

La dispute autour de l'arbre. Comment faire la linguistique historique ?

François Jacquesson

► **To cite this version:**

François Jacquesson. La dispute autour de l'arbre. Comment faire la linguistique historique ?. 2018. halshs-02925543

HAL Id: halshs-02925543

<https://halshs.archives-ouvertes.fr/halshs-02925543>

Preprint submitted on 30 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La dispute autour de l'arbre : Comment faire la linguistique historique ?

F. Jacquesson

Certains spécialistes de linguistique historique sont irrités qu'on conteste le bien-fondé de l'image de « l'arbre généalogique » des langues. Expliquons calmement de quoi il est question.

1. Les emprunts de mots

Les ressemblances entre deux langues peuvent être nombreuses mais, ici comme ailleurs, la ressemblance ne suffit pas à prouver la parenté ; et on peut aussi se demander ce que signifie vraiment « parenté » entre des langues. Deux langues peuvent avoir en commun un ou plusieurs mots (*peanuts, coca, hard*) sans que cela « ne prouve rien ». De nombreuses langues possèdent ces mots depuis cent ans parce que la puissance de certaines modes américaines, relayées par le cinéma ou autrement, fait que ces mots ont été souvent empruntés. Mais le fait que deux langues possèdent aujourd'hui ces mots ne prouve pas qu'elles ont une histoire commune ancienne ; il prouve qu'elles ont une histoire commune depuis cent ans. Ce n'est pas sans intérêt.

Mais si vous vous intéressez à l'histoire ancienne, ces mots-là ne conviennent pas. Le fait qu'on emprunte des mots à une autre langue ne date pas d'aujourd'hui et l'on peut repérer ce phénomène à toutes les époques. C'est intéressant, parce que cela « signe » une diffusion culturelle précise, et c'est important historiquement. Par exemple, dans les premiers siècles de l'ère chrétienne, on voit la langue égyptienne (dite alors : copte) pleine de mots grecs. S'il y avait eu des affiches dans la capitale copte de l'époque, elles auraient été pleines de mots grecs comme chez nous de mots anglais.

Cette diffusion peut aussi être vraie négativement. En Europe du Nord, à une date antérieure à la diffusion du christianisme, beaucoup de langues ont emprunté des noms de champignon, mais le nom pour « l'ours » est anormalement différent d'une langue à l'autre. On sait pourquoi. Ces champignons étaient utilisés comme drogues dans des rituels, et avaient une grande valeur : ils circulaient largement, avec les mots qui les désignent, à travers l'Eurasie froide. Inversement, l'ours était un animal puissant, semi-magique, et craint. On évitait de lui donner son vrai nom, qui était donc tabou ; et on lui donnait des surnoms, qui étaient différents d'une langue à l'autre. Ce fait est encore visible aujourd'hui, par exemple en russe où « ours » se dit *medved'* 'mangeur de miel'.

Ce qui irrite certains spécialistes de linguistique historique, dans de tels cas de mots (et de faits culturels) qui circulent, c'est que cette situation est le contraire de ce qu'ils cherchent. A leurs yeux, ces emprunts sont des phénomènes superficiels, qu'on doit étudier en effet, mais pour s'en débarrasser ; car ce qui compte, c'est ce qui est plus authentique.

2. La quête de l'authentique

Dans une certaine mesure, et à condition d'être très prudent avec ce mot « authentique », ils ont raison. Le mot *dancing*, en français, « dit très bien » qu'il est d'Après-guerre ; et *shopping* un peu plus tard, quand la prospérité revient. Il est clair qu'on ne trouverait pas ces mots chez La Fontaine.

Ces mots nouveaux sont évidemment intéressants, mais les mots anciens ? Le problème est qu'on imagine souvent que, au contraire des mots nouveaux qui viennent de partout, les mots anciens seraient « bien de chez nous » et tous sortis du même bocal à confiture « comme on faisait autrefois ». Mais ce n'est pas vrai, les confitures sont diverses, de même que les grands-mères, et les mots « bien de chez nous » ont, eux aussi, des histoires bizarres. Plus ils sont « chez nous » depuis longtemps, plus ils ont pris la même patine. Mais il faut y voir de plus près.

On dit que la plupart des mots français (on veut dire : des mots du français d'aujourd'hui) sont d'origine latine. C'est tout-à-fait vrai, même si notre degré de « sang romain » est à peu près nul, pour la plupart d'entre nous ! Par exemple, les mots *tout-à-fait vrai* sont tous d'origine latine. Ou plutôt : les morceaux qui les composent sont d'origine latine. Pour *vrai*, qui n'a qu'un morceau, il est le résultat français d'un mot latin (mais ce ne serait pas vrai pour le mot *très*). Quant à *tout-à-fait*, c'est une expression très ancienne en français, mais pas du tout latine : c'est une création française, dont on ne trouve pas l'équivalent en italien ou en castillan. Toutes les pièces sont d'origine latine, mais l'assemblage ne l'est pas. Cela signifie que les gens qui parlent entre eux refont les choses à leur manière. Ils sont bien contents d'avoir des mots à se mettre sous la dent, mais ils les mâchent à leur façon.

