

Linguistique comparée des langues modernes

Professeur : André THIBAULT

Semaine 2 : la Romania, hier et aujourd'hui ; facteurs de différenciation du latin vulgaire ; concepts de substrat, superstrat et adstrat.

1. Introduction

La semaine dernière, nous avons fait un survol des langues indoeuropéennes ; aujourd'hui, nous allons nous pencher plus précisément sur les langues romanes dans leur ensemble, et en particulier sur la langue dont elles proviennent, le latin. Les romanistes appellent « Romania » l'ensemble des territoires où, en Europe, on pratique encore aujourd'hui des langues et des parlers issus de l'évolution du latin oral de l'Antiquité.

2. La carte de la Romania

On va d'abord se reporter à la première carte de la Romania que vous pouvez télécharger sur le site (et qui est aussi présente dans le diaporama). Il s'agit d'une carte de l'Europe sur laquelle figurent les frontières linguistiques et politiques, qui justement ne coïncident pas toujours, ce dont il est très important de prendre conscience.

Les frontières **politiques** sont notées en pointillé fin, alors que les frontières **linguistiques** sont notées par des petits traits. L'idée que l'on se fait des États-nations, en Europe, nous entraîne souvent à croire qu'une langue correspond à un pays, tout simplement, et que les langues s'arrêtent à la frontière. Or, la situation est beaucoup plus complexe, comme on va le voir.

Commençons tout en haut de l'Hexagone, avec la Belgique. On y parle **flamand**, qui est une langue germanique, en **Flandre**, mais on y parle **français** ainsi que des patois **wallons** (donc galloromans, donc d'origine latine) en **Wallonie**. La capitale, Bruxelles, est une enclave francophone en Flandre ; c'est la raison pour laquelle, parmi les romanistes, on l'appelle le point sur le "i" de la Romania. On remarquera que la zone où l'on trouve des parlers flamands s'étend à l'extérieur de la Belgique, jusque dans la Flandre française, dans le département du Nord, autour de la ville de Dunkerque.

Entre la Wallonie et la Moselle, il manque le **Luxembourg**, qui ne figure malheureusement pas sur cette carte assez grossière (veuillez vous référer à la seconde carte à télécharger, plus détaillée). On y parle trois langues : le français, l'allemand, et le **luxembourgeois** (qui appartient à la famille des langues **germaniques**). On descend un peu vers le sud avec la **Moselle** et l'**Alsace**, où l'on parle, en plus du français bien sûr, des **dialectes mosellans et alsaciens**, qui eux aussi appartiennent aux langues **germaniques**. Mais attention : dans ce que l'on appelle la **Moselle romane**, ainsi que dans certaines **vallées des Vosges** jouxtant l'Alsace dialectophone, on n'a pas traditionnellement pratiqué des dialectes germaniques mais bien des **patois galloromans** ; les Alsaciens les appellent des patois "**welsches**". Ces parlers font partie des langues romanes, plus précisément du galloroman, plus précisément encore des **patois d'oïl**, qui occupent tout le nord de la Galloromania, mais qui ne survivent aujourd'hui que de façon précaire et lacunaire, en milieu rural surtout, chez les gens âgés essentiellement, car ils ont tous été éliminés par le rouleau compresseur du français, lequel s'est imposé partout.

Complètement à l'autre bout de la France se trouve la **Bretagne**, où l'on parle encore aujourd'hui **breton**, lequel, comme on l'a vu la semaine dernière, n'est pas une langue romane mais bien une langue celte, importée de Grande-Bretagne aux 5^e, 6^e et début du 7^e s. après J.-C. Elle aussi bien sûr cohabite avec le français, mais survit un peu mieux que la plupart des patois d'oïl.

Au sud-est de la zone d'oïl, on retrouve (ou on retrouvait, car ils sont moribonds) les parlers **francoprovençaux**, ainsi nommés parce qu'ils partagent des caractéristiques du français (et des parlers d'oïl) d'une part, et du provençal d'autre part. Il est important de remarquer que la **Suisse romande**, c'est-à-dire la partie de la Suisse où l'on parle français, est presque entièrement comprise dans la zone des patois francoprovençaux. Dans les Alpes Suisses, en particulier dans le Canton du Valais, les patois survivent encore dans quelques villages. En France, ils sont plutôt moribonds. On ne le voit pas sur la carte, mais une petite partie de la Suisse romande fait aussi partie des patois d'oïl, plus précisément des **patois francs-comtois** ; c'est le canton suisse qu'on appelle le **Jura**. On peut voir qu'une petite partie du francoprovençal déborde en Italie ; la carte est en fait très grossière ici, mais cette région s'appelle le **Val d'Aoste**. On y parle des **patois francoprovençaux** très étroitement apparentés aux patois parlés dans le canton du Valais, en Suisse romande.

