

SUR LA DÉLIMITATION ET L'INDIVIDUATION DES LANGUES

Avec des exemples pris principalement dans le domaine roman

Georg KREMNITZ
Universität Wien

Kann man sprachen zählen ?
Wolf Thümmel

1. LE PROBLÈME : « LANGUE » VS. « VARIÉTÉ »

On sait que jusqu'à présent les sciences du langage n'ont pas réussi à formuler une définition acceptable pour ce qu'est une *langue* individualisée ni à trouver des critères généralement admis pour différencier les termes *langue* et *variété* (anciennement, on a surtout employé des termes comme *dialecte* etc.; je ne parlerai ici que des *variétés* comme terme général désignant toutes les variations du langage humain, soit dans l'espace, dans l'emploi social, situationnel ou médial, tandis que je réserverai le terme de *dialecte*, si nécessaire, aux variétés dans l'espace, lui préférant toutefois « variété spatiale » ou « variété dans l'espace »). Dans les cours universitaires, les enseignants se contentent le plus souvent de donner des définitions provenant de la linguistique interne, sans en discuter les problèmes, et dans les publications, les auteurs les esquivent en général. Or, d'une part les sciences de la communication et du langage ont dépassé leur stade uniquement interne, la communication est un fait de société, de l'autre des arguments linguistiques apparaissant de plus en plus souvent dans les débats politiques, il serait donc utile –c'est peu dire– de disposer de définitions claires.¹ Cependant, le problème échappe aux sciences du langage, parce que ce n'est pas exclusivement un problème linguistique, ou, mieux dit, ce n'est pas un problème que les sciences du langage peuvent résoudre sous tous les angles.² Les présentes réflexions ne pourront donc pas proposer de solution du problème, elles se contenteront modestement d'en présenter les contours tels qu'ils m'apparaissent aujourd'hui et de montrer quelques conséquences que pourront entraîner certaines attitudes définitoires. Je présenterai d'abord les solutions que la romanistique a propo-

1. Je remercie mon ami Heinz Tichy, spécialiste des minorités linguistiques en Autriche, de m'avoir plusieurs fois posé le problème. Je remercie Barbara Czernilofsky et tout particulièrement François Pic d'avoir lu ce texte avec beaucoup de soin. Toutefois: toutes ses faiblesses restent les miennes.

2. On peut s'amuser en lisant mes efforts désespérés de trouver des critères valables et utilisables pour la délimitation de ces termes dans Kremnitz (1974 : 1-7, 38-46). J'y pêche par la négligence trop poussée du politique.

sées jusqu'à présent –les listes des langues, surtout dans les grands manuels– avant de me concentrer sur les stratégies employées pour trouver des définitions et les inconvénients que ces définitions peuvent poser. Comme je ne pourrai pas proposer de solution *scientifique* à un problème *social*, je me contenterai à la fin de quelques remarques pragmatiques. Je continue cependant à penser qu'une *rationalisation* des débats, qui tienne compte des avantages et des inconvénients des solutions proposées serait pour le moins utile.

2. EVOLUTION DU NOMBRE DES LANGUES ROMANES DANS L'HISTOIRE DE LA RECHERCHE

On sait que le premier à se préoccuper des langues vulgaires et notamment romanes a été Dante Alighieri (1265-1321) dans son texte *De vulgari eloquentia* (vers 1304). Il tente d'y esquisser le rôle futur d'un italien qui aurait des fonctions comparables à celles du latin de son temps, un italien normativisé, en termes modernes (*illustre, cardinale, aulicum et curiale vulgare*, I, XVI, 6). Dans ce contexte, il établit une première typologie des langues en Europe, et parmi elles des langues romanes (VIII,5) où il parle d'un *ydiuma tripharium*, un idiome qui apparaît sous trois formes : « nam alii *oc*, alii *oil*, alii *si* affirmando locuntur, ut puta Yspani, Franci et Latini ». Pour lui, trois variétés du vulgaire latin sont suffisamment différenciées pour être énumérées. Ce sont celles qui jouent un rôle, à cette époque, comme langues de pouvoir et de littérature, bien que la délimitation exacte reste un peu dans le vague (l'*oc* recouvre le sud de la France actuelle et la Péninsule Ibérique chrétienne, l'*oil* la France du nord et le *si* l'Italie, cf. VIII, 6). A partir de ce moment, le nombre des langues romanes pris en considération va régulièrement augmenter. Les représentants des Lumières connaissent les relations génétiques entre le latin et les langues romanes, mais nous savons que leurs connaissances étaient souvent intuitives et peu sûres. De là l'erreur de J. F. M. Raynouard (1761-1836) qui pensait que l'ancien occitan était à la base des langues néolatines (il ne tenait pas compte du roumain pour cette affirmation).

Friedrich Diez (1794-1876) voit, dans sa *Grammatik der romanischen Sprachen* en trois volumes (1836-43), ouvrage qui, dans un certain sens, pose les bases de la *romanistique* moderne, six langues romanes : « zwei östliche, die italiänische und walachische ; zwei südwestliche, die spanische und portugiesische ; zwei nordwestliche, die provenzalische und französische [deux à l'est, l'italien et le valaque ; deux au sud-ouest, l'espagnol et le portugais ; deux au nord-ouest, le provençal et le français, p.1] ». Ses critères essentiels sont les particularités grammaticales, la *distance* linguistique (avec les termes de Heinz Kloss : l'*Abstand*, cf. 3.2.) et l'*importance* littéraire. C'est pourquoi il n'énumère ni le rhéto-roman (« Churwälsch »), ni le sarde, et relègue le catalan à une place inférieure, comme variété de l'occitan (I, 3). Mais, plus tard, il constatera que cette répartition n'est pas entièrement satisfaisante, en ce qui concerne le catalan, et ajoutera : « Die katalanische Sprache steht zur provenzalischen nicht eigentlich im Verhältnis einer Mundart ; sie ist vielmehr ein selbständiges mit ihr zunächst verwandtes Idiom [la langue catalane ne se trouve pas vraiment dans une relation de dialecte avec la provençale ; elle est plutôt un idiome indépendant qui lui est intimement apparentée] » (I, 112). Ce qui est important dans cette énumération c'est que, dès le début, la *diversité des critères employés* empêche une définition claire. Il s'agit d'un problème dû essentiellement à la tentative d'utiliser un terme du langage quotidien pour une terminologie scientifique exacte. Mais derrière cette tentative il y a une incertitude profonde sur l'objet même de cette déno-

mination, parce qu'il n'y a pas de consensus social. Des problèmes comparables se retrouvent dans les autres philologies modernes qui voient le jour du vivant de Diez.

Wilhelm Meyer-Lübke (1861-1936), le plus important et influent romaniste de son temps, propose, presque un siècle plus tard, dans un ouvrage destiné à l'enseignement universitaire, une liste de neuf langues romanes, à savoir (de l'est à l'ouest) : le roumain, le dalmate, le rhéto-roman [comprenant l'ensemble des variétés rhétiques en Suisse, au Tyrol et dans le Frioul], l'italien, le sarde, le provençal, le français, l'espagnol et le portugais (1920, 16-17). Il justifie cette liste principalement par les situations politiques et littéraires. Plus loin (23-24), il présente une liste des variétés de ces langues qui contient beaucoup d'éléments qui apparaissent de nos jours comme langues indépendantes. Ses critères relèvent donc, entre autres, nettement de la politique linguistique (*ante litteram*) : il ne concède le grade de langue qu'aux variétés qui ont obtenu une reconnaissance officielle suffisante (à la seule exception du dalmate, définitivement éteint en 1898, mais « découvert » et décrit par son disciple Matteo Bartoli, 1906). Cette liste à neuf entrées, qui affecte au catalan le rang d'une variété du provençal qui avance consécutivement au recul de l'arabe (1920, 26-27), rencontra le mécontentement surtout des représentants de la *Renaixença* catalane de son temps (surtout Antoni M. Alcover), qui avaient espéré un soutien intellectuel pour leurs efforts. Ils fournirent beaucoup de matériaux à Meyer-Lübke, lui permirent de connaître le pays, et de publier en 1925, une monographie sur cette langue, complétant ainsi la liste de son manuel. Dorénavant, dans son entourage, le nombre des langues romanes était fixé à dix.

Un autre grand romaniste, Graziadio Isaia Ascoli (1829-1907), avait proposé, un peu plus tôt, de se baser uniquement sur des faits de linguistique *interne* ou *formelle* pour délimiter les langues, sans recours à des œuvres littéraires ou à des faits politiques. À l'aide de cette méthode, il avait isolé deux nouvelles unités qu'il considérait comme langues romanes, d'une part les variétés franco-romanes (que Meyer-Lübke retiendra pour sa liste) et de l'autre les variétés franco-provençales (que Meyer-Lübke considérera comme variétés du français). Mais il déclencha de cette façon une âpre polémique, surtout avec le linguiste français Paul Meyer, car si l'on utilisait ces critères, l'unité linguistique de la France était mise en question. Les critères scientifiques d'Ascoli se heurtaient à des considérations politiques (on sait que Meyer et son collègue Gaston Paris, qui occupaient les postes de commande de la philologie à Paris, ne voyaient pas non plus d'un bon œil les tentatives de démontrer la bipartition linguistique de la France, entre domaine d'oc et domaine d'oïl, entrepris presque au même moment par les félibres occitans Charles de Tourtoulon et Octavien Bringuier, bien que dans leur cas il s'agissait même d'une mission officielle ; par chance – si l'on peut dire –, Bringuier mourut et le projet ne put pas être terminé, cf. Tourtoulon/Bringuier 1876). Un des problèmes fondamentaux devient visible dès cet épisode.

Cette liste des dix ou onze unités se retrouvera dans les manuels de linguistique et philologie romanes pendant un bon moment. Si les *Eléments* de Bourciez, un des rares ouvrages de synthèse publiés en France, reste un peu en-deçà de cette liste, ne tenant pratiquement pas compte du catalan, du franco-provençal, du sarde ni du dalmate, les ouvrages de référence de Tagliavini (11), de Bal (11) ou de Renzi (11), par exemple, la maintiennent. D'autres, plus méfiants, ne dressent pas de liste complète des langues romanes. Les deux dernières sommes des langues romanes écrites par un seul auteur, celles de Lausberg (1956-62) et de Bec (1970-71), s'en tiennent à cette liste. Il est vrai que Lausberg ne donne pas de place à part au franco-provençal qu'il mentionne toutefois (se rapprochant sous cet aspect de Meyer-Lübke). Avant de dresser sa liste (I, 40-60), il indique comme critère principal l'emploi des langues dans l'écriture (officielle), mais il avoue que ce critère ne peut soutenir l'ensemble de sa liste. A la fin de cette discussion, il parle d'une « ungleichmäßig motivierte Serie von zehn romanischen Spra-

chen [une série inégalement motivée de dix langues romanes] ». Les ouvrages énumérés ici ne servent que d'exemples : la liste pourrait facilement s'allonger, le constat resterait sensiblement le même.

Cependant, depuis une trentaine d'années, le nombre des langues romanes semble littéralement exploser. La somme de Holtus/Metzeltin/Schmitt, le *LRL*, considère les unités suivantes comme langues romanes : le roumain, le dalmate et l'istrien, le frioulan, le ladin, le romanche (vol. III), l'italien, le corse, le sarde (vol. IV), le français, l'occitan, le catalan (vol. V), l'aragonais/navarrais, l'espagnol, l'asturien/léonais, le galicien, le portugais (vol. VI), auxquels s'ajoutent dans le vol. VII, sous le terme de langues en contact, d'autres unités encore, quoique avec un traitement différent. Le nombre des langues pleinement traitées s'élève de cette façon à 16. Ainsi, il est clair que les critères ne peuvent plus être les mêmes qu'auparavant. Il est vrai que certaines langues ont eu accès à une officialité relative, comme le galicien en Espagne, ou tout au moins à une reconnaissance partielle, plus ou moins implicite, comme le corse en France, le domaine du rhétique par contre a éclaté en trois unités différentes qui, elles aussi, ont toutes un certain statut dans leur territoire respectif, mais le dalmate reste une langue disparue quoique prise en compte, à la différence d'autres langues romanes disparues, et les variétés istriennes ne jouissent d'aucune reconnaissance officielle et paraissent être en voie de disparition. De même, les statuts de l'aragonais et de l'asturien ne sont guère assurés. D'autres langues avec un statut officiel ne trouvent qu'un traitement marginal : ainsi, le moldave, en 1989, quand paraît le vol. III, encore langue officielle et officiellement distinguée du roumain dans la République Soviétique Moldave, n'a droit qu'à un article historique, mais pas à un traitement linguistique (Heitmann 1989). Si l'on est de l'opinion que le moldave fait partie de l'ensemble roumain, cette solution s'accepte facilement, mais les critères de l'officialité et d'une politique linguistique officielle de plus d'un demi-siècle volent en éclats (on pourrait voir dans ce traitement également une prise de position idéologique).³ Les solutions adoptées par le *LRL*, si on peut les qualifier de pragmatiques, n'apportent pas de solution à la question de la définition de « la langue ».

Cette tendance à la multiplication des unités considérées comme « langues » s'accroît dans les dernières années. Quand en 1999 Bernard Cerquiglini, alors directeur de l'Institut national de la langue française, est chargé par le gouvernement de dresser une liste des « langues de la France », y apparaissent les *langues d'oïl* qui ont commencé à surgir dans les discussions publiques dans les années '80, à la suite du rapport Giordan (1982) qui, en tant que premier texte officiel, a soulevé le problème de la multiplicité des langues en France, et des discussions qui l'ont suivi (cf. Giordan 1984). Si dans Giordan (1984) ne sont mentionnés que le gallo et le picard, Cerquiglini énumère, sous le titre général des langues d'oïl, le franc-comtois, le wallon, le picard, le normand, le gallo, le poitevin-saintongeais, le bourguignon-morvandiau et le lorrain. Il est licite de penser que la forme de cette énumération ne concède pas à ces langues des places équivalentes à celles du français ou de l'italien, mais elles apparaissent comme « langues » (ce qui, soit dit en passant, suscite des soupçons chez les défenseurs des langues « régionales » *historiques* de la France). De plus, la liste contient les quatre créoles « à base lexicale française » parlés sur le sol français, à savoir le martiniquais, le guadeloupéen, le guyanais et le réunionnais.⁴ On sait qu'à l'intérieur du mouvement renaissantiste occitan il

3. Pour les relations entre roumain « roumain » et roumain « moldave » cf. maintenant les travaux qui se font autour de Klaus Bochmann à l'Académie Saxe des Sciences, et notamment la synthèse d'Urspring 2005 ; pour la genèse du problème, l'ensemble des travaux de Klaus Heitmann sur ce sujet.

4. Alors que le Ministère de l'Éducation nationale n'a créé qu'*un seul CAPES* de créole. On voit que les contradictions peuvent se produire à l'intérieur d'un même pouvoir politique.

y a des tendances à faire éclater l'ensemble occitan en plusieurs langues différentes –provençal, gascon, auvergnat, etc.– dont les défenseurs combattent souvent plus énergiquement les tenants d'une position différente dans la bataille des frontières que les progrès du français au dépens des variétés de l'occitan. Ces contradictions remontent dans l'histoire, mais la renaissance du XIX^e siècle partait toujours du principe de l'unité de la langue. Pourquoi ce changement de direction ?