Ici, le spécialiste des lois phonétiques risque de m'interrompre.

3. La machine à sons

Il va expliquer que ce n'est pas monsieur X ou madame Y qui choisit comment on parle ; que les mots que nous employons, dans la presque totalité des cas, nous les avons reçus en apprenant à parler, et que nous devons les prononcer d'une certaine façon pour être compris. Il va ajouter que, quand on observe le changement des mots à travers les siècles, on constate l'opération de phénomènes généraux qui n'ont rien à voir avec l'initiative personnelle, et que même si c'est triste, c'est comme ça.

Par exemple, la langue française du nord (la norme en France) est caractérisée par de nombreuses voyelles dites nasales, qu'on peut par exemple écrire *an* comme dans *passant* ou *allant* ; *on* comme dans *son*, *bon* ; *in* comme dans *infernal* ou *brin*. Ces voyelles sont difficiles pour les étrangers qui apprennent notre langue ; elles sont assez typiques du français ; il faut les apprendre de bonne heure pour les prononcer correctement. Or, personne n'a décidé que *vin*, *pin*, *lin*, etc. se prononçaient de cette façon ; et de même pour *pont*, *font*, *long* etc. C'est un processus régulier qui a agi à chaque fois qu'une voyelle était suivie d'une consonne nasale (n, m) en fin de mot. C'est tellement régulier que l'expression de « loi » (phonétique) n'est pas étrange.

On peut même voir quels faits ont bloqué cette transformation. On dit *bon* (avec voyelle nasale : on n'entend pas le « -n » final), mais on dit *bonne* (avec le « -n » final). C'est parce qu'autrefois, le « -e » du féminin était prononcé, exactement comme il est encore prononcé dans le Sud de la France. Conséquence : le « -n » n'était pas en position finale ; et conséquence : le « -o » de *bonne* est resté un « o normal ». La loi s'applique donc bien : les « -n » en position finale ont disparu en « nasalisant » la voyelle qui précédait, mais ceux qui n'étaient pas en position finale sont restés prononcés.

Il faudrait, pour être honnête, ajouter que tout a changé depuis cette loi, puisque (en français du Nord), il y a belle lurette que le « -e final » de *bonne* ne se prononce plus, et que donc le *bonne* d'aujourd'hui se trouve exactement dans la condition du *bon* d'autrefois : avec un « -n » final. Cela signifie donc que la loi ne marche plus. C'était peut-être une loi, d'accord, mais apparemment une loi liée à des circonstances particulières.

