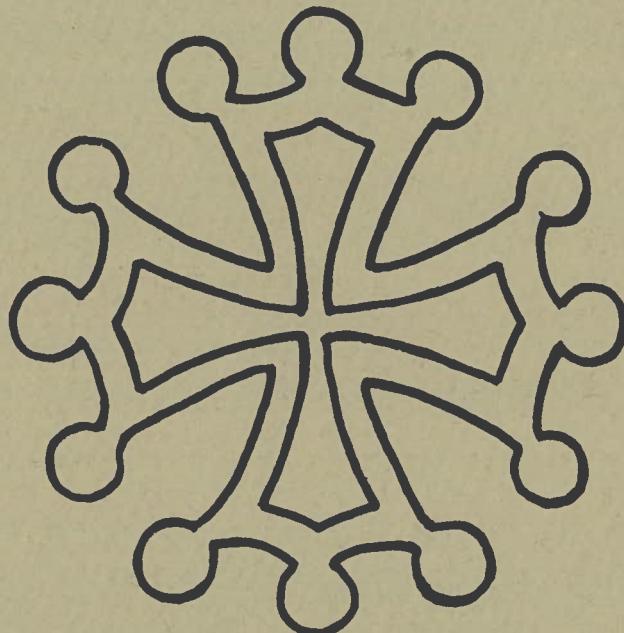


UNIVERSITÉ DE TOULOUSE II - LE MIRAIL

TRAVAUX DE L'INSTITUT  
D'ÉTUDES MÉRIDIONALES

E. R. A. 352 C. N. R. S.



**MICRODIALECTOLOGIE  
ET DIALECTOMÉTRIE  
DES PYRÉNÉES GASCONNES**

*par*

**Jean-Louis FOSSAT**

*avec la collaboration de*

**Dennis PHILPS** (Université de Londres)

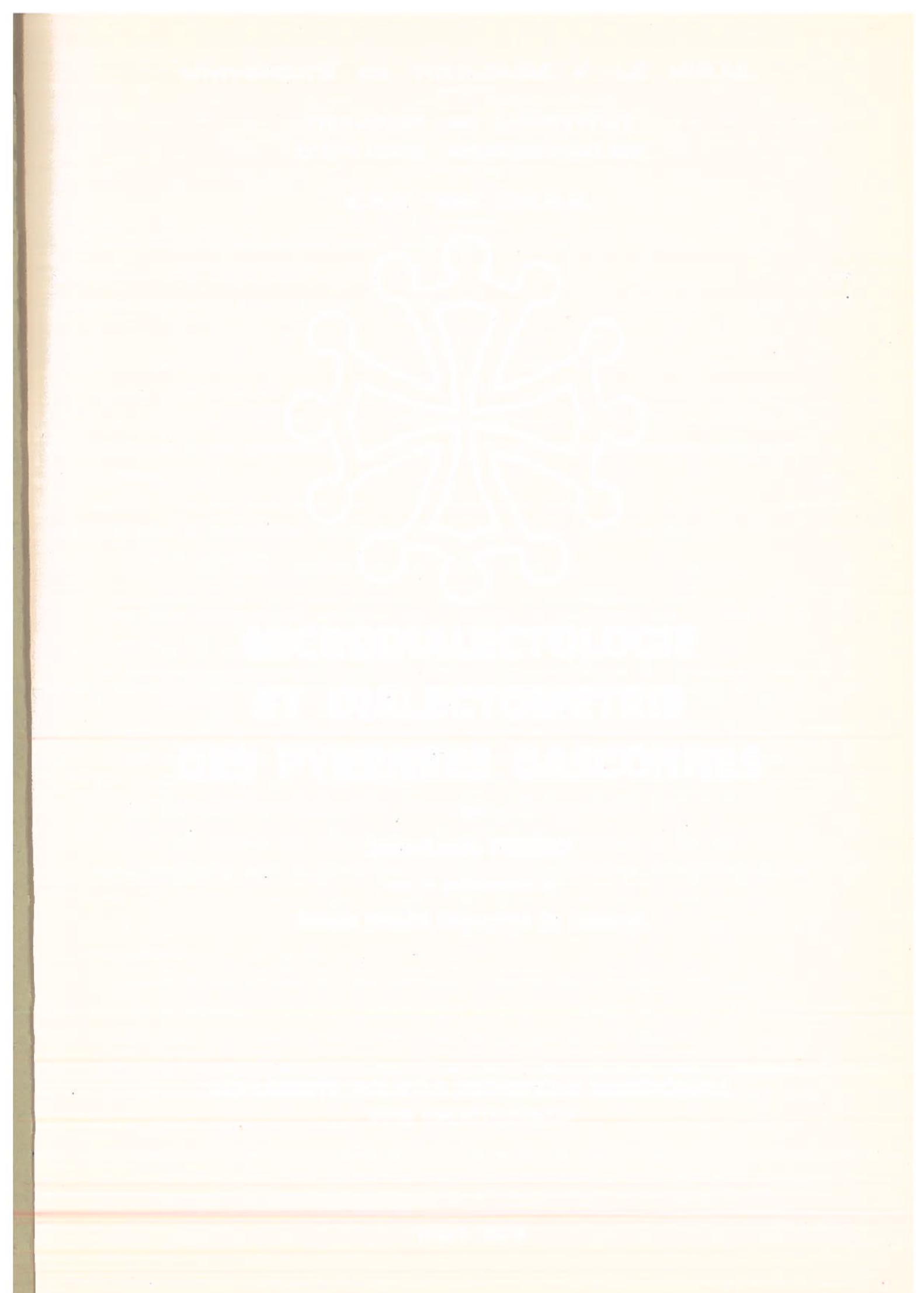
DOCUMENTS POUR LA RECHERCHE MÉRIDIONALE  
*SÉRIE DIALECTOLOGIQUE*

(mars 1976)



COPYRIGHT

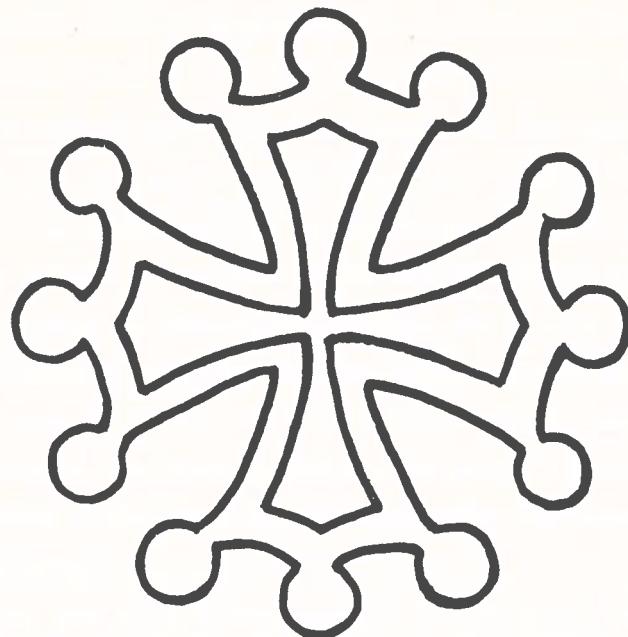
E.R.A. 352 CNRS  
Université de Toulouse Le Mirail  
109 bis, rue Vauquelin  
31081 TOULOUSE Cédex



UNIVERSITÉ DE TOULOUSE II - LE MIRAIL

TRAVAUX DE L'INSTITUT  
D'ÉTUDES MÉRIDIONALES

E. R. A. 352 C. N. R. S.



**MICRODIALECTOLOGIE  
ET DIALECTOMÉTRIE  
DES PYRÉNÉES GASCONNES**

*par*

**Jean-Louis FOSSAT**

*avec la collaboration de*

**Dennis PHILPS** (Université de Londres)

DOCUMENTS POUR LA RECHERCHE MÉRIDIONALE  
SÉRIE DIALECTOLOGIQUE

(mars 1976)

Petite introduction méthodologique  
aux études dialectométriques menées en Gascogne

par  
Jean-Louis FOSSAT  
E.R.A. 352  
C.N.R.S.

La présente étude est née de la réflexion largement collective au cours de la mise en place en 1976 du séminaire de 3ème cycle d'Etudes Romanes, à l'Institut d'Etudes Méridionales de l'Université de Toulouse Le Mirail; ont plus spécialement collaboré à cet effort d'analyse et de clarification pédagogique, Normand Beauchemin, de l'Université de Sherbrooke, Dennis Philps, de l'Université de Londres, M.H. Testa, étudiante à l'Université de Paris, M.T. Canguilhem, étudiante à l'Université de Toulouse Le Mirail; qu'ils soient tous vivement remerciés de leur apport; l'un des résultats les plus immédiats de cette réflexion sera le traitement automatique des données dialectométriques élaborées par Jean Séguay.

