

LANGUES :
HISTOIRES ET USAGES
DANS L' AIRE MÉDITERRANÉENNE

Langue et Parole
Recherches en Sciences du Langage
Collection dirigée par Henri Boyer

La collection *Langue et Parole* se donne pour objectif la publication de travaux, individuels ou collectifs, réalisés au sein d'un champ qui n'a cessé d'évoluer et de s'affirmer au cours des dernières décennies, dans sa diversification (théorique et méthodologique), dans ses débats et polémiques également. Le titre retenu, qui associe deux concepts clés du *Cours de Linguistique Générale* de Ferdinand de Saussure, veut signifier que la collection diffusera des études concernant l'ensemble des domaines de la linguistique contemporaine : descriptions de telle ou telle langue, parlure ou variété dialectale, dans telle ou telle de leurs composantes ; recherches en linguistique générale mais aussi en linguistique appliquée et en linguistique historique ; approches des pratiques langagières selon les perspectives ouvertes par la pragmatique ou l'analyse conversationnelle, sans oublier les diverses tendances de l'analyse de discours.

Il s'agit donc bien de faire connaître les développements les plus actuels d'une science résolument ouverte à l'interdisciplinarité et qui cherche à éclairer l'activité de langage sous tous ses angles.

Déjà parus

Alain COÏANIZ, *Langages, cultures, identités. Questions de point de vue*, 2005.

Aline GRANGE, *L'Europe des drogues*, 2005

Bernard LE DREZEN, *Victor Hugo ou l'éloquence souveraine*, 2005.

Henri BOYER, *De l'autre côté du discours*, 2003.

Alda MARI, *Principes d'identification et de catégorisation du sens*, 2003.

Mary ROWLAND-OKE, *Description systématique de la langue obolo-andoni*, 2003.

site : www.librairieharmattan.com

e.mail : harmattan1@wanadoo.fr

© L'Harmattan, 2005
ISBN : 2-7475-9056-9
EAN : 9782747590563

Teddy ARNAVIELLE
(Éditeur)

LANGUES :
HISTOIRES ET USAGES
DANS L' AIRE MÉDITERRANÉENNE

Préface de
Robert LAFONT

L'Harmattan

5-7, rue de l'École-Polytechnique ; 75005 Paris
FRANCE

L'Harmattan Hongrie
Könyvesbolt
Kossuth L. u. 14-16
1053 Budapest

Espace L'Harmattan Kinshasa
Fac. des Sc. Sociales, Pol. et
Adm. ; BP243, KIN XI
Université de Kinshasa – RDC

L'Harmattan Italia
Via Degli Artisti, 15
10124 Torino
ITALIE

L'Harmattan Burkina Faso
1200 logements villa 96
12B2260
Ouagadougou 12

Les communications rassemblées dans cet ouvrage ont été présentées au
Colloque International

La Méditerranée et ses langues

organisé à l'Université Paul-Valéry (Montpellier III)
les 20, 21, 22, 23 mars 2002
par le Laboratoire DIPRALANG (EA 739)

avec le soutien du
Conseil Régional du Languedoc-Roussillon,
du Pôle Universitaire Européen de Montpellier et du Languedoc-Roussillon,
et de la Délégation Générale à la Langue Française et aux Langues de France
(Ministère de la Culture et de la Communication)

Coordinateurs :

T. Arnavielle, H. Boyer

Comité scientifique :

A. Badia i Margarit (Université de Barcelone), C. Blanche-Benveniste (Université de Provence-Aix-Marseille), J.M. Blecua (Université Autonome de Barcelone), S. Chaker (Inalco), P. Charaudeau (Université de Paris Nord-Paris XIII), M. Doss (Université de Caire), F. Ferreres Maspla (Université de Barcelone), P. Garde (Université de Provence), P. Gardy (CNRS - Université de Montpellier III), G. Gouran (Université de Montpellier III), G. Kremnitz (Université de Vienne), R. Lafont (Université de Montpellier III), M.-P. Masson (Université de Montpellier III), H. Meschonnic (Université de Vincennes-Saint Denis - Paris VIII), H. Salah (Université d'Alger), P. Sauzet (Université de Vincennes-Saint Denis - Paris VIII), R. Simone (Université de Rome), Ch. Touratier (Université de Provence), T. Yucel (Université d'Istanbul).

Organisation du colloque

M.C. Alén Garabato (Université de St Jacques de Compostelle), T. Arnavielle (Université de Montpellier III), T. Bouguerra (Université de Montpellier III), H. Boyer (Université de Montpellier III), C. Camps (Université de Montpellier III), P. Dumont (Université de Montpellier III), C. Lagarde (Université de Perpignan), S. Miossec (Université de Montpellier III), J.-M. Petit (Université de Montpellier III), M. Verdelhan (Université de Montpellier III), J. Zenati (Université d'Alger - Université de Montpellier III).

Secrétariat :

K. Djordjević, J. Zenati

PRÉSENTATION

Carrefour, berceau, mère, creuset, foyer, ... : au-delà du poncif, la Méditerranée est bien tout cela à la fois. Des langues sont nées sur ses rives, comme l'écriture qui les a fixées. Certaines ont été effacées, que conserve la mémoire des érudits, d'autres s'y sont perpétuées par l'évolution, d'autres par la transformation et la diversification, par l'imprégnation de langues d'autres aires. Quelques-unes sont parties à la conquête du monde, auxquelles leur diffusion a donné un visage nouveau ; d'autres ont souffert ou souffrent des conflits des hommes, et des rapports de puissance, sans que leur validité, parfois précaire, soit éteinte : comptons sur la détermination des hommes pour assurer leur perpétuation. Toutes ont été le support de très anciennes et très riches cultures, toujours vivantes, ouvertes et créatives. Il n'en est pas qui n'aient été objet d'étude : pour l'historien des origines et des évolutions, le comparatiste, le sociolinguiste, l'ethnolinguiste, le géolinguiste ...

C'est cette approche multipolaire qui a été retenue comme projet fédérateur du Colloque tenu à Montpellier au printemps 2002 : *La Méditerranée et ses langues* : une rencontre résolument ouverte à la diversité des traitements disciplinaires qui constituent aujourd'hui le champ de la linguistique (et dont on voudrait ainsi mettre en évidence la complémentarité, revendiquée par le Laboratoire DIPRALANG, initiateur et support de la manifestation), « une rencontre constante du passé et du présent, le passage de l'un à l'autre ... (Fernand Braudel, *La Méditerranée*, 1985, Flammarion, « Champs », p. 7). Et aussi, à la charnière du siècle et du millénaire, une réflexion collective où quelques perspectives puissent être proposées.

Favoriser une rencontre et des échanges entre des spécialistes que leurs activités scientifiques séparent plus souvent qu'elles ne les rassemblent : tel est l'objet que s'est assigné l'équipe organisatrice du colloque. À cet effet, il a paru opportun de définir deux axes prioritaires autour desquels organiser les communications en ateliers et les conférences :

- un caractérisé par une orientation historique et génétique, sans exclusive (des études de synchronie moderne ont été bienvenues), où phonétique, écriture, morphologie, syntaxe, lexicque, ont eu leur part, soit dans le cadre d'une langue, soit dans celui d'une comparaison plus ou moins large ;

- un concernant les dynamiques géo et sociolinguistiques des langues en contact et/ou en conflit dans l'espace historico-culturel méditerranéen, les statuts, les enjeux politiques dont elles ont été ou sont les vecteurs privilégiés.

Pour des raisons de commodité, c'est en deux volumes que l'ensemble des contributions au colloque ont été réparties : celui-ci recueille les communications de l'axe « historico-linguistique », auquel, répond, dirigé par Henri BOYER, le volume consacré à l'axe « sociolinguistique », *Langues et contacts de langues dans l'aire méditerranéenne. Pratiques, représentations, gestions*, L'Harmattan (coll. « Sociolinguistique »), 2004. Découpage largement arbitraire sur le fond, commandé par des contraintes éditoriales : ce sont bien les deux volets d'une même entreprise qui sont présentés.

Que soient ici remerciés tous ceux qui ont permis que ces livres existent : contributeurs, présents ou empêchés ; conseillers scientifiques ; soutiens financiers : l'Université Paul Valéry, le Pôle Universitaire Européen de Montpellier et du Languedoc-Roussillon, le Conseil Régional du Languedoc-Roussillon, la Délégation générale à la Langue française et aux Langues de France (Ministère de la Culture et de la Communication) ; collaborateurs associés à l'organisation du colloque et à la préparation de ses Actes, tout particulièrement Madame Ksenia Djordjević, Maître de conférences, et Monsieur Jamel Zenati, Chargé de cours.

