Zeitschrift: Revue de linguistique romane
Herausgeber: Société de Linguistique Romane
Band: 51 (1987)
Heft: 201-202

Artikel: Étalonnage d'une méthode géolinguistique
Autor: Guiter, Henri
DOI: https://doi.org/10.5169/seals-399807

Nutzungsbedingungen
Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. 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. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

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. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

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. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 13.07.2025

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

ÉTALONNAGE D'UNE MÉTHODE GÉOLINGUISTIQUE

Dans notre contribution aux Mélanges offerts à Ramon Aramon i Serra (1) nous avons entrepris de rechercher quelle était l'articulation linguistique des divers parlers romans au long de la cordillère cantabro-pyrénéenne. Cette enquête concernait dix-huit provinces espagnoles (La Coruña, Pontevedra, Lugo, Orense, Oviedo, León, Santander, Palencia, Vizcaia, Burgos, Alava, Logroño, Navarra, Zaragoza, Huesca, Lleida, Barcelona, Girona) et huit département français (Pyrénées-Orientales, Aude, Ariège, Haute-Garonne, Gers, Hautes-Pyrénées, Landes, Pyrénées-Atlantiques) (2).

Nous nous proposions de relever au centre de chaque division administrative les traitements de quelques particularités phonétiques et lexicales qui devaient satisfaire à deux conditions :

1) présenter obligatoirement plusieurs solutions différentes dans le domaine étudié ; autrement dit, nous éliminions toute carte mononyme ;

2) être documentées sur tout le domaine, compte tenu de l'hétérogénéité des matériaux dont nous disposions. Cette documentation était très étendue pour les régions gasconne et catalane ; il n'en était pas de même ailleurs.

---

(1) Enric Guiter, « Aproximació lingüística a la cadena càntabro-pirinenca », *Miscel.lània Aramon i Serra* III, Barcelona, 1983, p. 247.

(2) A vrai dire, les limites linguistiques ne coïncident pas toujours avec les limites administratives. C'est ainsi que la frange orientale de la province de Huesca est nettement catalane, tandis que la partie sud-orientale, le long de la limite avec la province de Zaragoza, a subi la même influence castillane que cette dernière. Au nord des Pyrénées-Orientales, le Fenollet languedocien se rattache à l'Aude. L'Ariège se limite pour nous au comté de Foix, c'est-à-dire aux arrondissements languedociens de Foix et de Pamiers ; l'arrondissement gascon de Saint-Girons est rattaché à la partie méridionale de la Haute-Garonne, gasconne elle aussi. L'ouest de la Vizcaya et la majeure partie de l'Alava sont castellanisés.

Ces considérations nous amenèrent à retenir seulement dix-huit traitements phonétiques (*a* + *y* ; diphtong. *e, o* ; palatalis. *u* ; *a* final ; *a* final + *s* ; *-k'-* et *-ty-* ; *f-* ; *pl-* ; *l-* ; *k'l-* ; *-ll-* ; *-ssy-* ; *-sk'-* ; *-n-* ; *-(o)w* ; *-kw-* ; *-mb-* ; *-nd-*) et neuf traits lexicaux (*laeuu, cadere, cane, oue, domu, muliere, carne, fratre, occidere*).

Les réalisations de ces vingt-sept traits au centre des diverses divisions administratives contiguës nous ont permis d'établir entre celles-ci des distances de Hamming, que nous pouvons traduire en pourcentages, D. Ces pourcentages forment une suite discontinue, et pour en faire des classes, nous nous sommes laissé guider par les coupures plus importantes qui y apparaissaient naturellement. C'était là une démarche très empirique, qui ne s'appuyait sur aucun fondement théorique.

Les valeurs comprises entre 63 % et 78 % ne se rencontrent que sur les Pyrénées, entre aragonais et catalan, d'une part, gascon et languedocien, d'autre part.

Les valeurs comprises entre 37 % et 56 % apparaissent entre galicien, d'une part, asturien et léonais, d'autre part ; catalan, d'une part, aragonais et castillan, d'autre part ; gascon, d'une part, languedocien, de l'autre.

Les valeurs comprises entre 18 % et 33 % isolent l'asturien, le léonais et l'aragonais du castillan ; l'asturien, du léonais ; le catalan oriental, du catalan occidental ; l'Aude, de l'Ariège.

Nous avions tenu pour négligeables les valeurs inférieures à 15 %.

