

Zeitschrift: Revue de linguistique romane
Herausgeber: Société de Linguistique Romane
Band: 52 (1988)
Heft: 207-208

Artikel: Phonétique et glottochronologie
Autor: Guiter, Henri
DOI: <https://doi.org/10.5169/seals-399833>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 15.03.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

PHONÉTIQUE ET GLOTTOCHRONOLOGIE

Les diverses méthodes glottochronologiques qui ont été proposées, s'appuient toutes sur le lexique. On compte, par exemple, les pourcentages de signifiants d'un même signifié qui n'ont pas même étymon dans chacun des divers couples de langues confrontées.

Il semble que des données du même ordre pourraient être obtenues en établissant les pourcentages de divergences dans les traitements phonétiques.

Le but ultime de la glottochronologie serait de dresser un arbre généalogique du langage avec une date à chaque enfourchure. Mais le premier problème qui se pose est l'élaboration même de l'arbre, en l'adaptant aux éléments fournis par la comparaison des lexiques ou des traitements phonétiques.

Avant d'entrer dans le vif du sujet, il convient de se demander avec quelle précision sont définis les nombres que nous allons mettre en œuvre. Nous pouvons utiliser à cette fin la comparaison des listes lexicales de neuf langues romanes (portugais, espagnol, catalan, gascon, provençal, français, italien, sarde, roumain) et de huit langues germaniques (islandais, norvégien, danois, suédois, anglais, frison, néerlandais, allemand), dont nous avons déjà eu l'occasion de donner les résultats ⁽¹⁾. La divergence primitive d'une langue romane et d'une langue germanique ne peut que remonter à la séparation de l'italique commun et du germanique commun : on devrait donc retrouver le même résultat pour les 72 couples possibles confrontant une langue de chacune des deux familles. Or il n'en est pas rigoureusement ainsi.

(1) Henri Guiter, « Origines européennes et glottochronologie », *Quantitative Linguistics*, vol. 18, Bochum, 1983, 136.

Henri Guiter, « Corrélations de méthodes génético-linguistiques », *Cahiers de l'Institut de Linguistique XI*, Louvain, 1985, 125.

Henri Guiter, « Contraintes glottochronologiques », *Cahiers de l'Institut de Linguistique XII*, Louvain, 1986, 5.

La valeur moyenne du pourcentage de désaccords pour ces 72 mesures est 65 %. Mais l'une des mesures n'arrive qu'à 61 %, alors qu'une autre atteint 70 %. Cependant, sur ces 72 mesures, 58, soit 80 %, sont comprises entre 63 et 67, c'est-à-dire ne diffèrent pas de la valeur moyenne de plus de deux unités de pourcentage en plus ou en moins. (Et il y aurait 68 mesures, soit 95 %, entre les valeurs 62 et 68, c'est-à-dire 65 ± 3). Nous avons donc une probabilité de 80 % pour que deux mesures de la grandeur ne diffèrent pas entre elles de ± 4 , l'une pouvant être à $+ 2$ quand l'autre est à $- 2$, ou inversement.

Cette statistique, rendue possible par le nombre élevé de langues de chacune des deux familles apparentées, nous donne une idée de l'incertitude dont sont frappées les différences lexicales entre deux langues. Il serait, par conséquent, illusoire d'introduire des décimales dans les valeurs des pourcentages : nous nous en tiendrons toujours à des nombres entiers d'unités.

Comment s'expliquent les irrégularités entre les pourcentages de désaccords ? Elles tiennent aux innovations capricieuses des vocabulaires. Les héritiers du *gw*-indo-européen se retrouvent tous à l'initiale des noms désignant la « vache » en roman et en germanique. Par contre, si nous passons au « cheval », nous tombons sur des créations anarchiques, le roman avec les héritiers du bas-latin *caballu* (*cavalo*, *caballo*, *cavall*, *chibau*, *chivau*, *cheval*, *cavallo*, *kabaddu*, *cal*), le scandinave (isl. *hestur*, nor. dan. *hest*, suéd. *häst*), l'anglais (*horse*), le frison (*hyns*), le néerlandais et l'allemand (*paard*, *Pferd*). De même, le nom de la « brebis » repose en roman sur les étymons *ouicula*, *ueruex*, *foeta*, *pecora* ; s'il venait à être remplacé par celui de la « chèvre », ce serait l'unique étymon *capra* qui s'étendrait à tout le domaine. Ces imprévisibles irrégularités n'empêchent pas qu'il ne se dégage une norme d'ensemble ; mais elle est affectée d'inévitables fluctuations aléatoires.