Exprimons ici un regret : que Jacques Allières, Professeur émérite à l'Université de Toulouse, qui avait accepté avec enthousiasme de faire partie du Comité scientifique du colloque, n'ait pu, enlevé à nous avant sa tenue, nous faire bénéficier, avec sa gentillesse bien connue, de sa science de linguiste historien et comparatiste. C'est à lui que nous dédions ce volume.

Teddy ARNAVIELLE

PRÉFACE

LES MÉDITERRANÉES ET LEURS LANGUES

Robert LAFONT

Il est deux façons de considérer la Méditerranée. La première est la moins naturelle, mais c'est elle qui a cours généralement. Quand on dit que la Grande-Bretagne a eu et a une politique méditerranéenne, ou que la France défend ses intérêts en Méditerranée, on considère ce bassin maritime tout entier depuis un lieu, Londres ou Paris, qui lui est géographiquement étranger. On sanctionne ainsi une mise en dépendance historique de l'Europe du Sud par les pouvoirs de l'Europe du Nord, phénomène arbitraire, extrapolation du commandement de l'espace, qui n'a rien à voir avec la logique des relations humaines, qui s'est simplement imposé par les armes. Cette extrapolation est simplificatrice. Elle ne fait pas de différence entre Maghreb et Maéreq, entre Alger et Alexandrie ou Beyrouth. Mettre Méditerranée au singulier est encore l'habitude dans les conférences internationales consacrées à l'avenir socio-économique de la « région », dont l'une des plus importantes eut lieu récemment à Barcelone. Mais cette région existe-t-elle et peut-on dessiner une perspective englobante pour Marrakech, Sarajevo et Tel Aviv ?

En fait, si l'on prend les mesures de l'histoire longue chère à F. Braudel, on doit considérer non tant la Méditerranée en soi que deux mouvements d'occupation du continent européen (qui est en réalité une presqu'île de l'Asie) par les populations et les cultures successives. Le premier est le mouvement proprement méditerranéen. C'est celui qui prend origine au Proche-Orient dans une haute Antiquité et se trouve porté par Phéniciens et Grecs vers l'Ouest. Il remontera dans le Continent par des couloirs : l'un est la vallée du Rhône et l'autre le seuil du Lauragais dont le « pied maritime » est à Narbonne. Le relais historique est pris par Rome. Par ses conquêtes, situant le commandement du Bassin en son centre, elle en fait un *Mare Nostrum*, elle l'extrapole en tous sens. Mais l'extrapolation majeure est européenne. Elle remonte jusqu'aux bouches du Rhin et du Danube et ne se fixe sur ce *limes* que par l'échec de plus vastes ambitions. Quand le *limes* craque, en 406, les Barbares se précipitent. Il s'agit en fait d'un mouvement multiséculaire qui prend origine sur l'Oural et au-delà, où Celtes ont été suivis de Germains, Germains de Slaves, qui ont reçu la butée des peuples des steppes asiatiques. Il atteint la Manche et se recourbe vers le Sud sur la

basse Loire. On sait qu'il va occuper le Bassin méditerranéen occidental, les Vandales le portant jusqu'à Carthage en passant par l'Espagne.

Ainsi l'Europe se trouve occupée, dans sa longue histoire, par les deux branches d'une tenaille, la branche méditerranéenne et méridionale, la branche continentale et septentrionale. Cette occupation humaine correspond à l'évidence géographique de deux Europes, celle qui prolonge la plaine sibérienne sans obstacle majeur jusqu'au Pas-de-Calais, celle qui s'articule en cloisonnement de massifs et vallées sur la Méditerranée. Deux événements permettent de fixer le point de jonction des deux branches à l'ouest, le point d'écrasement. Le premier est l'affrontement de deux peuples germains, les Francs qui, passant le Rhin, s'établissent entre Somme et Loire, et les Wisigoths qui ont accompli un très long périple par l'Orient avant d'installer un royaume entre Loire et Guadalquivir avec Toulouse pour capitale. Ils représentent deux options culturelles contradictoires : les Francs ignorent l'écriture en leur langue, la vie urbaine, et restent fidèles à leur droit tribal, les Wisigoths ont conquis l'écriture, harmonisent leur droit propre au droit romain, et adoptent l'essentiel de la structure sociale romaine. Ajoutons le motif religieux : les Francs viennent de passer brutalement du paganisme au catholicisme romain, les Wisigoths ont porté avec eux l'arianisme oriental. On sait comment l'affrontement se dénoue militairement en 507 par la victoire des Francs à Vouillé, c'est-à-dire Poitiers, suivie d'une descente vers le sud.

Le second affrontement est celui du pouvoir franc et de la vague de remontée méditerranéenne sous bannière de l'Islam. Dans cette affaire on ne peut négliger les alliances temporaires que, par peur des Francs, Aquitains, Septimaniens et Provençaux, nouèrent avec les Musulmans. On en vient ainsi à la date de 732 et à la célèbre bataille de Poitiers, suivie d'une nouvelle descente franque, dévastatrice, celle de Charles Martel.

Les deux batailles de Poitiers, pour leur immense importance culturelle permettent de fixer sur le seuil du Poitou la jointure des deux Europes. Si l'on considère qu'une autre jointure est sur le confluent de la Saône et du Rhône, on voit se dessiner, en couronne au Massif que la géographie française dit Central, une vaste poche d'Europe méridionale, en dépendance relationnelle de la Méditerranée.

Du moins à l'ouest de la Mer et du Continent. Car, à l'échelle d'espace et d'histoire où nous avons analysé la dualité horizontale européenne, il nous faut nécessairement considérer aussi une distribution verticale. Il y a géographiquement de façon très claire deux Méditerranées séparées l'une de l'autre grosso modo par la chaîne des Apennins, l'Aspromonte et le canal de Sicile. L'orientale est faite d'Adriatique et d'Égée, de part et d'autre du chaos de la Péninsule balkanique. Elle fourmille de complications côtières et insulaires. Du

point de vue large des relations humaines, l'essentiel est qu'elle confronte à cette complexité les deux rebords continentaux, l'africain de Lybie-Égypte et l'asiatique de Palestine-Liban-Turquie. Par ces limites, qui la connectent avec des profondeurs continentales qui n'ont rien d'européen, elle s'amarre à cette région du monde qu'on appelle aujourd'hui Proche-Orient.

Au contraire, la Méditerranée occidentale présente une évidente simplicité. Au nord, un Arc, l'arc qu'on dit latin, de Gibraltar à Messine, au sud la côte maghrébine. Entre les deux, mais de part et d'autre, deux systèmes insulaires : les Baléares et les trois îles de Corse, Sardaigne, Sicile. La configuration est vraiment celle d'un bassin, et la relation obligatoire est Nord-Sud/Sud-Nord, côtière entre Marseille et Alger, plus profonde entre Loire et Sahara. Cette relation fut occupée par la colonisation française, Paris prenant son relais à Marseille, c'est celle de l'actuelle immigration Maghreb-Europe. Colonisation et immigration ont tissé deux sortes de liens, de sens inverse, mais additionnels, entre les populations.

Or, à l'échelle du monde actuel, on peut constater l'émergence planétaire de grandes régions qui restaurent des relations naturelles en traversant les frontières étatiques. Il est devenu banal de parler à cette échelle de la région du Proche-Orient. Il faudrait bien, de façon parallèle, employer une expression du genre de Proche-Occident ou Sud-Occident pour désigner l'ensemble articulé d'ouest Europe-Méditerranée, ou, tout aussi bien et d'un terme qui existe déjà, d'Euroméditerranée.

On désigne ainsi deux bassins de cultures et d'histoire qui n'ont pas grand-chose à voir l'un avec l'autre. L'Est s'est caractérisé par deux grands recouvrements successifs et totaux : celui de l'Empire byzantin et celui de l'Empire turc. Il en résulte deux patines culturelles fortes, celle de l'Église orthodoxe, et celle de l'Islam. Mais ces patines recouvrent une fragmentation ethnique exceptionnelle, source d'interminables et inexpiables conflits. Cette région orientale méditerranéenne, il faut bien le rappeler, s'est révélée récemment par les conflits qui ont suivi l'éclatement de la Yougoslavie et par celui de Palestine (mais on n'oubliera pas Chypre) peut-être la région la plus dangereuse du monde. Prendre en charge ces conflits dans le cadre d'une région méditerranéenne unique revient ainsi à faire peser sur la Méditerranée occidentale des problèmes d'une gravité extrême, qui ne sont pas les siens. C'est pourquoi je pense qu'il nous revient, là où nous sommes aujourd'hui, à Montpellier, de quitter le plan des responsabilités géopolitiques larges et arbitraires qu'assument les États, d'ailleurs fort mal, et très dangereusement, pour penser sérieusement à la tâche à horizons limités, mais puissamment définitionnels, de l'aménagement humain de la région dont

nous sommes le centre, horizontalement et verticalement. À en faire une aire de paix et de progrès.

