

Zeitschrift: Revue de linguistique romane
Herausgeber: Société de Linguistique Romane
Band: 30 (1966)
Heft: 117-118

Artikel: Quelques paramètres caractéristiques des systèmes vocaliques
Autor: Guiter, Henri
DOI: <https://doi.org/10.5169/seals-399368>

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: 14.03.2025

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

QUELQUES PARAMÈTRES CARACTÉRISTIQUES DES SYSTÈMES VOCALIQUES

Même lorsqu'il a acquis une syntaxe française absolument correcte, certaines singularités phonétiques permettent de préciser l'origine géographique d'un sujet s'exprimant en français : un Roussillonnais, par exemple, ne connaît qu'un seul type de *a*, de *o* ou de *e*, articule des *l* et des *s* « *pingues* » avec la langue fortement concave, etc. Au fond, l'« accent » est une manifestation substratique, une sorte de revanche de la langue atavique sur la langue nouvellement acquise ; il est des habitudes qui ne se perdent pas facilement, si tant est qu'elles se perdent jamais.

Cette revanche s'est produite maintes fois dans l'histoire du langage, et souvent ses manifestations diverses ont décelé une cohérence impressionnante. En français, lorsque nous constatons que le *u* de lat. *lūna* a abouti au *ü* de *lune*, le *o* de lat. *mōla* ou *flōre* au *ö* de *meule* et *fleur*, le *a* de *patre* au *e* de *père*, le *k* de *cattu* au *ç* de *chat*, etc., nous ne pouvons manquer d'être frappés par le fait que tout le système de la langue a glissé de l'arrière vers l'avant de la bouche. Des faits analogues s'étaient produits en grec, depuis l'attique jusqu'à la koiné byzantine, lorsque le *ā* était arrivé à *ē*, le *u* à *ü* puis à *i*, les diphtongues *ei* et *oi* à *i*, la diphtongue *ai* à *e*, le *ē* à *i*, avec, comme maigre compensation, le passage de la diphtongue *ou* à *u*.

C'est un phénomène d'un autre ordre qui se produit en sanskrit. Lorsque le *e* et le *o* de l'indo-européen commun vont rejoindre le *a*, nous assistons à un rassemblement des points d'articulation vocaliques vers le milieu du palais. Le point d'articulation moyen ne s'est pas déplacé vers l'avant ou l'arrière, mais la densité des phonèmes a crû au centre de la bouche.

Dans les mutations vocaliques du français, le caractère plus ou moins labial des voyelles s'est conservé ; en grec ou en sanskrit, au contraire, le passage de *u* à *i* ou de *o* à *a* correspond à une délabialisation.

Il nous a semblé intéressant de « chiffrer » ces tendances diverses ; mais nos tentatives n'ont pas été sans se heurter à de sérieuses difficultés, qui nous ont conduit à des retouches successives et à une réduction progressive de nos ambitions initiales.

En premier lieu, nous avons renoncé à l'étude du système consonnantique. Elle n'est certes pas impossible, mais elle est beaucoup plus difficile et plus méticuleuse que celle du système vocalique. Des consonnes produisant des impressions acoustiques voisines peuvent avoir des points d'articulation passablement différents, et inversement. Nous avons eu peur de nous noyer dans des nuances, qui auraient enlevé de sa netteté à ce premier essai.

Ensuite, nous avons renoncé à l'étude de toutes les voyelles, pratiquée tout d'abord, pour nous limiter aux voyelles toniques. Cet abandon de terrain vise à éliminer les difficultés tenant aux voyelles diminuées, voyelles neutres ou voyelles caduques, comme les finales de l'anglais ou de l'allemand, les *e* dits muets du français, les voyelles neutres du catalan et du portugais, etc. D'ailleurs, en étudiant des langues comme le latin, le castillan, l'italien ou le finnois, qui conservent à l'atone des timbres vocaliques nets, nous avons pu constater la coïncidence des statistiques portant sur toutes les voyelles ou seulement sur les voyelles toniques ; le système total est simplement un peu plus avancé que le système tonique, mais les positions relatives sont les mêmes. La restriction que nous avons introduite ainsi est donc plus apparente que réelle. Bien entendu, dans une langue sans accent tonique comme le basque, nous avons tenu compte de toutes les voyelles ; ceci modifie peu, ainsi que nous venons de le dire, sa position relative par rapport au castillan ou à l'italien.

Nous avons donc classé les voyelles toniques, labiales ou non, par points d'articulation ; dans un but de clarté, nous avons résolu de limiter à huit le nombre des groupes : *i, e, a, o, u, ü, ö, í*. Autrement dit, nous ne distinguons pas entre *e* ouvert, fermé, moyen, nasal, etc., il nous suffit de considérer un ensemble de phonèmes s'articulant sensiblement sur le même domaine du palais. Les frontières de ces divers domaines sont peut-être arbitraires ; nous nous sommes fié à notre oreille : par exemple la graphie *ä* de l'allemand nous semble correspondre à un *e*, mais la même graphie du finnois à un *a*.