#### 1. Généralités-définitions.

La dialectométrie, pour une aire donnée, mesure l'incidence de la variation linguistique dans l'espace; on retrouve dans l'espace la projection des dissemblances observées sur l'axe historique pour une langue donnée.

La dialectométrie mesure les différences: en cela elle est un instrument tout nouveau d'approche de la fonction dialectale et de la notion d'aire dialectale; cet instrument est encore en 1976 largement ignoré des chercheurs avancés qui se préoccupent à des titres divers d'analyser la fonction maximale d'intercompréhension et la fonction minimale de démarcation.

#### 2. Bibliographie des travaux de type dialectométrique.

Un état des travaux est prématûré; pour l'heure, les travaux se limitent aux analyses menées de part et d'autre aux Universités de Toulouse et de Montpellier, dans les différentes équipes et formations de dialectologie romane; nous présentons ici aux lecteurs l'état de notre réflexion en matière de dialectométrie pyrénéenne.

#### 3. Détermination du gradient de gasconité d'après les données de l'A.L.G. VI.

La carte 2531 de l'A.L.G. VI détermine 11 classes d'espace, définies quantitativement comme 11 zones de sécurité linguistique; de ce point de vue, qui pourra paraître suspect d'interprétation psychologique, on regroupera les 11 classes d'espace en trois zones, selon le critère de l'endémisme: endémisme maximal; endémisme moyen; endémisme minimal.

La formation pratique du dialectométricien consiste à s'accoutumer à reconnaître la classe d'espace d'une unité ponctuelle ou d'un groupe d'unités ponctuelles, à partir du dépouillement des matrices.

Soit par exemple l'unité ponctuelle 693 de l'A.L.G. = classe d'espace 3; 693NO = classe 4. On dira que 693NO et 685SE constituent une forme reconnue mathématiquement; on parle ici de reconnaissance des formes au sens où un ingénieur-système en automatisation parle de reconnaissance des formes.

Soit le couple géographique 7910 et 791 = classe 5, contre 790SE = classe 4. L'appréciation quantitative des différences est fondée sur la reconnaissance des formes ventilées par paramètres, phonétique, phonologique, morphologique, lexical, pour s'en tenir aux paramètres actuellement mesurés.

Le mode de détermination est variable selon les paramètres; la détermination du paramètre phonétique se fait selon les principes de la phonétique historique; le mode de détermination du paramètre phonologique obéit à des critères distributionnels; la distance morphologique s'établit selon le critère de la reconnaissance de marques différentes; la distance lexicale en système ouvert, s'établit par tirage au hasard de N% de cartes onomasiologiques, soit 1 sur 10 en moyenne.

La dialectométrie de première heure n'est en rien une mesure statistique du polymorphisme; elle n'intègre pas les données statistiques de l'A.L.G. VI.

#### 4. Considérations sur le cadre théorique des études dialectométriques.

La dialectométrie est une des branches fondamentales de la reconnaissance des formes élémentaires isolées, qui forment des classes d'agrégation, par paramètre. Cette pratique suppose la connaissance de quelques notions élémentaires en matière de théorie des codes et de théorie des signaux; ces notions complexes doivent être acquises au cours de stages pratiques d'informatique à l'Université de Toulouse Le Mirail, notamment par les chercheurs de D.E.A. d'Etudes Romanes désireux de traiter efficacement des données dans cette perspective.

Le point de départ est le cas de la transmission de l'information sans bruit, la voix n'apportant aucune perturbation; le symbole émis à l'entrée de la voix est interprété correctement à la sortie.

La dialectométrie est une application directe de cette théorie; elle met en jeu les concepts de code à décodage unique et de code à décodage multiple.

En code à décodage unique, le code est, à la sortie, 0 ou 1; il s'agit de code à décodage instantané; on appelle code à décodage instantané en information, un code où l'on peut décoder chaque mot sans avoir à se référer à d'autres caractères que ceux qui le composent; ainsi, / m a r a /, "bélier" n'a pas besoin d'être référé aux cooccurrences du champ lexical, c'est-à-dire de la pratique : / a m a r i / ; / t a r d a n / .