*

Vous vous demandez peut-être ce que cette déclaration d'intentions a à voir avec le titre que j'avais annoncé, qui comportait la mention de « langues », ce qui correspond à l'objet de ce colloque. Il me faut donc maintenant dire où allait ce préambule dévorant : à des réflexions qui paraîtront étonnantes seulement dans la mesure où l'officialité, la coutume acquise, l'effacement ordinaire ne nous ont pas permis de les penser.

J'ai tenu à définir cette « poche spatiale » où nous sommes, entre Loire, Lyon et mer, avec des arguments seulement spatiaux, et historiques en complément. Or espace et histoire ont précisément pris dans ce cadre une forme linguistique : la langue d'oc, qui l'occupe presque tout entier. Les évolutions du latin ont conduit là, au centre de l'Arc dit latin, à une stabilisation où trouvent leur carrefour les autres langues parentes et voisines. L'identification linguistique se fait là suivant les directions que nous avons mises en lumière. L'occitan n'est pas de l'Europe du Nord parce qu'il est au sud de cette ligne de jointure et de fracture qui passe par Lyon et Poitiers. Il appartient à l'aire méditerranéenne occidentale parce qu'il fait exactement face au Maghreb. En définitive, il aurait pu aussi bien que la géographie et l'histoire, dont il est le produit, servir à notre binarisation et de l'Europe et de la Méditerranée.

On voit vite combien est inadéquate à cette situation la définition qu'on donne le plus souvent de lui, de langue régionale de France. Une langue qui couvre un front européen de la nature du Portugal, et plus vaste que lui, et qui se trouve détenir les clefs des relations qui construisent un « proche-Occident » ne peut être dite régionale que si l'on prend, comme nous le proposons, pour la région une dimension planétaire.

Quant à son rapport avec la France, on sait ce qu'il en est. On ne peut ignorer que la culture qu'élaborait l'Occitanie au XII^e siècle se plaçait sur le front du progrès de l'intelligence et de la créativité européennes, et a été reçue comme telle par l'Europe. Une première réduction, catastrophique, a été la conquête par la France des terres de Toulouse. Elle a été suivie d'autres conquêtes, qui font que la France s'est incorporée, à l'exception du Val d'Aran et de quelques vallées italiennes, la totalité occitane. Elle a ainsi dévoré et invalidé une des composantes principales de l'Europe.

Elle l'a aussi évidée de sa langue. Les phases successives de la répression linguistique s'achèvent sous nos yeux. Une enquête récente fixe à une ou deux

génération après nous la fin de la transmission naturelle de l'occitan, et, comme on sait, les mesures compensatoires rencontrent dans l'appareil d'État une résistance impitoyable.

L'évidement linguistique ne va pas d'ailleurs seul. La négation structurelle socio-économique opère à partir de la seconde moitié du XIX^e siècle, quand le réseau de relations par route et par fer dessine sur le territoire français une toile d'araignée dont le centre est Paris. L'Occitanie devient alors illisible sur la carte, retroussée dans sa logique nord-méditerranéenne, et inconcevable, parce que dépouillée d'histoire explicative autant que de langue identifiante. Suit une crise industrielle qui se prolonge jusqu'aux années soixante-dix du XX^e siècle. Ces phénomènes conjoints, décrits il y a une trentaine d'années sous le nom de « colonialisme intérieur », ont créé une énorme béance sociologique et culturelle. Quand on fait du plein, on fait nécessairement aussi du vide. Sur le glacis entre Loire et mer, la concentration française au Nord a installé ce que vous me permettez d'appeler une « privation de sens ». D'un espace signifiant sur lui-même elle a fait une périphérie insignifiante.

En même temps que le colonialisme intérieur battait son plein au Nord de la Méditerranée, s'achevait la guerre d'Algérie. C'est une date très importante car le Maghreb peut désormais courir ses propres chances et entrer en dialogue avec la rive Nord hors des contraintes de la dépendance coloniale. L'Algérie l'a fait à travers les maladroites, c'est le moins qu'on puisse dire, d'un État centralisé de modèle français, auquel il a ajouté la tyrannie du parti unique. Mais il l'a fait. En même temps, une émigration accélérée créait au nord de la mer un autre Maghreb, l'islam devenant la seconde religion de ce pays de tradition chrétienne. Les liens des deux rives en étaient d'autant resserrés.

Le malheur, complémentaire des erreurs, est que l'opposition au régime installé à Alger, à son caractère policier et à sa corruption a été prise en charge par l'intégrisme islamique, obscurantiste et sanglant. Mais, outre la résistance à la double oppression, de l'État et de l'intégrisme ; de très importants milieux démocratiques, le Maghreb présente une originalité patente : l'existence au Maroc et en Algérie de la population amazigh (berbère). Celle-ci est une forte garantie identitaire, le pays n'ayant été jamais arabisé que par des effectifs faibles et le souvenir d'un peuple indépendant et courageux, fortement typé linguistiquement et culturellement, assurant depuis l'époque des Numides un sentiment national. Aujourd'hui la résistance amazigh est devenue l'élément principal de la lutte à la fois contre le totalitarisme d'État et le terrorisme islamique. Elle est l'arme qu'a le Maghreb pour entrer en démocratie et pour prendre donc sa place dans le proche-Occident, en dialogue avec le Sud européen.

Il est sans doute excessif de comparer des situations identitaires différentes et des oppressions de nature diverse, - l'Occitanie n'est pas Tamazgha - mais on ne peut s'empêcher de remarquer le parallélisme global des deux rives, qui présentent deux langues officielles importées au cours des siècles et qui se réservent tout l'espace de l'officialité, l'arabe et le français, par-dessus une authenticité d'origine. Quelles que soient les réserves qu'on doit apporter à un projet de restitution de leur espace aux deux langues, du fait pour l'occitan d'une exténuation d'usage, d'une fragmentation territoriale pour l'amazigh, il est bien évident que leur présence latente conjuguée, comme une hantise d'histoire volée, finit de donner cohérence à la région planétaire que nous avons essayée de définir. La référence faite à elles, comme minimum largement acceptable, est le gage le plus fort de l'entrée dans une cohérence moderne et productive d'une région du globe désarticulée par les conquêtes et l'exploitation décentrée. C'est un gage de démocratie territoriale et de « remise au monde » des espaces.

CONFÉRENCE D'OUVERTURE

LE POLYCENTRISME LINGUISTIQUE DANS LES BALKANS ET AUTOUR DE LA MÉDITERRANÉE

Paul GARDE

Le pourtour méditerranéen, du point de vue de ses langues, peut être sommairement divisé en trois grandes zones, dont deux relativement homogènes, au moins en apparence, et une troisième fortement diversifiée. Au nord-ouest une zone romane couvre l'Italie, la France et la péninsule ibérique. Au sud, du Maroc à la Syrie, une vaste zone sémitique, ou hamito-sémitique, est principalement recouverte par la langue arabe, mais aussi par le berbère et par l'hébreu. Au nord-est enfin la zone balkanique et anatolienne se caractérise au contraire par une grande variété de familles linguistiques : turque, répartie dans les deux péninsules ; arménienne, iranienne (représentée par le kurde) en Anatolie ; grecque, albanaise, slave, romane à nouveau (représentée ici par le roumain) dans les Balkans.

Mais tout linguiste sait bien que ce regard globalisant est trompeur. L'uniformité, là où le profane croit la déceler : dans le domaine présumé de chaque grande ou petite langue, n'est qu'apparente, et dans les faits Babel est partout. La liste des communications présentées à notre colloque nous le confirme : presque toutes, d'une façon ou d'une autre, sont consacrées à la diversité. Celle-ci n'est pas moindre au sein des domaines linguistiques les plus vastes, tel qu'« arabe », « espagnol » ou « français », que dans les ensembles de taille plus modeste.

Cette diversité est elle-même polymorphe, elle présente des visages très divers. Je veux ici, après avoir esquissé une brève typologie de la diversité linguistique, insister sur un type particulier que j'appellerai « polycentrisme », et dont je chercherai des exemples un peu partout autour de la Méditerranée, mais avant tout dans la zone elle-même la plus diverse, les Balkans. Je m'arrêterai pour finir sur un cas particulièrement complexe, celui de l'ensemble qu'on appelle souvent « serbo-croate ».