CHOIX DES PARAMÈTRES.

Un premier paramètre doit caractériser, à notre sens, la position d'ensemble, plus ou moins avancée, du système vocalique dans la bouche ; ce sera l'abscisse x , du centre de gravité du vocalisme tonique.

Pour déterminer numériquement la valeur de cette abscisse x , il importait de choisir une origine et une échelle. Nous avons pris comme origine ($X_A = 0$) le point d'articulation moyen du a ; l'échelle a été déterminée en fixant à 100 ($X_E = 100$) l'abscisse du point d'articulation moyen du e (et \ddot{o}), à 200 ($X_I = 200$) celle relative à i (et \ddot{u}), à -100 ($X_O = -100$) celle relative à o , à -200 ($X_U = -200$) celle relative à u (et i).

Chacun de ces points d'articulation moyens sera affecté d'un « poids » exprimé par le pourcentage de fréquence du phonème (ou des phonèmes) correspondant. Dès lors, nous déterminons l'abscisse X du centre de gravité du système en appliquant le théorème des moments : « Le moment par rapport à un point de la résultante d'un système de vecteurs est égal à la somme des moments de chacun des vecteurs composants par rapport au même point ». Il est particulièrement simple de prendre comme point de référence le point origine. Dès lors, si nous désignons par le symbole [] le pourcentage :

$$X \times 100 = X_E ([e] + [\ddot{o}]) + X_I ([i] + [\ddot{u}]) + X_O ([o]) + X_U ([u] + [i])$$

Un deuxième paramètre doit être lié à la dispersion ou au rassemblement relatifs des points d'articulation à l'intérieur de la bouche. Une langue dans laquelle le pourcentage de a est très élevé, et les pourcentages de i et de u très faibles, présentera une grande densité de réalisations au centre de la bouche ; et une grande dispersion dans le cas contraire. Pour fixer la valeur de ce paramètre y , nous avons d'abord songé à faire le rapport $[a]/([i] + [u] + [\ddot{u}] + [i])$. Mais cette représentation ne permettait pas de suivre simplement les variations de y en fonction des pourcentages. Après quelques tâtonnements, nous avons trouvé plus commode de choisir la correspondance :

$$y = [a] - ([i] + [u] + [\ddot{u}] + [i])$$

Caractériser chaque langue par ces deux paramètres x et y , nous

permet déjà de répartir sur un plan les points représentatifs des diverses langues, et cette première analyse nous a paru fort intéressante.

Mais pour répartir les points représentatifs, non pas seulement dans un plan mais dans l'espace, nous avons jugé convenable de choisir un troisième paramètre ζ en relation avec la plus ou moins grande labialisation des systèmes vocaliques :

$$\zeta = [o] + [u] + [\ddot{o}] + [\ddot{u}]$$

Ce troisième paramètre ζ est moins significatif, car il disperse beaucoup moins les points représentatifs que les deux premiers x et y .

PRÉPARATION DE LA MATIÈRE D'ÉTUDE,

Nous avons donc besoin de déterminer pour chaque langue étudiée, le pourcentage de fréquence des diverses voyelles toniques.

Ce comput doit porter sur un texte suffisamment long pour que la loi des grands nombres puisse jouer, et éviter les textes poétiques, où des effets plus ou moins artistiques peuvent être attendus du jeu des diverses voyelles. Si pour étudier le français on se limitait au fragment :

Rapide cliquetis du kirghise qui fuit...

ou, au contraire, à cet autre :

La pourpre du velours où coulent vos bijoux...

il est manifeste qu'on obtiendrait des résultats absolument contradictoires, parce que l'on se serait simultanément heurté aux deux écueils signalés plus haut.

Nous avons pu constater qu'en choisissant un texte suffisamment long, nous obtenions pour une même langue des résultats suffisamment constants, même si les dates des textes différaient autant que de l'époque classique à l'époque moderne.

En français, par exemple, la *Préface des fables* de la Fontaine nous a donné les résultats :

$$[a] = 28 \quad [e] = 27 \quad [i] = 16 \quad [o] = 12 \quad [u] = 5 \quad [\ddot{o}] = 7 \quad [\ddot{u}] = 5$$

$$\text{d'où} \quad x = 54 \quad y = 2 \quad \zeta = 29$$

et l'*Avenir de l'intelligence* de Charles Maurras :

$$[a] = 21 \quad [e] = 29 \quad [i] = 16 \quad [o] = 11 \quad [u] = 10 \quad [\ddot{o}] = 4 \quad [\ddot{u}] = 8$$

$$\text{d'où} \quad x = 50 \quad y = -13 \quad z = 33$$

En castillan les résultats fournis par le premier chapitre du *Don Quijote* de Cervantes sont les suivants :

$$[a] = 30 \quad [e] = 23 \quad [i] = 17 \quad [o] = 20 \quad [u] = 10$$

$$\text{d'où} \quad x = 17 \quad y = 3 \quad z = 30$$

et ceux de *La voz de la letra* d'Alonso Zamora Vicente

$$[a] = 31 \quad [e] = 25 \quad [i] = 15 \quad [o] = 19 \quad [u] = 10$$

$$\text{d'où} \quad x = 16 \quad y = 6 \quad z = 29$$

Devant la fidélité relative de ces données, nous n'avons pas hésité à moyenniser celles de textes classiques et celles de textes modernes pour obtenir les paramètres définitifs de chaque langue.