Que penser du niveau des frontières linguistiques correspondant aux trois premières catégories des distances de Hamming ?

On sait que nous avons proposé en 1971, au colloque de Strasbourg, une méthode globale de géolinguistique quantitative (3). Elle consiste à établir le pourcentage de différences, N, entre deux points contigus d'un atlas linguistique. L'emploi du module (N/100) ([100 — N]/100), à des puissances différentes, nous permettait d'établir une formule de correction de N en fonction de la densité des points de l'atlas (ce sont toujours les valeurs corrigés de N, que nous envisagerons désormais), et aussi de déterminer, en fonction des valeurs de N, des niveaux de

(3) Henri Guiter, « Atlas et frontières linguistiques », *Les dialectes romans de France à la lumière des atlas régionaux,* Paris, 1973, p. 61.

différences que, par convention, nous appelions niveau de langues au-dessus de N = 80, niveau de dialectes entre N = 50 et N = 80, niveau de sous-dialectes entre N = 30 et N = 50.

Il existait dans le domaine cantabro-pyrénéen des limites dont le niveau avait fait l'objet de déterminations par les deux méthodes ; la méthode globale nous permettait donc de fixer le niveau correspondant des distances de Hamming.

La frontière entre l'Aude et les Pyrénées-Orientales avait été évaluée comme frontière de langues par la méthode globale. Elle figurait dans la première catégorie de nos distances de Hamming ; cette catégorie (63 à 78) correspondait donc à des frontières de langues.

La frontière entre la province de Lleida et celle de Zaragoza avait été évaluée, elle aussi, comme frontière de langues, et nous la trouvions maintenant dans la deuxième catégorie des distances de Hamming ; nous en concluions que cette deuxième catégorie correspondait aussi à des frontières de langues, mais d'un niveau inférieur (37 à 55) à celui de la première catégorie. Dans un cas, il s'agissait de frontières entre l'ibéro-roman et le gallo-roman ; dans l'autre, de frontières à l'intérieur de l'ibéro-roman ou du gallo-roman.

La frontière entre la province de Lleida et celle de Barcelona correspondait à la frontière dialectale entre catalan occidental et catalan oriental. Elle figurait maintenant dans la troisième catégorie des distances de Hamming ; nous admettions donc que cette troisième catégorie (18 à 33) correspondait à des frontières de dialectes.

A la réflexion, nous n'étions cependant pas satisfait par ces évaluations sommaires et, surtout, par le découpage quelque peu arbitraire des catégories de distances de Hamming. Il devait y avoir possibilité d'une approche moins discutable.

Nous nous sommes donc proposé de bien préciser les évaluations chiffrées de quatre frontières provinciales particulières, et de reporter sur un graphique les valeurs de N et de D qui correspondent à chacune d'elles.

1° Frontière Aude-Pyr.-Or.

La valeur de D est 0,74. La valeur de N peut nous être fournie par l'application de la méthode globale aux inter-points recoupés dans l'ALPO par cette frontière ; les valeurs figurent dans les pages 67-79

de notre référence 3. Mais il convient de remarquer que de petits domaines situés le long de la frontière (Capcir, Bas-Fenouillet, Peyrapertusès) sont occupés par des dialectes de transition ; nous les avons étudiés dans notre communication au congrès de linguistique romane de Barcelone en 1953 (4). En évitant ces ilots, nous avons la possibilité d'effectuer sur l'ALPO (5) dix-neuf mesures, dix entre le Roussillon et le Narbonnais, neuf entre le Conflent et le Fenouillet. La valeur moyenne de ces dix-neuf mesures est 0,93.

Ce résultat est conforté par les trois mesures effectués sur l'ALF (6) à partir des points 787 et 793 (p. 89 de notre référence 3) ; elles fournissent une valeur moyenne de 0,92.

2° Frontière Lleida-Zaragoza.

La valeur de D est 0,56. La valeur de N résulte des éléments fournis par l'étude de l'ALC (7) et publiés par la revue *Miscellanea Barcinonensia* (8). Le bourg de Binéfar se trouve au sud de la province de Huesca, dans la frange qui partage avec la province de Zaragoza une forte influence castillane. Au contraire, la petite ville de Fraga dans la frange orientale de la même province, parle le même catalan que Lleida et Alcaraç, dans la province voisine. Nous pouvons donc effectuer trois mesures entre Binéfar d'une part, Fraga, Alcaraç et Lleida d'autre part ; elles nous donnent pour N une valeur moyenne de 0,85.