Petite typologie de la diversité

Les facteurs de diversité à l'intérieur de ce qui apparaît à tort ou à raison comme une même « langue » sont bien connus. On peut distinguer :

1° la diversité temporelle, la diachronie. Elle a été depuis deux siècles (depuis l'essor de la comparaison indo-européenne avec Bopp), et jusqu'à une date récente, le thème favori des études linguistiques. L'étude des changements, sur le court terme ou la longue durée, est un sujet de recherche inépuisable.

2° la diversité spatiale : dialectologie ou géographie linguistique. La diversité dialectale à l'intérieur d'un continuum linguistique est un phénomène universel, dont l'importance est apparue il y a environ un siècle (notamment avec l'Atlas de Gillieron), mais qui depuis a donné lieu à des recherches dans tous les domaines linguistiques. Dans certains territoires cette diversité apparaît même à deux niveaux, par diversification successive de deux langues communes, l'une ancienne, l'autre récente : en France par exemple, on peut étudier aussi bien les dialectes ruraux, issus de la diversification du latin, que les « français régionaux », résultats de celle du français.

3° la diversité fonctionnelle, objet de la socio-linguistique. Celle-ci a été définie comme discipline autonome vers le milieu du XX^e siècle seulement, avec notamment les travaux de Labov. On avait toujours distingué des « niveaux de langue », étudiés dans le cadre de la stylistique. On a plus récemment recherché de façon plus précise les phénomènes langagiers propres à telle ou telle situation, à tel ou tel milieu social (parfois appelés aussi « dialectes », terme donc ambigu).

Parmi les phénomènes socio-linguistiques, il en est un qui retient particulièrement l'attention : c'est celui de la standardisation, ou fixation et diffusion de « langues standard ». Nous préférons ce terme (cf. angl. *standard language*) à ceux de « langue écrite » (all. *Schriftsprache*) ou « langue littéraire » (russe *literaturnyj jazyk*) qui désignent la même réalité, mais peuvent créer des confusions. Le standard est la forme de langue, assez fortement codifiée, qui est employée dans les formes de communication les moins spontanées : expression écrite, textes officiels ou littéraires, discours public, enseignement, médias, etc., mais qui dans les sociétés modernes tend à supplanter les autres types de langue même dans l'usage courant, ou en tout cas exerce une forte influence sur eux.

Monocentrisme et polycentrisme

Le plus souvent, le domaine d'un standard donné correspond à un certain ensemble plus ou moins homogène du point de vue de la géographie linguistique. Le meilleur exemple est l'italien. Dans l'ensemble de la famille romane, on distingue une « langue italienne » subdivisée en un grand nombre de dialectes, certes très différents les uns des autres, mais possédant un assez grand nombre de traits communs qui les différencient des autres dialectes romans. La langue est alors conçue comme un système de dialectes ou

« diasystème ». Mais le terme « langue italienne » désigne aussi la langue standard, seule enseignée et habituellement écrite, connue dans toute l'Italie et parlée par tous, occasionnellement ou exclusivement selon les personnes. La « langue » ainsi comprise, même si elle repose à l'origine sur une base dialectale, définie d'ailleurs par des formules complexes (*lingua toscana in bocca romana*), ne se confond avec aucun des dialectes, mais s'oppose à eux.

Dans le cas de l'Italie, comme dans beaucoup d'autres, la « langue » comme système de dialectes et la « langue » comme standard couvrent, à quelques détails près, le même espace géographique : en gros, Italie plus Tessin. Les divergences sont minimales : Frioul et Sardaigne italiens par le standard, mais non par le dialecte ; Corse dans la situation inverse ; pas de divergence au contraire dans le Val d'Aoste et le Haut-Adige, où l'italien coexiste avec des dialectes et des standards (français ici, allemands là) qui lui sont également étrangers. On dira donc que le domaine italien est monocentrique : sur le territoire du diasystème s'est créé un seul standard.

Il n'en va pas toujours ainsi. La France nous offre l'exemple d'un standard fonctionnant sur le territoire d'au moins deux diasystèmes romans : français et occitan (et de plusieurs autres). Mais le cas contraire est très fréquent aussi : c'est celui où dans le domaine d'un seul diasystème fonctionnent deux ou plusieurs standards. Nous l'appellerons « polycentrisme », admettant que le standard est le « centre » d'un diasystème. Il s'agit donc d'un système à plusieurs centres.

La pluralité des standards est généralement facile à établir, parce qu'ils sont des unités discrètes, des systèmes homogènes, plus ou moins fortement codifiés. On peut les compter. L'unité du diasystème est plus difficile à prouver, parce qu'il s'agit d'ensembles flous dont la délimitation, au sein du continuum dialectal d'une grande famille linguistique (romane, slave, turque, arabe, etc.), suppose une part de convention. On pourra parler d'un diasystème unique si les limites suggérées par la dialectologie (faisceaux épais d'isoglosses) sont nettes. Mais il n'en est pas toujours ainsi ; nous aurons à évoquer quelques cas litigieux.

Les cas de polycentrisme, où la langue comme diasystème ne coïncide pas avec la langue comme standard, sont ceux où le profane hésite et se demande : « Y a-t-il une ou deux langues ? », et où il attend du linguiste qu'il lui fournisse une réponse à cette question. D'innombrables polémiques ont eu lieu à ce sujet, gâtées par la passion politique qui anime le plus souvent les tenants de l'une ou l'autre solution. Il me semble que le rôle du linguiste n'est pas de chercher à tout prix une réponse univoque, qui n'est généralement exigée que pour des raisons partisanses, et ne peut être obtenue qu'au prix de la simplification d'une réalité complexe. Il est au contraire de se donner les instruments conceptuels pour

rendre compte de cette complexité même. La notion de polycentrisme, et l'examen des diverses variantes qu'il comporte, me paraît être un de ces instruments.

La divergence entre deux standards coexistants dans un même diasystème peut être fondée sur l'un ou l'autre des trois types de diversité évoqués plus haut : fonctionnelle, temporelle ou spatiale. Nous les examinerons successivement, d'abord à partir des langues de la péninsule balkanique, mais sans nous interdire des incursions dans d'autres coins de l'espace méditerranéen.

Polycentrisme à base fonctionnelle (diglossie)

Il s'agit du phénomène bien connu qui, depuis l'article fondateur de Ferguson en 1959, a été abondamment étudié sous le nom de « diglossie ». Une même population, sur l'ensemble d'un même territoire, utilise deux formes de langue, génétiquement étroitement apparentées entre elles, et toutes deux plus ou moins codifiées, mais dont chacune est réservée à certaines situations de parole. On distinguera, selon la terminologie de Ferguson, une langue « haute » et « basse » (*high/low*), la première réservée à des situations plus formelles, la seconde à un usage plus familier. La spécialisation fonctionnelle des deux idiomes est telle qu'il est impossible d'utiliser l'une dans les situations propres à l'autre, et que par conséquent la traduction de l'une dans l'autre est impossible, à moins de vouloir produire un effet comique.

1° Grec. L'exemple le plus caractéristique se situe précisément dans les Balkans, c'est celui du grec moderne, du moins dans la situation qui a existé jusqu'en 1974, mais dont il subsiste encore d'importantes traces. On sait que cette langue apparaît sous deux formes, dites l'une « purifiée » (*katharevousa*), l'autre « populaire » (*dhimotiki*) correspondant respectivement aux niveaux « haut » et « bas » définis par Ferguson. La langue « haute » s'impose (ou s'imposait) dans les textes religieux, les discours officiels, les travaux scientifiques, l'enseignement, la langue « basse » dans la vie courante, mais aussi dans les œuvres littéraires : poésie, roman.

Ces formes sont codifiées toutes les deux, on a publié de nombreuses grammaires de l'une et de l'autre ; même la langue « populaire », écrite par les plus grands auteurs, ne se fonde pas dans la multiplicité des dialectes. Il s'agit bien de standards. Tous deux reposent bien sur un même système linguistique. Le système phonologique est exactement le même. Les différences concernent la morphologie, la syntaxe, et plus encore le vocabulaire, et tiennent à ce que la langue « purifiée » a conservé volontairement des formes, des tournures et des lexèmes archaïques, plus proches du grec ancien, tout en les enchâssant dans le système moderne.

