociale ou chronologique. Si ce programme est mis en œuvre, la description contrastive, si riche soit-elle, n'est pas encore achevée. Encore faudrait-il, comme le soulignait récemment Ambroise Queffélec, noter tout ce qui figure dans une description supposée fiable de la norme de référence et qui est inconnu ou non fonc-

tionnel dans l'usage décrit. Cette description par défaut peut aboutir à des suppressions ou à une marque spécifique dans un dictionnaire fonctionnel (du type « français de France » dans un dictionnaire du français de Belgique ou « français européen » dans d'autres cas), mais est difficilement gérable dans un inventaire différentiel qui travaille en principe sur des éléments repérés par excès. De toutes façons, l'inconvénient majeur de la description contrastive, c'est de dépendre de la norme ou de l'usage de référence, lequel n'est accessible que par des descriptions déjà imparfaites, ou du moins travaillées par une volonté normative. Dans les dictionnaires existants, la faiblesse des éléments régionaux français, l'excès, d'ailleurs volontaire, des éléments littéraires, parfois véhiculés par la pédagogie, suffisent à gauchir le sentiment différentiel. Cependant, à condition de savoir que l'on part d'un artefact, ces dictionnaires constituent une approximation raisonnable de la norme de référence.

Pour accéder à une description convenable d'un usage quelconque du français, il faut évidemment se dégager du concept de norme extérieure unique de référence, même si son reflet lexicographique est un point de départ contrastif utile, et pour le moment inévitable.

Cela requiert un approfondissement du concept d'usage et une étude du passage des usages en coexistence plus ou moins pacifique avec des normes. Celles-ci peuvent en effet être « endogènes », problème dont va nous parler M. Manessy, mais je pense que des éléments extérieurs, imposés par les contacts historiques, sont toujours à l'œuvre dans leur élaboration (norme exogène ne signifiant pas forcément référence unique à un français standard). Ces normes ont évidemment des traits phonétiques, morphosyntaxiques et graphiques essentiels ; mais seuls leurs caractères lexicaux fournissent un point d'observation privilégié, donnant d'ailleurs accès à d'autres types d'information, par la sémantique et le rapport au monde.

On peut en effet induire de l'observation d'un lexique spécifique des traits, phonétiques et phonologiques, à l'évidence morphologiques, très largement syntactiques, comme le manifestent les modèles descriptifs proposés par Maurice Gross et leurs « grammaires-lexiques ». Cela n'empêche pas d'envisager une étude lexicale incluse dans la description globale, sachant que la démarche est alors beaucoup plus coûteuse. Les mêmes objets lexicaux ont l'avantage d'apporter obligatoirement des implications sémantiques, en signification et en désignation (par les terminologies), et des implications pragmatiques, donc psychosociales. On peut certes décrire un usage par d'autres voies, mais l'approche lexicale est centrale s'il s'agit d'aborder le passage des usages locaux, qu'ils soient « nationaux » ou « régionaux », à une norme institutionnelle. La norme pédagogique (scolaire) est en général un compromis entre la norme de référence externe du « français universel » et une norme endogène plus fragile. L'une des questions majeures, pour cette norme de l'école, est de gérer les conflits avec les autres usages institutionnels du français, observables dans le discours administratif, politique ou médiatique. Il est clair que l'observation de ces contradictions et donc les chances de les résoudre passent par la connaissance du niveau lexical (incluant la phraséologie et la syntagmatique) plutôt que de tout autre.

En outre, les interventions volontaires du groupe social sur les usages sont en pratique restreintes au lexique et à la graphie lexicale (ce qu'on appelle l'« orthographe d'usage »). Sauf circonstances très particulières, ces interventions de nature politique jouent un rôle modeste par rapport aux évolutions sociolinguistiques spontanées, par rapport à la politique pédagogique, enfin par rapport aux jugements de valeur qui sous-tendent les attitudes sociales et sont eux aussi spontanés (les conflits entre la spontanéité du discours et la spontanéité métalinguistique ont été mis en évidence par l'école de Labov).

La constitution de corpus

L'étude des variétés du français passe par des enquêtes sociolinguistiques et par la constitution de corpus (ou corpora). Ces derniers relèvent de choix et leur représentativité est fonction des objectifs. En fait, la situation qui correspond à la préparation d'un inventaire lexical, d'un dictionnaire contrastif (on a vu qu'il pouvait être plus ou moins compréhensif) ou d'un dictionnaire fonctionnel de l'usage ou des usages à décrire est déjà réductrice. Par ailleurs, viser un corpus très large, très important, implique des investissements considérables.