Nous avons pu élaborer une matrice de 253 cases pour comparer les lexiques de 23 langues du groupe *kentum* ⁽²⁾. Mais la perspective de manier les données beaucoup plus lourdes de la phonétique ⁽³⁾ nous rend prudent : nous nous limiterons maintenant à sept langues romanes de

(2) Henri Guiter, « Arborescences linguistiques », *Quantitative Linguistics*, vol. 35, Bochum, 1988, 171.

(3) Enric Guiter, « Afinitats fonètiques a la Romània contínua », *Homenaje a Alonzo Zamora Vicente I*, Madrid, 1988, 135.

la « Romania continue » : portugais, espagnol, catalan, gascon, provençal, français et italien.

Construction d'un arbre lexical

La comparaison des listes lexicales des 21 couples que peuvent constituer les sept langues romanes considérées, nous fournit les valeurs de leurs pourcentages de désaccords, que nous rassemblons dans le tableau synoptique suivant :

Espagnol	7					
Catalan	11	9				
Gascon	18	19	17			
Provençal	22	23	19	10		
Français	26	28	27	17	13	
Italien	16	18	15	17	17	19
	Portugais	Espagnol	Catalan	Gascon	Provençal	Français

Nous remarquons une triade de pourcentages particulièrement faibles entre le portugais, l'espagnol et le catalan. C'est par eux que nous allons commencer l'élaboration de l'arbre (fig. 1).

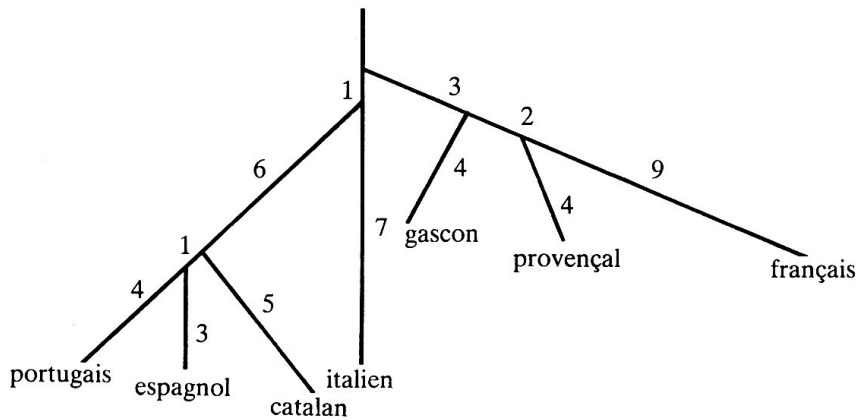


Fig. 1

Les deux rameaux correspondant au portugais et à l'espagnol doivent totaliser la valeur 7. Comment doit être partagée cette valeur ? Le catalan va nous fournir une indication : il doit être plus éloigné de deux points du portugais (11) que de l'espagnol (9). Mais 7 ne peut pas être partagé en deux nombres entiers différant de 2 ; nous nous contenterons donc d'une différence de 1 point, et adopterons 4 pour le rameau portugais et 3 pour le rameau espagnol.

Le catalan dispose donc encore de 6 points ($9 - 3$) à répartir entre son propre rameau et la distance des enfourchures. Pour fixer notre choix, nous allons évaluer la distance moyenne entre les quatre langues non ibéroromanes et le portugais (20,5) ou le catalan (19,5). Le catalan devrait en être plus rapproché que le portugais d'un point. Cette répartition n'est pas possible, car, selon le partage du nombre 6, nous obtenons une différence de 0 ou de 2. Nous optons pour la différence 0, ce qui nous conduit à donner la longueur 5 au rameau catalan, et la longueur 1 à la distance entre les deux enfourchures.

Mesurées sur l'arbre, la distance entre espagnol et catalan est 9 ($3 + 1 + 5$), et celle entre portugais et catalan 10 ($4 + 1 + 5$).

Une autre triade de pourcentages encore faibles, bien que supérieurs aux précédents, se rapporte au gascon, au provençal et au français. Entre le français et le provençal nous avons une distance de 13. Pour la répartir entre les deux rameaux, considérons les distances respectives avec le gascon, 17 et 10. Le français doit en être éloigné de 7 points de plus que le provençal, ce qui suppose que 13 est partagé en 10 et 3. Il resterait 7 points disponibles pour le gascon, ce qui lui permettrait d'être à la distance 17 ($7 + 10$) du français et 10 ($7 + 3$) du provençal.