Dans le sud de la Galloromania, nous avons l'**occitan**, appellation qui désigne un ensemble de parlers du sud de la France dont certains ont connu, au moyen âge, le statut de langue littéraire, à l'époque des troubadours provençaux ; aujourd'hui, toutefois, l'occitan ne survit que sous la forme d'une myriade de parlers locaux, assez différenciés les uns des autres, et qui de toute façon ne se parlent plus qu'en milieu rural et chez les gens âgés, comme les patois du nord de la France. Selon les publications les plus récentes (cf. Chambon / Greub), le gascon, traditionnellement considéré comme une subdivision de l'occitan, est en fait un ensemble individualisé depuis les tout débuts du moyen âge. On remarquera que l'aire de l'occitan déborde en Italie ; il s'agit des **parlers provençaux du Piémont**. Enfin, il faut mentionner la **Corse** qui, bien que politiquement française, pratique traditionnellement des **parlers italiens**, lesquels se rattachent aux dialectes **toscans** de la Péninsule Italique.

À l'extrémité sud-ouest de l'Hexagone, on retrouve le **Pays Basque français**, où l'on parle une langue qui, comme on l'a dit la semaine dernière, n'est même pas indoeuropéenne : il s'agit bien sûr du **basque**. On notera que l'aire du basque enjambe la frontière et se continue de l'autre côté des Pyrénées, dans le **Pays Basque espagnol**, qui est en fait beaucoup plus important du point de vue démographique et plus étendu géographiquement.

Revenons un peu vers le centre-sud. On aborde l'aire linguistique du **catalan**, avec d'abord le **Roussillon** (principale ville : Perpignan), qui fait partie de la France mais où la langue traditionnellement parlée par le peuple est le catalan. Juste à côté, à l'ouest du Roussillon, on retrouve la Principauté d'**Andorre**, dans les Pyrénées, qui est trilingue dans les faits (français, espagnol, catalan) mais dont la seule langue officielle est le catalan. Puis, plus au sud, on retrouve cette partie de l'Espagne où l'on parle le catalan, sous différentes formes. D'abord, la **Catalogne** proprement dite (capitale *Barcelona*) ; ensuite, le **Pays Valencien** (capitale *Valencia*) ; puis, les **Îles Baléares** (*Mallorca, Menorca, Ibiza* ; en catalan, on dit *Eivissa*).

Le centre de la **Péninsule Ibérique** est occupé par l'**espagnol**, dans toutes ses nombreuses variétés régionales, dont l'andalou est la plus caractérisée. À l'ouest, on retrouve le **galicien** au nord, en **Galice**, puis le **portugais** au centre et au sud, au **Portugal**. À eux deux, ils constituent un groupe appelé par convention le *gallego-portugués*. Il y a huit siècles, le galicien et

le portugais n'étaient qu'une seule et même langue. Ils se sont différenciés peu à peu au fil du temps, de part et d'autre de la frontière politique séparant l'Espagne du Portugal.

Si l'on se déplace maintenant du côté de l'Italie et de la Suisse, on a dans le nord le groupe **rhétoroman** avec le **romanche** (dans le Canton des **Grisons**, en **Suisse**), le **ladin** (dans les **Dolomites**, en Italie) et le **frioulan** (dans le **Frioul**, en Italie). Ces différents parlers romans alpins n'ont en commun qu'une chose : le fait d'être suffisamment différents des parlers italiens septentrionaux pour ne pas être considérés comme en faisant partie. On ne parle donc de « groupe rhétoroman » que par convention, mais on aurait tort de les considérer comme ayant déjà connu une quelconque unité linguistique. Puis, nous avons rapidement mentionné la semaine dernière le **dalmate**, langue aujourd'hui éteinte qui se parlait en Dalmatie, région qui fait aujourd'hui partie de la **Croatie**. Un autre groupe à part entière au sein du groupe italo-roman est constitué par le **sarde**, en **Sardaigne** (attention, le sarde n'est pas une variété dialectale d'italien, c'est une langue romane à part entière). Tout le reste constitue le **complexe** (ou **conglomérat**) **dialectal italien**. Attention : ne pas oublier la **Corse**, que nous avons déjà mentionnée ci-dessus, qui fait politiquement partie de la France mais dont le parler ancestral fait linguistiquement partie du complexe dialectal italien ; ne pas oublier non plus que le Canton du **Tessin**, en Suisse, est italianophone (il existe aussi deux petites vallées italianophones dans le Canton des Grisons).

Pour compléter cette carte, il reste à se reporter complètement à l'est, où l'on retrouve l'aire linguistique du **roumain**, séparée du reste de la Romania, qui englobe la Roumanie et la Moldavie. On remarque qu'il se trouve une enclave en plein centre de l'aire du roumain. Ce sont en fait des régions où survivent des parlers hongrois (langue non-indoeuropéenne).