Dans l'optique du dictionnaire fonctionnel, le corpus clos exploité philologiquement (type TLF) représente la charge la plus lourde. Il n'est envisageable qu'exceptionnellement, car son rendement s'affaiblit à mesure que les relevés augmentent. L'expérience du TLF, celle du *Co-Build* britannique, avec un corpus plus réduit, sont à cet égard éloquentes.

À côté de ce type de relevés exhaustifs portant sur un grand nombre de sources, d'autres corpus finalisés, exploitables sélectivement, et plus économiques sont possibles. Ils peuvent concerner de l'écrit normalisé, littéraire, journalistique, pédagogique, techno-scientifique, administratif, etc., tous matériels relativement peu rentables en termes différentiels, mais aussi de l'écrit spontané, notamment du matériel écrit scolaire permettant d'observer des écarts de différente nature par rapport à une norme, qui dans ce cas doit être la norme locale, non une norme abstraite, de l'oral institutionnalisé (radio, télévision, discours officiels), en fait « oral-écrit », enfin de l'oral spontané. Dans chaque secteur, les choix et leur dosage orientent les résultats, mais la méthodologie des échantillons pertinents est bien connue en sociologie. À mon sens, du matériel psycholinguistique serait aussi requis, notamment les associations lexicosémantiques. Dans tous les cas, mais surtout s'agissant de français non maternel, le niveau de maîtrise des locuteurs est une donnée essentielle, ainsi que leur pratique linguistique (multilinguisme, pratique orale et écrite d'une ou plusieurs langues, etc.) ; la nature des autres langues pratiquées est évidemment pertinente et rend compte d'effets de contact qui peuvent éclairer sur les particularités de l'usage français étudié.

L'un des problèmes de méthode les plus ardu est celui du niveau d'abstraction de l'objet visé. Les dialectologues européens ont à cet égard une expérience de l'enquête de près d'un siècle et tous les chercheurs de terrain sont tentés de définir des entités claires et précises, alors que l'usage du français à décrire est bien souvent déjà très complexe, géographiquement et sociologiquement. La difficulté est alors de conserver des méthodes homogènes, pour éviter de juxtaposer des données obtenues de manière diverse rendant difficile leur interprétation. Alors que les situations sociolinguistiques sont très diverses à l'intérieur d'un même usage national du français, le besoin d'une description unifiée est évident et requiert un équilibrage délicat entre ces situations : villes et campagnes, milieux agraires, commerçants ou prolétariés, niveaux sociaux et culturels, etc.

Or, les français de France, de Belgique, de Suisse, du Québec sont déjà bien loin de bénéficier d'une description de ce type, tenant compte des variations internes, des conflits d'usage, des niveaux de normalisation, etc. Les descriptions en usage relèvent de compromis entre tradition, contraintes économiques et novation. Dès lors, il peut paraître absurde de requérir, même à propos d'objectifs prioritaires, une telle richesse d'investigation. Il convient pourtant de garder ces exigences en vue, pour rentabiliser au maximum un matériel toujours insuffisant, mais dont l'interprétation peut apporter des informations indispensables. Cet apport est plus précieux que l'abondance d'un matériel incontrôlé.

Conclusion

Il n'y a aucune contradiction entre le désir d'une étude sociolinguistique en profondeur des usages du français et la description par voie de dictionnaire ou de manuel des éléments spécifiques et différentiels du lexique. Ce dernier est au centre du rapport entre langue et pratiques sociales. En tout état de cause, rien n'est plus sociologique en linguistique que la lexicologie.

Les inventaires différentiels, aussi réducteurs soient-ils, apportent à cette étude un matériel absolument indispensable. Par ailleurs, la normalisation des nomenclatures et terminologies est, elle aussi, indispensable à toute pédagogie.

Enfin, l'optique du dictionnaire permet de refléter toute problématique scientifique, en la projetant sur une intention multiple, qui peut être économique (éditoriale) et politique, et qui correspond à une idéologie. Il faut s'en souvenir, et on est en droit de s'en méfier, mais il reste que les idéologies politiques sont, avec la pédagogie scolaire, les leviers qui permettent d'agir sur la construction de ces normes sociales du langage qu'il s'agit de libérer, d'assurer et de démocratiser, car elles sont indispensables à la survie des langues dans les sociétés.