Toutefois, avant d'adopter ce schéma, évaluons les distances moyennes de chacune de ces trois langues aux quatre langues non gallo-romanes. Nous trouvons 17,75 pour le gascon, 20,25 pour le provençal et 25 pour le français. Le français n'excède que de 4,75 la distance moyenne du provençal ; si nous maintenons entre la longueur des rameaux de ces deux langues une différence de 7 ($10 - 3$), il nous sera difficile de les assembler avec le reste du groupe. Nous allons donc ramener cette différence à 5, en donnant la valeur 9 au rameau du français et 4 à celui du provençal.

La distance moyenne du gascon aux quatre langues non gallo-romanes est plus faible que celle du provençal de 2,5 points. Pour obtenir une différence de 2, il nous suffira de répartir 6 points du gascon en 4 points pour le rameau et 2 entre les deux enfourchures.

Avec les transactions qui se sont ainsi imposées, les distances entre les trois langues de ce groupe deviennent 13 entre le français et le provençal, 15 (au lieu de 17) entre le français et le gascon, 10 entre le provençal et le gascon.

Ce seul écart de 2 points est relativement modeste par rapport à l'incertitude qui peut raisonnablement peser sur les mesures.

Il faut maintenant assembler les deux groupes que nous venons d'organiser. La distance du français au portugais est 26, celle du français à l'italien 19, celle du portugais à l'italien 16. Il n'est pas possible de partager 26 en deux nombres entiers différant de 3 (19 — 16) ; adoptons donc une différence 4, ce qui nous donne 15 pour le français et 11 pour le portugais jusqu'à l'enfourchure centrale.

Nous pourrions faire partir la branche ibéroromane et la branche gallo-romane du même niveau, de part et d'autre de l'axe central. Mais, pour tenir compte des réfections qui nous ont été imposées dans un ensemble linguistique plus étendu, nous ferons diverger le gallo-roman une unité de pourcentage avant l'ibéroroman. Ceci est en accord avec le fait que l'invasion franque avait précocement isolé la Gaule du Nord, de l'Italie ; elle doit correspondre à une enfourchure antérieure.

Reste à apprécier la longueur du rameau italien. Pour s'accommoder au mieux avec les segments déjà mis en place, c'est la longueur 7 qui semble préférable.

Nous pouvons dresser à nouveau le tableau des distances entre les langues, en utilisant maintenant le données de l'arbre :

Espagnol	7					
Catalan	10	9				
Gascon	19	18	19			
Provençal	21	20	21	10		
Français	26	25	26	15	13	
Italien	18	17	18	15	17	22
	Portugais	Espagnol	Catalan	Gascon	Provençal	Français

La comparaison des deux tableaux nous montre entre leurs termes les écarts suivants : deux fois + 3, deux fois + 2, cinq fois + 1, six fois 0, une fois — 1, trois fois — 2 et deux fois — 3.

Aucun écart n'est donc supérieur à 3 ; ce résultat est meilleur que ce que nous laissaient attendre les prévisions d'incertitude. En outre 17 mesures, soit 80 %, ne s'écartent pas de plus de deux unités.

Il est donc relativement facile de construire un arbre lexical, et les mesures de différences lexicales se prêtent bien à une présentation sous la forme d'arbre.

Notre article de la RLiR « Datation de divergences romanes » (4) donne la liste des signifiants français utilisés. Les exemples de datations qui illustrent cet article, permettent de garnir cinq des six encheures rencontrées : 1000 pour portugais-espagnol, 750 pour portugais-catalan, 525 pour ibéro-roman-italien, 375 pour français-italien, 675 pour français-provençal. Le seul repère que nous n'avions pas envisagé comme exemple, est celui de français-gascon, qui se situe vers 475.

Construction d'un arbre phonétique

Les matériaux mis en œuvre sont beaucoup plus volumineux que dans l'étude lexicale. Même en nous limitant à la comparaison de sept langues romanes, nous avons pris un certain nombre de précautions. D'abord, nous nous bornerons à distinguer huit timbres vocaliques : *i, e, a, o, u, ö, ü, ë* (voyelle neutre). L'usage de *r* apical ou *r* uvulaire, de *s* apico-alvéolaire ou *s* dorso-alvéolaire, n'ayant pas valeur phonologique dans les langues considérées, nous ne tenons pas compte de ces différences. Par ailleurs, nous laissons de côté les phonèmes romans qui ont même traitement dans les sept langues choisies, par exemple *p* initial ou *a* tonique entravé (v.g. *parte*), puisqu'ils n'apportent aucun facteur de différenciation.