L'Italie a promulgué, en 1999, une loi sur la protection de ses minorités linguistiques, incluant la liste des douze langues minoritaires qu'elle reconnaît, avec quelques différences de détail toutefois. Depuis ce moment, certaines régions italiennes luttent pour la reconnaissance de *leur* variété linguistique spécifique, notamment le Piémont pour le piémontais et la région Vénétie pour celle du vénitien. Or, dans tous les manuels classiques ces variétés apparaissent comme dialectes, variétés spatiales de l'italien, souvent regroupés sous le terme de variétés gallo-italiennes, parce que ce sont des parlers de transition vers la Gallo-Romania, donc essentiellement vers le franco-provençal et l'occitan, le français mis à part. Le zèle subit du personnel politique régional, on s'en doute, n'est pas dû (uniquement) à une prise de conscience linguistique différentielle, mais surtout au fait que ces régions espèrent de cette façon profiter des crédits (pourtant modestes) que Rome verse aux minorités reconnues. Si jamais une de ces régions obtient gain de cause, on pourra parier, je pense, sur la naissance d'autres langues nouvelles ...

Dans la Péninsule ibérique, de nouvelles variétés romanes aspirent à une reconnaissance officielle comme langues –nous avons vu que l'aragonais et l'asturien ont fait leur entrée dans le *LRL*–, tandis que dans d'autres cas des luttes internes menacent une fragmentation du domaine –je pense évidemment aux débats autour de la place du valencien. Surtout dans ce dernier cas, il est évident que l'opinion unanime des chercheurs est confrontée à une volonté politique contraire, qui semble actuellement trouver auprès de la population un certain appui qui toutefois est loin d'être général. Le nouveau statut de la Catalogne de 2006 reconnaît l'aranais comme langue, précisant toutefois qu'il s'agit d'une variété de l'occitan (gascon) ; mais on assiste à une certaine tendance publique à l'autonomisation de cette variété (art. 6 § 5 ; 11 ; 36, etc.). Dans une récente anthologie des langues minoritaires de la Péninsule, les auteurs comptent douze langues romanes, dont en plus de celles déjà cités l'*extremeño*, *a fala de Xálima* (proche du galaïco-portugais), le murcien, le mirandais et le portugais d'Olivenza, mais ne reconnaissent pas de différence entre catalan et valencien –les deux éditeurs sont d'éminents catalanistes– question qui en ce moment est la plus âprement débattue dans l'Etat espagnol, sur le plan politique (Raddatz/Torrent-Lenzen 2006). Même le Portugal, que tous les anciens manuels considéraient comme un des rares Etats européens sans minorité autochtone (avec l'Islande), a fini par en trouver une (et même deux, si l'on pense à la langue des signes) ; il a fini par donner un statut officiel au *mirandais* qui, linguistiquement parlant, n'est rien d'autre que la prolongation de l'asturien vers le sud (cf. par exemple Merlan 2004).

En Roumanie, les locuteurs de l'aroumain aspirent à la reconnaissance de cette langue comme langue minoritaire, alors que la politique officielle la classe comme variété du roumain, en dépit d'une certaine tradition écrite. Là encore, les raisons ne sont pas uniquement scientifiques, mais également matérielles. Les discussions sur le rapport entre roumain et moldave ont déjà été évoquées; la dénomination différente des deux variétés n'empêche pas qu'elles utilisent la même grammaire normative, celle de l'Académie de Bucarest.

En Belgique, une revue se publie depuis une quinzaine d'années, *micRomania*, dédiée à la poésie dans les petites (*micRo*) langues romanes. Chaque numéro contient une anthologie poétique précédée d'un article défendant la diversité des langues romanes. Comme elle publie

sous cet aspect des textes de variétés linguistiques ayant parfois une aire d'emploi très restreinte, le lecteur a du moins l'impression que la revue défend les aspirations de quelques auteurs du Tessin ou de l'Istrie à élever leur parler en langue romane. Certes, on se trouve en domaine poétique, mais les positions de la revue dans le domaine de la définition de ce qu'elle considère comme « langue » sont peu claires. Elles semblent appuyer des positions fragmentaristes. Le lecteur a l'impression que la volonté de (quelques) locuteurs suffit à reconnaître une variété comme langue. Ce qui est parfaitement licite en poésie, peut-il servir de ligne directrice pour une politique linguistique ?

Une mention à part mérite le catalogue des langues du monde dressé régulièrement par le *Summer Institut of Linguistics* (SIL), une association intégriste protestante aux Etats-Unis, qui, dans sa dernière édition, énumère exactement 6912 langues parlées dans le monde (bien que les éditeurs relativisent quelque peu ce chiffre dans les parties introductives). D'une part cette liste semble reproduire, bien qu'involontairement (?), le « comique de précision » cher à Rabelais. D'autre part, on parvient à ce chiffre en faisant, dans le domaine allemand la distinction entre bavarois, haut- et bas-saxon et colonais (tandis que le viennois est absent), scindant ainsi en plusieurs éléments un espace linguistique qui fonctionne comme une unité. Dans le domaine roman, il n'y a pas d'entrée « occitan », mais seulement « provençal », « limousin », « languedocien », « gascon », « auvergnat » et même « shuadit » (ou judéo-provençal). Un tel assemblage, dépourvu de critères clairs, ne mériterait pas une ligne, si le SIL ne tendait pas à promouvoir internationalement ces conceptions et si certaines bibliothèques n'avaient pas tendance à les accepter, obligés par ce que l'on ne peut appeler qu'un impérialisme culturel de la frange intégriste des Américains du nord (cf. Gordon ¹⁵2005)⁵ –ou alors une incompétence professionnelle remarquable.

Il serait aisé de multiplier les exemples. Il est certain qu'un consensus des spécialistes –établi, il est vrai, sur des critères très mous– qui a guidé l'opinion publique dans ce domaine pendant un bon siècle, ne fonctionne plus. Le personnel politique, les journalistes, d'autres intermédiaires interviennent dans les discussions, souvent sans connaissance des mécanismes de la communication, parlant uniquement « avec les tripes » (quiconque a une pratique des réunions des « mouvements de défense et promotion des langues minoritaires » connaît de telles déclarations). On peut considérer cela comme un mouvement vers la démocratisation, mais une telle démocratisation est-elle possible dans tous les domaines sans connaissances spécifiques ? Peut-on par exemple décider par consultation populaire le traitement médical qu'on appliquera à un malade ? Quand on pose la question de cette manière, elle paraît ridicule, mais les problèmes de la communication, se laissent-ils régler par des procédés semblables ? Ne nécessitent-ils pas un minimum de connaissances ? Et une séparation stricte entre l'engagement personnel et l'analyse scientifique ? Il est vrai qu'il ne s'agit pas *uniquement* d'un problème scientifique, mais en même temps d'un problème politique qui fait que toutes les prises de positions scientifiques risquent presque inévitablement de tomber sous le soupçon de l'*idéologie*. D'autre part, la communication et la linguistique ont leurs règles, et il peut être utile –voire indispensable– d'en tenir compte.

Je me propose à présent de passer en revue les critères qui ont servi à l'individuation des langues avant de regarder quelles conséquences on peut en tirer.

5. Je dois le détail de ces données au mémoire de mon étudiant Clemens Sander sur le SIL et ses pratiques linguistiques.

3. CRITÈRES POSSIBLES POUR DÉLIMITER DES LANGUES

Avant d'entrer dans les détails, je voudrais présenter quelques réflexions générales et quelques éléments à propos de la terminologie employée. Je rappelle seulement pour mémoire l'importance des termes de *statut* et *prestige*, le premier désignant la situation légale d'une langue ou variété (mais dès que cette situation est assurée par un statut d'officialité, l'unité en question est considérée comme *langue*, dans son territoire respectif, –sauf en cas de forte crise), le deuxième désignant la reconnaissance dont jouit une langue, soit parmi ses locuteurs (*prestige interne*) soit à l'extérieur (*prestige externe*) ; surtout les variétés qui n'ont pas d'Etat à leur disposition dépendent, pour la réalisation de leur potentiel communicatif, fortement du prestige externe. Entre ces deux termes se glisse celui de *valeur d'usage* ou *valeur communicative*, proposé par Roberto Bein (2001) et moi-même (2003) ; il tente de désigner l'emploi réel d'une langue ou variété dans une société, indépendamment de toute promotion ou répression officielles. Il semble fort difficile de délimiter les contours de cet usage réel, mais surtout depuis que l'immigration se multiplie une autre fois (la mémoire collective ne retient plus les vagues antérieures), partout en Europe, on constate un abîme entre les pratiques communicatives déclarées ou promues et les réalités communicatives observables. Des politiques qui vont souvent à l'encontre des pratiques visibles peuvent avoir des conséquences sociales, essentiellement négatives jusqu'à présent. C'est pourquoi nous considérons les deux termes de *statut* et *prestige* comme insuffisants désormais.

Je rappelle également pour mémoire les termes de *normativisation* et de *normalisation* de la tradition de la sociolinguistique catalane dans l'acception que j'ai moi-même proposée (Kremnitz 2003).

3.1. COMMUNICATION ET DÉMARCACTION

Saussure a rendu célèbre les deux faces du signe linguistique, à savoir son image acoustique et le concept qui apparaissent nécessairement ensemble, comme les deux faces d'une pièce de monnaie. La communication connaît une paire comparable, dont les composantes apparaissent toujours ensemble : chaque fois que nous parlons, ou, plus généralement dit, quand nous nous trouvons dans une situation de communication, nous activons une tension entre une *fonction communicative* et une *fonction démarcative* ou *identitaire* de notre discours. Nous parlons pour communiquer avec autrui, mais en même temps, par les particularités de notre façon de parler, nous nous démarquons des autres. Les différents cercles concentriques de facilité de la communication désignent en même temps les différents degrés de *l'identité collective*. Dans la communication quotidienne, quand nous évoluons dans notre entourage habituel, ces particularités sont insignifiantes. Pourtant, dans ma jeunesse, un locuteur expert de la variété souabe de l'allemand –dans laquelle j'ai grandi– pouvait reconnaître que son interlocuteur venait d'un endroit à dix ou quinze kilomètres de distance et souvent lui dire de quel endroit il était originaire. J'ai fait des expériences semblables dans d'autres domaines linguistiques, elles ont été démontrées jadis de façon exemplaire par deux articles complémentaires de Jean Séguy (1973) et Xavier Ravier (1973). De nos jours et en raison d'une part de la généralisation des médias oraux et visuels et de l'autre de la mobilité croissante, les locuteurs perdent beaucoup de cette faculté (de plus, elle est devenue moins fonctionnelle qu'auparavant). Mais nous reconnaissons toujours que quelqu'un parle la même langue que nous, qu'il la parle avec la facilité d'un locuteur natif ou avec un accent étranger. La démarcation n'est pas uniquement

spatiale, elle peut être également sociale, sexuelle, etc. Elle consiste en traits collectifs, mais aussi en traits individuels ; elle peut concerner des habitudes phonétiques, articulatoires en général, mais aussi le vocabulaire, voire des traits grammaticaux (il suffit de penser à un Argentin parlant castillan). Peut-être les particularités lexicales, à l'intérieur d'un continuum langagier, acquièrent-elles une importance particulière, parce qu'elles peuvent, semble-t-il, troubler la communication plus facilement que d'autres particularités. Ces traits démarcatifs peuvent concerner tous les niveaux du discours. On pourrait parler de *rayons de communication*. Or, les locuteurs sont habitués à distinguer entre les traits démarcatifs de leurs interlocuteurs, parmi les messages émis ; ils ont une *praxis* de distinction et d'interprétation acquise au même moment que le langage et exercée quotidiennement. La *réception* linguistique n'est pas, contrairement à une idée reçue et encore assez vivace, un fait passif et mécaniciste, mais un travail d'interprétation fort complexe des chaînes sonores (pour simplifier beaucoup, on pourrait dire en termes structuralistes que nous évoluons dans le domaine de la *parole* quand nous parlons, mais dans celui de la *langue* quand nous entendons). On pourrait appeler cette activité le *travail de compréhension* (ce terme renvoie au livre de Robert Lafont, *Le travail et la langue*, 1978). Si les chaînes qui nous arrivent sont très diverses de celles que nous avons habitude d'interpréter, il peut y avoir un point à partir duquel le travail de compréhension ne donne que des résultats partiels voire nuls. A ce point, la démarcation devient plus forte que la communication : nous sommes arrivés aux limites de notre potentiel communicatif. Ces limites peuvent être influencées jusqu'à un certain degré : nous pouvons entraîner nos oreilles à capter des sons peu familiers (car même la capacité de discernement acoustique diminue quand les sons que nous captons nous sont étrangers), ou notre intérêt pour le sujet et notre connaissance du contexte sont si grands que nous comprenons des choses qui nous échapperaient si nous étions fatigués ou si elles nous étaient indifférentes ou si tout simplement nous ne voulions pas comprendre.

Il y a des situations historiques qui provoquent un affaiblissement des rayons de communications existants : quand de nouvelles frontières se dressent, quand des moyens de transport disparaissent, etc. Si cette situation perdure pendant suffisamment de temps, les gens perdent l'habitude de cette communication, qui a demandé un effort dès le départ : de nos jours, par exemple, les jeunes Tchèques qui vont à l'école comprennent de moins en moins le slovaque, langue très proche du tchèque qui, uniquement pour des raisons historiques, possède une forme référentielle différente de celle du tchèque. Mais, aujourd'hui le slovaque est peu présent en République tchèque, les jeunes perdent donc ou ne prennent pas l'habitude de l'entendre ou lire et par conséquent celle de le comprendre. Une nouvelle frontière linguistique commence à se dresser qui mènera presque nécessairement à une connaissance moindre de la Slovaquie et sa culture par les voisins. Le problème se pose (encore ?) beaucoup moins pour les jeunes Slovaques parce qu'il y a de nombreuses émissions tchèques à la télévision slovaque (le contraire étant bien plus rare), les oreilles maintiennent par conséquent l'habitude d'interpréter ces sons un peu étranges. Dans l'exemple précité, la *démarcation* l'emporte sur la communication : là où jadis, il y avait intercompréhension assez générale, une frontière linguistique est en train de s'affirmer. Il serait facile d'allonger la liste des exemples et de l'étendre à des situations romanes.

Au contraire, si des frontières, qu'elles soient artificielles ou qu'elles soient naturelles, disparaissent et que des nouvelles relations se créent, de nouvelles habitudes d'interprétation et de compréhension peuvent se développer. C'est le processus qui s'est produit dans les grandes sociétés européennes où des langues étatiques (« nationales ») se sont imposées. Il ne faut pas croire que les différentes variétés de l'allemand ou du français étaient mutuellement accessibles, voilà deux siècles, au commun des mortels, analphabète et (plus ou moins) ad-

scrit à la glèbe. Le cas est bien sûr différent pour les couches plus élevées de ces sociétés (et pourtant : il suffit de penser aux souvenirs de Madame de Sévigné, 1626-1696, pendant son voyage en Provence pour rendre visite à sa fille, mariée à un haut dignitaire provençal ; elle arrivait à peine à communiquer avec la plus haute noblesse provençale), et comme nous dépendons surtout de leurs témoignages pour interpréter le passé, nous nous faisons souvent une impression trop moderne des relations de communication. Il existe un témoignage contraire, par exemple, de quatre soldats allemands, rescapés du naufrage de la grande armée napoléonienne de 1812/13, qui se sont rencontrés au hasard de leur retrait précipité et qui ne pouvaient pas communiquer entre eux, leurs variétés de l'allemand étant trop différentes et eux n'ayant pas l'habitude de les interpréter ; ils eurent finalement recours à un procédé fort compliqué à plusieurs niveaux, où intervenait même le français en tant que *lingua franca* partielle. En réalité, ces quatre jeunes hommes ne pouvaient pas directement communiquer entre eux. De ma propre expérience, je connais des récits sur les grandes difficultés communicatives à l'intérieur des administrations françaises dans des entretiens téléphoniques entre Paris et des villes occitanes, encore dans les années 1970. Ce n'est qu'au cours des deux derniers siècles, avec la scolarité obligatoire et notamment à partir de la pratique (passive) de variétés linguistiques différentes par la généralisation de la radio et les brassages internes des populations que les capacités communicatives des gens ordinaires ont énormément augmenté, en même temps que les variétés employées à l'intérieur d'un Etat se sont considérablement rapprochées (cf. Kremnitz 1999).