Nous n'avons pas recherché systématiquement des textes de « bonne langue » : des articles de journaux ou des guides de conversation usuelle, conviennent parfaitement pour le but que nous nous proposons.

Des enquêtes sur des textes de longueur suffisante ont été effectuées pour quarante-six langues vivantes ou mortes. Pour soixante-six autres langues nous n'avons disposé que de textes très courts, soit parce qu'on n'en connaît pas de longs passages (c'est le cas, par exemple, des vieilles langues ibériques), soit parce qu'il nous a été impossible — même en écrivant à des spécialistes — de nous procurer un fragment assez considérable. Notre curiosité nous a alors poussé à utiliser les quelques lignes que nous possédions, par exemple celles fournies comme échantillons par les « *Langues du Monde* ». Bien entendu, la valeur de ces derniers témoignages est fort légère, et ils sont entachés d'une grande indétermination. Ils sont dans la situation d'analyses chimiques effectuées sur des prises d'essai trop faibles. Nous souhaitons de tout coeur que des chercheurs mieux placés que nous, puissent faire dans ces cas des enquêtes plus vastes et plus solides, et, pour ne pas alourdir exagérément cet article, nous nous abstenons de donner les résultats qui leur sont relatifs, en dépit du grand intérêt de certains d'entre eux. Il est par

exemple frappant que le mandchou et le mongol voisinent avec le turc ; le samoyède et le guiliak, avec le lapon, etc..

Les axes que nous avons été amenés à tracer sur la fig. 2, l'ont été en tenant compte des cent douze points d'ensemble, qui donnaient, bien entendu, une densité beaucoup plus élevée que les quarante-six points auxquels nous nous limitons ici.

Nous hésitons à donner la longue liste des textes dépouillés pour la classification de ces quarante-six langues, étant donné l'intérêt mineur de cette énumération ; cependant, si plusieurs lecteurs en manifestaient le désir, nous pourrions la publier ultérieurement en appendice.

Remercions ici les personnes de bonne volonté qui ont bien voulu collaborer à ces décomptes fastidieux ;

M^{lle} Gusta Sigurd, assistante de philologie germanique à Montpellier, et islandaise d'origine, nous a fourni les éléments relatifs aux quatre langues scandinaves ; nous avons aussi tenu compte de ceux qu'elle nous a communiqués pour l'anglais et l'allemand, en les associant à nos propres enquêtes.

M. M. Yves Coutant, de Courtrai, et Paul Van Schoors, d'Ostende, nous ont transmis leurs résultats relatifs au néerlandais-flamand, que nous avons conjugués avec les nôtres. Nous avons bien séjourné dans tous les pays germaniques (sauf l'Islande), mais il nous a été précieux de bénéficier de l'aide de spécialistes.

M^{lle} Flora Aurima, venue de Tahiti à Montpellier pour la poursuite de ses études, nous a donné les décomptes concernant sa langue maternelle, le tahitien.

Aux principales langues romanes littéraires (portugais, castillan, catalan, gascon, languedocien, provençal, français, italien, roumain) nous avons joint le romanche et le savoyard (le terme de « franco-provençal » nous semble mal choisi, à moins que l'on ne se décide à baptiser le castillan « catalano-portugais », ou le français « occitano-flamand »...), et aussi deux dialectes qui nous sont particulièrement proches géographiquement, le roussillonnais (catalan du Nord-Est), et le languedocien maritime (parlé de Leucate à Aigues-Mortes ; cf. Ronjat, *Gram. ist.*, I, 129) baptisé par souci de concision « narbonnais ». Le roussillonnais diffère du catalan central par la fermeture en *u* de tous les *o* ; les *o* entravés par une nasale s'y étant fermés, *ponte* et *punctu* y sont devenus homophones, *punt*. Le narbonnais diffère du languedocien commun par la palatalisation du *ū* en *ö* et non en *ü*.