En dépit de ces simplifications, il nous restera à étudier le sort de 50 phonèmes (ou groupes de phonèmes) vocaliques et 100 consonantiques. Nous en donnons la liste en note (5).

(4) Henri Guiter, « Datation de divergences romanes », *Revue de Linguistique Romane* (48), Strasbourg, 1984, 269.

(5) -á-, á + y, á + w, á + w + l, a + w, á + l, á + ne, á + n, (k +) a, -a, -é-, é entrav., é (+ y), é (+ w), é (+ ll), é + n, -ê-, ê (+ palat.), ê (+ y), ê (+ l), ê (+ n final), ê (+ n + cons.), ê (+ n + vel.), (k +) é, e-, -e, i (+ l), i (+ w), i (+ n fin.), -ô-, ô entrav., ô (+ y), ô (+ w), ô (+ l + cons.), ô (+ l + y), ô (+ n + cons.), -ô-, ô entrav., ô (+ y), ô (+ l + t), ô (+ n + cons.), ô (+ n + y), ô (+ n + vel.), ô (+ ne), o-, -o, u, u + y, u + l, u + n fin.
pl-, -p-, -pp-, -pr-, -pl-, -p (+ t)-, -py-, -p(u), -t-, -tr-, ty-, 'ty-, -ty', -ty, -t'k, -te, k (+ a)-, k (+ e, i)-, -k (+ o, u)-, -k (+ e, i)-, -ke, -k (+ y)-, -kye, kl-, -k'l-, kw (+ a)-, kw (+ e, i)-, -kw-, -kt-, -ks-, -k(u), bl-, -b-, -bb-, -br-, -b'l-, -b (+ t), -by-, -by (+ u), -d-, -dr-, dy-, -dy-, -dy (u), -de, g (+ a)-, g (+ e, i), -g (+ o, u)-, -g (+ á)-, -g (+ e, i)-, -g (+ e, i), gl-, -g'l-, -gr-, gw-, -gw-, f-, fl-, -f-, s-, -s-, 'sy-, -sy', -ssy-, -sk (+ e, i), -sty-, -s, -my-, -mn-, -mb-, -n-, -nn-, -ny-, -ng'l-, -nd-, -ns-, -ne, l-, -l-, -ll (a), -ll (u), -l (+ y)-, -l (+ sonor.), -l (+ sourd.), -l, r-, -rr-, -aria, -ariu, -rs-, -(e)re, y (+ o, u)-, y (+ á)-, y (+ a, e)-, -y-, -y (u), w-, -w-, -wy-, -we.

On aura une idée du travail par les résultats des deux premières lignes du tableau des voyelles et de celui des consonnes :

Le *a* tonique libre garde son timbre partout, sauf en français où il devient *e*.

Le *a* tonique suivi d'un groupe *-kt-* prend le timbre *e*, sauf en provençal et en italien.

Le groupe *pl* initial devient *ch* en portugais, *ly* en espagnol, *py* en italien et *pl* ailleurs.

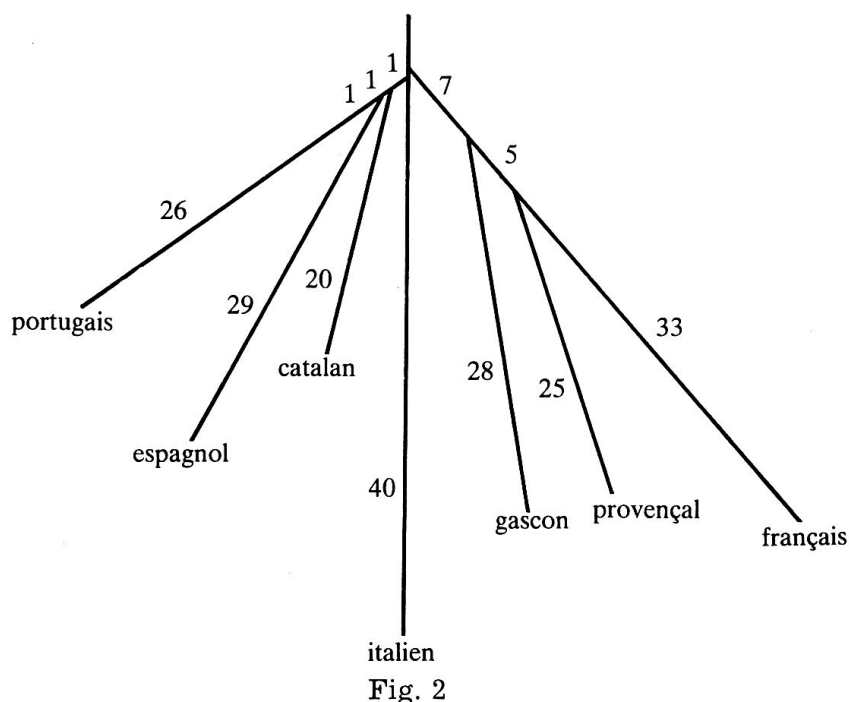
Le *p* intervocalique subsiste en italien, devient *v* en français, et *b* partout ailleurs.