On appelle code morphosémantique non instantané le code où l'on ne peut décoder un mot sans faire appel à d'autres caractères que ceux qui le composent; ainsi le mot / k a n t è k / PT1 se décode nécessairement par opposition à PT3/k a n t è t / en analyse classique; en dialectométrie stricte, il ne saurait être question que de décodage phrastique en sémantique de mot; la notion de décodage transphrastique paraît largement échapper aux techniques dialectométriques, jusqu'à l'heure; ce

n'est pas un pur hasard si la syntaxe est si peu représentée dans les mesures dialectométriques actuelles.

Selon une telle procédure, on aboutit à une définition du code de Hamming, très simplifiée par Jean Séguy dans ses premières mesures, aux dires des mathématiciens consultés; le code de Hamming est un code de détection et de correction des erreurs; une erreur est soit simple, soit binaire, soit ternaire; supposons qu'en cours de transmission linéaire, à partir d'une source S, une erreur affecte une position j; alors 0 se transforme en 1 et 1 se transforme en 0; ce code de Hamming intervient pour corriger une erreur qui intervient sur n'importe laquelle des positions d'un mot-code; pour corriger N erreurs, il faut disposer de K positions de contrôle; d'où le concept de code détecteur d'erreurs par position de contrôle paramétrique. Tel est le raisonnement qui a servi de base à la mise en place des matrices de Hamming élaborées par Jean Séguy pour l'A.L.G.VI. On doit à la vérité de dire que Jean Séguy nous avait de longue date suggéré une telle pratique pour mesurer la distance lexico-sémantique en corrélation avec l'indice de position sociale d'un informateur, dans les enquêtes sociolinguistiques et ethnolinguistiques menées en Gascogne.

Un des faits les plus inquiétants de la technique est la mise en évidence de la non-transitivité de la distance, d'où la nécessité de mettre en place des matrices de contrôle, et des tests de mesure de la fonction d'intercompréhension.

Dans ces conditions s'est posée la question de savoir si oui ou non le modèle explicatif proposé vaut uniquement pour des conditions historiques et géographiques précises en Europe, par exemple pour les Pyrénées, ou s'il vaut en règle générale, c'est à dire pour mesurer les variations dialectales du Quechua, deuxième langue officielle du Pérou, ou les variations dialectales du basque et du catalan, secondes langues nationales enseignées en Espagne. On ne sait pas assez bien encore la portée de la dialectométrie en ce qui concerne l'enseignement des langues régionales; la notion de région a fait jusqu'à présent l'objet d'approches soit félibréennes soit suspectes.

Les premières applications automatisées de type dialectométrique ont commencé à apparaître avec la pratique de la cartographie de densité, appliquée par H. Goebel à divers domaines romans; ce traitement statistique, au sens informatique du terme traitement, est une des bases de la dialectométrie et constitue le tiers du volume VI de l'A.L.G.; de telles mesures permettent de réexaminer le concept de frontière dialectale et celui d'aire dialectale; c'est en cela que la dialectométrie de première génération est un instrument entièrement nouveau; pratiquement, la méthode d'évaluation de la distance selon le code de Hamming est une méthode d'évaluation du degré de ressemblance entre deux objets ou formes, définis dans une structure à treillis, une structure à grille.

##### 5. La procédure dialectométrique.

La dialectométrie se fait par étapes à partir de la récolte de corpus dialectaux

cartographiés manuellement ou triés selon des méthodes de traitement de l'information. Sur la base des cartes classiques singulières, on pratique la cartographie cumulative, et ceci pour chaque item de chaque paramètre; sur cette base, se pratique l'encodage en matrices paramétriques de type primaire; sur ces premières matrices à items, lues en colonne par la perforatrice, on pratique les matrices de type secondaire ou de type 2: les matrices des distances paramétriques, appréciées soit en chiffres, soit en pourcentages; l'addition des moyennes paramétriques conduit à l'établissement de la moyenne linguistique globale, appréciée soit en chiffres, soit en pourcentage. On sait ainsi que pour les Pyrénées Centrales la moyenne linguistique globale de la distance de chacun des points à la norme générale est de 26,6%; il s'agit ici de distance linguistique non corrigée, avec des éléments de pondération qui interviennent plus ou moins arbitrairement pour certains items.