Voici le tableau de nos résultats, pourcentages et paramètres :

	[a]	[e]	[i]	[o]	[u]	[ö]	[ü]	[ɨ]	x	y	z	Groupe	$\frac{60x-24y-10z}{10}$
Allemand.....	41	11	25	8	7	5	3		50	6	23	A	2626
Anglais.....	20	31	21	12	12	4			41	-13	28	A	2492
Basque.....	36	21	17	15	11				18	8	26	B	632
Breton.....	30	40	7	11	3	3	6		52	14	23	A	2554
Castillan.....	31	27	14	24	4				23	3	28	B	1028
Catalan.....	29	31	14	17	9				24	6	26	B	1036
Chinois.....	13	31	23	26	7				37	-17	33	A	2298
Cornique.....	20	29	15	16	1	19			60	4	36	A	3144
Danois.....	16	37	14	18	6	5	4		48	-8	33	A	2742
Finnois.....	29	18	17	20	10	2	4		22	-2	36	B	1008
Français.....	25	29	16	11	8	5	6		51	-5	30	A	2280
Gaélique (Erse)...	36	20	17	10	8	9			37	11	27	B	1684
Gallois.....	28	18	10	11	5	23	5		50	8	44	A	2364
Gascon.....	30	29	12	9	14		6		28	-2	29	B	1438
Gothique.....	46	11	28	5	10				42	8	15	A	2174
Grec (commun) ..	22	32	9	34	3				10	10	37	C	-10
Grec (attique)....	22	32	9	34			3		22	10	37	B	710
Grec (byzantin) ..	15	24	29	21	8		3		51	-25	32	A	3340
Guarani.....	28	27	7	21	12			5	-14	4	33	C	-1266
Indo-européen....	25	29	9	34	3				7	13	37	C	-262
Irlandais.....	36	30	19	10	3	2			54	14	15	A	2754
Islandais.....	27	38	10	14	3	8			46	14	25	A	2174
Italien.....	24	29	15	24	8				19	1	32	B	796
Japonais.....	34	12	22	21	11				13	1	32	C	416
Languedocien....	35	27	12	8	13		5		27	5	26	B	1240
Lapon.....	38	29	5	17	11	0			0	22	26	C	-788
Latin.....	25	24	20	18	13				20	-8	31	B	1086
Narbonnais.....	35	27	12	8	13	5			22	10	26	B	920
Néerlandais.....	24	31	17	19	3	3	3		49	1	28	A	2636
Norvégien.....	22	31	15	19	2	8	3		52	2	32	A	2752
Ombrien.....	29	37	10		24				9	-5	24	C	420
Osque.....	29	26	19		26				12	-16	26	B	844
Polonais.....	27	20	19	20	8			6	10	-6	28	C	464
Portugais.....	29	33	13	17	8				26	8	25	B	1118
Provençal.....	29	35	11	11	10		4		34	4	25	B	1694
Romanche.....	30	36	15	5	14				33	1	19	B	1766
Roumain.....	29	24	17	6	17			7	4	-12	23	C	298
Roussillonnais....	29	31	14	7	19				14	-4	26	B	676
Russe.....	34	18	7	22	11			8	-28	8	33	C	-2206
Sanskrit.....	67	6	24	9	4				17	49	13	C	-286
Savoyard.....	29	31	9	24	4	1	2		22	14	31	B	674
Serbe.....	31	27	20	13	9				36	2	22	A	1892
Suédois.....	20	28	12	24	2	12	1		38	5	39	A	1770
Tahitien.....	33	21	17	20	9				17	7	29	B	662
Tchèque.....	31	22	9	24	6			8	-12	8	30	C	-1216
Turc.....	26	20	25	7	8	2	5	7	45	-19	22	A	2936

Précisons que les données relatives au grec (commun, attique et byzantin) et à l'indo-européen reposent sur l'examen des mêmes textes (Démosthène, Platon, et Xénophon). Pour le grec commun υ a la valeur u ; pour le grec attique, \ddot{u} ; pour le byzantin α prend la valeur e ; ϵ , η , \omicron , la valeur i ; ω , la valeur u . Pour l'indo-européen, nous sommes partis du grec commun, et avons restitué au timbre a le contingent des anciens α longs passés à η ; pour déterminer ce contingent, nous avons tenu compte des \bar{a} conservés en latin.

Nous constatons que la majorité des langues a un point d'articulation moyen se situant entre 0 et 50, c'est-à-dire entre a et le milieu du trajet de a à e . Il existe cependant des langues articulées en arrière de a , et jusqu'à mi chemin de a et o ($x = -50$), et aussi des langues articulées beaucoup plus en avant, jusqu'au voisinage du point d'articulation de e ($x = 80$).

ÉTUDE DE LA DISPERSION DES LANGUES ROMANES DANS UN DIAGRAMME EN XY.

Bien entendu, nous accorderons un intérêt particulier aux langues romanes, dont nous avons étudié treize échantillons différents. (Le sarde n'entre pas en ligne ici, car les fragments que nous en possédions dans *La lingua sarda* de M. L. Wagner étaient trop courts.)

Nous voyons apparaître de façon frappante la division ternaire en roumain, Romania continue et français.

Si nous nous représentons la répartition spatiale projetée sur les deux plans de la figure 2, nous voyons le roumain à gauche et en bas ($x = 4$, $y = -12$, $z = 23$), les langues de la Romania continue au centre, dans un parallépipède de coordonnées moyennes $x = 26$, $y = 4$, $z = 26$ et d'arêtes de longueurs 18 en x , 20 en y , 13 en z , le français à droite et en haut ($x = 51$, $y = -5$, $z = 30$).