On ne traduit pas d'une de ces « langues » dans l'autre. L'opinion unanime est qu'il s'agit de deux formes d'une seule et même langue, car la foi en l'unité de l'hellénisme est très forte chez tous les Grecs, elle est un élément essentiel de leur sentiment national.

Sans doute faut-il parler de cette diglossie au passé. La concurrence entre les deux standards a été vive en Grèce au cours des deux derniers siècles, et s'est vite politisée. La défense intransigeante de la langue « purifiée » était le fait de la droite, et son élimination au profit de la langue « populaire » revendiquée par la gauche. La dictature des colonels de 1967 à 1974 a été marquée par un retour offensif et autoritaire de l'usage puriste. Aussi, après la restauration de la démocratie, le nouveau gouvernement a proclamé son abolition officielle, et désormais la langue « populaire » est seule enseignée et, en principe, seule écrite. Je laisse aux spécialistes le soin d'examiner quelles traces de la diglossie ancienne subsistent encore un quart de siècle après son abolition.

2° Slaves orthodoxes. Les Serbes, les Bulgares et les Russes, ont connu autrefois la même diglossie que les Grecs entre une langue « haute », le slavon, et leurs langues parlées servant de variante « basse » : *Loquendum est russice, scribendum autem slavonice*, écrivait au XVII^e siècle le grammairien Ludolf. Ils ont échappé à cette situation plus tôt que les Grecs, et par des voies diverses : les Russes au XVIII^e siècle par un compromis entre les deux langues, les Serbes et les Bulgares au XIX^e par le triomphe de la langue « basse ». Ici on parle bien dans chaque cas de deux « langues » ; on a beaucoup insisté à l'époque soviétique sur le caractère non-russe du slavon.

3° Arabe. Sur l'autre rive de la Méditerranée au contraire la diglossie a duré jusqu'à nos jours et n'est pas près de disparaître : il s'agit des pays de langue arabe. Dans chacun d'eux, par exemple l'Égypte, coexistent l'« arabe littéral », reproduisant la langue du Coran, ou cherchant à la reproduire, quitte à la moderniser ; et l'« arabe dialectal » parlé dans le pays considéré (par exemple « arabe égyptien »). Le premier est seul écrit et enseigné, le second réservé à la conversation courante entre compatriotes. Ici, comme en Grèce, l'unité est fortement ressentie, et le nom de la langue reste le même.

Polycentrisme à base temporelle

Il s'agit d'un phénomène plus rare et plus circonscrit. Nous en trouvons un exemple dans le turc. Cette langue a subi à la suite de la révolution kémaliste, dans les années 20 et 30, une réforme volontariste et radicale : non seulement elle a changé d'alphabet en 1930 (latin au lieu d'arabe), mais son lexique a été « purifié » autoritairement par l'élimination de nombreux emprunts arabes ou persans, et leur remplacement par de vieux mots turcs, ou d'innombrables

néologismes formés à partir d'éléments autochtones. Du fait d'une volonté politique forte, de l'enseignement, des médias, ce nouveau vocabulaire s'est bien implanté en Turquie, il y est aujourd'hui universellement employé. Mais l'ancien usage, le turc « ottoman » d'avant la réforme, a continué à être utilisé dans des communautés turcophones subsistant en dehors de l'État turc et n'ayant que peu ou pas de relations avec lui. On peut en citer deux : les Turcs de Bulgarie, et parmi les Arméniens de la diaspora, ceux qui avaient conservé l'usage du turc. Ainsi, ce sont deux variantes de cette langue, séparées l'une de l'autre par une réforme vieille de trois-quarts de siècle, qui coexistent sur des territoires différents. Il s'agit, bien sûr, d'un phénomène résiduel : les Turcs de Bulgarie, avec la fin du communisme, ont commencé à rétablir les relations avec le pays voisin (comme l'avaient fait déjà ceux de Grèce), et quant aux Arméniens, les générations nouvelles n'ont aucune raison de conserver en exil une langue à quoi rien ne les rattache, sauf de très amers souvenirs. Mais, quoique temporaire et en voie de disparition, cette coexistence de deux stades successifs d'une même langue méritait d'être signalée.

Polycentrisme à base spatiale

C'est le cas le plus fréquent. Deux standards se sont formés sur deux territoires distincts, entre les dialectes desquels on ne peut pourtant déceler aucune différence déterminante : soit que vraiment aucun faisceau d'isoglosses significatif ne les sépare, soit qu'une telle limite puisse être décelée, mais ne coïncide pas avec celle des deux standards.

1° Albanais- Un tel phénomène a existé quelque temps chez les Albanais. Leur territoire linguistique est divisé en deux ensembles dialectaux bien distincts : le tosque en Albanie du Sud, et le guègue qui couvre l'Albanie du Nord et les régions albanophones d'ex-Yougoslavie, notamment le Kosovo. Dans la première moitié du XX^e siècle plusieurs variantes avaient été utilisées, mais en 1952 le régime d'Enver Hoxha fit adopter un standard fondé presque exclusivement sur le tosque. Cette mesure fut appliquée d'abord seulement en Albanie, mais pas au Kosovo où l'on continua pendant plus de vingt ans à écrire une langue fondée sur le guègue. C'est seulement en 1968 que les intellectuels et les dirigeants du Kosovo décidèrent d'adopter l'albanais officiel, celui de Tirana, bien qu'il fût assez éloigné de leur usage local. Ainsi, pendant un quart de siècle, deux standards, l'un basé sur le tosque, l'autre sur le guègue, ont coexisté. Pourtant leur limite n'était pas celle, linguistique, entre les deux dialectes, elle coïncidait avec la frontière politique entre Albanie et Yougoslavie.

C'est justement pour cette raison que les dirigeants albanais du Kosovo (poussés par la contestation étudiante) ont décidé d'abolir cette divergence. Quand, vers la fin des années 60, le pouvoir yougoslave relâche sa pression et se libéralise quelque peu, les Albanais ont le souci d'affirmer l'unité de leur nation en dépit des frontières qui les divisent. Ils acceptent d'en payer le prix en adoptant un standard assez éloigné de leur dialecte natal.

2° Roumain et moldave. En 1922, lors de la formation de l'URSS, le pouvoir soviétique détache une petite portion de territoire ukrainien au bord du Dniestr, où vit une insignifiante minorité roumaine, pour y créer une minuscule République soviétique de « Moldavie ». En 1940, cette entité multiplie sa superficie par vingt du fait de l'annexion à l'URSS de la Bessarabie, entre Dniestr et Prut, enlevée à la Roumanie et peuplée d'une nette majorité roumaine. Ladite Moldavie, dès les années 20, s'était vue dotée d'un nouveau standard, la « langue moldave », écrite en alphabet cyrillique (comme l'avait été le roumain jusqu'en 1866). Les principes orthographiques sont différents de ceux du roumain et inspirés de ceux du russe, ils reposent sur une autre interprétation phonologique des diphtongues roumaines. Le vocabulaire est chargé de mots d'origine russe pour nommer les réalités soviétiques. Quelques particularités dialectales moldaves sont censées être prises en compte. Sous cette forme, ce standard sera seul utilisé (conjointement avec le russe), aussi bien dans la « petite » Moldavie de l'entre-deux-guerres que dans la « grande » après 1945 : seul imprimé dans toutes les publications, seul enseigné, fixé dans des grammaires et dictionnaires.

Cependant les bases linguistiques de la différenciation sont inexistantes. Aucune limite dialectale significative ne coïncide avec le Dniestr, où s'arrêtait la « petite » Moldavie, ni avec le Prut, frontière de la « grande ». Les divergences orthographiques reposent sur une théorisation différente des mêmes faits, non sur des réalités distinctes. Les divergences de vocabulaire ne concernent qu'une couche superficielle du lexique. Les bases humaines sont plus faibles encore. Roumains et Moldaves ont toujours été de même culture et même religion, et se considèrent comme une seule nation.

Aussi dès la fin de l'URSS et l'indépendance de la Moldavie en 1991, l'alphabet cyrillique et l'orthographe qui va avec ont été abandonnés, les emprunts russes se sont raréfiés, et les Moldaves ont recommencé à écrire exactement comme les Roumains, et à appeler leur langue, dans l'usage courant, du nom de « roumain ». Celui de « langue moldave » a toutefois été conservé comme terme officiel, pour satisfaire les importantes minorités non-roumanophones (Russes, Ukrainiens, Gagaouzes, Bulgares, etc.) qui s'opposent avec vigueur à tout ce qui pourrait laisser présager une réunification avec la Roumanie.