Les 3.150 comparaisons de résultats ($21 \times [50 + 100]$) nous permettent de dresser le tableau des pourcentages de désaccords vocaliques, celui des pourcentages de désaccords consonantiques, et celui des moyennes de ces deux pourcentages.

C'est ce dernier tableau qui va nous intéresser.

Espagnol	55					
Catalan	48	50				
Gascon	66	68	54			
Provençal	60	79	61	47		
Français	69	83	72	77	57	
Italien	68	59	58	78	81	86
	Portugais	Espagnol	Catalan	Gascon	Provençal	Français

Nous constatons que les distances phonétiques sont beaucoup plus grandes que les distances lexicales : la valeur moyenne de celles-ci était 17,5 ; elle atteint maintenant 65,6, soit près de quatre fois plus. Le maniement de ces nombres plus importants va être plus compliqué. Nous allons essayer de calquer notre démarche sur celle que nous avons suivie précédemment (Fig. 2).



La distance de l'espagnol et du portugais est 55 ; pour la répartir entre les deux rameaux, nous notons que la distance du catalan à l'espagnol dépasse de deux unités celle du catalan au portugais. La partition de 55 entre deux nombres entiers excluant la différence 2 entre ceux-ci, nous adopterons une différence 3 avec un rameau de 26 pour le portugais et 29 pour l'espagnol.

Il reste 21 (50 — 29) unités à utiliser pour le catalan. La distance moyenne du portugais aux 4 langues non ibéro-romanes est 65,75, celle du catalan 61,25. Cette différence, inférieure à 5, nous amènerait à faire partir le rameau catalan de l'enfourchure espagnol-portugais, ce qui est peu satisfaisant, puisque historiquement la séparation catalan-portugais a précédé la séparation espagnol-portugais. Nous transigerons en situant l'enfourchure catalan-portugais très légèrement au-dessus de l'enfourchure portugais-espagnol, par exemple à un point de distance, et il nous reste 20 points pour la longueur du rameau catalan.

Mesurées sur l'arbre, la distance espagnol-portugais est 55 (pour 55), la distance espagnol-catalan 50 (pour 50), et la distance catalan-portugais 47 (pour 48).

Les choses seront moins aisées du côté gallo-roman. Le provençal et le français ont une distance de 58 points à répartir entre leurs rameaux respectifs.

Pour cette répartition, nous songeons à nous orienter avec les distances provençal-gascon (47) et français-gascon (77). Ces distances présentent entre elles une différence de 30 points ($77 - 47$) ; il faudrait donc que le rameau du français eût une longueur de 44, et celui du provençal, de 14. Il resterait pour le gascon 33 points à répartir entre la longueur du rameau et la distance des deux enfourchures ; nous retrouverions ainsi les distances français-provençal 58 ($44 + 14$), français-gascon 77 ($44 + 33$), provençal-gascon 47 ($14 + 33$).

Mais la prudence nous invite à calculer la distance moyenne de chaque langue gallo-romane aux quatre langues non gallo-romanes. Nous trouvons ainsi 66,5 pour le gascon, 70,25 pour le provençal, 77,5 pour le français. Les distances du français et du provençal aux quatre langues non gallo-romanes ne diffèrent que d'un peu plus de 7 ; il sera impossible de construire un arbre si nous acceptons entre ces deux langues une distance de 30, comme nous le suggérait la considération du seul groupe gallo-roman.