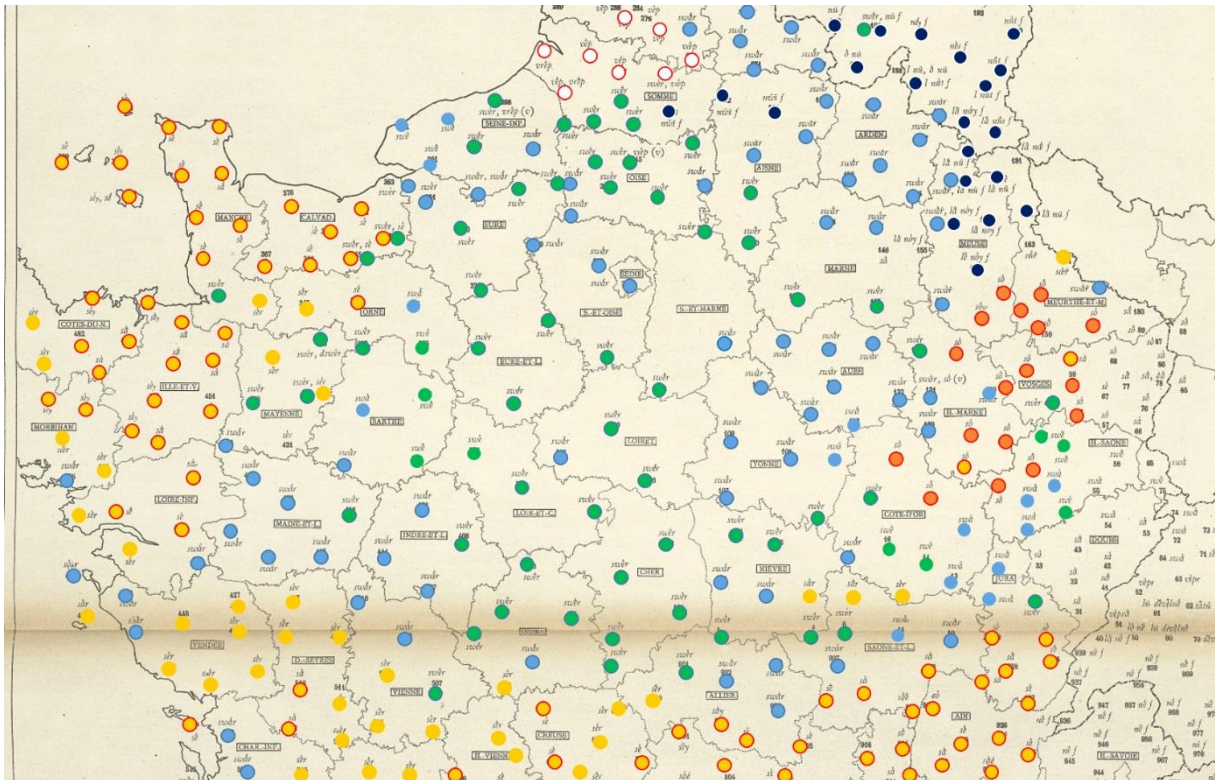
4. La « loi » et ses circonstances

Il existe en effet deux façons, toutes deux utiles, pour observer de tels phénomènes. La première façon est de dégager des correspondances régulières. Par exemple on peut observer qu'il existe une ressemblance régulière entre latin et le français moderne, comme le montre le tableau suivant.

latin	français
sēro	soir
fēriae	foire
vidēre	voir

Tableau 1 : une correspondance régulière entre latin et français

On a donc un son /war/, écrit 'oir', et qui « remonte » à un segment latin -ēr-. C'est intéressant, et on peut expliquer comment il a pu se faire que les sonorités latines, via les sonorités anciennes dont l'orthographe 'oir' témoigne, aient abouti à la prononciation d'aujourd'hui. C'est bien un fait régulier. On peut l'appeler une « loi » si l'on y tient, mais alors il ne faut pas oublier de rappeler que c'est une loi locale et limitée. En effet, non seulement les autres langues comme l'italien et le castillan, le portugais, l'occitan etc. n'ont pas du tout la même « loi », mais l'extension de cette prononciation en français est récente. Voyez la « photographie » de la situation juste avant la Guerre de 14. Il s'agit du mot « soir ».



Carte 1 : Le mot pour « soir » selon l'Atlas Linguistique de la France, carte n°1238.

●	swèr	●	swè	●	< nuit
●	swar	●	swa	○	< vèpres
●	sèr, sar	●	sè, sa		
		●	so		

On voit que la forme qui est devenue dominante en français actuel, la forme /swar/ était à cette époque (il y a environ cent ans) répandue dans le centre de la France du nord, où elle était d'ailleurs presque partout en concurrence avec /swèr/, très répandue également. On trouve des formes différentes dans l'ouest du pays, depuis la Normandie /sèr, sar/ et même /sè, sa/ ; formes qu'on retrouve dans le sud et, à un moindre degré, dans l'est où l'on a aussi souvent /so/. Dans le nord de l'actuel territoire français et en Belgique, la carte montre des formes tout à fait différentes, dérivées des mots « nuit » et « vèpres » : le mot « soir » n'est pas pratiqué.

Cette carte montre deux choses décisives. D'abord et surtout, elle nous montre que, si personne n'avait fait d'enquêtes avant 1914, nous croirions certainement en toute bonne foi que les choses sont simples : « en français, on dit /swar/ » ! Or, les enquêtes nous montrent que c'est faux. Il faut dire : « aujourd'hui, en français courant, on dit /swar/ ». Car cette carte nous prouve qu'il y a cent ans, le « paysage linguistique » était beaucoup plus subtil et différencié. Les raisons qui font que les formes

dont témoigne cette carte ont (presque) disparu, sont bien connues : les guerres d'abord, qui ont amené des déplacements de population importants, et une rupture de la vie en villages ; les moyens de communication ensuite (routes, chemins de fer, radio, télé) qui ont assez rapidement imposé l'impression d'une langue unique.

Mais cette carte nous montre aussi, en conséquence, que les correspondances régulières du petit tableau n°1 étaient limitées à un dialecte particulier. Il se trouve que ce dialecte est devenu le français courant d'aujourd'hui, parce que c'était celui de la région parisienne, d'accord, mais ce n'était pas « le français » tout seul. Et donc, l'effet impressionnant de la « loi » qui montre la correspondance des sons « entre latin et français » nous trompe. En réalité, il s'agit d'une correspondance entre latin et un dialecte particulier ; il y en avait d'autres. C'était seulement une loi limitée et locale. Et qui, comme pour les voyelles nasales, repose sur des phénomènes de changement qui n'ont duré qu'un temps.