La divergence linguistique était réelle chez les Albanais, fictive dans le cas roumain. Mais pour les uns et les autres c'est le sentiment national des locuteurs qui a été déterminant pour refuser le polycentrisme et y mettre fin.

3° Bulgare et macédonien - Au XIX^e siècle, lors des « renaissances nationales » dans les Balkans, les Slaves de Macédoine s'étaient pour la plupart considérés comme Bulgares, et avaient adopté la langue bulgare moderne nouvellement codifiée. Dans l'ensemble bulgare-macédonien des différences dialectales existent certes (on relève des traits proprement « macédoniens » même dans certains vieux manuscrits slaves), mais pas d'opposition clairement tranchée qui permette de parler de frontière linguistique. L'intercompréhension des Macédoniens avec les Bulgares est totale (elle est excellente aussi avec les Serbes).

Cependant, quand la Bulgarie devient indépendante en 1878, la Macédoine, elle, reste ottomane, puis, à partir de 1913, elle est rattachée à la Serbie, plus tard Yougoslavie. Elle vit donc d'une vie propre. Dès cette époque quelques auteurs avaient commencé à écrire dans les dialectes locaux. Ce ne fut donc pas une innovation absolue que la décision prise en 1945, par le pouvoir titiste, de codifier et de promouvoir une « langue littéraire » macédonienne, fondée sur ces dialectes, avec certains traits phonologiques et morphologiques caractéristiques, un bon nombre de lexèmes régionaux, et une orthographe phonétique comme celle du serbe, et non pas traditionnelle comme celle du bulgare. Pour les autorités yougoslaves il s'agissait, bien entendu, de faire taire l'irrédentisme en faveur du pays voisin. Mais de nombreux intellectuels et linguistes macédoniens adhèrent à cette entreprise avec enthousiasme.

Depuis, ce standard est seul enseigné dans les écoles et utilisé dans les publications. Il a été bien accepté par la population : l'identité macédonienne, dont la langue est le symbole, lui permet d'échapper à la nécessité de se déclarer soit bulgare, soit serbe, au gré des fluctuations politiques ; et l'identité bulgare est d'autant moins attractive que pendant les cinquante ans de communisme, la Bulgarie a toujours été moins prospère et moins libre que la Macédoine. Aussi le standard nouvellement créé est-il exclusivement utilisé, non seulement sur place, où l'on pourrait soupçonner des pressions politiques, mais même en émigration : les 300 000 Macédoniens d'Australie l'emploient aussi, et publient aux antipodes des journaux dans leur langue.

Il semble donc bien que l'opération qui a échoué en Moldavie ait réussi en Macédoine : non seulement parce que les spécificités proprement linguistiques étaient plus réelles, mais surtout parce que les circonstances historiques, politiques et économiques ont favorisé l'intériorisation de la nouvelle identité. Ce

sont donc, dans ce cas comme dans les précédents, les facteurs politiques et culturels, et non pas linguistiques, qui ont été déterminants.

4° Hors des Balkans - En s'éloignant de la péninsule, mais sans quitter les régions méditerranéennes, on pourrait trouver d'autres applications du même principe.

À Malte, les Britanniques, cherchant à contrer l'influence italienne, ont, entre les deux guerres, favorisé la codification du dialecte arabe local, écrit en caractères latins, et devenu la langue maltaise. Elle est aujourd'hui pleinement reconnue et utilisée comme langue nationale d'un État indépendant. Contrairement aux autres formes d'arabe dialectal, celle-là ne vit pas en état de diglossie avec l'arabe littéral, qui est ignoré parce que, pour une population catholique, il est dépourvu de prestige religieux. Si une autre langue concurrence aujourd'hui le maltais, c'est l'anglais.

En Corse, on écrit et enseigne aujourd'hui la langue corse, avec les péripéties que l'on sait. Linguistiquement, il s'agit d'un dialecte toscan, beaucoup plus proche de l'italien standard que ne le sont le napolitain ou le piémontais. Mais la revendication identitaire corse, qui se fait contre l'uniformisation française, ne se réclame en rien de l'Italie. C'est pourquoi la seule solution était de promouvoir un standard proprement corse, ce qui a été fait avec, semble-t-il, un fort assentiment de la population.

En Galice, la langue galicienne, aujourd'hui officielle, est linguistiquement un dialecte portugais. Si frontière dialectale il y a, elle est entre Portugal du nord et du sud, et non entre Portugal et Galice (cas semblable, linguistiquement, à celui du Kosovo). Mais les Galiciens, qui vivent depuis toujours dans le cadre de l'Espagne et se déterminent essentiellement contre le souvenir du centralisme franquiste, ne sont nullement attirés par le Portugal et préfèrent promouvoir leur propre idiome.

Cas inverse : la langue arménienne se présente avec deux variantes bien distinctes linguistiquement : arménien oriental de l'Arménie indépendante (hier soviétique) et arménien occidental de la diaspora (autrefois, avant les massacres, de l'Arménie turque). Il y a entre les deux des différences phonologiques assez nettes, avec une mutation consonantique systématique, et des divergences lexicales. Toutes deux sont codifiées, couramment écrites et enseignées, l'une dans le pays, l'autre chez les exilés. Mais cette différence est très rarement mentionnée, elle est presque tabou : le sentiment de l'unité du peuple arménien, et donc de sa langue, est plus fort que tout.

Faut-il ajouter ici l'occitan ? Une unité linguistique, dialectologique, bien nette. Deux standards, deux orthographes : l'une, « mistralienne », monodialectale et à

préhension phonétique, l'autre, « occitane », pandialectale et donc quelque peu archaisante. Aux yeux de certains, la répartition aurait pu un temps paraître géographique, de part et d'autre du Rhône, mais en fait les facteurs de choix ont été, comme on sait, multiples et complexes. Quoi qu'il en soit, ici aussi l'unité de l'ensemble, qu'on l'appelle « provençal » ou « occitan », n'est contestée par aucun des deux camps. Et l'un et l'autre rencontrent une pesante indifférence des populations. Faut-il s'en plaindre ? De petits groupes passionnés débattent, mais les foules ne s'entretuent pas. Nous sommes, heureusement, bien loin du Kosovo !

Nous avons énuméré huit cas, dans les Balkans et autour de la Méditerranée, où l'on pourrait se poser la question : une langue ou deux ? Et dans chaque cas la réponse est presque unanimement donnée par les usagers d'au moins un des idiomes en cause : « une seule langue » pour l'albanais, le roumain, l'arménien et l'occitan ; « deux » pour macédonien/bulgare, maltais/arabe, corse/italien, galicien/portugais. Chaque fois, la réponse est fournie par la *vox populi*, pas par les linguistes. Les critères purement linguistiques seraient incapables de suggérer une solution univoque.

Polycentrisme complexe : l'ensemble serbo-croate

Je traiterai de ce dernier cas à part, car il diffère de tous les précédents par sa complexité. D'abord, dans ce qu'on a appelé longtemps « territoire linguistique serbo-croate » (Bosnie-Herzégovine, Croatie, Monténégro, Serbie), on distingue aujourd'hui trois standards, et pas seulement deux : serbe, croate et bosnien (*bosanski*). Ensuite, la division n'est pas vraiment géographique : s'il y a toujours eu des régions « purement serbes » ou « purement croates », il y en avait presque autant où les populations étaient mêlées, et il en reste encore quelques îlots même après les guerres de la décennie écoulée. Enfin et surtout, la séparation entre plusieurs standards n'est pas récente, elle ne date pas du XX^e, ou au mieux du XIX^e siècle, comme dans tous les autres cas examinés jusqu'ici. Au contraire, elle a toujours existé, depuis les plus anciens textes médiévaux, mais elle a bien des fois changé de forme.

Le territoire considéré est assez homogène linguistiquement, et les principales divisions dialectales qu'on y observe ne coïncident pas avec les limites des « peuples » serbe, croate et bosniaque (division à base confessionnelle) : le principal dialecte, dit « stokavien », est commun aux trois peuples, les autres (*kajkavien* et *cakavien*) ne se rencontrent que chez une partie des Croates.