Les mouvements de fragmentation d'un espace communicatif ou de fusion de deux espaces peuvent s'expliquer en termes de communication et de démarcation : « normalement », il existe un équilibre entre ces deux tendances. Mais quand l'importance de la démarcation augmente, une fragmentation de l'espace communicatif peut se produire (la masse de la communication diminue et les locuteurs des deux côtés perdent l'habitude de s'interpréter mutuellement jusqu'à ne plus se comprendre ; en même temps, les variétés langagières évoluent de manière différente et s'éloignent peu à peu l'une de l'autre). Si, par contre, l'importance de la communication augmente, on peut assister à la fusion de formes jusque-là divergentes. Ces mouvements ont en même temps des répercussions sur l'identité collective que ressentent et assument les locuteurs. Chaque fragmentation amoindrit en même temps la valeur communicative des deux langues qui naissent : leurs espaces se rétrécissent, les locuteurs trouvent moins d'interlocuteurs possibles ; finalement, on doit s'attendre à ce que la somme de la communication dans les deux nouvelles langues n'atteigne pas celle que la seule langue qui existait au départ a assurée, sauf si pour des raisons idéologiques (« identitaires ») une langue acquiert une attractivité nouvelle. Comme leur espace communicationnel s'est rétréci, les locuteurs doivent avoir recours à d'autres stratégies langagières pour la communication à plus large échelle qui auparavant a pu se faire dans la langue unique. Ceci rend nécessaire l'emploi de langues auxiliaires, *linguae francae*, qui peuvent à la longue se substituer partiellement ou même entièrement aux langues préexistantes. Bien sûr, il s'agit, au moins dans le passé, d'une évolution sur la longue durée (encore faut-il faire des restrictions à cette affirmation car il existe des récits sur des substitutions générales assez rapides dans des sociétés où les langues ne sont guère ou pas idéologisées). L'accélération de toutes les évolutions communicatives, dues aux mass media modernes, laisse cependant prévoir que ces mouvements pourront s'intensifier considérablement dans un avenir proche. Il suffit de penser que les grands mouvements de substitution linguistique en Europe, qui se produisaient depuis des siècles, ont tous accéléré leurs progrès d'une façon spectaculaire depuis environ 1945 (ce qui explique en bonne partie le processus complémentaire de succès des mouvements de défense des langues minoritaires, ou, par euphémisme, « moins utilisées »). Il faut ajouter que les reculs ne se font normalement

pas selon des courbes linéaires mais selon des courbes exponentielles. Il y a une sorte de dialectique entre accentuation de la démarcation et baisse de la pratique communicative –celle-ci peut se réduire à certaines situations-phares. On peut arriver à des cas limites comme aujourd'hui ceux de certaines variétés de l'occitan où la volonté démarcative de certains (prétendus ?) locuteurs est si grande qu'ils considèrent des variétés proches des leurs comme si différentes qu'ils ne veulent pas y voir la même langue. Dans la situation précaire où se trouve l'occitan cela veut dire qu'ils ne trouvent finalement presque pas d'interlocuteur avec lequel ils pourraient communiquer dans « leur » langue. La démarcation peut ainsi arriver à un point maximal : la langue imaginée de cette façon ne possède plus qu'un seul locuteur, sa valeur communicative est tombée à zéro. Il n'est pas nécessaire d'insister sur le fait que de cette façon la langue en question n'a plus de circulation sociale.

3.2. *ABSTAND ET AUSBAU*

Le sociolinguiste allemand Heinz Kloss (1904-1986) proposa, voilà plus d'un demi-siècle (cf. Kloss 1952, 15-24; 1969, 74-77; Haarmann 2004), deux termes pour la discussion (socio-) linguistique qui eurent un succès incontestable. Il distingue entre *Abstandsprachen* et *Ausbausprachen*, langues par distance et langues par élaboration. Par cette différence, il entendait distinguer les langues qui par leur seule forme linguistique se distinguaient suffisamment des langues les plus proches pour être reconnues comme « langues » et pas comme variétés d'une autre langue. La langue *Abstand* par excellence de cette manière est la basque pour laquelle on n'a pas encore trouvé de parenté génétique quelque peu sûre. D'autre part, la distance est suffisamment importante pour distinguer entre les grandes branches de langues, par exemple, à l'intérieur des langues indo-européennes, les différentes familles que constituent les langues romanes, germaniques, slaves, etc. L'intercompréhension entre langues appartenant à des familles différentes est très limitée. Mais à l'intérieur de ces familles, les distances sont souvent peu importantes, surtout si l'on prend en considération des variétés voisines. Ainsi, entre les variétés de l'est de la France et celles du nord de l'Italie, les différences sont presque insensibles, on peut parler d'un *continuum* géographique (cette idée présidait déjà aux conceptions linguistiques de Gaston Paris et Paul Meyer, au ^{xix}^e siècle, qui de cette façon entendaient dépasser la conception de la bipartition linguistique de la France –lancée jadis par quelques savants surtout occitans et repris par les félibres provençaux et qui menaçait de déboucher sur des revendications politiques). Les différences entre les langues constituées ne sont pas nécessairement dues aux différences linguistiques, mais à la volonté de certains locuteurs de se différencier des voisins (cf. paragraphe précédent). Ce n'est que lorsqu'on compare des variétés géographiquement très éloignées, on se rend compte que les différences peuvent être considérables (en plus, il faut se rappeler que les variétés romanes étaient au Moyen Âge bien plus proches les unes des autres qu'elles ne le sont aujourd'hui ; la même remarque s'applique aux langues slaves ; leur éloignement croissant est dû à des facteurs internes de communication, comme à des facteurs externes de politique). Dans ces cas, la création d'une langue référentielle repose sur les différences avec « l'autre » langue, ces différences sont accentuées –par des réalisations phonétiques différentes, par des inventaires phonologiques divergents, par l'introduction de néologismes particuliers, etc.– et si les modèles différents sont acceptés par les sociétés en question, deux langues de référence différentes se dressent sur un espace où jadis il y avait un continuum de communication. La condition essentielle pour la revendication de l'existence d'une langue par élaboration est par conséquent qu'il y ait (eu) des efforts de normativisation et que ces efforts aient (eu) au moins un certain succès dans la ou

les sociétés. De cette façon, par exemple les langues scandinaves –danois, suédois, norvégien (sous deux variétés)– sont des langues par élaboration tellement proches que les locuteurs s'entendent mutuellement assez bien, quand chacun parle la sienne (le cas est différent pour l'islandais) ; mais elles ont développé quatre formes référentielles différentes, surtout pour se démarquer politiquement. Un autre cas, dans lequel un ancien continuum s'est scindé en deux langues référentielles est la séparation du néerlandais et du (bas-) allemand. Depuis quelques décennies, on voit surgir une nouvelle langue germanique qui se situe, dans un certain sens, entre l'allemand et le néerlandais, c'est le luxembourgeois qui pendant des siècles a été considéré et a fonctionné comme variété de l'allemand, mais qui, à la suite de la dernière guerre mondiale, pour des raisons de démarcation, a été élaboré lentement en langue distincte. Aujourd'hui, c'est la langue officielle du Grand Duché de Luxembourg, mais son élaboration est encore modeste ; elle progresse cependant.

Il va de soi que les activités d'élaboration dépendent d'options idéologiques et politiques. Or, celles-ci peuvent changer. Depuis le début du XIX^e siècle, pour ne citer qu'un exemple, il y avait de fortes tendances à former à partir du continuum des variétés slaves du sud, dont certaines avaient des traditions écrites, une seule langue référentielle. Très vite, il fallait sacrifier certaines zones marginales : le bulgare au sud ne participait jamais vraiment à ces efforts, et le slovène au nord reprenait ses traditions indépendantes (après 1944, le bulgare employé dans la nouvelle République yougoslave subit un processus d'élaboration sous la dénomination de « macédonien », alors que le bulgare/macédonien fortement opprimé en Grèce, maintient la dénomination « bulgare » ; le macédonien est aujourd'hui partout –sauf dans un irrédentisme bulgare– considéré comme langue différente du bulgare bien que l'intercompréhension soit toujours très bonne). Mais les variétés parlées en Serbie, Croatie, Monténégro (Crnagora) et Bosnie-Herzégovine (pour employer des termes actuels) subissaient une forte poussée unificatrice, due à des facteurs politiques, et à partir d'une réunion à Vienne (Wien), en 1850, ces tendances se concrétisaient. C'est surtout le travail du grammairien Vuk Stefanović Karadžić (1787-1864) qui concrétisa ces aspirations. Les variétés serbes et croates furent interprétées comme variétés d'une *seule* langue, le serbo-croate, parfois le croato-serbe, selon des règles grammaticales relativement unifiées, mais écrite en Serbie avec des caractères cyrilliques, en Croatie avec des lettres latines. Ce processus de rapprochement fonctionnait, bon gré, mal gré, jusqu'à l'éclatement de l'Etat yougoslave, en 1991. Depuis lors, des tendances opposées ont apparu : l'indépendance croate fut à l'origine d'un travail d'élaboration acharné qui tendait à éliminer le plus possible les traits communs ; une partie considérable du vocabulaire (surtout administratif et politique, mais aussi technique) a été recréée en Croatie, dans le but d'une démarcation du serbe, de telle façon que pour un locuteur ordinaire du serbe, un texte administratif croate pose aujourd'hui des problèmes de compréhension (la même remarque s'applique aux émigrés croates qui reviennent au pays). Ce mouvement d'élaboration de références différentes et –dans le meilleur des cas– mutuellement opaques se poursuit aujourd'hui : une forme référentielle bosniaque est en train de se constituer en se différenciant des traditions serbes, et même le monténégrin tourne le dos à Belgrade et commence à élaborer sa propre forme référentielle. Dans un espace qui avait développé une grande aire de communication se dressent aujourd'hui des langues concurrentes (cf. par exemple Okuka 1998) et une identité complexe, pourtant assez développée, se scinde en quatre identités collectives et politiques. Des remarques comparables, quoique moins dramatiques s'appliquent aux relations entre tchèque et slovaque, initialement variétés très proches l'une de l'autre que les fondateurs de l'Etat tchécoslovaque entendaient rapprocher de manière à créer, au bout d'un certain temps, une seule langue référentielle ; ces deux langues s'orientent depuis 1993 dans des directions opposées. Un cas plus ancien concerne l'espace de l'Empire ottoman, la Turquie actuelle et les pays li-

mitrophes, où s'étendait un continuum turc, avec des traditions écrites parfois multiséculaires. Vers la fin de la première guerre mondiale, le mouvement (impérialiste) pan-touran caressait l'idée d'un grand Etat allant de la Roumélie européenne jusqu'à la frontière mongole et construisant une seule langue référentielle. On sait que ce mouvement n'a pas eu le succès escompté. D'autre part, la répartition actuelle entre les langues turques officielles quelque part, comme le turc, l'azéri, le kazakhe, l'ouzbek, le turkmène, le tatar, le tchéchène, etc. repose en grande partie sur des décisions plus ou moins volontaristes (turques et soviétiques) qui ne tenaient pas nécessairement compte des réalités communicatives de cet espace. Entre temps, des traditions se sont établies, mais ne sont sans doute pas suffisamment anciennes pour ne pas subir des changements en cas de forts bouleversements politiques ou sociaux. Il serait facile de multiplier les exemples, mais ceux que je viens d'esquisser devraient suffire pour éclairer les tendances dont il est question.

De cette manière, les termes d'*Abstand* et d'*Ausbau* désignent les conséquences des tendances communicatives et démarcatives.

3.3. LE PROBLÈME DE LA CONSCIENCE COLLECTIVE

Un terme souvent employé dans ces débats est celui de *conscience* collective (des locuteurs) qui rejoint, en partie, celui de *prestige interne*. Tous les locuteurs se font des idées sur ce qu'ils parlent, et il résulte de ce qui vient d'être dit que ces idéologies jouent un rôle considérable pour le sort des langues en question. Le locuteur monolingue qui ne se trouve pas dans une situation diglossique –où une autre langue se superpose à celle qu'il utilise–, aura normalement un rapport positif à sa langue qui pour lui désigne aussi les limites de ce qui peut être dit (pour s'apercevoir de lacunes linguistiques d'une langue il faut en général être au moins bilingue). Or, sauf dans des isolats rares, une telle situation est exceptionnelle : en général, les locuteurs sont confrontés à la présence de plusieurs langues qui jouissent de prestiges différents. Si les limites de la pratique langagière coïncident plus ou moins avec les limites sociales, les relations entre les deux peuvent, par cette surdétermination, se compliquer. Il va de soi que normalement les langues dominantes sont les langues des dominants (les exceptions apparentes à cette règle reposent habituellement sur une interprétation trop superficielle des données). Les mouvements de renaissance linguistique se produisent en général dans des situations où cette dominance est remise en question, pour des raisons souvent difficiles à cerner. Là où les clivages sociaux (sentis) ne recourent pas les clivages linguistiques la naissance de conflits linguistiques est peu probable. Là où la situation d'oppression semble sans issue, les gens tendent à se conformer au modèle dominant ; seulement à partir du moment où des solutions plus favorables aux opprimés semblent se dessiner à l'horizon, ceux-ci adoptent des attitudes plus revendicatives. Ainsi, les peuples indiens de beaucoup d'Etats d'Amérique du Sud ne semblent pas très convaincus à l'idée de doter leurs langues –le quechua, l'aymara par exemple– de normes référentielles tant que ces langues restent dominées et tant qu'une carrière, aussi modeste soit elle, dans la société rend indispensable la maîtrise du castillan. L'emploi de ces langues équivaut, pour leurs locuteurs, à les confiner à une place sous-privéligiée dans la société ; c'est pourquoi ils considèrent parfois les tentatives de promotion de « leurs » langues comme des activités néo-colonialistes. Par conséquent, ils aspirent à une maîtrise aussi poussée que possible (ce qui n'est pas évident dans des systèmes éducatifs souvent fort déficients) du castillan, et cela malgré leurs problèmes patents de communication. Ce n'est que dans les sociétés, comme en Équateur –où cette mise à l'écart ne fonctionne plus systématiquement et où un éveil politique des peuples autochtones se traduit aussi par une remise en va-

leur de leurs langues— que des travaux de normativisation ont plus de chances d’aboutir socialement.

Il faut insister sur le fait que ce « savoir linguistique » n’est ni clair ni explicite ; Coseriu, s’appuyant sur Leibniz, le caractérisa jadis comme « *claro confuso* » (Coseriu 1958,33), terme que Brigitte Schlieben-Lange reprit et développa plus tard (1975). Cette pénombre dans laquelle évolue la conscience linguistique la rend fortement influençable et, finalement, vulnérable.