On a donc, de l'ouest à l'est, les **quatre** sous-ensembles de la **Romania** : l'Ibéroromania, la Galloromania, l'Italoromania et la Dacoromania (j'aurais pu dire "Balkanoromania", mais la Dacoromania est de très loin sa composante la plus importante, comme nous l'avons déjà dit la semaine dernière). Je rappelle que ces grands ensembles ne correspondent pas seulement à des critères linguistiques objectifs, mais aussi – et surtout – à des considérations politiques et géographiques (la langue catalane, par exemple, est plus proche du languedocien que de l'espagnol, mais fait tout de même partie de l'Ibéroromania, ce qui montre bien le caractère jusqu'à un certain point subjectif et arbitraire de ces grands regroupements, qui sont surtout utilisés par commodité).

Cette carte nous donne l'occasion d'introduire le concept de "**Romania Continua**". C'est ainsi qu'on appelle l'ensemble formé de l'Ibéroromania, de la Galloromania et de l'Italoromania, puisqu'elles se touchent, qu'elles forment un continuum géographique. La Dacoromania, en revanche, a été séparée du tronc commun par les aléas de l'histoire (plus précisément, par les invasions slaves), et ne fait donc pas partie de ce qu'on appelle la Romania Continua.

Les spécialistes parlent aussi de "**Romania Nova**" et de "**Romania Creolica**". La Romania Nova est l'ensemble des territoires en dehors de l'Europe où des langues romanes ont été exportées et se sont enracinées (le français au Canada, l'espagnol et le portugais en Amérique latine, etc. etc.). La Romania Creolica est l'ensemble des territoires où existent des créoles à base lexicale romane (essentiellement les créoles français et les créoles portugais).

3. Les cartes de la Romania au début du III^e siècle et aujourd'hui

Nous allons maintenant passer aux cartes intitulées “La Romania au début du III^e siècle” et “La Romania actuelle”). La première donne un aperçu de l'extension géographique maximale des territoires où l'on parlait latin **au début du 3^e siècle après J.-C.** J'attire votre attention sur les faits suivants :

- le **Pays Basque** n'était pas latinisé ; on y parlait bien sûr une langue qui constitue l'ancêtre des dialectes basques d'aujourd'hui.
- l'est de la Sicile, ainsi que la pointe de la botte de l'Italie et son talon, n'étaient pas latinisés non plus, car on y parlait **grec** ;
- certains territoires n'étaient qu'**incomplètement romanisés** ; c'était le cas de la Grande-Bretagne, ainsi que de régions correspondant approximativement au sud-ouest de l'Allemagne, au nord de la Suisse et à l'est de l'Autriche ; toutes ces régions ont cessé de parler le latin vulgaire, après avoir été envahies, occupées et peuplées par des envahisseurs germaniques (les Angles et les Saxons pour la Grande-Bretagne, les Alamans et les Bavarois pour l'Allemagne et la Suisse).
- Certaines régions **entièrement romanisées** ont tout de même **cessé de parler latin** vulgaire ; on voit par exemple que la Flandre belge et le sud des Pays-Bas, ayant été envahis par les **Francs**, ont passé au francique ; que la Bavière et l'Autriche, envahies par les **Bavarois**, ont aussi hébergé un dialecte germanique, qui est devenu aujourd'hui les actuels dialectes bavarois en Bavière et les actuels dialectes autrichiens en Autriche ; que le territoire qui correspond grosso modo à l'ex-Yougoslavie et à la Bulgarie a été quasi entièrement délatinisé, suite à des invasions **slaves** ;
- le nord de l'Afrique était romanisé ; on a cessé d'y parler latin, où ce qu'était alors devenu le latin, lors des invasions **arabes** ; ces invasions ont traversé le détroit de Gibraltar et remonté jusque dans la Péninsule Ibérique, celle-ci ayant été occupée pendant huit siècles (de 711 à 1492) par les Arabes (ou plus précisément des Arabo-Berbères), qui en ont été chassés peu à peu par les populations chrétiennes du nord de la Péninsule, dans un long processus historique que l'on a appelé “la Reconquête” (esp. *la Reconquista*) ;
- les envahisseurs germaniques appelés les **Francs** ont conquis la Gaule, mais leur nombre ne fut pas assez important pour qu'ils arrivent à la délatiniser, contrairement à ce qu'on observe pour la Flandre et les Pays-Bas. On a continué d'y parler le latin vulgaire de basse époque (non sans emprunter bon nombre de mots au dialecte francique, ainsi que, probablement, certaines habitudes articulatoires). De même, les **Lombards**, envahisseurs germaniques appartenant au groupe des Ostrogoths, ont envahi le nord de l'Italie, mais sans pouvoir en déloger le latin ; leur langue s'est perdue.
- La comparaison de ces deux cartes nous permet d'introduire un nouveau concept : celui de **Romania Submersa** (c'est-à-dire “submergée, immergée”). Ce mot désigne toutes les régions qui étaient romanisées au début du 3^e siècle de notre ère, et qui, depuis, ont accueilli une langue d'origine non romane (comme l'anglais en Grande-Bretagne, divers dialectes germaniques dans le sud-ouest de l'Allemagne, des langues slaves dans les Balkans, l'arabe dans le nord de l'Afrique).
- La carte de la Romania au début du 3^e siècle introduit encore deux autres concepts : **Romania occidentale** (rayures obliques) et **Romania orientale** (rayures verticales). On voit que la frontière se situe entre le nord de l'Italie et la Péninsule Italique proprement dite, ainsi qu'entre les territoires correspondant à la Slovénie et à la République Tchèque d'une part, et au reste des Balkans romanisés d'autre part (correspondant grosso modo aux pays de l'ex-Yougoslavie et à la Roumanie). Les romanistes ont en