A la gauche et en avant des langues romanes groupées se situe le latin ($x = 20$, $y = -8$, $z = 31$), ainsi que ses proches parents italiques, l'osque et l'ombrien.

Nous nous sommes posé la question de savoir si les paramètres des diverses langues romanes pouvaient se déduire des paramètres du latin, en tenant compte des diverses modifications phonétiques subies par les voyelles toniques du latin. Remarquons bien qu'il ne s'agit pas là d'une étude diachronique, mais de la comparaison de deux systèmes, chacun dans sa perspective synchronique.

Nous constatons que le *o* libre du latin, long ou bref, est représenté par un *ō* français (meule, fleur); peu nous importe que l'on passe d'un terme à l'autre par toute une suite de diphtongues: nous laissons de côté ces équilibres instables ou métastables intermédiaires pour ne nous intéresser qu'aux équilibres stables du début et de la fin. Le paramètre *x* a varié de — 100 à 100 pour un pourcentage déterminé de phonèmes; le paramètre *y* est resté le même puisque *o* et *ō* ont des points d'articulation également proches de *a*; *z* n'a pas changé non plus, les deux phonèmes s'accompagnant d'arrondissement labial.

Lorsque le *ū* latin est remplacé par *ü*, même invariabilité de *y* et de *z*, mais *x* varie de — 200 à 200 pour les phonèmes intéressés par le changement; malgré leur fréquence modeste, l'amplitude du bond effectué entraîne une augmentation sensible de l'abscisse du centre de gravité.

Lorsque le *a* libre passe à *e*, le phonème n'avance son point d'articulation que de 0 à 100; en même temps *y* diminue du pourcentage correspondant, et il n'y a pas de modification de *z*.

Nous avons recherché, en nous limitant à un diagramme plan *x y* (fig. 1), quelles variations du point représentatif entraînait chaque grande modification affectant le vocalisme tonique du latin, et la position, résultant de la somme de ces modifications, du point représentatif de chaque langue romane. Ce point « calculé » à partir du latin et des changements intervenus, coïncide-il avec le point « mesuré » directement sur un texte de la langue correspondante ?

Empressons-nous de remarquer qu'une coïncidence rigoureuse est hautement improbable. L'analogie ou les fantaisies lexicales, les emprunts, se sont chargés de modifier les contingents des divers phonèmes d'une langue à l'autre.

Pourquoi l'italien a-t-il conservé le classique *dīctū*, alors que d'autres romans ont adopté *dīctū*? Pourquoi le castillan a-t-il substitué *nōce* à *nūce*? Le décompte rigoureux de tous ces caprices est pratiquement irréalisable.

Nous pouvons simplement espérer que les anomalies soient assez fortement minoritaires, pour que les points « calculés » se situent dans le voisinage des points « mesurés »; que les dispersions des deux séries de points présentent des répartitions comparables.

Les résultats ont dépassé nos espérances. Pour le roumain, l'italien, le portugais, le provençal, le français, la concordance est excellente. Encore bonne pour le castillan, elle l'est moins pour le catalan, le languedocien et le gascon.