Ces relations compliquées peuvent s’illustrer par quelques situations prises dans les langues romanes en Europe. Quand Ascoli proposa de regrouper les trois variétés romanes des Alpes —le rhétique des Grisons, le ladin des Dolomites et le Frioulan— en une seule langue romane, il ne tint pas compte des consciences collectives des trois groupes concernés. Ce qui était (et est) vrai sur le plan de la linguistique interne —il s’agit sans aucun doute des restes d’une Romania alpine fragmentée par les avances des peuples germaniques—, ne rencontra presque pas d’écho dans les consciences des populations concernées, celles-ci n’ayant plus, depuis longtemps de communauté de communication (si jamais elles en ont eue). La définition linguistique et la définition communicative des espaces ne concordait pas. Dans le cas du franco-provençal, l’analyse d’Ascoli différait de celle de la plupart des locuteurs en ceci que là où lui voyait une langue différente du français et de l’occitan, ceux-ci, dans leur grande majorité, ne voyaient qu’une variété du français. L’affaire se complique un peu par le fait que le franco-provençal étend son domaine sur trois États, la France, la Suisse et l’Italie, et que dans les deux derniers la conscience différentielle est un peu plus vive qu’en France. Mais nulle part, sauf dans le Val d’Aoste, des revendications linguistiques n’ont eu quelque succès. En Suisse, les Romands s’orientent depuis longtemps beaucoup plus vers Paris que vers des formes régionales et évacuent de cette façon les traits franco-provençaux de leur français. La force assimilatrice de l’Etat français —et l’élargissement de l’espace communicatif— ont eu raison des différences. De même, on pourrait réduire les différends entre roumain et moldave à une différence de conscience collective, fortement influencée par deux siècles d’histoire qui se sont déroulés très différemment pour les deux groupes (la question se complique encore si l’on tient compte des locuteurs du roumain/moldave en Ukraine), les différences linguistiques internes étant trop insignifiantes pour expliquer les contradictions. Dans d’autres domaines, les mouvements de renaissance linguistique seraient impensables sans une affirmation de la conscience linguistique, au moins dans une partie de la société concernée (naturellement, nous avons là affaire à des avant-gardes, tout le monde ne peut pas concevoir un travail de création de norme référentielle, mais si ce travail n’est pas accepté par de larges couches de la société concernée il reste sans effet et tombe dans l’oubli).

Pour expliquer de telles situations, des notions comme *aliénation* (Robert Lafont, 1967) ou *auto-odi* (la sociolinguistique catalane, cf. Ninyoles 1969) ont été introduites dans le débat (Kremnitz 1987, 1990 ; récemment Max Doppelbauer a pu cerner davantage la préhistoire de ce terme, cf. Doppelbauer 2005). Cette analyse et ce terme qui remontent au philosophe Theodor Lessing ont encore une autre racine, moins discutée, dans le terme marxiste de *fausse conscience*. Cependant, il faut se poser la question de savoir si une conscience peut être *fausse*, si ce terme n’est pas, à son tour, chargé d’idéologie et induisant par l’abandon d’une démarche matérialiste, une incursion dans l’idéologie (cf. les réflexions brèves mais intéressantes de Cichon 2005). Ce qui est sûr c’est que l’idéologie que se forge une communauté sur sa langue joue un rôle considérable pour l’évolution de la communication dans cette langue. Elle dépend de nombreuses variables, ainsi il est difficile de prédire les conséquences sociales des activités que des individus peuvent entreprendre.

3.4. CHANGEMENTS DES RAYONS DE COMMUNICATION. LA PLASTICITÉ DE LA COMMUNICATION ET LE PROBLÈME DES NORMATIVISATIONS MULTIPLES POUR UNE LANGUE

Les langues, en tant que moyens de communication, peuvent changer leur rayon de communication et le font effectivement. Cela se constate à de nombreux niveaux et il n'est pas nécessaire de penser immédiatement à des catastrophes historiques comme les grandes guerres du xx^e siècle qui ont bouleversé les rayons de communication de plus d'une langue. Cependant, c'est à ce niveau que les changements sont les plus clairs à observer : depuis 1945 nous savons que le yiddish ne gagnera pas la compétition contre l'iwrit, d'une part parce que la plupart de ses locuteurs ont été exterminés, d'autre part parce que l'iwrit a pu recueillir une partie du prestige de l'hébreu, tandis que le yiddish souffre de son passé de langue des incultes des *stetl* de l'est. En même temps, le rôle de l'allemand en tant que *lingua franca* dans un grand rayon de l'Europe centrale et de l'est a pratiquement disparu et les quelques progrès qu'il fait depuis 1990 ne suffisent pas à récupérer le terrain perdu ; il a en plus pratiquement disparu de nombreuses régions où il était la première voire la seule langue parlée comme la Prusse Orientale (surtout la partie soviétique, aujourd'hui russe), la Poméranie ou la Silésie. Le polonais a été transplanté à l'ouest, mais le russe qui était au départ le grand vainqueur de ces opérations, a perdu, à son tour, énormément après les évolutions de 1989/90. D'une part, il a perdu du terrain comme langue officielle et parlée, mais il en a surtout perdu comme langue étrangère apprise dans les écoles, et même comme objet de recherche dans les universités. De nombreux instituts universitaires de recherche russes ou slaves ont dû fermer leurs portes, d'autres sont retombés dans l'insignifiance pratique. Il serait facile d'allonger cette liste. Pareillement, les événements de 1989/90 ont promu un nombre considérable de langues jusqu'alors reléguées au second plan : par exemple celles qui sont devenues langues d'Etat comme le slovène ou le letton, etc. D'autres ont pratiquement commencé leur existence, comme les héritiers du serbo-croate.

Il y a d'autres processus, beaucoup plus longs et moins visibles, comme le lent déclin du français à partir (au plus tard) du début du xix^e siècle, comme *lingua franca* et langue dominante en Europe. Le français avait supplanté le castillan dans cette position, plus ou moins à partir de la paix de Westphalie en 1648. Il était devenu la langue des cours européennes, de la noblesse et, en grande partie, de la recherche scientifique où il avait sérieusement mordu sur le latin. Les insuccès politiques et militaires de la France au xviii^e et les progrès scientifiques et philosophiques ailleurs, notamment en Angleterre, ont freiné une expansion ultérieure du français ; la Révolution et ses conséquences, le nationalisme naissant et la relative stagnation de la France ont fait que cette prééminence se réduit de plus en plus. Si les progrès de l'anglais étaient au départ fortement concurrencés par ceux de l'allemand, celui-ci est le grand perdant du xx^e siècle. Mais le mouvement de recul du français continue devant nos yeux, il est vrai à petits pas, pas qui deviennent cependant sensibles quand le regard de l'observateur se limite à quelques décennies.

Ce qui se présente ici à un niveau très élevé, vaut également à des échelles beaucoup plus modestes. Des variétés locales peuvent changer d'orientation : on sait que la limite entre provençal et languedocien s'est déplacé vers l'ouest, à partir du xvii^e siècle, de manière que la limite entre les deux variétés occitanes, autrefois sur le Rhône, s'est déplacée sur le Vidourle, à quarante kilomètres plus à l'ouest, et que la *Renaissance* occitane, dans le département du Gard, est provençale et non languedocienne. L'implantation récente de quelques entreprises de la Suisse alémanique dans les environs de Fribourg a provoqué l'installation de nombreuses familles de langue allemande dans quelques communes de ce canton, créé un petit noyau de langue allemande et restreint le monopole (jusqu'alors indiscuté) du français dans ces com-

munes. La lente pénétration de l'italien référentiel dans les campagnes de la Péninsule a fortement influencé les relations linguistiques, mettant surtout en situation difficile des variétés régionales ou locales qui jusqu'alors maintenaient leur domaine de communication sans grandes pertes et qui, presque d'un coup, perdaient une grande partie de leur valeur communicative. L'exemple le plus criant est l'érosion massive du sarde depuis le fascisme et l'avènement de la République. Il n'est pas du tout sûr que les actuelles mesures en faveur de cette langue suffisent à lui assurer un avenir. Que dire des conséquences linguistiques, souvent indirectes, de l'immigration récente de langues allemande et anglaise à Majorque ?

Il n'est pas possible de dresser ici un bilan des changements des rayons de communication, même pas pour les langues romanes. Ces mouvements sont trop complexes voire contradictoires, souvent perceptibles uniquement pour des observateurs très attentifs (ils échappent souvent même à l'attention des linguistes). Leurs raisons peuvent être multiples : des changements dans des données matérielles de la vie, des changements sociaux ou politiques, mais aussi des réflexes collectifs ou identitaires (bien que ceux-ci n'arrivent que rarement à contrecarrer, à longue échéance, les données matérielles ; exemple l'insuccès linguistique du mouvement d'indépendance irlandaise, dont le succès politique –la République– n'a pas pu empêcher l'érosion du gaélique, parce que, jusqu'à un passé très récent, la nouvelle République ne pouvant pas assurer du travail sur place, la connaissance de l'anglais était donc fonctionnellement nécessaire à tous ceux qui craignaient devoir émigrer à la recherche de leur pain quotidien). Ces mouvements peuvent cependant fortement influencer le prestige des variétés concernées ; ils peuvent être au départ d'un *Ausbau* de variétés (par exemple la création des langues référentielles estoniennes et lettones à partir du XIX^e siècle) comme ils peuvent être responsables de l'abandon de langues qui ont un passé culturel important (comme l'occitan ou le gaélique).

Un élément peut fortement influencer de tels mouvements, à savoir l'existence de formes référentielles divergentes pour une même langue. Nous savons que ce fait se produit assez souvent dans des États qui étaient sous domination coloniale jusqu'à une date récente –ainsi pour de nombreuses grandes langues véhiculaires africaines (par exemple le haoussa), il y a des mises en écriture à partir de systèmes anglais, français, parfois encore portugais ou arabe. De même, en Amérique latine, des langues comme le quechua connaissent plusieurs tentatives de normativisation qui reposent parfois sur des bases linguistiques et idéologiques assez différentes (parfois elles pâtissent d'un simple manque de savoir professionnel). Or, de telles tentatives, surtout si elles obtiennent un certain succès social dans des sociétés qui s'alphabetisent, peuvent contribuer à une fragmentation d'un espace communicationnel jusqu'alors unique, par des mouvements d'*Ausbau* divergents qui peuvent entraîner le sentiment chez les locuteurs que l'« autre » variété n'est pas une variété différente mais une autre langue que ces locuteurs ont l'impression de ne pas maîtriser, voire de ne pas comprendre. Si ces politiques linguistiques introduisent en plus des choix différents pour l'élargissement des lexiques, l'impression d'étrangeté peut rapidement devenir telle que les flux communicationnels se réduisent (ce phénomène peut devenir plus fort encore, si des frontières étatiques étanches entraînent la réduction des flux communicationnels) et que peu à peu nous assistons à la naissance de deux nouvelles langues, dont la valeur communicative se rétrécit évidemment.

De tels faits ne s'observent pas nécessairement dans des situations ou contextes lointains. Elles existent également en Europe, d'une part comme résultats de mouvements anciens, ainsi la rupture entre allemand et néerlandais (qui n'est pas complète, les variétés voisines se comprenant facilement), mais aussi celle entre tchèque et slovaque qui ne s'explique que par le fait que ces deux communautés slaves de l'ouest étaient rattachées administrativement à deux centres différents, à savoir Vienne (Wien) et Budapest. Actuellement, on observe des tendances à créer des formes référentielles divergentes chez un certain nombre de langues minoritaires.

Les discussions en breton sont anciennes et connues : à un certain moment quatre systèmes orthographiques (et en partie linguistiques) étaient en concurrence (cf. Merser 1989) ; il semble que la situation se soit actuellement apaisée. Dans ce domaine, le cas de l'occitan est particulièrement compliqué : à une opposition ancienne, qui repose sur des différences idéologiques et sociolinguistiques, entre système mistralien et « classique » –celui-ci qui a trouvé sa meilleure élaboration dans la grammaire de Lois Alibert dans les années 1930–, se sont surajoutées récemment plusieurs contradictions supplémentaires qui ont abouti à des grammaires référentielles partielles (plus ou moins à l'intérieur du système alibertin). De cette façon, nous avons aujourd'hui une grammaire référentielle générale, celle Alibert, avec quelques apports de ses successeurs, une grammaire référentielle partielle pour le Val d'Aran (en Catalogne) qui s'appuie sur la référence générale mais note plusieurs traits particuliers de l'endroit, une grammaire référentielle partielle pour les vallées occitanes d'Italie qui elle aussi note certaines particularités (qui n'ont pas toutes une valeur phonologique), le système mistralien implanté en Provence et accrédité par plusieurs tentatives latérales, par exemple pour le gascon qui parfois n'ont que peu de consistance linguistique (et peuvent changer selon les évolutions linguistiques de leurs auteurs). Si l'on ajoute que derrière ces différentes propositions se cachent des conceptions parfois très différentes du domaine géographique et social de la/des langue/s, on devine la complexité de la situation. Il est vrai que l'officialisation partielle de l'occitan en France par la création d'un concours d'enseignement de cette langue (CAPES) ne reconnaît que deux des systèmes, l'alibertin et le mistralien, et clarifie ainsi un peu la situation. Mais, d'une part, à l'intérieur du système alibertin, il y a aujourd'hui certaines incohérences qui ne sont pas toutes dues à des analyses approfondies, mais parfois simplement à la mauvaise maîtrise du principe. Ensuite, la création de systèmes différents dans les différents États contribue à un éloignement mutuel des locuteurs –on voit déjà paraître des publications, même à prétention scientifique, où l'Aranais apparaît comme « langue romane » sans mention d'appartenance à l'ensemble occitan– et les différences entre les systèmes insécurisent profondément les locuteurs. Dans une situation où l'occitan, langue fortement en péril, ne bénéficie que d'une présence publique modeste, de telles divergences réduisent sa valeur communicative et peuvent avoir des conséquences néfastes pour l'avenir de son emploi. J'ai discuté cet exemple un peu plus longuement pour montrer les conséquences (possibles) de telles discordes ; il est clair que partout où ce n'est pas un pouvoir politique qui appuie une tentative de normativisation, mais des particuliers, des problèmes de ce type peuvent se poser (cf. pour plus de détails Kremnitz 1974, 2001 ; Sumien, sous presse).

Ce qui précède ne masque pas le fait qu'il s'agit d'un problème réel. De surcroît, une forme référentielle qui s'éloigne beaucoup de la *praxis communicative* des locuteurs pose à son tour de nombreux problèmes. L'irrationalité de l'orthographe de l'anglais en est le plus bel exemple. Et même les quelques simplifications qui se sont imposées depuis les États-Unis ne remédient pas aux problèmes de fond. Ce n'est pas par hasard que les intellectuels de l'Amérique du Sud ont discuté, pendant tout le XIX^e siècle, la question de savoir si la forme référentielle européenne pouvait suffire aux besoins des Américains ou non. On sait que, surtout dans le *Cono Sur*, des réformes (il est vrai assez timides) ont été tentées qui au Chili ont pu se maintenir pendant huit décennies, mais que finalement le souci de l'union référentielle a pris le dessus sur les préoccupations pédagogiques. Ce sont les enfants de l'Amérique qui en payent le prix : pour eux, l'acquisition du système orthographique est nettement plus difficile que pour les Castellans de Castille. Dans ce contexte, le développement de formes référentielles des grandes zones américaines est sans doute, du point de vue de la démocratisation des sociétés, un progrès ; il faudra voir s'il doit se payer un jour par des problèmes d'*Ausbau* divergents (cf. Oesterreicher 2000 ; Torrent-Lenzen 2006).