effet constaté que le latin de la fin de l'Antiquité et du début du moyen âge s'est vite différencié en deux grands ensembles, qui se caractérisent encore aujourd'hui par un certain nombre de traits différenciateurs. Mais avant d'aborder ces traits, nous allons d'abord nous poser une ou deux questions, et tenter d'y apporter des éléments de réponse.

4. Problématiques

a) Pourquoi le latin parlé s'est-il différencié d'une région à l'autre ?

b) Quelle sorte de latin a été exporté dans les différentes régions de la Romania ?

Avant même de commencer à répondre à ces questions, il faut bien prendre conscience d'une chose : ce n'est pas le latin classique, tel qu'il nous est parvenu sous sa forme évidemment écrite à travers la grande littérature latine et les traités de grammaire de l'époque, qui est à la base des langues romanes. La différence entre ce latin châtié, littéraire et codifié d'une part, et le latin parlé par le peuple d'autre part, était aussi grande que celle qui existe aujourd'hui entre le français écrit et le français parlé. Or, c'est justement le latin dans sa version parlée, dans sa variété orale, qui est à la base de ce qui allait devenir les différentes langues romanes encore parlées de nos jours. C'est cette variété qu'on a longtemps appelée le **latin vulgaire** (mais qu'il conviendrait plutôt d'appeler aujourd'hui le **protoroman**). Rien à voir avec la vulgarité ; le mot "vulgaire" vient ici du latin VULGUS, qui voulait tout simplement dire "le peuple, la foule, le commun des mortels". Si l'on considère l'expansion de l'Empire romain, il s'agit, en proportions variables selon les régions, du latin parlé par les colons, les soldats, les administrateurs et les négociants.

Or, contrairement à la langue classique qui nous apparaît comme un objet relativement figé, le latin vulgaire était une langue **vivante**, donc soumise à la **variation**, dans le temps, dans l'espace et selon les classes sociales, les registres et les niveaux de langue, etc. Il est donc tout naturel qu'en se diffusant sur des territoires immenses, le latin vulgaire se soit différencié ; c'est le propre de toute langue vivante de grande extension que de se morceler, en particulier si elle est parlée par un très grand nombre de locuteurs, sur une étendue immense, dans des contrées où l'on parlait auparavant d'autres langues qui ont laissé des traces dans le latin local.

4.1. Les cartes des concepts « manger » et « tête »

4.1.1. « manger »

Le document téléchargeable (et aussi présent dans le diaporama) sur lequel figurent deux cartes tirées d'un ouvrage classique sur les langues romanes, *Le Origini delle Lingue Neolatine* de C. Tagliavini, illustre bien la différence entre le latin classique et le latin vulgaire. Deux concepts y sont illustrés : dans la première carte, celui du verbe exprimant le concept de "manger", et dans la seconde, l'idée de "tête". En latin classique, le mot pour dire "manger" était **EDERE** ; il n'a survécu nulle part, dans aucune langue romane¹. Que trouve-t-on à la place ? En espagnol et en portugais, le verbe *comer* vient d'un dérivé préfixal de **EDERE**, **COMEDERE**, qui voulait dire "dévorer, bouffer" ; donc, un mot plus imagé. C'est

¹ Mais on en trouve la trace dans le mot anglais *edible*, adapté du latin EDIBILIS, qui signifie "comestible".

comme si, en français de l'an 2500, plus personne ne disait *manger*, le verbe *bouffer* étant devenu le mot neutre pour exprimer la notion en question. Dans la plupart des langues romanes, on a un mot qui vient du latin **MANDUCARE**, qui voulait dire "mâcher". Il s'est répandu dans le nord de la Sardaigne, en Dacie et en Gaule, d'où il s'est propagé au catalan et à l'italien. Enfin, on retrouve deux synonymes beaucoup moins diffusés géographiquement : **PAPPARE**, mot onomatopéique appartenant au langage enfantin, qui a donné le sarde *pappai*, et **MAGULARE**, qui a donné *maglier* en romanche, et qui serait dérivé d'un mot de latin tardif, **MAGULUM**, qui voulait dire "bouche".