3° Frontière Lleida-Barcelona.

La valeur de D est 0,22. Ce sont encore les données de l'ALC que nous allons mettre à profit. La frontière du catalan oriental et du catalan occidental coupe onze inter-points, et leur valeur moyenne de N est 0,57.

---

(4) Enric Guiter, « Els altres Capcirs », *Actes du VIIe Congrès International de Linguistique Romane* II, Barcelona, 1955, p. 707.

(5) Henri Guiter, *Atlas linguistique des Pyrénées orientales,* Paris, 1966.

(6) Gilliéron et Edmond, *Atlas linguistique de la France,* Paris, 1902-1915.

(7) Antoni Griera, *Atlas lingüístic de Catalunya,* Barcelona, 1924 et sq.

(8) Sor Anna Sardà i Enric Guiter, « L'atlas lingüístic de Catalunya i la fragmentació dialectal del català », *Miscellanea Barcinonensia* XIV, Barcelona, 1975, p. 93.

4° Frontière Girona-Pyr.-Or.

La valeur de D est 0,11. Nous en revenons à l'ALPO ([5]) et à son exploitation ([3]). La frontière politique y coupe quarante-deux inter-points pour lesquels la valeur moyenne de N est 0,354. En associant les données de l'ALC ([7]) et l'ALPI ([9]) au long de cette même frontière (pp. 92-94 de notre référence 3), nous disposons de onze inter-points avec une valeur moyenne de 0,344 pour N. Nous tiendrons compte de ces diverses mesures en retenant 0,35 comme valeur de N.

Nous pouvons ainsi porter sur un graphique quatre points dont les coordonnées respectives sont : 0,74 et 0,93 ; 0,56 et 0,85 ; 0,22 et 0,57 ; 0,11 et 0,35.

A première vue, ces points semblent se situer sur un arc de parabole. Il est aisé de justifier cette impression en effectuant pour chacun d'entre eux le rapport D/N^2. Ces rapports se dispersent peu autour d'une valeur moyenne qui est 0,80. Si nous traçons la courbe représentative de la fonction $D = 0{,}8\ N^2$, nous observons qu'elle suit à peu près le trajet jalonné par les points expérimentaux.

Le calcul du coefficient de Bravais-Pearson γ permet de préciser l'adaptation de la fonction choisie à la représentation des points expérimentaux. On sait que ce coefficient peut varier entre + 1 et — 1, + 1 si l'adaptation de la fonction est parfaite, — 1 si le choix de la fonction a été totalement aberrant. Nous désignerons par $y = 0{,}8\ N^2$ les valeurs qui seraient celles de D, si tous les points expérimentaux se trouvaient exactement sur la courbe théorique.

D	$D - \bar{D}$	$(D - \bar{D})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$	$(D - \bar{D})(y - \bar{y})$
11	— 29,75	885	9,8	— 30,9	955	919
22	— 18,75	351	26,0	— 14,7	216	276
56	+ 15,25	232	57,8	+ 17,1	292	261
74	+ 33,25	1.105	69,2	+ 28,5	812	945
163		2.573	162,8		2.275	2.401

$\bar{D} = 40{,}75 \quad \sqrt{\Sigma (D - \bar{D})^2} = 50{,}72 \quad \bar{y} = 40{,}7 \quad \sqrt{\Sigma (y - \bar{y})^2} = 47{,}7$

$$\gamma = \frac{\Sigma (D - \bar{D})\ (y - \bar{y})}{\sqrt{\Sigma (D - \bar{D})^2}\ \sqrt{\Sigma (y - \bar{y})^2}} = \frac{2.401}{(50{,}72)\ (47{,}7)} = 0{,}992$$

(9) Günther Haensch, *Las hablas de la alta Ribagorza*, Zaragoza, 1960.

La valeur très élevée de ce coefficient de corrélation manifeste l'accord des résultats expérimentaux et de la fonction représentative.

On peut maintenant parler d'étalonnage de la méthode des distances de Hamming (avec les 27 traits particuliers que nous avons choisis) par la méthode globale, et nous pourrons corriger les insuffisances de l'étalonnage simpliste que nous avions d'abord envisagé.