3.5. CONCEPTIONS FRAGMENTARISANTES VS. CONCEPTIONS SYNTHÉTISANTES

En ce qui concerne l'individuation des langues, on peut observer deux tendances contradictoires ; l'une consiste à reconnaître le moins possible de langues et l'autre tend à accroître leur nombre. Ces tendances peuvent se retrouver aussi bien quand on tente d'identifier les langues par des critères internes ou linguistiques que par des critères externes ou sociolinguistiques. Elles reposent essentiellement sur des positions communicatives différentes, se référant au critère souvent évoqué de l'*intercompréhension*. Je reprends et complète ici quelques remarques faites surtout sous 3.1.

Ce critère veut circonscrire le domaine d'une langue par le fait que des locuteurs s'entendent mutuellement –tant que cela est le cas, on se trouve dans le domaine de la même langue. Or, Carlo Tagliavini a qualifié très élégamment de « piuttosto vago » (1972, 352) ce critère, étant donné qu'il ne peut guère être objectivé. A partir des modèles de la communication, nous savons que nous ne nous comprenons toujours que partiellement, que la compréhension est nécessairement fragmentaire, l'interlocuteur et le locuteur ne possédant pas les mêmes connotations ni les mêmes expériences, ce qui fait que même deux personnes qui parlent la même variété d'une langue peuvent être loin de se comprendre (cf. p. ex. Kremnitz 1997, 23-25). Par ricochet, on peut constater qu'il y a toujours un degré minimal de compréhension, tout au moins dans la communication directe, qui ne doit pas, cela est vrai, passer nécessairement par la communication langagière. D'autre part, la compréhension de variétés qui diffèrent de celle du locuteur est influencée par un grand nombre de variables comme les expériences communicatives de l'individu (si j'ai la pratique d'une autre variété, je l'entends mieux que si je l'entends pour la première fois), sa culture (non uniquement) langagière, mais aussi de faits tout à fait accidentels comme une fatigue momentanée ou une surdité plus ou moins accusée. La variation spatiale peut jouer un grand rôle : on comprend souvent des variétés voisines d'une langue, mais des variétés plus éloignées posent beaucoup plus de problèmes. Si un Bavaïrois et un Wurtembergois, qui sont géographiquement voisins, se comprennent sans problème, c'est bien plus difficile pour un Tyrolien et un Saxon. Cela est valable au-delà des frontières linguistiques reconnues : un Gascon et un Aragonais se comprennent plus facilement qu'un Picard et un Toulousain. Les conditions spécifiques d'émission et de réception peuvent être pour beaucoup dans la compréhension. On peut aller plus loin : la volonté joue un grand rôle ; si je m'efforce de comprendre je comprends mieux, si je ne veux pas comprendre je suis sûr de ne rien comprendre. Cela peut prendre des formes caricaturales : des enfants peuvent fort bien communiquer entre eux, même s'ils parlent deux langues différentes, mais à partir du moment où ils *savent* qu'ils parlent deux langues différentes, ils abandonnent leurs efforts devant les difficultés qui ont tout d'un coup surgi (c'est une observation que j'ai faite personnellement). La multiplication des contacts par la télévision et par le tourisme dans les dernières trente années ont contribué à une meilleure intercompréhension des locuteurs des différents créoles des Antilles ; alors qu'ils déclaraient naguère ne pas se comprendre, malgré la proximité linguistique évidente de leurs variétés, les médias recourent aujourd'hui plutôt au créole pour les communications, surtout entre îles à langue officielle française et anglaise, parce que la compréhension est meilleure. L'intercompréhension n'est pas nécessairement réciproque ; elle peut être, pour des communautés entières, unilatérale. Cela est normalement le cas entre castillan et portugais : un locuteur du portugais comprend nettement mieux ce que dit un locuteur castillanophone (européen) qu'inversement, cela est dû surtout aux accidents phonétiques du portugais. De même, les Bulgares comprennent en général assez bien les Russes, tandis que dans le sens inverse la compréhension est beaucoup plus limitée. Un critère si peu sûr ne peut donc pas jouer un grand rôle pour résoudre le problème.

Si l'on considère que de menues différences entre deux variétés linguistiques justifient leur reconnaissance comme deux langues différentes, on adopte une démarche *fragmentarissante* et de cette façon, le nombre des langues sur cette terre augmente considérablement. Ces langues ont nécessairement moins de locuteurs chacune (que si les domaines respectifs étaient plus grands) et risquent de limiter le rayon de communication des locuteurs. Or, nous savons que les rayons de communication changent, pas uniquement dans nos sociétés modernes où la mobilité a beaucoup augmenté, mais aussi dans des sociétés agro-pastorales où elle peut également être grande, à cause de l'épuisement des sols ou de la rareté des vivres. De cette manière, beaucoup de langues apparentées que l'on retrouve dans les listes de certains pays, en général à économie « archaïque », sont le fruit de migrations anciennes plus ou moins forcées. Une démarche fragmentarissante les classera comme langues différentes et ne tiendra pas compte, de cette façon, de leur proximité interne. Celle-ci pourrait permettre aux locuteurs de s'entendre mutuellement, sans trop d'efforts, et si l'on réfléchit à la création de systèmes graphiques et à des normativisations, ils pourraient, par leur rapprochement, élargir le rayon de communication des locuteurs. Une démarche fragmentarissante par contre insiste sur les différences, les considère comme non réductibles à des formes communes. Sur le plan de la description momentanée, elle peut être plus précise, mais elle néglige les possibilités des dynamiques communicatives. Elle limite, pour ainsi dire, les locuteurs dans la situation communicative du moment, sans ouvrir des chemins vers un élargissement du rayon de communication.

Dans l'élaboration des langues écrites de référence européennes, à partir de la Renaissance, ce sont en général des critères *synthétisants* qui ont joué le rôle le plus important. Les créations de ces langues référentielles partaient toujours du principe d'agrandir le rayon de communication de chaque langue écrite – en bonne partie pour des raisons pratiques et matérielles : les éditeurs avaient tout intérêt à promulguer des langues dans un grand rayon de communication, parce que cela leur permettait d'augmenter leurs chiffres de vente, et de même les souverains voulaient disposer de formes référentielles uniques pour mieux communiquer avec leurs sujets. La poussée normativisatrice à partir de la Renaissance se laisse globalement expliquer par cette coïncidence d'intérêts, légèrement idéologisée par des auteurs littéraires qui luttèrent pour la promotion des langues vivantes face au latin. Et il ne surprend pas de constater que cette poussée est perceptible surtout dans les domaines linguistiques où un pouvoir relativement stable permettait une propagation efficace de ces nouvelles normes : dans les monarchies anglaise, française et castillane, alors que les processus en Italie et dans les pays de langue allemande seront beaucoup plus lents et complexes. Cette tendance devient encore plus forte au moment où, à partir de la Révolution française, les pouvoirs commencent à se démocratiser, car à partir de ce moment-là la communication à l'intérieur de l'ensemble de la société étatique acquiert une nouvelle importance. En même temps, c'est le moment où commence l'oppression systématique des minorités linguistiques. Cette démarche rendait évidemment nécessaire des pressions assimilatrices. Il suffit de penser qu'au moment où Luther traduisait la Bible, il y avait encore cinq traditions écrites (et imprimées) dans l'Empire (c'était déjà une réduction par rapport à la période de l'écriture manuscrite tardive), qui se réduisirent rapidement à deux, pour des raisons religieuses, dont les différences étaient toutefois insignifiantes, sur le plan de la communication. Si les normativisations de la Renaissance n'avaient pas reposé essentiellement sur des critères synthétisants, on peut soupçonner que l'Europe occidentale aurait connu, à partir de ce moment, un nombre nettement plus grand de langues écrites, avec des assises territoriales bien plus petites.

Dans toute tentative d'évaluer le nombre des langues du monde et de les individualiser on peut apercevoir ces deux tendances.

4. CRITÈRES INTERNES / FORMELS / LINGUISTIQUES / « OBJECTIFS »

Quiconque fait appel à ces critères, se réfère à des données de la linguistique formelle ou systématique, soit phonétiques, phonologiques, morphologiques, syntaxiques, à un moindre degré lexicales (surtout dans le domaine de la formation des mots). Il reste, pour ainsi dire, dans le cadre de la linguistique *stricto sensu*. Cette démarche est en général suffisante pour distinguer des variétés linguistiques qu'une forte distance sépare : elle suffit pour faire la différence entre des variétés germaniques et romanes, ou romanes et basques (et encore ; sous le franquisme, il y a eu des tentatives –rares il est vrai– de considérer le basque comme une variété interne basse du castillan). Le problème devient aigu au moment où les différences n'atteignent pas un certain seuil critique, grosso modo à l'intérieur de ce que nous appelons des familles ou branches de langues. Mais c'est principalement entre ces variétés rapprochées que le problème est souvent le plus aigu.

Avant l'avènement de la linguistique historique et comparative, au début du XIX^e siècle, les critères employés pour identifier et séparer les langues étaient peu clairs. Il y avait un peu de sociolinguistique *ante litteram*, mais aussi de considérations linguistiques. Le premier qui se soit limité à des considérations de « linguistique interne » fut, je l'ai mentionné, Ascoli. Mais ces critères furent encore choisis de manière assez impressionniste : c'étaient ceux qui paraissaient les plus importants pour justifier certaines différenciations (comme c'étaient en général des critères saillants, ces choix se montraient souvent bons). Si les critères phonétiques ont en général une grande importance, la morphologie, même la syntaxe et dans une certaine mesure le lexique, peut intervenir. Mais l'ensemble reste impressionniste, et si, dans certaines présentations, on changeait quelques critères employés pour d'autres, on pourrait facilement arriver à des configurations différentes. La question « À partir de quand les différences constatées suffisent-elles pour parler de deux langues différentes ? » n'est presque jamais explicitement posée. Souvent, le lecteur a l'impression que l'existence de telle ou telle *unité préétablie* devait ensuite être justifiée par un choix judicieux de critères qui appuyait la thèse présentée. Ce que je viens de dire ne veut nullement discréditer les auteurs qui s'y sont employés ; ils ont essayé de justifier avec *une* série d'arguments des réalités complexes qui suivaient des lignes de force très variées. Ils ont, à partir des convictions scientifiques de leur temps, tenté de formuler la *quadrature du cercle* et ils n'ont nécessairement eu qu'un succès limité. On peut le voir quand on prend un exemple particulièrement soigné de cette argumentation, à savoir la liste des dix-neuf critères qui séparent l'occitan (le provençal selon Ronjat) des langues qui l'entourent. Cette liste, établie par Jules Ronjat, sera plus tard reprise et légèrement amplifiée par Pierre Bec ; elle est toujours fort commode pour mettre en relief les traits saillants de cette langue ; mais elle est établie *après coup* pour justifier un choix idéologique (que l'auteur de ces lignes partage d'ailleurs). De ces 19 critères, 4 différencient l'occitan du catalan, 7 du castillan, 8 de l'italien, 12 du franco-provençal et 16 du français ; ils sont donc fort utiles pour défendre la thèse de la bi- voire la tripartition linguistique de la Gaule, mais ils ne peuvent pas la *prouver* comme une loi scientifique (il faut dire que des positions différentes, comme celles de Paul Meyer ou Gaston Paris ou aujourd'hui quelques-uns de leurs épigones lointains, se trouvent confronté au même dilemme ; cf. Ronjat I, 6-9 ; Bec ⁶1995, 23-31, ¹1963, 24-33). On peut imaginer que dans tous les cas parallèles la démarche est à peu près la suivante : un espace langagier relativement unitaire est le point de départ, soit qu'il s'agisse d'un espace dont les locuteurs ont toujours été considérés comme appartenant à une même langue –c'est le cas de l'occitan, où la conscience d'une *certaine* unité linguistique remonte au Moyen Âge (et n'a jamais tout à fait disparu du savoir du peuple, comme le prouvent par exemple certaines ré-

ponses à l'enquête de l'abbé Grégoire après 1790)–, soit qu'il s'agisse d'une unité non encore identifiée –ce cas se présente plutôt en dehors de l'Europe. Si l'on veut appuyer une conscience populaire, on fait appel à des données externes, sociolinguistiques que l'on tente ensuite de prouver par des arguments internes, linguistiques. Implicitement, la manière de procéder rend compte de la difficulté du problème et de la nécessité de tenir compte d'arguments d'ordres très divers, mais en posant ainsi le problème, on ne le résout pas encore, puisque les critères sociologiques n'apparaissent que subrepticement.

Revenons à la liste de Ronjat/Bec, qui chez ce dernier fait suite à une explication historique et géographique assez développée : elle contient onze critères phonétiques, cinq morphologiques (auxquels Bec ajoute un sixième), un syntaxique et deux lexicaux. Nous avons par conséquent affaire à une liste particulièrement soignée et équilibrée. Il faut dire que cette précision est nécessaire puisque les distances entre les langues traitées sont petites. Mais trois objections demeurent : d'abord, ces critères sont choisis selon le savoir et l'expérience d'un spécialiste, mais ils reposent sur un choix arbitraire. Une liste différente ne changerait pas profondément le résultat mais pourrait apporter des nuances. Ensuite, les critères s'appliquent aux formes de référence du français, italien, catalan et castillan (pour le franco-provençal cela n'est pas possible, faute d'une forme référentielle de cette langue), mais ils ne tiennent guère compte de la variation spatiale de ces langues qui fait que les variétés voisines de l'occitan et d'une des autres langues sont en général beaucoup plus rapprochées que les normes générales (Bec le montre d'ailleurs dans ses discussions des différents critères). On n'obtient la clarté des différences qu'en ne tenant pas compte des formes intérimaires. Finalement, les observateurs ne déclarent pas à partir de quelle distance on a affaire à deux langues différentes. Les quatre traits distinctifs entre catalan et occitan suffisent-ils, ou faut-il au moins les sept qui séparent l'occitan du castillan, voire les huit qui le séparent de l'italien (et dans ce dernier cas on sait que l'italien standard tire fort vers la *Romania orientale*, alors que les variétés gallo-italiennes accusent nettement moins de traits différentiels par rapport à l'occitan) ? Certains problèmes restent. A la fin de cette discussion je voudrais insister sur le fait que mon intention est nullement de dénigrer le travail de Ronjat et Bec, chercheurs pour lesquels j'ai la plus haute estime et dont l'un est un très bon ami de longue date. J'ai pris cet exemple parce qu'il me semble être un des mieux élaborés dans le cadre des langues romanes et parce que je pense que les résultats sont solides. Mais ils restent dans une certaine mesure vulnérables, parce qu'ils restent essentiellement dans le linguistique, n'utilisant les données sociales que comme point de départ.

Toutes les tentatives d'isoler et de délimiter des langues à l'aide de critères de linguistique interne fonctionnent, bien sûr avec des variantes, selon ce modèle, que ce soit en Europe ou ailleurs. Si elles ne partent pas d'une présupposition, comme les travaux d'Ascoli –qui de cette façon a évité un des éléments non-linguistiques de la démarche– elles ne rencontrent pas l'appui des locuteurs, donc elles ne « fonctionnent » pas socialement. Cette remarque s'applique aux deux langues qu'Ascoli a identifiées avec sa méthode rigoureuse. Quand on tient compte des variétés régionales limitrophes qui réduisent le nombre des traits différentiels, on ne tient pas compte des langues de référence fonctionnant socialement et ayant une importance communicative bien plus importante. Qu'est-ce que je prouve, si je montre que deux parlants voisins appartenant officiellement à des langues différentes sont plus proches que leurs formes référentielles ? D'autre part, le nombre des traits différentiels est lui-même très relatif. Il serait facile de confronter des variétés à l'intérieur d'une langue historique, reconnue comme unité, qui montreraient plus des quatre différences que l'occitan et le catalan accusent, mais la communication à l'intérieur de cette langue ne cesserait de fonctionner, si les conditions externes ne changeaient pas.