4.1.2. « tête »

La seconde carte nous montre le cas des mots exprimant le concept de "tête". On avait en latin classique **CAPUT**. Ce mot survit beaucoup mieux que **EDERE** ; en effet, on le retrouve, sous la forme phonétiquement évoluée *cap*, dans les parlers (gascons, languedociens) du sud-ouest de la France², en catalan et en roumain. La Toscane et le sud de l'Italie ont gardé le mot sous la forme *capo*. En espagnol et en portugais, le mot latin qui a donné naissance à *cabeza* (esp.) et *cabeça* (port.) est un dérivé de **CAPUT** : **CAPITIA**. Le sens premier de ce mot en latin était "ouverture qu'ont les tuniques pour y passer la tête". En galloroman ensuite, le mot *tête*, ou *testo* en provençal, vient du latin **TESTA** qui voulait dire "pot en terre cuite". C'était donc une appellation facétieuse, plaisante, comme si l'on disait aujourd'hui *la citrouille*, *la cruche* ou *la casserole* pour désigner la tête (une tête vide, plus précisément !). Ce type lexical s'est ensuite étendu à plusieurs régions de l'Italie, sous la forme *testa*. On trouve enfin deux autres étymons, c'est-à-dire deux autres mots latins ayant donné lieu à des formes relevées dans les langues romanes modernes : **CONCA**, qui s'est conservé tel quel en sarde, et **COCCIA**, que l'on retrouve dans cette région de l'Italie qu'on appelle les Abruzzes. Ces deux mots voulaient dire "coquille", "coquillage" (donc, encore une fois, un objet creux).

4.2. Les facteurs de différenciation du latin vulgaire (ou « protoroman »)

Ces exemples attirent notre attention sur le phénomène de la **différenciation** du latin vulgaire, lequel a évolué différemment dans chaque région de la Romania pour devenir les langues que nous appelons aujourd'hui *langues romanes* ou *néo-latines*. Nous allons maintenant examiner différents facteurs qui expliquent ces phénomènes de différenciation.

4.2.1. La chronologie

On considérera comme premier facteur de différenciation la **chronologie** de la romanisation. En effet, Rome n'a pas exporté la même variété de latin au 2^e s. avant J.-C. et au 2^e s. après ; le latin a évolué pendant cette période de quatre siècles. À titre de comparaison, il est facile de constater que l'anglais exporté par l'Angleterre en Amérique n'est pas l'anglais exporté par ce même pays en Australie (et que l'anglais australien ressemble davantage, malgré tout, à l'anglais d'Angleterre que l'anglais nord-américain). Pourquoi ? Parce que la colonisation de l'Australie a eu lieu à une date beaucoup plus récente (fin 18^e puis 19^e s.) que la colonisation de l'Amérique (17^e et 18^e s.). (Cela dit, bien sûr, la provenance régionale et surtout sociale des colons fut aussi très importante, comme du reste pour le latin vulgaire.) Pour prendre un exemple plus extrême, le français exporté en Nouvelle-France (c'est-à-dire au Canada) au 17^e

² Le français l'a emprunté dans l'expression *de pied en cap* « des pieds à la tête ». À vrai dire, il survit également dans le mot français *chef*, mais l'évolution phonétique l'a rendu méconnaissable et il a connu en outre une spécialisation sémantique par métaphore (le *chef* est celui qui se trouve à la *tête* d'un groupe).

siècle n'était pas le même français que celui exporté au Maghreb au 20^e siècle, ce qui fait que les Canadiens ne parlent pas français comme les Marocains, et que le français des Marocains parfaitement bilingues est plus proche du français de France que ne l'est le français du Canada. On assiste au même type de phénomène pour le latin vulgaire, qui s'est diffusé, dans le temps, de cette manière :

- D'abord, bien sûr, dans la **Péninsule italique**, à partir de Rome, qui étendit son dialecte à toute la province du Latium (4^e-3^e siècle avant J.-C.) ;
- Puis, à la **Sardaigne** (conquise en -238) ;
- Ensuite, à la **Péninsule Ibérique** (conquise environ de -218 à -12)
- Puis à la **Gaule** (Marseille déjà entièrement romaine à partir de -125 ; conquête définitive de toute la Gaule par César entre -58 et -50 ;
- Enfin, à la **Dacie** (au début du 2^e siècle après J.-C. ; elle reçut alors de nombreux colons romains).