5. Le « modèle de l'arbre » et ses limites

Que ces phénomènes de changement aient été très locaux dans l'espace et dans le temps, on aurait pu s'en douter. Car s'ils avaient été des lois générales et d'application continue, nous parlerions tous la même langue dans l'Europe latine. Le latin aurait sûrement changé, mais il aurait partout changé de la même façon ; et on voit clairement que ce n'est pas vrai.

C'est là qu'intervient le « modèle de l'arbre ». Certains linguistes disent : Oui, le changement des langues est fait de changements locaux. Un même tronc se divise en branches, puis en petites branches, selon que les changements singularisent telle partie des gens. Ils ajoutent que ce sont ces changements locaux qui permettent de faire la généalogie des langues, qui forment donc des familles et, avec le temps, des sous-familles.

Le problème, comme le montre aussi notre carte ci-dessus, c'est que ce n'est pas vrai. Car cela supposerait que la « filiation » est linéaire, et que chaque mot français (ou italien etc.) descend d'un mot latin, exactement comme un individu sexué descend à chaque fois de la recombinaison des gènes de deux parents. Un individu ne va pas chercher les gènes d'un 3^e ou d'un 4^e ! Or, c'est exactement ce qui se passe avec le vocabulaire et la grammaire. Le mot français actuel *soir* (prononcé /swar/) vient bien du latin *sēro*, mais en latin ce mot signifie 'tard, plus tard, trop tard'. Le mot latin qui signifiait 'le soir' est *vesper*, celui-là même qui a donné le mot français moderne *vêpres*.

Dans les enquêtes faites avant 1914, ce mot « vêpres » est très minoritaire : il apparaît sur un rebord du domaine francophone, dans une partie du département de la Somme et dans celui du Pas-de-Calais. Beaucoup de gens, non seulement en France, mais dans d'autres pays « latinophones » aussi, ont donc remplacé un mot par un autre, en changeant la signification. La filiation n'est pas linéaire.

En outre, on sait bien que dans de nombreuses langues, on assimile 'matin' et 'demain', comme en allemand avec *Morgen*. C'est très compréhensible. Mais en anglais on a fini par faire une différence entre *morrow* (comme dans *to-morrow* 'demain') et *morning* ('matin'), quoique ces deux mots soient étroitement liés. Mais pas en français où le mot *demain* est un rafistolage du latin *mane* 'demain', tandis que *matin* vient de *matutinu* qui est à l'origine un adjectif évoquant la déesse romaine Matuta, équivalent de l'Aurore.

Les mots ont des formes, c'est certain ; et heureusement. Mais ils ont aussi un sens. Si l'on ne veut suivre que le changement des formes, on ne comprend presque rien, en tout cas on ne comprend pas *soir* puisqu'en latin *sēro* ne veut pas dire 'soir' ; on ne comprend pas pourquoi *matin* peut être traduit par *Morgen*, parce que ces mots ont eu dans le passé des significations tout à fait différentes. Cela signifie que les « arbres généalogiques » forment une image approximative pour suivre les formes, mais certainement pas une bonne image pour comprendre les significations.

6. Le linguiste doit être aussi intelligent que les gens qui parlent les langues !

Ceux qui parlent, c'est-à-dire nous, doivent faire avec les mots qu'ils trouvent. Il est exact qu'il est très difficile d'inventer des mots vraiment nouveaux, et c'est pourquoi le plus souvent on en emprunte aux langues étrangères, quitte à en déformer complètement le sens. Mais on a vu les modes faire des choses étonnantes... Qui aurait parié sur le succès du mot *cinéma* ? Ou sur *ordinateur* - qui maintenant est presque systématiquement *ordi* ? Qui aurait prévu le succès d'une expression comme *je m'en tape* ou comme *tu me lâches* ? ?

Les linguistes doivent comprendre que, s'ils veulent comprendre les langues, ou l'histoire des langues, il leur faut essayer d'être aussi astucieux que les gens qui les parlent. On comprend que les linguistes aient été fascinés par les « lois », surtout à une époque où les lois de l'astronomie et de la physique passaient pour le summum de la science. Mais les « lois » des linguistes sont locales, limitées, et très nombreuses. Ils ont raison de chercher des correspondances régulières, bien sûr ; mais on ne voit pas pourquoi il faudrait s'arrêter là, tout essoufflé.

Il faut trouver comment elles agissent, pourquoi elles sont limitées, pourquoi il y en a d'autres. Et surtout, il faut comprendre que ce qui motive les gens qui parlent, comme Saussure l'expliquait très bien (on lui a beaucoup reproché d'être si clair !), ce n'est pas seulement les formes, c'est le sens. Si on laisse le sens de côté, on pourra peut-être grimper dans les arbres, mais on aura caché la forêt.

Vincennes, le 22 mars 2018