4.1. LE PROBLÈME DE « MESURER » DES DISTANCES LANGAGIÈRES — LA DIALECTOMÉTRIE

Il va de soi que je ne suis pas le premier à pointer ces problèmes. De nombreux chercheurs ont tenté d'éviter les écueils cités. Les stratégies employées consistaient, d'abord et en général, à multiplier le nombre des éléments observés et les points d'observation. Quand on ne traite plus 19 phénomènes, mais par exemple quelques centaines, ce qui est aujourd'hui possible, le risque de passer à côté du ou des critères essentiels diminue prodigieusement et le risque de l'artificialité des critères disparaît *pratiquement*. En même temps, ces recherches renouvelées partent en général des données des atlas linguistiques à petite échelle qui tiennent rigoureusement compte de la variation dans l'espace ; aujourd'hui, grâce à l'électronique, il existe des répertoires encore plus importants. Mais cette multiplication des données ne résolvait pas le problème de décision, à savoir à partir de quel niveau seuil on pouvait parler de langues différentes. D'autre part, un problème nouveau surgit tout d'un coup : les séries observées étaient des séries dialectologiques, alors que la question est une question —pour le moins— également sociale. Des données dialectologiques peuvent-elles apporter des réponses à des questions de société ?

La tentative la plus conséquente fut l'élaboration de la dialectométrie, due au linguiste toulousain Jean Séguy (1913-1973), qui, à la fin du dernier volume de son *Atlas linguistique et ethnographique de la Gascogne*, tente de « faire tenir le gascon dans une formule ou un schéma » (Ravier 1976, 391, qui reprend une formule de Séguy lui-même). Il veut de cette façon donner au gascon une condensation mathématique et préciser les distances entre les différentes variétés du gascon à l'aide d'indices numériques. On sait que Séguy ne put pas continuer ses recherches, car il décéda peu de temps après avoir terminé ce dernier volume de l'atlas. La méthode fut reprise, surtout par les soins de Hans Goebl qui pouvait faire appel, dès le départ, à l'aide de l'ordinateur et qui, de cette manière, obtient de masses de données bien plus grandes que son prédécesseur (Goebl 1982, 2001). Cependant, si elles servent admirablement à déterminer les distances linguistiques avec une précision impensable jusqu'alors, elles ne sont pas d'une très grande utilité pratique pour délimiter les limites entre les langues ou les variétés. Elles peuvent suggérer un *potentiel* communicatif, mesurant et indiquant la distance linguistique, mais elles ne peuvent pas traduire la *pratique communicative* actuelle en chiffres —les distances vérifiables sont des distances qui se sont créées dans l'histoire— ni numériser les consciences collectives. Si la dialectométrie a, de cette manière, acquis une grande importance pour les travaux de typologie linguistique, elle ne peut pas résoudre le problème auquel nous sommes confronté ici.

Pourtant, Henri Guiter (1909-1994) a tenté dans plusieurs travaux (surtout Guiter 1973) d'utiliser les données des atlas linguistiques pour délimiter des variétés et langues entre elles. Pour ce faire il a établi des seuils de différences (1973, surtout 79), mais ces seuils étaient, à leur tour, jusqu'à un certain degré, subjectifs. Le résultat le plus intéressant de cette démarche est sans doute un essai tendant à montrer que les frontières politiques n'entraînent pas nécessairement un alignement des frontières dialectales, ici appliqué à la frontière franco-espagnole et l'existence d'une variété roussillonnaise du catalan (Sarda Roqueta 1976). Mais si les résultats sont intéressants, un certain doute reste sur deux points : se pourraient-ils vérifier en les appliquant à d'autres éléments de la langue ? Et à quel point les seuils établis sont-ils vérifiables dans la communication ?

L'accumulation de données de linguistique interne, « objectives », opérationnelle pour distinguer des variétés montrant des différences formelles considérables, ne semble pas suffire, dans le cas de variétés apparentées, à délimiter les langues entre elles. Même la multiplication des données opérée par les recherches récentes n'y arrive pas, puisqu'elle ne peut pas résoudre le problème des seuils à appliquer (qui, à mon avis tout à fait personnel, se révéleraient

très mouvants et donc prêtant de nouveau à débat). En plus, les données de communication ne se laissent formaliser que de manière très approximative ; comme j'ai essayé de le montrer.

5. CRITÈRES EXTERNES / SOCIOLINGUISTIQUES / REPRÉSENTATIONNELS / DÉCLARATIFS / « SUBJECTIFS »

C'est pourquoi des chercheurs, provenant en général des disciplines sociales ou de la sociolinguistique, proposent depuis un certain temps de tenir compte, surtout ou exclusivement, des données sociales, surtout de la conscience des locuteurs en question et, pour reprendre le terme de Benedict Anderson, leurs *imaginings* (Anderson ³1996). La sociolinguistique l'avait précédée en parlant de *représentations* (cf. par exemple Lafont 1980). D'après cette conception « les locuteurs savent ce qu'ils parlent » ; eux ont le pouvoir de nomination. Un tel critère n'est pas concevable sans certaines présuppositions idéologiques, dont le premier est une conception de base démocratique de la/des société/s. Or, on sait que ce critère n'est pour le moins pas applicable à tous les domaines de la recherche scientifique. L'est-il ici ?

Pour réfléchir sur ce problème, il faut se demander : que savent les locuteurs des langues qu'ils emploient ? Les discussions sur ce sujet sont anciennes, ainsi que les différentes terminologies proposées. De toute façon : les pratiques communicatives sont –du moins en partie– fonction des évaluations des langues faites par le sujet. Il règle son comportement langagier entre autres sur les expériences, personnelles ou collectives (transmises). Il a appris que les variétés linguistiques diffèrent entre elles et que ces différences sont (ou peuvent être) des différences de valeur (sociale). On a proposé pour cette jonction de termes différents (et je ne pourrai esquisser ici qu'une des filières), d'abord celui de *conscience linguistique* (par exemple Schlieben-Lange 1971, le terme trouve un ancêtre dans la *conscience populaire occitane* qu'évoque Robert Lafont en 1952 ; en cherchant bien, on trouverait sans doute des usages plus anciens encore de ce concept). Ce terme est évidemment en relation avec la tradition marxiste de la *conscience de classe* proposée par Marx et Engels, et utilisé surtout dès 1923 par Georg Lukács ; c'est de cette tradition marxiste que naît l'idée d'une conscience *fausse* ou *malheureuse*. Glück (1979, 67-80) la critique assez sévèrement. Schlieben-Lange lui préfère plus tard le terme de *savoir linguistique* (1975), insistant sur les éléments de connaissance positive des locuteurs. J'ai proposé jadis de séparer ce qui est savoir objectif (et identifiable) de ce qui ne l'est pas et de distinguer entre *savoir linguistique* et *idéologie linguistique* (Kremnitz ²1994, 55-57), toutefois avec des problèmes de délimitation des deux domaines. La sociolinguistique nord-américaine se contente du terme des *attitudes*, faisant de cette façon l'impasse sur tous les éléments idéologiques. Peter Cichon (1998) a récemment essayé de construire un modèle plus complexe de ces relations.

Ces différents termes ne s'excluent pas nécessairement, car ils accentuent des aspects divers d'un ensemble complexe. Des éléments de savoir, aussi de savoir objectivement falsifiable, d'évaluation et conscience (donc : d'idéologie) y entrent dans une relation à plusieurs entrées. Le locuteur tente toujours de relier les informations qu'il reçoit, les expériences qu'il fait, les évaluations qu'il apprend pour en tirer une interprétation d'ensemble aussi cohérente que possible. Cependant, là où des variétés différentes se rencontrent dans une formation sociale, cette interprétation sera marquée par des contradictions internes, des expériences partielles qui ne rentrent pas dans une seule figure d'explication, voire des évaluations qui se contredisent. Car dans de nombreuses situations de contact, un *savoir* social (et par là une *estimation* sociale) se dresse contre un *autre* ; ainsi, il sera difficile de trancher.

De plus, ce savoir n'est pas clair ; je renvoie au terme de « clair-confus », employé par Co-seriu et Schlieben-Lange, à la suite de Leibniz (cf. 3.3.). Une interprétation freudienne montrerait en outre qu'il ne se situe pas sur le niveau du conscient, mais plutôt sur celui du pré-conscient (pour plonger, au moins sous certains aspects, carrément dans l'inconscient). Etant donné que le sujet ne tient pas à sa disposition tous les éléments de ce savoir, il n'en possède pas l'architecture et il est toujours confronté à des bribes. Bien sûr, la ou les sociétés qui l'entourent lui donnent des éléments de synthèse –mais ces éléments peuvent être contradictoires entre eux ou ils peuvent être en contradiction avec l'expérience vécue du sujet. Celui-ci est alors obligé de procéder à une *réinterprétation* de ce qu'il pensait acquis, et il n'est pas sûr que la nouvelle synthèse atteigne toujours un niveau de cohérence suffisant. En plus, la construction prend peu à peu la forme d'un bâtiment avec beaucoup d'annexes qui ne sont pas nécessairement en relation entre elles. Au bout d'un certain temps et d'une certaine expérience de vie, le sujet a à sa disposition des fragments de savoir qui répondent à des impressions très différentes et tentent de les expliquer. Il peut faire appel à ces éléments pour interpréter des situations différentes, voire contradictoires, selon les stimulations qu'il reçoit. Si des conditions externes, même insignifiantes, changent, d'autres pièces d'interprétation sont mises en avant et de cette manière le sujet tente de répondre à la nouvelle situation. Nous possédons tous comme un dépôt de théories et d'explications pour toutes sortes d'expériences, pas uniquement langagières, et nous utilisons celles qui semblent les plus adaptées au moment. Cette plasticité est une faculté importante que nous, les hommes, avons pour nous accommoder à des changements de notre condition de vie. Il est vrai que ces nouvelles interprétations reposent souvent sur des expériences très limitées, elles ne sont pas nécessairement très sûres.

Cela veut dire que les locuteurs sont facilement influençables ; ils n'ont que peu de critères pour examiner leurs (nouvelles) expériences. Celles-ci peuvent rapidement être remplacées par d'autres qui semblent mieux répondre à la réalité. Et elles peuvent être manipulées par des groupes de pression, qu'il s'agisse de gouvernement ou de groupes plus ou moins privés. Nous savons que les gouvernements ne se privent pas de telles actions, surtout depuis que, à l'époque moderne, les langues sont devenues ouvertement des objets de la politique (cf. Maas 1989). Comme ces interventions rencontrent des individus qui n'ont que des idées assez vagues sur le fonctionnement de la communication –nous venons de le discuter– ceux-ci sont exposés à des arguments dont ils ne peuvent pas nécessairement évaluer la valeur réelle.

Pour notre propos, cela veut dire que les sujets sont relativement influençables, surtout là où leur propre expérience est limitée. Cela veut dire aussi que pour délimiter des variétés comme langues différentes, le critère des représentations seul ne résout pas le problème, le savoir des locuteurs étant souvent fortement influencé par des manipulations. Celles-ci peuvent entraîner des changements de positions parfois spectaculaires –j'ai déjà fait allusion à la situation en ex-Yougoslavie qui offre plusieurs exemples assez éloquentes. Il va de soi que l'opinion de locuteurs est une donnée importante, mais pour résoudre le problème que nous traitons ici, elle ne peut suffire.

Cependant, souvent cet argument est mis en relation avec un autre, à savoir la question de l'existence d'une forme référentielle d'une langue. Si une telle forme existe –c'est la thèse– il s'agit bien d'une langue, dans le cas contraire c'est moins sûr.

5.1. ARTIFICIALITÉ DE TOUTE NORME

Les formes référentielles des langues, telles que nous les connaissons sont relativement récentes. Leur émergence est liée à l'invention et l'utilisation capitaliste de l'imprimerie, liées à leur tour au développement de l'État moderne, à partir de la Renaissance environ (cf. les

deux synthèses récents de van Creveld 1999, et Reinhard ²2000). Les transformations du concept d'État lors des siècles suivants ont contribué aux changements des formes de référentialité langagière. Ce n'est pas par hasard que les langues deviennent ouvertement objets de politique à partir de ce moment ; cette politique ne se contentera plus d'interdire ou de prescrire telle ou telle langue, mais elle intervient dans les contenus internes des formes référentielles. Les conceptions des normes ont encore changé à la suite de l'invention et surtout l'imposition des nouveaux médias, à partir du XIX^e siècle.

Il est vrai que le latin impérial a joui d'une certaine référentialité : les grammaires de Donat et Priscien en portent témoignage. Mais elles ne régissent que l'écriture manuelle. Toutefois, leur succès était tel que tout le moyen âge durant les termes de *lingua latina* et *grammatica* fonctionnaient comme synonymes et opposaient le latin, langue à règles, aux idiomes parlés qui le plus souvent ne recevaient pas la dénomination de « langue ». La diglossie médiévale est sérieusement battue en brèche à partir du moment où certaines sociétés commencent à élaborer des formes référentielles pour les langues parlées qui doivent, de cette façon, remplacer le latin dans l'emploi officiel et surtout écrit. Pour les langues qui sont concernées en premier lieu, le castillan, le français, l'anglais –l'allemand est dans une situation particulière à cause de l'absence d'un pouvoir central suffisamment fort– ce sont des processus pluriséculaires. Dans chaque cas, le latin revient par la porte de service, car les nouvelles langues de pouvoir manquent de terminologie justement dans les domaines où elles devraient remplacer le latin. Tout de même, les nouvelles normativisations qui commencent à s'élaborer se configurent selon des critères politiques et sociaux : ce sont les centres qui imposent, en grande partie, leurs variétés, les langues dominantes ressembleront beaucoup aux langues des dominants. Les nouvelles formes référentielles qui voient le jour ne doivent pas leurs formes, *en premier lieu*, à des critères linguistiques, mais à des critères politiques ; elles sont, dans une large mesure, arbitraires du point de vue de la langue. Nebrija, le premier grammairien du castillan, l'avait bien vu quand il considérait la grammaire de cette langue comme *arte*. Il est vrai que la longue durée du processus, surtout pour les premières langues concernées, et les changements sociaux qui entraînent des réajustements des formes référentielles, cachent le volontarisme de l'opération. Les corpus de textes devenant de plus en plus importants mettront bientôt l'accent non pas sur le caractère arbitraire des opérations de normativisation, mais sur le poids des traditions qui voilent le regard des usagers. « La force normative du factuel » pèsera lourdement sur toute tentative de réforme de la forme référentielle, et cette attitude conservatrice s'accroît encore dans les sociétés complètement alphabétisées (ou peu s'en faut), où la connaissance du code, surtout orthographique, est un capital symbolique que les usagers n'entendent pas exposer à une dévalorisation par une réforme orthographique, serait-elle de loin supérieure à l'ancienne norme. Cela explique (en partie) que les dernières grandes réformes linguistiques et orthographiques qui ont réussi sont celles de l'Union Soviétique et de la Turquie après la Première Guerre Mondiale, deux sociétés peu alphabétisées à l'époque, tandis que la « réformette » orthographico-grammaticale proposée en France par le gouvernement Rocard (en 1991) a rondement échoué et celle, également limitée, de l'orthographe allemande à la fin du XX^e siècle crée toujours des débats.