4.2.2. La nature scolaire ou spontanée de la transmission et de la diffusion du latin

Un autre facteur de différenciation semble avoir joué un rôle important : il s'agit de la nature plus ou moins **scolaire** ou, au contraire, plus ou moins **spontanée** de la **transmission** du latin et de sa **diffusion** dans la population. C'est ce facteur qui nous permet de revenir à la carte de la Romania au début du 3^e siècle, et aux concepts de "**Romania occidentale**" et "**Romania orientale**". Selon les régions de la Romania, le latin a été diffusé par l'administration et les institutions romaines, qui l'ont transmis, à travers l'école, mais aussi à travers l'exemple de locuteurs appartenant à l'administration, à l'armée, aux entreprises commerciales, etc., à des populations qui, à la base, ne parlaient pas latin ; ou alors, en revanche, le latin a été diffusé par des colons de la péninsule italique, eux-mêmes latinophones, mais qui ont exporté leur propre latin populaire ou rural, celui qu'ils parlaient dans la vie de tous les jours, sans aucune influence scolaire ou normative. Si l'on se reporte à la carte, on dira que la Romania occidentale correspond plutôt au modèle de diffusion de la langue par le biais de l'école et de structures permettant la romanisation de populations non latinophones à la base, alors que la Péninsule italique, la Sicile et les Balkans ont été colonisés, et donc romanisés, par des colons qui parlaient déjà latin à la base, mais un latin non bridé par l'école.

Cette situation a eu, bien sûr, des répercussions linguistiques. La grande différence entre les langues de la Romania orientale et celles de la Romania occidentale réside dans la **formation du pluriel**, ainsi que celle de la deuxième personne du singulier des verbes. Cette différence, on va le voir, a été raisonnablement expliquée par la **chute du -s final** dans le latin parlé de la Romania orientale, alors que la Romania occidentale a bien maintenu ce -s final.

Pour mieux comprendre cette situation, nous pouvons établir un parallèle avec l'espagnol moderne. En espagnol standard, le -s final de syllabe se prononce très clairement. On dira par exemple *¿Cómo está Usted?* Or, dans le sud de l'Espagne, cette consonne a tendance à tomber, à ne pas être articulée. On dira donc : *¿Cómo e'tá U'te(d)?* Eh bien, on peut s'imaginer approximativement la même situation pour le latin il y a plus ou moins deux mille ans.

Le latin vulgaire de l'Italie et de la Dacie connaissait un amuïssement (affaiblissement) du -s final, ce qui n'était pas le cas à l'ouest, où l'enseignement du latin comme langue d'abord étrangère aurait favorisé, selon certains chercheurs, le maintien de ce -s final ; comme résultat, la formation du pluriel, ainsi que celle de la 2^e pers. du singulier des verbes, sont différentes.

Le **pluriel** des langues romanes de l'ouest se fait **en ajoutant un -s**, ce qui est un héritage de l'accusatif latin, alors que le pluriel en italien et en roumain se fait **en changeant la dernière voyelle**, ce qui a longtemps été expliqué comme un héritage du nominatif latin, bien que cette hypothèse ait été rejetée récemment au profit d'une explication plus complexe faisant intervenir des considérations d'ordre phonétique³ :

- latin (acc. pl.) LUPOS > port., esp. *lobos*, fr. *loups* mais latin (nom. pl.) LUPI > ital. et roum. *lupi*

- latin (acc. pl.) CAPRAS > port., esp. *cabras*, fr. *chèvres* mais latin (nom. pl.) CAPRAE > ital. roum. *capre*

La deuxième personne du singulier des verbes se termine en *-s* à l'Ouest, en *-i* à l'Est ; cela ne peut être que le résultat d'une évolution phonétique particulière qui distingue encore une fois la Romania occidentale de la Romania orientale :

- latin CANTAS > port., esp. *cantas*, fr. *chantes* mais ital. *canti*, roum. *cânti*.

Quant au pronom personnel latin **NOS** (qui signifie "nous"), il correspond à *noi* en italien et en roumain, mais a gardé son *-s* dans les langues de la Romania occidentale (fr. *nous*, catalan *nos[altres]*, espagnol *nos[otros]*, portugais *nós*). Encore une fois, on voit que le *-S* final du latin est devenu un *-i*.

On remarque aussi que le latin **DUOS** (qui signifie "deux") correspond à *doi* en roumain, et *due* en italien, mais *dos* en catalan et en espagnol, *deus* en ancien français (la graphie *deux* est postérieure, et la présence du *x* s'explique par le fait que cette consonne équivalait à *us*), *dois* en portugais, *duos* en sarde (qui, en raison de son grand archaïsme, n'a été classé ici ni dans la Romania orientale ni dans l'occidentale).