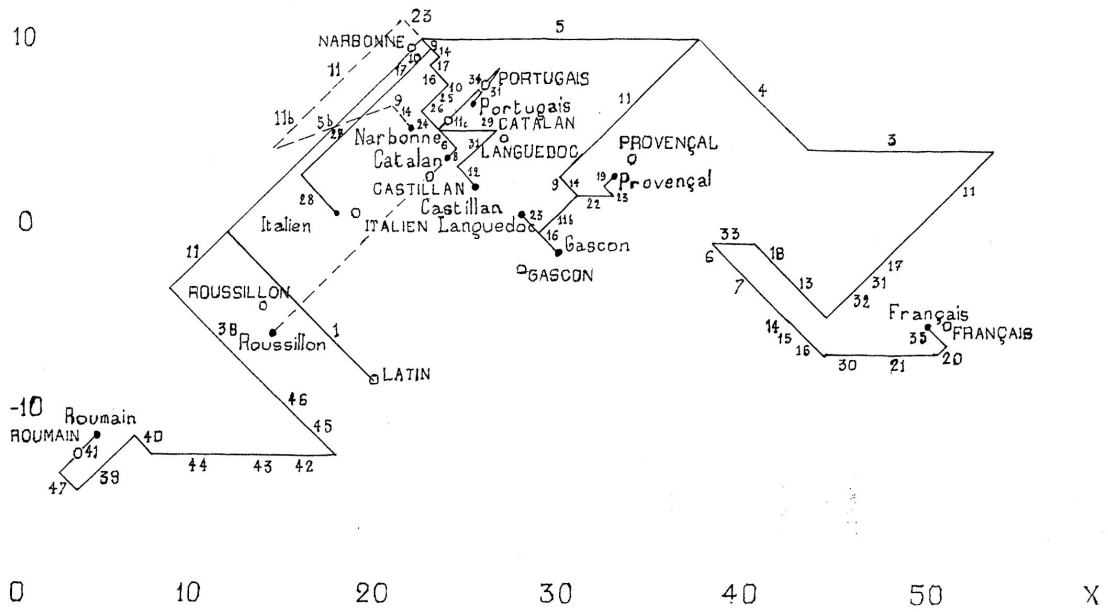
		PORTUGAIS	CASTILLAN	CATALAN	ROUSSILLON	GASCON	LANGUEDOC	NARBONNAIS	PROVENÇAL	FRANÇAIS	ITALIEN	ROUMAIN
1	i > e	↔↑	↔↑	↔↑	↔↑	↔↑	↔↑	↔↑	↔↑	↔↑	↔↑	↔↑
2	û > o	↑↔	↑↔	↑↔	↑↔	↑↔	↑↔	↑↔	↑↔	↑↔	↑↔	
3	o - > ô											
4	a - > e									↓↔		
5	û > ü					→	→		→	→		
5 b	û > ô							↑↔				
6	e + y > i			↓↔	↓↔					↓↔		
7	o + y > wi									↓↔		
8	o + y > u			↔↓	↔↓							
9	e (+a) > i	↓↔	↓↔	↔↓	↔↓	↔↓	↔↓	↔↓	↔↓		↔↓	
10	o (+a) > u	↔↓	↔↓	↔↓	↔↓						↔↓	
11	o > u					↔↓	↔↓	↔↓	↔↓	↔↓		
11 b	o (+n) > o				↔↓		↔↓	↔↓				
11 c	o > we		→									
12	ye (+ll) > i		↑↔									
13	e - > wa											
14	e (+ . i) > i	↓↔	↓↔	↓↔	↓↔	↓↔	↓↔	↓↔	↓↔	↔↑		
15	(k +) e > i									↔↑		
16	a (+ y) > e	↓↔	↓↔	↔↓	↔↓	↔↓				↔↑		
17	a (+ w) > o	↔↓	↔↓	↔↓	↔↓					↔↑	↔↓	
18	e (+ y) > wa									↔↑		
19	o > wa								↑↔			
20	o (+ v) > wa									↑↔		
21	u (+ y) > wi									↑↔		
22	o (+ y, w, k) > we								→			
23	i (+ w) > ye					↔↑	↔↑	↔↑	↔↑			

		PORTUGAIS	CASTILLAN	CATALAN	ROUSSILLON	GASGON	LANGUEDOC	NARBONNAIS	PROVENÇAL	FRANÇAIS	ITALIEN	ROUMAIN
24	ɛ (+ y) > i	↔	↔	↔	↔							
25	ɔ (+ y) > u	↔	↔	↔	↔							
26	ɔ (+ ɲ) > u	↔	↔	↔	↔							
27	ɔ (+ l, nk) > u											
28	ɛ (+ ɲ, l, nk) > i											
29	oy > we		→									
30	ɔ (+ y + n) > we											
31	a (+ l) > o	↔	↔									
32	o (+ l) > u											
33	ɛ (+ l) > o											
34	o (+ n) > a	↔										
35	e (+ n) > a											
36	i (+ n) > e											
37	ü (+ n) > ö											
38	ɛ (+ n, m) > i											
39	a (+ n, m) > i											
40	e (+ ..a) > ea											
41	o (+ ..a) > oa											
42	(r̄ +) i > i											
43	(lab., s, cons. + r) + i (+ n) > i											
44	i (+ n + ..u) > i											
45	(palat. +) u > i											
46	(palat. +) a > e											
47	(lab. +) ɛ (+ ..u) > ā											

1. DISPERSION DES LANGUES ROMANES DANS UN DIAGRAMME EN XY

Y

o PORTUGAIS : Point mesuré
 • Portugais : Point calculé



La séquence des divers vecteurs représentant les changements n'implique nullement une succession chronologique : nous avons disposé les vecteurs de manière à avoir une figure claire, évitant que les divers cheminements se recoupent et s'entrecroisent.

Les numéros adjoints aux vecteurs correspondent à une numérotation arbitraire des divers changements de timbres vocaliques que nous avons rassemblés en tableau.

Dans chaque case de ce tableau, nous indiquons schématiquement par deux flèches, l'une verticale, l'autre horizontale, les sens dans lesquels varient respectivement y et x pour chaque langue affectée par la mutation.

RÉPARTITION SPATIALE DE QUELQUES LANGUES DANS UN DIAGRAMME EN XYZ.

Utilisons maintenant les trois paramètres, et faisons correspondre à chaque langue le point admettant ces trois coordonnées dans un diagramme en XYZ (fig. 2).

Nous représentons graphiquement cette répartition spatiale par ses projections sur les deux plans XY et XZ, selon les procédés de la géométrie descriptive. (Le trièdre XYZ est direct.)

Les deux projections nous montrent que les quarante-six points (et le fait est encore plus net si nous y ajoutons les soixante-six autres points correspondant aux langues dont les échantillons étaient trop courts) se répartissent à l'intérieur d'un ellipsoïde de révolution autour de son grand axe (forme de l'enveloppe d'un ballon de rugby).