Le caractère volontariste des formes référentielles des langues devient plus évident encore quand on prend en considération les formes référentielles de certaines langues qui ont vu le jour au XIX^e ou au début du XX^e siècle : souvent, ces normes ont été élaborées et proposées aux sociétés respectives par des individus ou des petits groupes de savants (le plus souvent autodidactes) qui ont pris des décisions de politique linguistique et qui ont eu la chance de rencontrer un écho populaire qui a adopté ces propositions ; cela bien plus pour des raisons idéologiques ou politiques que pour des raisons linguistiques (il est vrai qu'il y a eu des « ratés »,

mais encore une fois, pas pour des raisons linguistiques mais pour des raisons politiques). On y voit souvent un très haut degré de volontarisme qui affecte même l'individuation des langues : la décision de créer deux normes différentes pour le tchèque et le slovaque est surtout politique et historique, les deux groupes vivant au XIX^e siècle sous domination différente (Autriche et Hongrie), les situations de plusieurs autres langues slaves sont comparables. Dans chaque cas, une vision fragmentarisante a pris le dessus ; elle renforce la valeur démarcative au détriment de la valeur communicative des langues ainsi « créées », tout en garantissant une cohésion interne plus grande. Cela devient particulièrement évident dans les cas où la direction d'une normativisation a changé de direction en cours de route.

Il en résulte que les formes référentielles des langues sont beaucoup plus fonction de décisions politiques et idéologiques que de raisons linguistiques. Leur succès ou leur échec ne se doit pas en premier lieu à leur qualité linguistique mais à l'appui de la société, des pouvoirs en question. Cela fait que ces formes référentielles peuvent être remaniées, si les bases idéologiques de la société en question changent ; ces remaniements sont plus faciles dans des sociétés peu alphabétisées. Et toutes ces opérations obéissent à des présuppositions idéologiques qui souvent ne sont pas élevées au niveau de la conscience collective. Il est surprenant –et décevant– que les sciences du langage n'aient guère dirigé leur intérêt dans cette direction,⁶ restant le plus souvent dans des considérations de linguistique interne. De cette façon, les sciences du langage se privent, à mon avis, d'une route prometteuse pour résoudre un certain nombre de problèmes de fond et d'approcher le vrai centre de la discipline que devrait être la communication. Il en résulte aussi que les représentations des locuteurs sont des bases bien incertaines pour décider de la délimitation des limites linguistiques, d'autant plus qu'elles sont hautement exposées à des courants d'opinion voire des modes qui ne reposent guère sur des bases sûres. Elles sont très manipulables.

5.2. POLYCENTRISME LINGUISTIQUE

C'est surtout le transfert de plusieurs langues européennes sur d'autres continents, à la suite du colonialisme européen qui a suivi les « découvertes », et le maintien de ces langues comme langues officielles ou du moins co-officielles qui a obligé la linguistique à revoir certains de ces concepts. L'anglais est resté la langue officielle des États-Unis après leur indépendance, mais dans la mesure où le nouvel État s'est aussi idéologiquement libéré de l'ancienne métropole, les choix référentiels s'en sont de plus en plus séparés. De même le castillan est resté la langue officielle des nouveaux États en Amérique latine après la séparation de l'Espagne. Le cas le plus intéressant est celui du Brésil, car l'ex-colonie a, dès le départ, plus de poids que le Portugal. En plus, une politique culturelle très restrictive du Portugal a contribué à une présence extraordinairement faible de locuteurs maîtrisant la forme référentielle de la langue, de manière que dès les premières décennies de l'indépendance des formes référentielles brésiliennes voient le jour qui se condensent assez rapidement en une norme qui concurrence celle de l'Europe. Les variétés du français n'emprunteront ce chemin de l'indépendance que tardivement ; cela est dû au fait qu'au Canada le français glisse d'une position dominante dans une situation de diglossie qui ne sera remise en question que vers le milieu du XX^e siècle, et que la politique coloniale de la France tente de faire de la langue le lien qui unit les groupes collaborationnistes à l'(ancienne) métropole.

6. Je dois les réflexions qui précèdent en grande partie à une remarque de Klaus Bochmann, il y a quelques années ; je l'en remercie. Les développements ultérieurs et les erreurs possibles restent cependant les miens.

Le fait qu'une langue puisse disposer de plus d'un centre référentiel a frappé l'attention des chercheurs assez tardivement bien que le phénomène existe (et soit observable) depuis le début de la création de formes référentielles modernes. Comme ce processus signifiait au départ et pour longtemps la réduction du nombre des normes concurrentes (particulièrement bien observable en domaines allemand et italien à partir du moment de l'invention de Gutenberg), la coexistence de formes référentielles était une donnée négative sans intérêt. Finalement, c'est Heinz Kloss qui semble être le premier à employer le terme de *langues pluricentriques* (« plurizentrische Hochsprachen », Kloss ²1978, 66-67) pour désigner ce phénomène. Le terme sera repris par Michael Clyne dans son ouvrage *Pluricentric Languages* (1992). Depuis ce moment, il est entré dans le langage courant des sociolinguistes. Il tente de saisir la situation où deux ou plusieurs groupes partagent la même langue, mais avec des variantes plus ou moins sensibles dans les formes référentielles respectives. Si au départ les langues concernées étaient surtout les (anciennes) langues coloniales, le terme s'est peu à peu élargi à d'autres cas, où une même langue est langue officielle dans plusieurs États. C'est notamment le cas de l'allemand, officiel en Allemagne, en Autriche, au Liechtenstein et co-officiel en Suisse, au Luxembourg et dans plusieurs autres États, mais aussi du néerlandais, officiel aux Pays-Bas et en Belgique. Souvent, de telles langues existent encore dans d'autres États, mais avec des statuts différents, ce qui risque de compliquer les descriptions.

S'il y a des communautés linguistiques, comme l'anglaise, qui semblent accepter la coexistence de formes référentielles différentes sans trop de problèmes, d'autres acceptent cette variation beaucoup plus difficilement, comme la française. Ce n'est que récemment que les institutions françaises ont commencé à accepter l'existence d'une norme canadienne ; pour la conscience collective il semble qu'il y a encore beaucoup de chemin à parcourir. Le cas du castillan semble se situer entre les deux extrêmes, mais là aussi l'observateur attentif rencontre des réticences nettes. Depuis, le terme a été appliqué également à des langues dominées comme le corse (p. ex. Thiers 1993, le premier à le faire à été Jean-Baptiste Marcellesi quelques années auparavant) ou l'occitan (Sumien, sous presse). Là, le cas est quelque peu différent, car ces langues n'ont pas réussi à se doter d'une forme référentielle qui fasse l'unanimité à un moment où cela semblait encore possible. Les propositions de solutions pluricentriques tentent de tenir compte du manque de consensus référentiel à l'intérieur d'une communauté qui se considère comme une unité de communication, mais où les différents sous-groupes n'acceptent pas de sacrifier sur l'autel de la communication élargie leurs particularités langagières senties comme importantes pour la cohésion interne du sous-groupe. Il va de soi que dans le cas de langues menacées une telle stratégie n'est pas sans risques. Mais elle est de toute façon préférable, du point de vue des chances d'emploi et de survie, à des conceptions identitaires plus rigoureuses encore qui tentent de trancher tout lien de communication entre les variétés comme le font quelques représentants de la renaissance occitane (qui vont jusqu'à refuser la dénomination de la langue) ou comme le fait une frange de valencianistes face au catalan.

La stratégie employée consiste à accepter les différences entre variétés à un niveau pratique pour maintenir une unité de principe. Elle répond de cette façon à un certain pragmatisme linguistique qui accepte cet axiome de la communication à savoir que les formes de la communication (langagière) sont en changement constant et que ces changements s'intensifient entre les sous-groupes d'un groupe de locuteurs qui sont peu en relation entre eux. C'est donc une stratégie de politique linguistique qui sacrifie une unité étroite n'existant (plus) pour maintenir des liens plus lâches qui facilitent l'accès aux différentes variétés. De cette façon, elle tente de maintenir la communication entre les différentes branches d'une langue. Et il semble que des succès communicatifs apparaissent. Elle montre une fois de plus à quel point les politiques linguistiques sont volontaristes et peuvent influencer les représentations.

6. VERS UNE SOLUTION SCIENTIFIQUE ?

Il ne semble pas, après ce que nous venons de voir, que les sciences du langage puissent répondre à la vieille question : qu'est-ce qu'une langue ? La linguistique interne pourrait donner une réponse cohérente si elle pouvait trouver un accord sur le degré de différence nécessaire pour que deux variétés parentes soient considérées comme deux langues différentes. Mais même si elle trouvait une réponse satisfaisante au problème, ces définitions ne fonctionneraient pas nécessairement dans les sociétés concernées pour les raisons que nous avons évoquées. La sociolinguistique pourrait chercher du côté des représentations, mais les résultats ainsi obtenus seraient fort précaires, facilement influençables voire manipulables par des groupes d'intérêts et de pression, tout d'abord les États existants. Nous en avons vu des extrêmes dans les toutes dernières décennies.

Le problème est qu'à côté d'éléments objectifs (ou objectivables) entrent en compte des traits subjectifs pour arriver à des définitions. Des deux côtés, on évolue sur du sable mouvant : les humains peuvent changer leur pratique langagière (et ils le font effectivement assez souvent) et ils peuvent changer leurs convictions sur des pratiques langagières. En cas de discordance les groupes opposés passent trop facilement d'un niveau à l'autre, s'ils peuvent espérer en tirer un gain argumentatif. Cette ambivalence se trouvait toujours dans le langage, structure et comportement à la fois. Mais à partir du moment où les *praxis* langagières ont reçu une forte poussée d'investissement idéologique, donc avec l'avènement du nationalisme moderne qui s'est généralisé en Europe à la suite de la Révolution française, cette tension s'est intensifiée. En plus, les deux perspectives peuvent se « prouver » jusqu'à un certain degré : je comprends (plus facilement) ce que je veux comprendre et je me refuse à ce que je déclare ne pas comprendre. Le comble de l'absurde est atteint quand un Catalan(iste) et un Valencian(iste) se parlent en anglais, parce que ce dernier déclare ne pas comprendre ce que dit le premier et que celui-ci ne veut pas parler « valencien ». ⁷ Mais, si les deux sociétés persistent dans leurs attitudes, les locuteurs, d'ici un certain temps, ne se comprendront effectivement plus, parce que le manque de communication commune fera d'une part que les deux variétés évoluent dans des sens différents et d'autre part les locuteurs n'auront plus l'habitude de s'entendre. Le grand gagnant sera le castillan plus payant sur le plan de la communication.

Il semble que la conception de la linguistique, depuis le début du XIX^e siècle, de prendre comme objet de ses recherches *les langues* et non pas *la communication* a fortement contribué à cette aporie. « Les langues » ont été réifiées, et l'on est passé à côté des dialectiques de la communication. Aujourd'hui, le mal est fait. Il ne semble pas possible de faire disparaître, du jour au lendemain, les investissements idéologiques que cette opération a entraînés. Certes, les tendances vers la globalisation ont de fortes répercussions sur nos conceptions langagières, mais on ne peut pas revenir à l'innocence d'antan. Les langues restent un lieu de forte charge idéologique, ce ne sont que les idéologies qui peuvent changer (parfois presque du jour au lendemain, comme nous l'avons vu).

Cela veut dire, en même temps, que ni la linguistique formelle ni la sociolinguistique/sociologie du langage ne pourront offrir une solution au problème, (au moins) deux conceptions de *langue* étant socialement en jeu. Les acteurs sociaux les utilisent, pas toujours de bonne foi, et il ne sera pas possible de les faire disparaître de si tôt. La seule sous-discipline des sciences du langage qui tente de tenir compte de la complexité des données et des contradictions est la *politique linguistique*, en France parfois appelée *glottopolitique* ; en Allemagne, on fait parfois

7. Cela est effectivement arrivé.

la différence entre *Sprachpolitik* et *Sprachenpolitik* (politique linguistique et politique des langues). Cette discipline tente de voir sous quelles conditions et dans quelles formes naissent les idéologies langagières et prennent leur forme (cf. Bochmann, 1987 ; Bochmann, dir. 1993 ; Maas 1989 ; en France les travaux autour de Robert Lafont et Jean-Baptiste Marcellesi). Elle pourrait à la longue proposer des critères plus fiables. Mais en même temps, elle sera sujette aux décisions des locuteurs qui ne sont pas toujours les meilleures, ni pour eux ni pour les langues qu'ils parlent ou qu'ils abandonnent. Les conceptions fragmentarisantes peuvent, sous certaines conditions, diminuer la valeur communicative des langues/varieties concernées, les conceptions synthétisantes, si elles ne rencontrent pas d'écho populaire, restent inopérantes. Les politiques peuvent tenter de créer un consensus qui entraîne un usage plus intense d'une langue/varieté, mais si un tel consensus ne peut pas être atteint, pour des raisons souvent incalculables, les locuteurs chercheront d'autres solutions. Mais, l'exemple de la Moldavie le prouve, ils peuvent être fort hésitants et se contredire dans un laps de temps assez bref. Toutefois, il pourrait être prometteur d'intensifier les réflexions autour des conceptions polycentriques : celles-ci pourraient contribuer à dédramatiser les débats et tenter de maintenir le potentiel communicatif des langues ainsi définies sans opprimer des sous-groupes qui cherchent à se démarquer. On pourrait tenter la chance.

Si au début de mes réflexions j'avais un espoir –certes mince– de trouver l'œuf de Colomb, j'ai dû rapidement déchanter. A présent, et pour un avenir prévisible, les sciences du langage ne pourront pas définir l'objet « langue » de manière satisfaisante, et il ne semble pas que cette situation change de si tôt. La définition de langue dans la société étant un élément idéologique, la linguistique a beau dire que les locuteurs se trompent –ceux-ci passeront outre, même à leurs dépens.