4.2.3. L'influence des langues de substrat

Nous avons donc vu jusqu'à maintenant deux facteurs de différenciation : la chronologie, et la nature de la diffusion du latin (scolaire ou spontanée). Nous allons maintenant aborder un autre facteur, celui de la langue de **substrat**. Nous avons vu que dans une grande partie de la Romania, le latin avait en fait le statut de langue apprise, de langue seconde, de langue qu'un peuple acquiert peu à peu, d'une génération à l'autre, en passant par des étapes de bilinguisme généralisé. Or, la langue maternelle de ces peuples romanisés a laissé des traces dans le latin qu'ils se sont mis à parler, collectivement. Il est très facile de comprendre cette situation en comparant, par exemple, avec les Alsaciens. Autrefois, les Alsaciens parlaient essentiellement le dialecte alsacien (qui appartient à la grande famille des langues germaniques, plus précisément au conglomérat dialectal allemand). Or, en quelques générations, ils se sont mis à parler majoritairement le français. Toutefois, plusieurs Alsaciens, en particulier les personnes âgées, se signalent par un accent alsacien plus ou moins fort lorsqu'ils parlent français. C'est normal, puisque c'est une langue étrangère qu'ils ont dû apprendre. Il faut bien comprendre que les populations celtiques, en particulier les Gaulois en Gaule et les Celtibères dans la Péninsule ibérique, ont aussi laissé dans le latin qu'ils apprenaient la marque, l'empreinte de leur langue maternelle. Ainsi, plusieurs chercheurs attribuent au **substrat** celte ce qu'on appelle la sonori-

³ Vincenzo FARAONI, *L'origine dei plurali italiani in -e e -i*. Alessandria, dell'Orso, 2018.

sation des consonnes sourdes intervocaliques ($p > b$; $t > d$; $k > g$), facteur supplémentaire de division entre la Romania orientale et la Romania occidentale. Voici quelques exemples pour illustrer le phénomène :

- It. LEPORE > ptg. *lebre*, esp. *liebre*, fr. *lièvre* mais it. *lepre*, roum. *iepure*.

- It. MUTARE > ptg. esp. *mudar*, fr. *muer* (en français, le phénomène est allé jusqu'à la chute) mais ital. *mutare*, roum. *muta*.

- It. IOCARE > ptg. *jogar*, esp. *jugar*, fr. *jouer* (ici encore, le phénomène est allé jusqu'à la chute) mais ital. *giocare*, roum. *juca*.

On peut citer un autre exemple d'influence des **substrats** dans la Romania :

- Le passage de *f-* à *h-* à \emptyset (c'est-à-dire la chute, la disparition) en espagnol (ainsi qu'en gascon), phénomène d'origine probablement basque ou en tout cas cantabrique (exemple : latin **FILIUM** > catalan *fill*, portugais *filho* mais espagnol *hijo*) ; en basque, le son *f-* à l'initiale n'existe pas non plus.

4.2.4. L'influence des langues de superstrat

Au phénomène des substrats est venu se superposer un autre phénomène de contact de langues connu sous le nom de **superstrat**. Si l'on revient à la carte "La Romania actuelle", on voit que des flèches ont été tracées, avec le nom de groupes d'invasisseurs germaniques. S'ils ont réussi à imposer leur langue aux Pays-Bas, en Allemagne, en Suisse alémanique ou en Autriche, ils n'ont tout de même pas réussi à la faire triompher dans plusieurs autres régions qu'ils ont occupées et dominées militairement. Par exemple, on a continué de parler latin vulgaire dans le nord de la Gaule, où les Francs se sont mêlés aux Galloromains ; de même, on a continué de parler latin vulgaire dans le nord de l'Italie, malgré l'invasion des Lombards, qui eux aussi se sont mêlés aux populations locales. Il y a eu bien d'autres envahisseurs qui ne sont pas représentés sur cette carte : les Burgondes en Bourgogne, les Wisigoths en Espagne, etc. Or, il ne faudrait pas croire que le fait que leur langue ne se soit pas imposée signifie qu'elle n'ait eu absolument aucune influence sur le latin vulgaire local. Lorsqu'une population suffisamment nombreuse vient s'installer au sein d'une communauté linguistique donnée, il y a de fortes chances pour qu'elle transmette un certain nombre d'éléments (qui peuvent être phonétiques, syntaxiques, morphologiques, sémantiques ou lexicaux, etc.) à la langue du pays en question. Par exemple, dans le cas du francique dans le nord de la Gaule, de nombreux mots ont été empruntés par les Galloromains à la langue des Francs, dont plusieurs se sont perpétués jusqu'à nos jours. C'est le cas par exemple du mot *fauteuil*, attesté en 1611 sous la forme du moyen français *faudeteuil*, puis plus anciennement encore, en 1080, sous la forme de l'ancien français *faldestoel* ; ce mot vient du francique *°faldistôl* qui signifie "siège pliant" ; on peut comparer les éléments formateurs de ce mot avec l'anglais *to fold* et l'allemand *falten*, qui veulent dire "plier", ainsi que l'angl. *stool* signifiant "tabouret" et l'allemand *Stuhl* qui veut dire "chaise". Le fauteuil était donc, au début, un siège pliable.