Nous allons donc nous engager dans des transactions qui bouleversent la belle ordonnance d'abord envisagée pour ce groupe. Il faut limiter à 8 points la distance entre le provençal et le français ; ceci nous conduit à partager 58 en accordant 33 au rameau français et 25 au rameau provençal. Nous devons maintenant attribuer au gascon un nombre de points supplémentaires qui nous permette de partager les dégâts entre le français et le provençal ; si nous adoptons la valeur 33, la distance du gascon au français sera de 66 ($33 + 33$) au lieu de 77, celle du gascon au provençal de 58 ($33 + 25$) au lieu de 47 : 11 points de moins dans un cas, de trop dans l'autre. Reste à fixer la position de l'enfourchure du gascon ; elle devrait nous permettre d'avoir par rapport aux langues non gallo-romanes un écart de 11 ($77,5 - 66,5$) entre gascon et français. Il n'est pas possible d'obtenir 11, il faut opter entre 10 et 12 ; si nous adoptons 10, les deux enfourchures devront être à une distance de 5 unités et la longueur du rameau gascon est 28, alors que la distance du français à l'enfourchure est 38 ($33 + 5$).

Il faut maintenant connecter entre elles les branches gallo- et ibéro-romanes. La longueur moyenne des segments de l'ibéro-roman à partir de leur enfourchure supérieure est 26, celle des segments du gallo-roman, 33. La moyenne des neuf distances entre langues gallo- et ibéro-romanes est 68. Il faut donc des segments intermédiaires d'une longueur de $68 - (26 + 33) = 9$.

Comment répartir ces 9 unités à partir de l'enfourchure commune ? La distance moyenne de l'ibéro-roman à l'italien (61,66) est beaucoup plus petite que celle du gallo-roman (81,66). Dans le partage des 9 unités, il faudra donc réduire au minimum, c'est-à-dire 1, la part de l'ibéro-roman, et attribuer 8 au gallo-roman, ces derniers points pouvant être répartis en un segment de 1 et un segment de 7.

Reste à déterminer la longueur du rameau italien. La distance du portugais à l'italien est 68, et la distance du portugais à l'enfourchure 28 (26 + 1 + 1) ; il resterait donc 40 pour l'italien. La distance du français à l'italien est 86, et la distance du français à l'enfourchure commune est 46 (33 + 5 + 7 + 1) ; il reste encore 40 pour le rameau italien, valeur que nous adoptons.

Nous pouvons maintenant dresser le tableau des distances entre les langues, telles qu'elles se mesurent sur l'arbre :

Espagnol	55					
Catalan	47	50				
Gascon	64	67	57			
Provençal	66	69	59	58		
Français	74	77	67	66	58	
Italien	68	71	61	76	78	86
	Portugais	Espagnol	Catalan	Gascon	Provençal	Français

Si nous comparons les deux tableaux de nombres, nous constatons que les transactions auxquelles nous avons été contraint, ont provoqué quelques écarts importants. Nous trouvons une fois — 12, une fois — 11, une fois — 6, une fois — 5, deux fois — 3, cinq fois 0, 2 fois 1, 3 fois 2, une fois 3, une fois 5, une fois 6, une fois 10 et une fois 11. Nous sommes loin de la cohérence de résultats que nous donnait la recherche lexicale.

Avouons d'ailleurs que nous ne serions jamais arrivé à construire un arbre rappelant vaguement l'aspect de l'arbre lexical, si nous ne nous étions laissé guider par l'ordonnance de celui-ci. Dès le départ, il était anormal de construire la fourche portugais-espagnol (distance 55), puisque la plus petite distance se rencontrait entre le portugais et le catalan (47). Ces deux distances impliquaient que la séparation portugais-espagnol avait précédé la séparation portugais-catalan, ce qui historiquement est une absurdité.

Par ailleurs, nous avons dû jongler avec les nombres pour tenter d'obtenir un tableau qui ne fasse pas apparaître des écarts monstrueux

entre les valeurs résultant des mesures directes, et celles comptées sur l'arbre.

On doit admettre que l'évolution phonétique est beaucoup plus capricieuse que les changements lexicaux. Son irrégularité rendrait même impossible la construction d'un arbre, si le nombre des langues mises en jeu ne restait pas très réduit, comme nous nous le sommes imposé présentement.

En conséquence, la réponse au problème posé au début de cet article est très simple : les changements phonétiques ne peuvent pas se substituer aux changements lexicaux pour l'élaboration d'une méthode glottochronologique.

Montpellier.

Henri GUITER