7. BIBLIOGRAPHIE UTILISÉE

- ALIBERT, Loïs (²1976) [1935] : *Gramatica occitana, segon los parlars lengadocians*. Montpelhièr : Centre d'Estudis Occitans [réimpressions ultérieures].
- ALIGHIERI, Dante (1995) [ca. 1304] : *De vulgari eloquentia*. Edició bilingüe. Traducció de Pep Gómez Pallarès i introducció de Mirko Tavoni. Vic : Eumo.
- AMMON, Ulrich / DITTMAR, Norbert / MATTHEIER, Klaus J. / TRUDGILL, Peter (ed.) (²2004-2006) : *Sociolinguistics/Soziolinguistik. An International Handbook of the Science of Language and Society. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*. Berlin ; New York : de Gruyter. 3 vol. [HSK ²3].
- ANDERSON, Benedict (³1996) [1983] : *Imagined Communities*. London : Verso.
- ASCOLI, Graziadio Isaia (1873) : « Saggi ladini ». *Archivio glottologico italiano*, I, 1-556.
- (1878) : « Schizzi franco-provenzali ». *Archivio glottologico italiano*, III, 71-120.
- BAL, Willy (1966) : *Introduction aux études de linguistique romane, avec considération spéciale de la linguistique française*. Paris : Didier.
- BARTOLI, Matteo (1906) : *Das Dalmatische. Altromanische Sprachreste von Veglia bis Ragusa und ihre Stellung in der apennino-balkanischen Romania*. Wien : Hölder. 2 vol.
- BEC, Pierre (⁶1995) [1963] : *La langue occitane*. Paris : PUF (Que sais-je ?).
- (1970-71) : *Manuel pratique de philologie romane*. Paris : Picard. 2 vol.
- BEIN, Roberto (2001) : « Die Wechselwirkung Prestige/Gebrauchswert des Französischen : früher erste, heute dritte Fremdsprache in Argentinien ». BORN, Joachim (éd.) : *Mehrspra-*

- chigkeit in der Romania. Französisch im Kontakt und in Konkurrenz zu anderen Sprachen.* Wien : Ed. Praesens, 82-90.
- BERRUTO, Gaetano (2004) : « Sprachvarietät–Sprache (Gesamtsprache, historische Sprache », *HSK* ²³. Vol. I, 188-195.
- BOCHMANN, Klaus (1987) : « Sprachpolitische Forschung : Theoretische Prämissen, Gegenstände, Methoden ». *Linguistische Arbeitsberichte* (Leipzig), no. 62, 2-14.
- BOCHMANN, Klaus (dir.) (1993) : *Sprachpolitik in der Romania. Zur Geschichte des sprachpolitischen Denkens und Handelns von der Französischen Revolution bis zur Gegenwart.* Berlin ; New York : de Gruyter.
- BOURCIEZ, Édouard (⁵1967) [1910] : *Éléments de linguistique romane.* Paris : Klincksieck.
- CERQUIGLINI, Bernard (1999) : *Les langues de la France. Rapport aux Ministres de l'Education Nationale et de la Culture.* <http://www.culture.fr/culture/dglff/>.
- CICHON, Peter (1998) : *Sprachbewusstsein und Sprachhandeln. Romands im Umgang mit Deutschschweizern.* Wien : Braumüller.
- (2005) : « 'Falsches' und 'fehlendes' Sprachbewusstsein – brauchbare oder irreführende Erklärungsbegriffe ? ». *Quo vadis, Romania ?* (Wien), no. 25, 13-19.
- CLYNE, Michael (ed.) (1992) : *Pluricentric Languages. Differing Norms in Different Nations.* Berlin ; New York : Mouton de Gruyter.
- COSERIU, Eugenio (1958). *Sincronía, diacronía e historia. El problema del cambio lingüístico.* Montevideo : Universidad de la República.
- DAHMEN, Wolfgang / KRAMER, Johannes (2001) : « Sprachgrenzen ». HOLTUS, Günter / METZELTIN, Michael / SCHMITT, Christian (Hg.) : *Lexikon der romanistischen Linguistik (LRL).* Tübingen : Niemeyer. Vol. I, 2, 501-512.
- DIEZ, Friedrich (1836-43) : *Grammatik der romanischen Sprachen.* Bonn : Weber. 3 vol. [rééditions et traductions].
- DOPPELBAUER, Max, 2005. « 'Selbsthass – Auto-odi'. Zur (Vor-) Geschichte eines Begriffs in der Soziolinguistik ». CICHON, Peter / CZERNILOFSKY, Barbara / TANZMEISTER, Robertt / HÖNIGSPERGER, Astrid (éd.) : *Entgrenzungen. Für eine Soziologie der Kommunikation.* Festschrift für Georg Kremnitz zum 60. Geburtstag. Wien : Ed. Praesens, 68-76.
- GIORDAN, Henri (1982) : *Démocratie culturelle et droit à la différence.* Rapport au ministre de la culture. Paris : La Documentation française.
- (dir.) (1984) : *Par les langues de France.* Paris : Centre Georges Pompidou.
- GLÜCK, Helmut (1979) : *Die preußisch-polnische Sprachenpolitik.* Hamburg : Buske.
- GOEBL, Hans (1982) : *Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie.* Wien : Österreichische Akademie der Wissenschaften. 3 vol.
- (2001) : « Dialektometrie ». *LRL*, vol. I, 2, 856-874.
- GORDON, Raymond G. jr. (ed.) (2005) : *Ethnologue : Languages of the World.* Dallas : SIL International ; également : <http://www.ethnologue.com>.
- GUITER, Henri (1973) : « Atlas et frontières linguistiques ». STRAKA, Georges / GARDETTE, Pierre (éd.) : *Les dialectes romans de France à la lumière des atlas régionaux.* Colloque de Strasbourg, 1971. Paris : Ed. du CNRS, 61-109.
- HAARMANN, Harald (2004) : « Abstandsprache – Ausbausprache ». *HSK* ²³. Vol. I, 238-250.
- HEITMANN, Klaus (1989) : « Rumänisch : Moldauisch ». *LRL*. Vol. III, 508-521.
- HOLTUS, Günter / METZELTIN, Michael / SCHMITT, Christian (ed.) (1988-2005) : *Lexikon der romanistischen Linguistik.* Tübingen : Niemeyer. 8 volumes en 12 unités [=LRL].
- JANICH, Nina / GREULE, Albrecht (éd.) (2002) : *Sprachkulturen in Europa. Ein internationales Handbuch.* Tübingen : Narr.

- KLOSS, Heinz (²1978) : *Die Entwicklung neuer germanischer Kultursprachen seit 1800*. Düsseldorf : Schwann [1952, München : Pohl].
- (1969) : *Grundfragen der Ethnopolitik im 20. Jahrhundert. Die Sprachgemeinschaften zwischen Recht und Gewalt*. Wien ; Stuttgart ; Bad ; Godesberg : Braumüller Wissenschaftliches Archiv.
- KREMnitz, Georg (1974) : *Versuche zur Kodifizierung des Okzitanischen seit dem 19. Jahrhundert und ihre Annahme durch die Sprecher*. Tübingen : Narr (TBL).
- (1987). « Auto-odi (Selbsthass). Zur Aufnahme und Verwendung des Begriffes in der Sprachwissenschaft ». *Semiotische Berichte* (Wien), XI, 419-437 [= *Doxa* (Budapest), XIII, 159-177].
 - (1990) : « Sur l’auto-odi (Selbsthass) ». *Per Robert Lafont*. Estudis oferts a Robert Lafont per sos collègas e amics. Montpelhièr : Universitat Paul-Valéry ; Nîmes : Vila de Nîmes, 197-207.
 - (²1994) [1990] : *Gesellschaftliche Mehrsprachigkeit. Institutionelle, gesellschaftliche und individuelle Aspekte. Ein einführender Überblick*. Wien : Braumüller [trad. catalana (1993) : *Multilingüisme social*. Barcelona : Edicions 62].
 - (1997) : *Die Durchsetzung der Nationalsprachen in Europa*. Münster ; New York, etc. : Waxmann.
 - (1999) : « L’évolution du ‘français régional’. Quelques observations ». *Mélanges de linguistique, sémiotique et narratologie dédiés à la mémoire de Krassimir Mantchev à l’occasion de son 60^e anniversaire*. Sofia : Colibri, 139-148.
 - (2001) : « Le travail normatif en occitan ». BOYER, Henri / GARDY, Philippe (coord.) : *Dix siècles d’usages et d’images de l’occitan*. Paris : L’Harmattan, 21-42.
 - (2003) : « Le concept de ‘conflit linguistique’ aujourd’hui. Essai d’une mise à jour. Avec une annexe : Quelques remarques sur le terme de ‘valeur communicative’ des langues ». *Lengas*, no. 54, 7-22.
 - (sous presse) : « Normes, variation linguistique, frontières et communication » [paraît dans les actes du Colloque TIDE à Bordeaux], 10 p.
- LAFONT, Robert (1952) : « Remarques sur les conditions et les méthodes d’une étude rationnelle du comportement linguistique des Occitans ». *Annales de l’Institut d’Etudes Occitanes*, no. 11, 41-45.
- (1967) : « Sur l’aliénation occitane ». *Le Fédéraliste* (Lyon/Turin), IX, 107-138.
 - (1978) : *Le travail et la langue*. Paris : Flammarion.
 - (1980) : « La spectacularisation de l’occitanophonie dans l’enquête sociolinguistique : la fonction du ‘retour’ ». *Lengas*, no. 7, 71-77.
- LAUSBERG, Heinrich (1956-62) : *Romanische Sprachwissenschaft*. Berlin : de Gruyter. 3 tomes en 4 volumes (le dernier tome n’a jamais paru).
- LUKÁCS, Georg (1923) : *Geschichte und Klassenbewusstsein. Studien über marxistische Dialektik*. Berlin : Malik.
- MAAS, Utz (1989) : *Sprachpolitik und politische Sprachwissenschaft*. Frankfurt am Main : Suhrkamp.
- MERLAN, Aurelia (2004) : « Dynamik des Sprachkontakts im Nordosten Portugals ». *Beiträge zur Fremdsprachenvermittlung* (Konstanz), no. 42, 3-23.
- MERSER, A[ndreo] ar (1989) : *Les orthographes du breton*. Brest : Brud Nevez.
- MEYER-LÜBKE, Wilhelm (³1920) [1901] : *Einführung in das Studium der romanischen Sprachwissenschaft*. Heidelberg : Winter.
- (1925) : *Das Katalanische. Seine Stellung zum Spanischen und Provenzalischen sprachwissenschaftlich und historisch dargestellt*. Heidelberg : Winter.
- micRomania*, depuis environ 1990. Charleroi.

- NINYOLES, Rafael Lluís (1969) : *Conflicte lingüístic valencià*. Barcelona : Edicions 62 (rééditions).
- OESTERREICHER, Wulf (2000) : « Plurizentrische Sprachkultur – der Varietätenraum des Spanischen ». *Romanistisches Jahrbuch*, LI, 287-318.
- OKUKA, Miloš (1998) : *Eine Sprache – viele Erben. Sprachpolitik als Nationalisierungsinstrument in Ex-Jugoslawien*. Klagenfurt ; Celovec : Wieser.
- RADDATZ, Hans-Ingo / TORRENT-LENZEN, Aina (éd.) (2006) : *Iberia polyglotta. Zeitgenössische Gedichte und Kurzprosa in den Sprachen der Iberischen Halbinsel. Mit deutscher Übersetzung*. Titz : Axel Lenzen Verlag.
- RAVIER, Xavier (1973) : « L'incidence maximale du fait dialectal ». STRAKA, Georges / GARDETTE, Pierre (éd.) : *Les dialectes romans de France. Actes du colloque national du CNRS sur les dialectes romans à la lumière des atlas linguistiques régionaux*. Colloque de Strasbourg, 1971. Paris : Ed. du CNRS, 43-59.
- (1976) : « Jean Séguéy et la traversée du langage gascon. Réflexions sur une topogenèse géolinguistique ». *Revue de Linguistique Romane*, XL, 389-402.
- REINHARD, Wolfgang (²2000) [1999] : *Geschichte der Staatsgewalt. Eine vergleichende Verfassungsgeschichte Europas von den Anfängen bis zur Gegenwart*. München : Beck.
- RENZI, Lorenzo (1976) : *Introduzione alla filologia romanza*. Bologna : Il Mulino.
- RONJAT, Jules (1930-1941) : *Grammaire istorique [sic] des parlers provençaux modernes*. Montpellier : Société des Langues Romanes. 4 vol.
- SARDA ROQUETA, Anna (1977) : « Le mythe du dialecte roussillonnais ». *Revue de Linguistique Romane*, XLI, 154-161.
- SCHLIEBEN-LANGE, Brigitte (1971) : « La conscience linguistique des Occitans ». *Revue de Linguistique Romane*, XXXV, 298-303.
- (1975) : « Metasprache und Metakommunikation. Zur Überführung eines sprachphilosophischen Problems in die Sprachtheorie und in die sprachwissenschaftliche Forschungspraxis ». EADEM (éd.) : *Sprachtheorie*. Hamburg : Hoffmann und Campe, 189-205.
- SÉGUY, Jean (1954-1973) : *Atlas linguistique et ethnographique de la Gascogne*. Paris : Ed. du CNRS. 6 vol. (vol. IV par Séguéy en collaboration de Xavier Ravier, vol. V par Jacques Allières).
- (1973) : « La fonction minimale du dialecte ». STRAKA, Georges / GARDETTE, Pierre (éd.) : *Les dialectes romans de France. Actes du colloque national du CNRS sur les dialectes romans à la lumière des atlas linguistiques régionaux*, Colloque de Strasbourg, 1971. Paris : Ed. du CNRS, 27-42.
- SUMIEN, Domergue (sous presse) : *La standardisation pluricentrique de l'occitan. Nouvel enjeu sociolinguistique, développement du lexique et de la morphologie*. Turnhout : Brepols.
- TAGLIAVINI, Carlo (⁶1972) [1949] : *Le origini delle lingue neolatine. Introduzione alla filologia romanza*. Bologna : Pàtron.
- THIERS, Ghjacumu (1993) : « Language contact and Corsican polynomial ». POSNER, Rebecca / GREEN, John N. (ed.) : *Trends in Romance Linguistics and Philology*. Vol. 5 : *Bilingualism and Linguistic Conflict in Romance*. Berlin / New York : Mouton de Gruyter, 253-270.
- THÜMMEL, Wolf (1977) : « Kann man sprachen zählen ? ». *Osnabrücker Beiträge zur Sprachtheorie*, no. 4, 36-60.
- TORRENT-LENZEN, Aina (2006) : *Unidad y pluricentrismo en la comunidad hispanohablante. Cultivo y mantenimiento de una norma panhispánica unificada*. Titz : A. Lenzen.
- TOURTOULON, Charles de / BRINGUIER, Octavien (1876) : *Étude sur la limite géographique de la langue d'oc et de la langue d'oïl*. Paris : Imprimerie nationale [Reprint : Masseret-Meuzac : IEO, 2004].

- URSPRUNG, Daniel (2005) : « Umdeutung eines Helden : Tradition von Erfindung und nationale Identität in der Republik Moldau ». BINDER IJIMA, Edda / DUMBRAVA, Vasile (éd.) : *Stefan der Große – Fürst der Moldau*. Leipzig : Leipziger Universitätsverlag, 15-60.
- VAN CREVELD, Martin (1999) : *The Rise and Fall of the State*. Cambridge : Cambridge University Press.
- WATZLAWICK, Paul / BEAVIN, Janet D. / JACKSON, Don D. (1967) : *Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York : Norton & Co.

RÉSUMÉ

L'article tente de voir quels sont les critères dont disposent les sciences du langage pour faire la différence entre les termes de *langue* et de *variété* et pour individuer les langues, le tout à l'aide d'exemples pris parmi les langues romanes. Il observe l'augmentation du nombre des unités considérées comme *langues* depuis Diez. Ensuite, il fait la différence entre critères linguistiques (formels) ou internes et critères sociolinguistiques ou externes, tente d'isoler les variables qui peuvent jouer un rôle pour cette délimitation, et il arrive à la conclusion que les deux séries ne suffisent pas pour décider des cas vraiment litigieux. La question de la délimitation des langues ne peut donc pas (encore?) être résolue par les sciences du langage, elle reste un problème politique. Néanmoins, les décisions *politiques* ont (ou peuvent avoir) des conséquences *linguistiques*. Parfois, le concept de langue poly- ou pluricentrique peut atténuer les tensions.

MOTS-CLEFS: langue, variété, critères internes, critères externes, langue polycentrique.

ABSTRACT

The article is an attempt to investigate which criteria can be used by language science in order to distinguish between the terms *language (langue)* and *variety (variété)* and identify languages with the help of examples from the Romance languages. It notes the increase in the number of speech forms recognized as languages since Diez. Then it establishes a difference between linguistic (formal) or internal criteria and sociolinguistic or external criteria, endeavours to isolate the variables that may play a part in this delimitation, and reaches the conclusion that the two categories are not sufficient to resolve highly contentious cases. The question of what constitutes a language can therefore not be resolved (yet) by language science alone; it is still a political issue. Nevertheless, political decisions have (or may have) *linguistic* consequences. Sometimes, the concept of polycentric or pluricentric language can alleviate the conflict.

KEY WORDS: language, variety, internal criteria, external criteria, polycentric language