La non transitivité de la distance conduit, après test d'intercompréhension, à introduire un chiffre correcteur d'erreur; on obtient ainsi une distance corrigée; c'est ainsi que nous avons procédé pour certaines des données de l'enquête de J. Cremona (1956) pour les parlers du haut gascon des vallées d'Aure et du Louron; la moyenne de distance établie par J. Cremona, ou plus exactement par nous sur la base des calculs de Cremona, est de 38,05 %. La moyenne crédible pour l'ensemble des Pyrénées est de 26,6 %; dans ces conditions, il est normal d'introduire un chiffre correcteur d'erreur de 11,45 % qui permet d'apprécier l'erreur de Cremona, c'est-à-dire le coefficient d'intercompréhension fondé sur les similitudes de système, ou similitudes de structure: par exemple -N- 0 à 75 % etc...; nous nous expliquons sur cette procédure dans le compte rendu critique des analyses de J. Cremona; il nous faudra aussi, un jour très proche, sur de telles bases, reprendre dans leur ensemble les données de l'abbé Lalanne pour le gascon maritime; cela pourrait utilement constituer la matière du stage pratique de D.E.A. d'Etudes Romanes pour l'année 1976-1977.

6. Pourquoi faire la dialectométrie, et comment améliorer, si besoin est, cet instrument ?

On observe que cet instrument est un remarquable outil d'appréciation quantitative des pratiques, à partir d'atlas de pratiques, de quelque type que ce soit, linguistiques ou économiques.

Nous appelons pratiques phonétiques, pour un groupe culturel, le fait de partager un certain nombre de traits tels que ceux de l'inventaire du gascon classique : -N-> 0 (lua) ; -LL->-r- (aquera) ; -LL>-t (aquet) etc...; l'A.L.G. arrête à 67 critères ou items le total de la distance phonétique; dans la pratique, on peut se demander à partir de quelle batterie lexicale on peut obtenir le résultat crédible pour N groupes culturels.

Nous appelons pratique phonologique pour un groupe culturel l'existence dans le parler de chaque individu du groupe de tel phonème à telle position dans la structure syllabique; le phonème est ainsi un objet positionné; ce paramètre, nécessaire-

rement succinct, est binaire.

Nous appelons pratique morphologique une pratique taxonomique qui permet de mesurer le degré d'organisation grammaticale dans un groupe culturel donné; on a déjà dit que l'information syntaxique reste peu approchable avec nos techniques actuelles; ce ne sera pas nécessairement perdre son temps, dans ces conditions, que de créer des méthodes de mesure à partir des textes de français ou d'occitan parlé, dans l'étiquette non écrite; c'est pourquoi, parallèlement aux mesures dialectométriques, notre groupe de recherches accélère ses analyses en matière de textes oraux, pour les langues régionales, minoritaires ou non.

Nous appelons pratique lexicale, pour un groupe culturel, la pratique, dans une situation donnée, de vocables utiles, à fonction référentielle, associée à une fonction d'intégration sémantique; la pratique lexicale permet ainsi de mesurer, dans un groupe culturel, le degré d'intégration sémantique d'un individu à la personnalité collective du groupe; si on procède par tirage au hasard de 10 % des cartes onomasiologiques, on a de très fortes chances de couvrir tous les champs lexicaux et sémantiques susceptibles d'attirer la curiosité de l'observateur; chaque référent, pour une pratique ethnographique donnée, engendre un nombre N de signifiants; la typologie des signifiants permet d'organiser la matrice primaire, en colonnes polyonymes ou mononymes.

Il suffirait, à ce stade, de disposer d'un atlas des pratiques proprement ethnographiques (forme des toits ou carte des croyances, peu importe) et de l'intégrer aux précédentes mesures; croire est une pratique; à tel lieu, on croit ou on ne croit pas aux sorciers; à tel lieu, on pratique ou non l'emploi de la brique chauffante, dans telle ou telle condition explicite; l'une et/ou l'autre de ces pratiques agit comme facteur d'intégration sémantique; chaque pointage de pratique concourt à l'établissement d'un indice d'intégration sémantique; c'est pourquoi aucun item sémantique ne doit être ni négligé, ni arbitrairement privilégié; le type de tonte des moutons dans un groupe économique importe au moins autant que tel type de croyance dans le même groupe.