Mais jusqu'au milieu du XIX^e siècle les usages écrits des Serbes et des Croates se sont développés séparément, sans le moindre contact. Les Serbes ont connu un polycentrisme fonctionnel, une diglossie de type grec, entre le slavon, langue

savante, et le serbe parlé ; les Croates, un polycentrisme spatial entre plusieurs standards régionaux, réduits en dernier lieu, au XVIII^e siècle, à deux : kajkavien de Zagreb et stokavien de Slavonie, Bosnie, Dalmatie. L'unification du standard de chacun des deux peuples s'est faite dans la première moitié du XIX^e siècle autour de deux formes très voisines de stokavien, serbe et croate, par élimination du slavon savant chez les Serbes et du kajkavien régional chez les Croates. Dans la deuxième moitié du siècle, sous l'influence des idées « yougoslaves », des efforts conscients furent déployés, surtout du côté croate, pour unifier intégralement les deux standards du point de vue morphologique et orthographique, mais non lexical. C'est alors que s'introduit dans l'usage savant, mais pas dans la langue courante, le terme composite de « langue serbo-croate ».

L'unification politique de 1918 entraîne des effets contradictoires : à la fois un rapprochement dans la vie courante, mais bientôt aussi une réaction de rejet, devenue vite presque unanime, chez le peuple le moins nombreux des deux : les Croates. Ceux-ci avaient été jusque-là les plus chauds partisans de l'unité, mais désormais ils la récusent parce qu'ils craignent de voir leur usage propre éliminé par celui des Serbes. Ce refus mènera à l'affirmation publique de la spécificité d'une « langue croate » à deux reprises : en 1941, sous l'« État Indépendant de Croatie » installé sous l'occupation allemande, et à partir de 1967 (début de la libéralisation du régime titiste), avec une Déclaration signée par la quasi-totalité des intellectuels croates. À partir de là, la « langue croate » est reconnue par la Constitution de 1974. Avec l'indépendance de 1991, sa spécificité est proclamée avec plus de force encore. Celle du « bosnien » s'affirme aussi, et de même celle du serbe. Il ne se trouve presque plus personne pour accepter le terme de « serbo-croate ».

Comme on le voit, l'affirmation de l'unité ou de la pluralité a suivi les vicissitudes politiques. Qu'en est-il sur le plan linguistique ? Les deux standards, autrefois entièrement distincts, avaient été rapprochés de très près, mais ne s'étaient jamais complètement confondus. On pouvait dire « je parle serbo-croate » au sens virtuel (je maîtrise cette langue), mais non pas au sens actuel (je suis en train de la parler), car si je parle effectivement, j'émettrai forcément soit du serbe, soit du croate (soit du bosnien, sur lequel nous reviendrons). Il y a un système linguistique serbo-croate, mais il ne peut y avoir un texte serbo-croate : tout énoncé concret appartient à l'un ou l'autre des standards.

En quoi consiste la différence ? Pas dans l'alphabet : le croate ne s'écrit qu'en lettres latines, mais le serbe peut utiliser les deux graphies, cyrillique et latine ; pas dans la prononciation : le seul trait phonologique qui paraît discriminant : *e* en serbe pour *je* en croate, ne l'est pas, parce que les formes en *je* sont possibles également dans une variante régionale du standard serbe. Les

différences sont essentiellement lexicales, mais concernent un nombre considérable de mots, souvent des séries entières, formant d'importants micro-systèmes. Parmi les mots différents on relève relativement peu de lexèmes concrets, conservant des divergences d'origine dialectale ; mais davantage de termes abstraits, culturels, institutionnels, scientifiques, techniques, témoignant des voies différentes de pénétration dans les deux peuples des influences culturelles, et de la disparité ancienne de leurs institutions. Les lexèmes et traits différents ne sont pas d'une très forte fréquence et n'empêchent pas une intercompréhension parfaite, mais ils suffisent à marquer bien nettement l'appartenance de chaque énoncé.

La séparation politique de 1991 n'a donc pas créé ces différences, elles existaient déjà. Elle a eu seulement pour résultat qu'on les a proclamées plus fort, au lieu de les minimiser, et qu'on a cherché, surtout en Croatie, à les exacerber. Mais les tentatives de certains politiciens pour multiplier en Croatie les archaïsmes et les néologismes, afin de mieux se distancier du serbe, n'ont eu aucun succès dans le public et sont en voie de résorption. Seule subsiste la volonté affirmée de défendre sa langue, dans une attitude qui n'est pas sans rappeler celle des francophones en lutte contre le « franglais ».

Le standard bosnien, lui, ne s'est affirmé comme tel que dans ces dernières années, bien que le nom même (*bosanski*) ait été courant au XIX^e siècle. Mais il continue l'usage qui avait cours en Bosnie depuis plusieurs décennies. Il apparaît comme intermédiaire entre le croate et le serbe. Il coïncide avec le croate pour l'usage de l'alphabet latin et des formes en *je*, et en est plutôt proche pour les mots courants, mais il suit le serbe pour les mots culturels, et comporte en outre un peu plus d'emprunts turcs ou arabes.

Dépasser la fausse question

Cette exploration pourrait être poursuivie au-delà du monde méditerranéen, et nous mener vers des masses humaines qui se comptent par centaines de millions : anglais/américain, portugais/bésilien, hindi/urdu, malaysien/indonésien, avec la réponse populaire : une seule langue dans chacun des deux premiers cas, deux dans chacun des deux derniers... Je préfère conclure sur une question de méthode.

Le rôle de la linguistique n'est pas de fournir un arsenal d'arguments aux politiques qui discutent avec tant de passion pour savoir combien il y a de langues dans tel ou tel territoire, alors que leur opinion et celle de la plupart de leurs compatriotes est déjà faite, pour des raisons qui n'ont rien à voir avec la science. Ils seraient bien heureux de pouvoir raisonner par analogie et dire au linguiste : « Dans le cas A (pays lointain opportunément choisi) vous avez

démontré qu'il y a une (ou deux) langue(s), donc dans le cas B (ma chère patrie) il y en a aussi une seule (ou deux), comme je vous l'ai toujours dit ». Il faut leur répondre bien clairement que la question posée est ambiguë ; que le mot « langue », en tant que désignant un objet dénombrable, n'est pas un terme linguistique. Quel linguiste saurait déterminer sans arbitraire combien il y a de langues romanes, slaves, turques, combien de langues dans le monde ? Le mot langue est employé couramment au sens de « système de dialectes » et au sens de « standard », et en outre la différence entre langue et dialecte n'est pertinente que dans des pays développés où s'affirment nettement des standards. Le dénombrement est en tout cas impossible.

« Une ou plusieurs langues ? » est la fausse question. La plupart de ceux qui font mine d'y chercher une réponse ont leur conviction faite d'avance, pour des raisons qui n'ont rien de scientifique, et leur examen des faits est biaisé au départ par le souci qu'ils ont de fournir des arguments à leur solution toute prête. Ce n'est pas ainsi que progresse la science.

Mais l'existence même, dans les Balkans, le monde méditerranéen, et un peu partout sur la planète, de si nombreux cas où le dénombrement est en défaut, où le terme « langue » manifeste son ambiguïté, de tous ces cas de « polycentrisme », ouvre au contraire à la vraie recherche linguistique un champ d'exploration fructueux, à condition qu'il soit abordé pour lui-même, sans préoccupation politique ou patriotique préalable, et avec le seul souci de rendre compte avec rigueur de la complexité des faits. J'ai esquissé ici quelques procédés d'analyse et de classement des phénomènes de ce genre, aux franges de la linguistique et de la sociolinguistique. De nombreux travaux ont déjà été consacrés un peu partout à ces questions. Il est souhaitable que les recherches de ce type soient systématiquement poursuivies et élargies.

Puisse la science ne plus jamais être une arme au service des affrontements haineux, mais au contraire un moyen de les dépasser, et de faire servir les faux débats à une vraie compréhension du monde.

COMPARATISME

LA SYLLEPSE : LE CAS DES LANGUES ROMANES

Eulàlia VILAGINES SERRA

La syllepse, qui a intéressé depuis toujours les grammairiens, est le jeu possible entre la forme et le sens pour le singulier et le pluriel. Grevisse et Goose (1991), par exemple, en donnent la définition suivante :

« La syllepse consiste à faire l'accord d'un mot, non avec le mot auquel il se rapporte selon les règles grammaticales, mais avec le terme qu'on a dans l'idée (accord avec le sens, *constructio at sensum*, ou accord logique). »

Par exemple, dans (1) :

(1) Une dizaine de personnes gesticulaient devant la porte du jardin, mais à distance. (écrit, lit.)

C'est l'accord entre « une dizaine de personnes », syntagme nominal considéré comme singulier, et le verbe au pluriel « gesticulaient ». L'inverse reste possible. Dans l'exemple (2), le même type de nom, « une trentaine de Musulmanes », est suivi d'un verbe au singulier, « était venue » :

(2) La première fois, une trentaine de Musulmanes *était venue* et pas une Européenne. (écrit, lit.)