Nous sommes frappés de voir les langues slaves, sauf le serbe, se rassembler vers la gauche, les langues romanes de la Romania continue, au centre, les langues germaniques vers la droite, les langues celtiques encore plus à droite ; les langues italiques anciennes se situent entre les langues slaves et les langues romanes. Sur la figure, nous avons rejoint les points représentatifs de chacun de ces groupes par des traits pointillés, afin de souligner la localisation de chaque famille.

Pour mieux classer les langues selon les affinités manifestées par les trois paramètres retenus, nous avons résolu de diviser l'ellipsoïde de révolution en trois fractions par deux plans perpendiculaires à son axe.

L'axe a été tracé expérimentalement, et choisi de manière que les

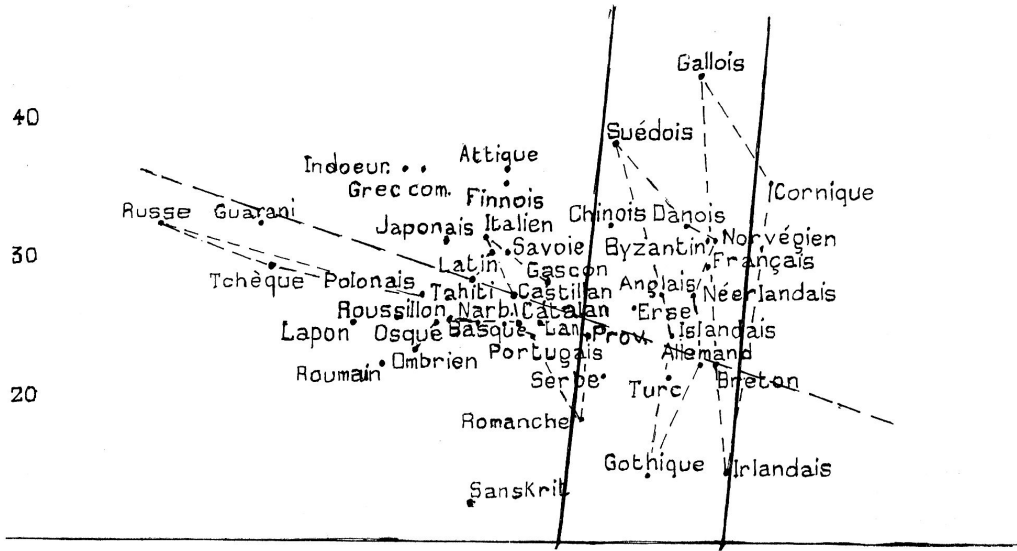
Z

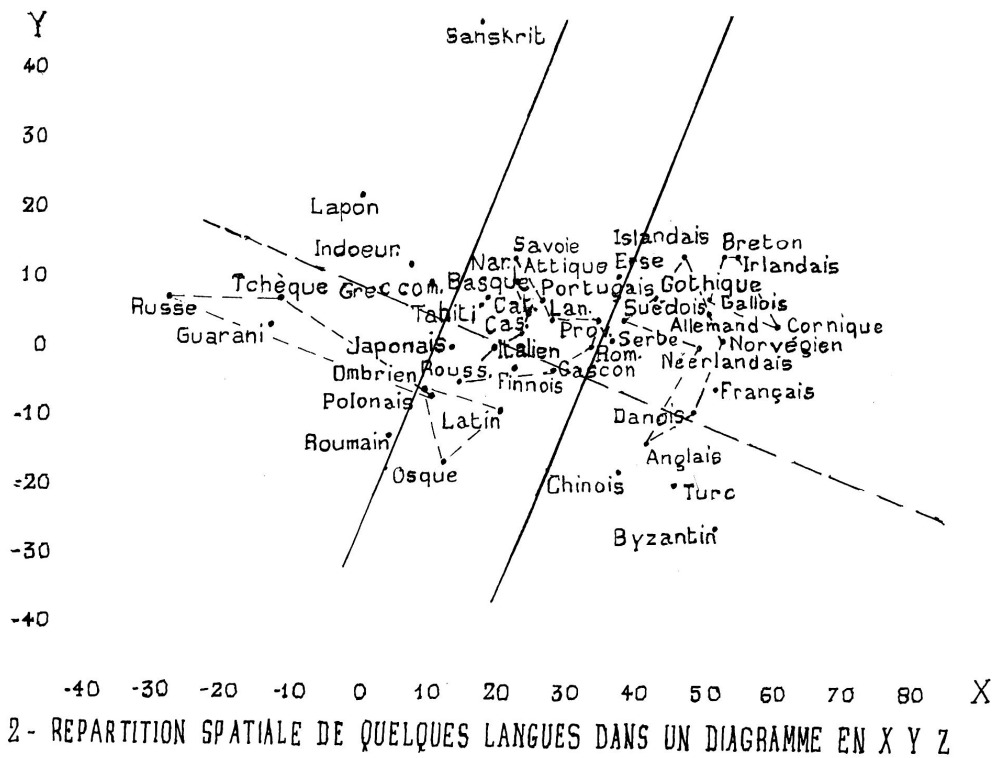
50

40

30

20





coefficients de ses équations ne soient pas trop compliqués. Nous le définissons par l'équation de la trace du plan le projetant sur XY :