Mais il nous est apparu que cette mesure n'est pas du même ordre que la mesure objective du fait dialectal; elle nous renseigne sur l'indice de position sociale et ethnique d'un individu informateur considéré, pour un groupe culturel donné; jusqu'à présent seules quelques monographies ethnographiques très isolées nous donnent quelques indices épars mais précieux, en ce qui concerne la position sociale des informateurs; dans ce mouvement, il faut rendre hommage à la perspicacité des ethnographes de notre laboratoire, D. Fabre et J. Lacroix, à R. Nelli, leur maître, à H. Polge; plus récemment J.C. Dinguirard, dans une étude lexicale et dialectale, qui ne néglige pas d'examiner le folklore, établit clairement la relation de covariance entre distance lexicale dans le canton de St-Béat et le fait que certains individus de la communauté présentent le trait

colporteurs; si l'explication avait quelque valeur générale, sinon universalisante, elle devrait pouvoir s'appliquer en d'autres endroits, au moins pour la culture pyrénéenne; les ethnographes professionnels auraient ainsi tout intérêt à coordonner leurs recherches en vue de l'établissement d'un atlas des pratiques. Sans aucun doute, il faudrait s'expliquer moins succinctement sur ce point ; mais l'orientation commune devrait être la recherche de la détermination d'indices de position sociale des informateurs définis selon les critères classiques de la démographie, de l'histoire et de la géographie humaine.

TABLEAU DES CLASSES D'ESPACE

N° de la classe d'espace	Nombre de traits de ressemblance	Endémisme
classe N° 1	39-40	maximal
classe N° 2	35-38	
classe N° 3	30-34	
classe N° 4	24-29	
classe N° 5	18-23	moyen
classe N° 6	15-17	
classe N° 7	12-14	
classe N° 8	10-11	minimal
classe N° 9	6-8	
classe N° 10	4	
classe N° 11	2	

Selon cette structure établie par Jean Séguy, selon ses propres critères, les parlers des vallées d'Aure et du Louron sont de classe N° 1; c'est-à-dire qu'ils appartiennent à la zone de gradient maximal de gasconité; ce qui était déjà clairement pressenti par l'enquête de Cremona, comme il ressort de nos calculs.

TRAVAUX DE L'INSTITUT D'ETUDES MERIDIONALES

---

D. E. A. 3ème cycle ETUDES ROMANES

LA MICRODIALECTOMETRIE: PROBLEMES ET METHODES

par

Jean-Louis FOSSAT

E. R. A. 352

C. N. R. S.

Cette étude est réalisée sur la base des données établies par J. Cremona(1956), dans son étude des parlers pyrénéens des vallées d'Aure et du Louron.