Par usage de la fonction $D = 0{,}8\ N^2$, ou même par simple examen de la courbe, il nous est possible de constater qu'à la valeur 0,80 de N correspond une valeur 0,51 de D ; à la valeur 0,50 de N, une valeur 0,20 de D ; à la valeur 0,30 de N, une valeur 0,07 de D. Ces trois valeurs de D devraient limiter les domaines des frontières de langues, de dialectes et de sous-dialectes ; on s'aperçoit que ces valeurs sont différentes de celles que nous avions précédemment adoptées. Voyons la hiérarchie de frontières qui en résulte.

Frontière de langues :

Ariège-Pyr.-Or. 0,78 ; Aude-Pyr.-Or. 0,74 ; Lleida-Hte-Garonne 0,70 ; Huesca-Hte-Garonne 0,66 ; Lleida-Ariège 0,63 ; Huesca-Htes-Pyr. 0,63 ; Huesca-Pyr.-Atlant. 0,63 ; Zaragosa-Lleida 0,56.

Frontières de dialectes :

Ariège-Hte-Garonne 0,48 ; Lugo-León 0,44 ; Orense-León 0,44 ; Lugo-Oviedo 0,40 ; Huesca-Lleida 0,37 ; Oviedo-Santander 0,33 ; Lleida-Girona 0,26 ; Lleida-Barcelona 0,22 ; Oviedo-León 0,22 ; Zaragoza-Huesca 0,22 ; Navarre-Huesca 0,22 ; Aude-Ariège 0,22.

Frontière de sous-dialectes :

León-Palencia 0,19 ; León-Santander 0,14 ; Santander-Palencia 0,14 ; Santander-Burgos 0,14 ; Santander-Bilbao 0,14 ; Landes-Gers 0,14 ; Landes-Pyr.-Atlant. 0,14 ; Girona-Pyr.-Or. 0,11.

Les frontières de langues les plus élevées se situent entre catalan et languedocien (0,78 à 0,63), puis entre aragonais et gascon (0,70 à 0,63), enfin entre catalan et castillan (0,56). Cette dernière frontière n'est nette qu'en domaine de « reconquête » ; la méthode globale nous avait déjà montré que le « câble » frontalier s'effilochait en de multiples fibres de niveau dialectal ou sous-dialectal, lorsqu'on approchait des Pyrénées.

Le passage progressif du catalan à l'aragonais y a été minutieusement étudié par G. Haensch ([9]).

Nous passons ensuite à des frontières dialectales, certaines de niveau très élevé, entre languedocien et gascon (0,48) ou entre galicien et asturo-léonais (0,44 à 0,40) ; d'autres de niveau plus moyen, entre aragonais et catalan (0,37) ou asturien et montagnais (0,33) ; d'autres enfin de niveau modeste entre catalan oriental et occidental (0,26 à 0,22), entre asturien et léonais, castillan et aragonais, languedocien de l'Aude et de l'Ariège (0,22).

Des frontières sous-dialectales de niveau 0,19 séparent le léonais du castillan ; de niveau 0,14 elles entourent la Montaña de Santander au sud et à l'est ; elles séparent aussi les Landes du Gers et des Pyrénées-Atlantiques. Une frontière de niveau 0,11 s'interpose entre Girona et les Pyrénées-Orientales.

Pour nous résumer, les éléments de comparaison que nous avions pu retenir pour caractériser chacune des 26 provinces (ou départements) de la région cantabro-pyrénéenne, se limitaient à 27 traits. C'était comme si nous avions disposé d'un atlas de 27 cartes avec 26 points d'enquête. Aussi bien le nombre trop réduit des cartes, que la trop faible densité des points, nous interdisaient l'emploi de la méthode globale.