Cet apparent décalage entre grammaire et sens se retrouverait dans plusieurs types d'accord : la syllepse de genre, la syllepse de nombre, la syllepse de personne, la syllepse selon le distributif.

Je n'aborderai dans mon travail que l'accord de syllepse en nombre, et lorsque je parlerai d'accord de syllepse, je sous-entendrai toujours « de nombre ».

Selon Nathalie Fournier, en français, la syllepse du nombre, qui est usuelle au XVI^e siècle, est un des points les plus débattus par les grammairiens au XVII^e siècle, dans un sens qui va vers la fixation du principe de l'accord grammatical et la réglementation des cas autorisés de syllepse. Par ailleurs, Bernard Colombat montre dans son article « L'accord dans la grammaire latine de la Renaissance à Beauzée »¹ que pour les auteurs du Moyen Age et l'Humanisme, les figures de construction ont été un moyen de trouver des solutions à ce genre de soucis.

« C'est précisément pour répondre à ce type de problème que le Moyen Age, suivi par l'Humanisme, a mis au point un arsenal de figures de construction. Pratiquement toutes concernent l'accord : (...) Autrement dit, avant même la

¹ Dans *Faits de langue* n° 8.

formulation des règles concernant l'accord, les grammairiens disposaient de règles pour décrire les exceptions à cet accord » (1996 : 21-2).

Il s'agissait donc de chercher des solutions à un souci pratique.

Comme il arrive souvent dans les travaux de recherche, il n'a pas été simple de se faire une idée de ce qui avait été écrit sur la question de la syllepse. Il y a deux types de travaux. Dans la perspective classique, la syllepse est encore présentée comme un manquement aux règles. C'est le cas de la position de Martínez dans Bosque et Demonte (1999 : 2701) :

« Se denomina 'silepsis' a la falta de concordancia entre unidades lingüísticas, que no obstante resulta aceptable o admisible por diferentes razones ». (Martínez, dans Bosque et Demonte, 1999, Vol.2, p. 2701)

Dans la perspective moderne, il y a une autre optique qui fait de la syllepse un révélateur de la classe nominale. C'est le cas de la notion de perspective, c'est-à-dire qu'il y a deux façons de voir le référent : soit le groupe est considéré comme un corps unique, soit il est considéré comme une collection d'individus. On trouve cette idée dans les grammaires de Riegel (1994), Bosque et Demonte (1999) ou Badia i Margarit (1994). La syllepse est aussi analysée en utilisant la notion de tête de syntagme, aujourd'hui acquise par la communauté scientifique.

J'ai choisi l'optique lexico-grammaticale, car je croyais que pour traiter la syllepse, on devait combiner le point de vue grammatical et le point de vue sémantique. Le corpus d'exemples le plus important sur lequel s'est basée l'analyse correspond au cas typique d'accord entre un syntagme nominal source d'accord et un verbe cible d'accord (cf. exemples (1) et (2)). Cela dit, je prendrai le phénomène de la syllepse au sens le plus large. Des phénomènes qui ne figurent pas dans la définition classique de la syllepse, et qui ont été rarement pris en compte par les grammaires, tels que les accords à distance, sont ici intégrés.

Comme la question se pose également dans d'autres langues, il semblait avisé d'examiner quelles étaient les limites de la norme et de l'usage dans d'autres langues romanes. J'ai choisi des langues de la Romania occidentale et centrale, le français, le catalan, l'espagnol et le portugais. L'analyse s'est basée sur un corpus d'environ 4 000 exemples écrits et oraux, 741 pour le portugais, 923 pour l'espagnol, 979 pour le catalan et 1 342 pour le français.

Dans cet article, je propose seulement une présentation sommaire de deux points abordés dans mon travail de thèse : la nature du syntagme nominal susceptible de déclencher l'accord de syllepse et le rôle du contexte syntaxique dans l'accord de syllepse.

1. Les traits du syntagme nominal source

Les accords de syllepse n'apparaissent pas de manière arbitraire. Seuls certains syntagmes nominaux considérés comme singuliers ont la possibilité de déclencher un accord au pluriel. J'ai travaillé sur deux types de syntagme :

- type *N*
- type *N₁ de N₂*

Dans le premier cas, le substantif est un collectif (type *une foule, la troupe*) ou un quantifieur (comme *une trentaine, 20%*). Il arrive que ces substantifs soient suivis d'un substantif pluriel (type *une foule de personnes, une trentaine de Musulmanes*). C'est le type *N₁ de N₂*.

La liste de lexèmes concernés est la même dans les quatre langues. Cette partie est consacrée à définir les collectifs et les quantifieurs.

1.1. Collectifs : définition générale

La notion de collectif, très ancienne, a reçu récemment des définitions plus élaborées. Les collectifs sont des noms à morphologie de singulier « dénotant des entités composées de parties distinctes préalablement constituées, homogènes entre elles, mais autonomes et hétérogènes par rapport au tout » (Flaux : 1999, p. 472).

Une « *population* » est faite de parties distinctes qui sont les personnes, « Ensemble des personnes qui habitent un espace, une terre. » (*Le Petit Robert*). Toutes les personnes sont homogènes entre elles, dans tous les traits qui en font la classe humaine (ce sont des humains et non des chats ou des fleurs). Chaque personne existe indépendamment de la « *population* », mais aucune personne n'est à elle seule une population.

Tous les collectifs ne se réalisent pas dans la langue d'une même manière. Dans l'usage on observe des différences. Par exemple, il arrive que le collectif n'accepte pas le complément désignant les parties constitutives, car ce serait tautologique. C'est le cas de « *la gent(e)* » (C-E-P).

D'autres l'acceptent seulement si celui-ci est à son tour qualifié :

- *un personnel de gens incapables* (F)
- *una tribu d'hommes forts* (C)
- *un jurado de hombres buenos* (E)
- *una familia d'homens e mulheres* (P)

Dans d'autres cas, ils acceptent facilement le complément désignant les parties constitutives :

- *une foule de malades* (F)
- *un equip de metges* (C)
- *una panda de paletos* (E)

- *un grupo de jóvenes (P)*

Quand il y a le complément désignant les parties constitutives, le collectif opère une sélection lexicale sur ce complément. Une équipe, par exemple, est un « groupe de personnes unies dans une tâche commune. » (*Le Petit Robert*). On s'attend donc à un substantif qui corresponde à cette idée, comme « *un equip d'infermeres* » (C) ou « *un equipo de médicos* » (E).

1.2. Quantifieurs

Une très grande quantité d'études ont été consacrées aux quantifieurs. Je n'ai pris ici que les acceptions en relation avec le problème de la syllepse. À cause de sa vaste étendue, la notion de quantifieur² est plus vague que celle de collectif dans les ouvrages. Pour Creissels (1991 : 143) la quantification est une opération d'« extraction », par laquelle « l'énonciateur envisage le référent visé comme un sous-ensemble d'un ensemble ou comme fragment d'une masse ». Les systèmes mis au point pour quantifier diffèrent non seulement d'une langue à l'autre, mais à l'intérieur d'une même langue.

Dans mon étude, je ne me suis pas intéressée à tout type de quantification. J'ai retenu seulement les quantificateurs qui ont une forme nominale « dét. N de », avec une opération d'extraction évoquée par Creissels (1991) et les expressions de pourcentage : « *une douzaine de députés* », « *la majorité des députés* », « *la moitié des enseignants* », « *trente pour cent des enseignants* ». Benninger (2001 : 25) appelle ces termes « substantifs quantificateurs ».

Les quantifieurs, néanmoins, n'expriment pas tous la quantité de la même manière, certains le font par un type de nombre (*douzaine, moitié*), certains expriment une partie du tout (*la plupart, une partie*), d'autres par la succession dans la quantité (*série, succession*).

Le quantifieur en construction *N de N* n'opère pas de sélection lexicale sur le complément :

- *la majorité des bactéries, la moitié des effectifs (F)*
- *una mayoría de proyectos, la meitat de ferits (C)*
- *la mayoría de puertas blindadas, la mitad de las lenguas (E)*
- *a maioria dos hipocampus, a metade dos inquiridos (P)*

Quand un quantifieur est à lui seul représentant du syntagme (type *la majorité, la moitié*), il a une valeur anaphorique. C'est-à-dire que, bien qu'il soit syntaxiquement autonome, sémantiquement il ne l'est pas. Cela fait une grande différence par rapport aux collectifs, entièrement autonomes, aussi bien du point de vue syntaxique que sémantique.

² Trois termes sont utilisés : quantifieur, quantificateur et quantifiant.