On parle de **superstrat** dans ce cas, car il s'agit d'une couche linguistique qui est venue se superposer (alors que dans le cas du substrat, il s'agissait d'une couche linguistique sous-jacente).

4.2.5. L'influence des langues d'adstrat

Il reste enfin à introduire un dernier concept, celui de langue d'**adstrat**. Contrairement aux concepts de substrat et de superstrat, qui impliquent que des peuples se sont fondus pour ne plus en faire qu'un seul, le concept d'adstrat renvoie à des situations où deux peuples ont été en contact, se sont côtoyés pendant une certaine période, d'une façon assez suivie et assez intense pour s'échanger de nombreux mots. On parle dans ce cas de langue d'**adstrat**, mot formé avec un préfixe issu de la préposition latine **AD**, qui dans ce cas précis veut dire "à côté" ; ce n'est pas une strate linguistique sous-jacente, ni superposée, mais latérale.

On peut citer comme exemple le rôle de l'**arabe** dans l'histoire de la langue espagnole (et, en fait, du portugais et du catalan). En effet, les Arabes ont occupé, à tout le moins en partie, la Péninsule Ibérique pendant près de huit siècles (de 711 à 1492). Au fur et à mesure que les Royaumes Chrétiens du nord arrivaient à reconquérir la Péninsule en avançant vers le Sud, les Arabes reculaient et refluaient vers le Maghreb. Pendant ces nombreux siècles de cohabitation dans la Péninsule, qui n'ont pas toujours été exclusivement belliqueux, loin de là, les Espagnols et les Arabes ont cohabité, s'échangeant de nombreux mots ; plusieurs d'entre eux étaient même bilingues. Mais comme cette cohabitation n'a pas eu pour conséquence une fusion des deux peuples, comme ce fut le cas des Francs et des Galloromains en France par exemple, on ne parle pas ici de superstrat mais bien d'adstrat. On pourrait dire la même chose des langues **slaves** pour le roumain. Le roumain a été en contact ininterrompu, depuis le 6^e siècle après J.-C. déjà, et ce jusqu'à nos jours, avec les langues slaves, et leur a emprunté énormément de mots. On dira donc que les langues slaves sont une langue d'adstrat pour le roumain.

Quelques exemples de mots espagnols et/ou portugais (il y en a plusieurs centaines) d'origine arabe : esp. *alubia* 'haricot', de l'arabe *lūbiyā*, de même sens, lui-même d'origine perse ; esp. *ojalá* et port. *oxalá* 'pourvu que, espérons-le', de l'arabe andalou *law šà llàh* correspondant à l'arabe classique *law šaʔa llah* qui signifie littéralement « si Dieu le veut » ; esp. *hasta*, port. *até* 'jusqu'à', de l'arabe andalou *hattá*, lui-même de l'arabe classique *hattá* ; esp. *almacén*, port. *armazém* 'magasin', de l'arabe andalou *almaxzán*, de l'arabe classique *maxzan*.⁴

Quelques exemples de slavismes en roumain, tous issus du vieux bulgare : *da* 'oui' (plusieurs langues slaves) ; *glas* 'voix' (cf. bulgare *glas*, id.) ; *prieten* 'ami' (cf. bulgare *priyatel*) ; *a iubi* 'aimer' (cf. bulgare *obicham*) ; *lopată* 'pelle' (cf. bulgare *lopata*).

4.2.6. Le « stratum »

On pourrait croire qu'il n'y a que des substrats, des adstrats et des superstrats ; mais le cœur même de la langue, qu'on peut appeler le « stratum », est constitué par le latin parlé ; c'est le fond de la langue, c'est l'essentiel de l'origine des langues romanes.

⁴ Référence bibliographique : CORRIENTE, Federico, *Diccionario de arabismos y voces afines en iberorromance*, Madrid, Gredos, 1999.

Donc, schématiquement :

- substrats
- stratum
- adstrats et superstrats

La différence entre les deux derniers est qu'une langue d'adstrat peut continuer d'exister, alors qu'une langue de superstrat a fini par céder la place.