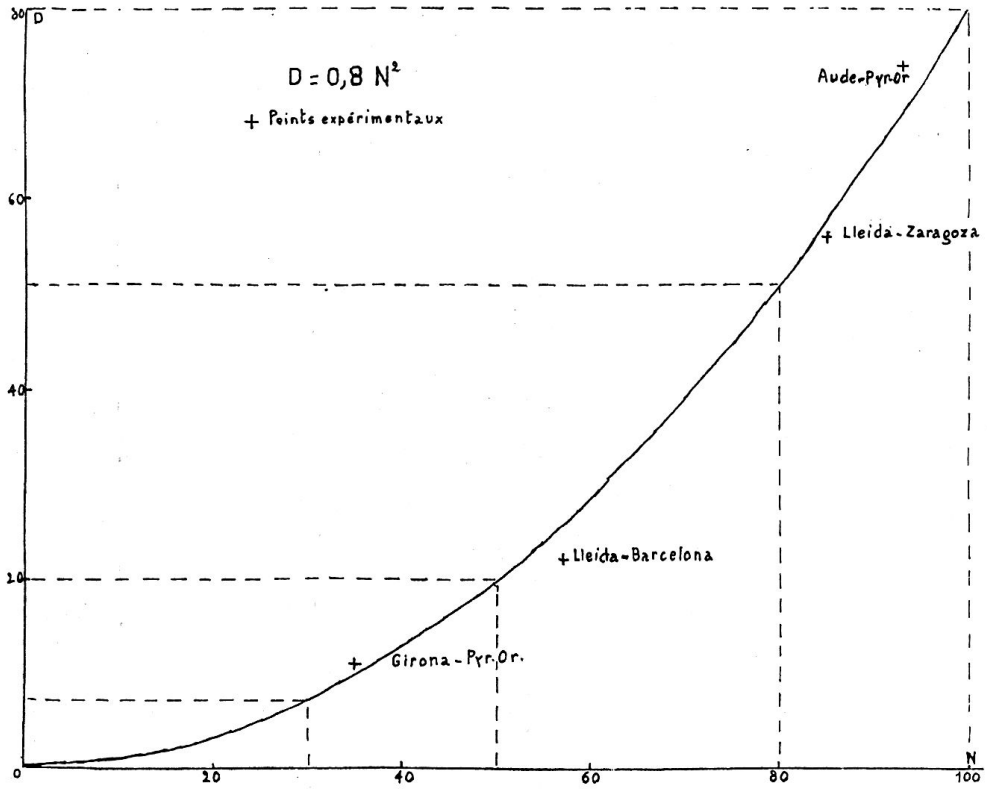
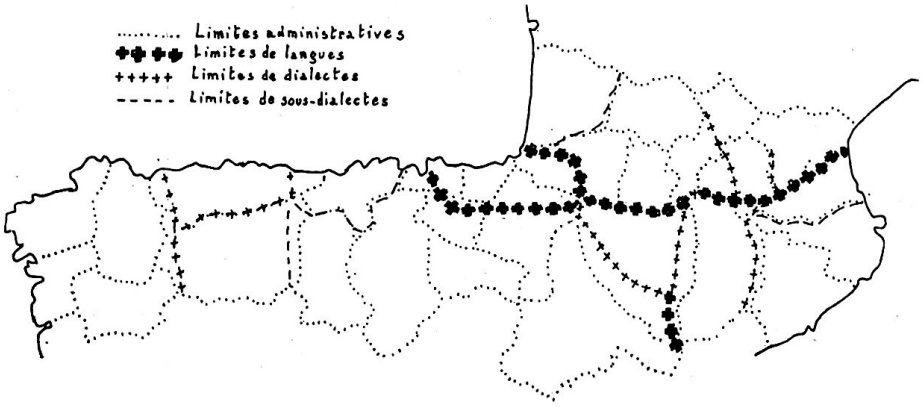
Nous avons alors songé à exprimer (en pourcentages) des distances de Hamming entre tous les couples de provinces contiguës.

Les frontières de certains de ces couples de provinces avaient déjà fait l'objet de mesures par la méthode globale. D'où la possibilité d'apprécier au moyen de cette dernière les niveaux de quelques distances de Hamming. Le classement qui en résultait n'était certes pas absurde, mais les délimitations entre les niveaux successifs présentaient un caractère arbitraire.

Pour y remédier, il nous est venu l'idée de rechercher s'il n'existait pas une fonction mathématique reliant les mesures fournies par la méthode globale, aux distances de Hamming. Cette fonction existe et elle est relativement simple. Elle nous permet de transposer dans l'échelle des distances de Haming, les limites de niveaux déterminées par la méthode globale.

Il faut bien préciser que ce résultat est valable seulement pour une recherche effectuée avec les 27 traits retenus. Si les traits venaient à

être choisis d'une manière différente, l'équation serait vraisemblablement différente, elle aussi. Il est curieux de constater que la comparaison des méthodes exclut des pourcentages de distances de Hamming entre provinces contiguës supérieurs à 80 %. Pratiquement la valeur maxima que nous avons rencontrée était 0,78.

Le but visé par cette étude était d'établir une relation mathématique entre deux méthodes géolinguistiques différentes ; elle nous a permis d'affiner les résultats obtenus par des recherches antérieures.

Montpellier. Henri GUITER