$$(1) \quad 4x + 10y - 1 = 0$$

et par celle de la trace du plan le projetant sur XZ :

$$(2) \quad 10x + 60z - 19 = 0$$

On sait qu'un plan perpendiculaire à une droite d'équations

$$ax + by + d = 0 \quad \text{et} \quad ax + cz + e = 0$$

a une équation de la forme

$$Ax + By + Cz + D = 0$$

dans laquelle les coefficients A, B, C, satisfont aux relations :

$$Aa = -Bb = -Cc$$

Dans le cas présent, ceci nous donne :

$$(3) \quad -60x + 24y + 10z + D = 0$$

Le paramètre D permet, selon les valeurs qu'il reçoit, de déplacer le plan parallèlement à lui-même.

Proposons-nous d'abord de choisir une première position du plan située le plus à droite possible et qui laisse à droite toutes les langues germaniques. En remplaçant les variables de l'équation (3) par les coordonnées des points représentatifs, le premier membre de cette équation doit prendre une valeur négative ; cette condition nous amène à choisir $D = 1770$. Le plan d'équation

$$(4) \quad -60x + 24y + 10z + 1770 = 0$$

dont nous avons indiqué les traces sur les plans de référence, va isoler à sa droite des langues à point d'articulation moyen très avancé (groupe A).

Déterminons maintenant une deuxième position du plan située le plus à gauche possible et qui laisse à gauche les langues slaves, russe, tchèque et polonais. En remplaçant les variables de l'équation (3) par les coordonnées des points représentatifs, le premier membre de cette équation doit prendre une valeur positive ; cette condition nous amène à choisir $D = 500$. Le plan d'équation

$$(5) \quad -60x + 24y + 10z + 500 = 0$$

dont nous avons aussi indiqué les traces sur les plans de référence, va

isoler à sa gauche des langues à point d'articulation moyen reculé (groupe C).

Entre les deux plans définis par les équations (4) et (5) se trouvent les langues à point d'articulation moyen modérément avancé (groupe B).

Les langues du groupe A ont en général y et z petits, c'est à dire qu'elles ont des points d'articulation dispersés dans la bouche et qu'elles sont peu labialisées. Le groupe C appelle des remarques contraires, et le groupe B présente des caractères intermédiaires.

Dans le groupe A nous trouvons les langues germaniques, les langues celtiques (sauf l'érse), le français, le grec de la koiné byzantine, le serbe, le turc, le chinois, soit dix-sept langues; dans le groupe B, les langues romanes (sauf le français et le roumain), l'érse, les langues italiques anciennes (sauf l'ombrien), le grec attique, le finnois, le basque, le tahitien, soit dix-huit langues; dans le groupe C, les langues slaves (sauf le serbe), le roumain, le grec commun, l'indo-européen, le sanskrit, l'ombrien, le japonais, le lapon, le guarani, soit onze langues sur quarante-six.

Quand nous étendons notre statistique aux soixante-six langues insuffisamment documentées, les trois groupes deviennent sensiblement égaux.

Pour classer une langue dont on a déterminé les trois paramètres x , y , et z , il suffit de remplacer par les valeurs numériques de ceux-ci les trois variables respectives du trinôme :

$$60 x - 24 y - 10 z$$

Si la valeur numérique obtenue est supérieure à 1770, la langue se classe dans le groupe A; si cette valeur est comprise entre 500 et 1770, dans le groupe B; si elle est inférieure à 500, dans le groupe C.

Ainsi, ce trinôme prend la valeur 2280 pour le français, 1028 pour le castillan, 796 pour l'italien, 298 pour le roumain. Le romanche avec 1766 est très voisin du groupe A, et ceci ne saurait surprendre étant donné sa position géographique.

Le basque (632) et le finnois (1008) attestent pour le paléo-européen la situation en groupe B; le lapon (-788) et le japonais (416), pour le paléo-asiatique, la situation en groupe C, qui est aussi celle du slave et des vieilles langues indo-européennes; le turc (2936), pour l'altaïque, la situation en groupe A.

L'appartenance au groupe B du latin et de la majeure partie des langues romanes héritières peut être due à un effet de substrat paléo-européen. L'appartenance au groupe A des langues germaniques et celtiques pour-

rait dépendre d'un influx altaïque, à mettre en relation avec la poussée des brachycéphales du néolithique à travers la plaine européenne.

Les aperçus suggérés par l'emploi des paramètres vocaliques que nous avons définis, pourraient être indéfiniment étendus ; il nous suffira d'avoir montré leur efficacité tant dans la classification des langues romanes que dans celle de quelques autres groupes de langues.

Henri GUITER.