## LA MICRODIALECTOMETRIE : PROBLEMES ET METHODES

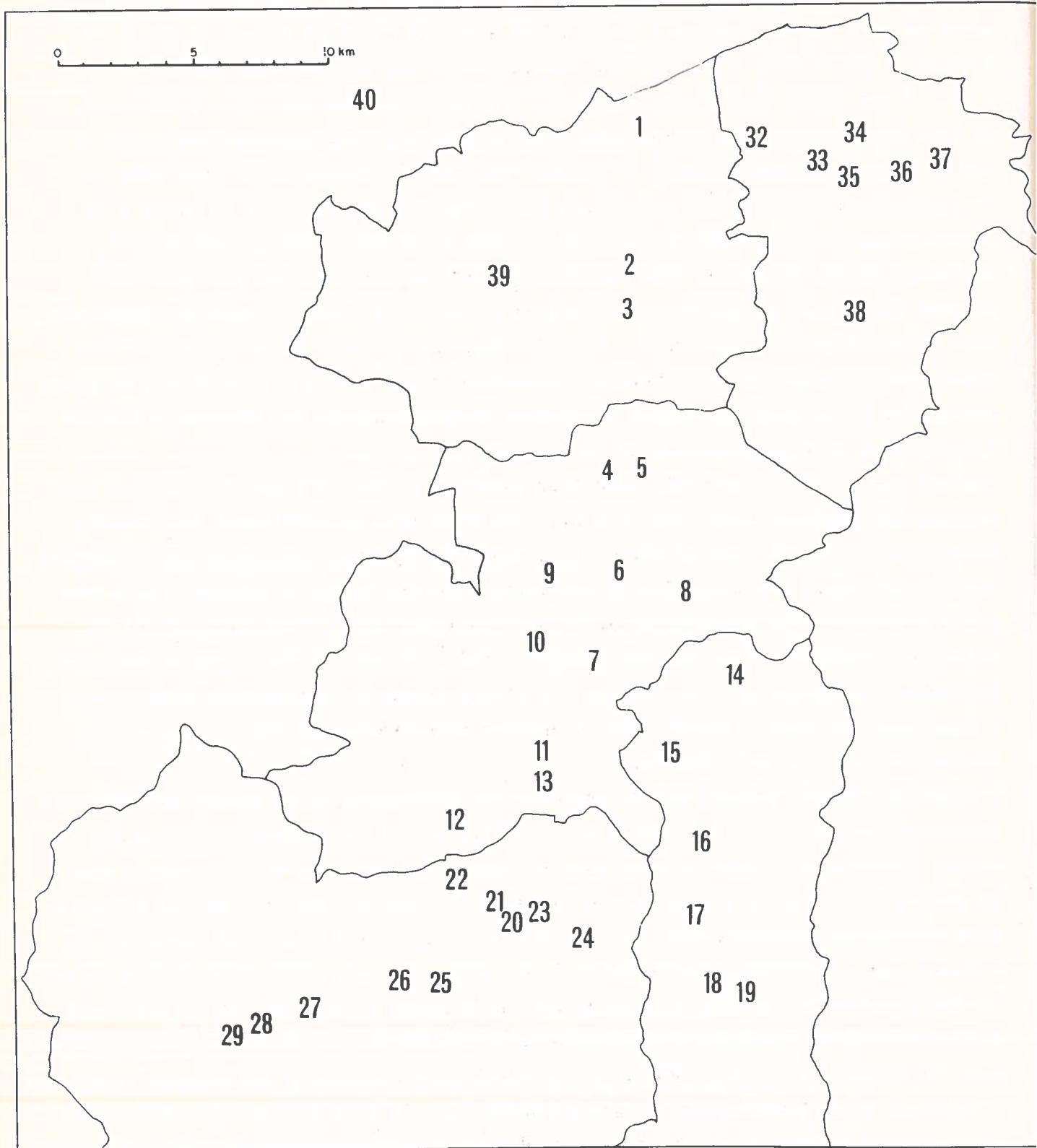
Le point de départ de notre analyse et de notre réflexion est l'enquête microdialectologique menée par des équipes de méridionalistes dans les Pyrénées, ces vingt dernières années. Tous les romanistes dialectologues connaissent le gascon pyrénéen par l'oeuvre magistrale de G. Rohlf ; mais peu de dialectologues et encore moins de linguistes ont eu connaissance des résultats de l'enquête pyrénéenne menée en 1956 par J. Cremona ; cette thèse a été le point de départ de notre réflexion ; sur la base d'une lecture classique, ce qui est attendu, puisqu'il s'agit d'une étude microdialectologique, j'ai pratiqué une relecture artisanale, de type dialectométrique ; pratiquée à cette échelle, la dialectométrie porte le nom de microdialectométrique ; il existe très peu d'analyses de ce type dans la littérature dialectologique européenne ; la première du genre et la seule à ma connaissance, est la thèse de doctorat de 3ème cycle de D. Philps, Balaguères, Bethmale, Biros, Toulouse, 1975, premier travail de vérification des hypothèses de Jean Séguy ; on se réfère volontiers et à juste titre aux travaux de vulgarisation scientifique de Jean Séguy ; mais J. Séguy est à juste titre universellement connu pour sa création de la méthode dialectométrique ; le point de départ de toute entreprise dialectométrique pour le domaine gascon aquitain est désormais l'Atlas linguistique et ethnographique de la Gascogne (A.L.G.VI).

Pendant que D. Philps poursuit dans leur ensemble et mène à terme les mesures concernant les Pyrénées Centrales, l'E.R.A. 352 Cnrs mène parallèlement une enquête ethnolinguistique en domaine méridional, sans coordination apparente avec l'enquête dialectométrique. C'est une des originalités essentielle de la méthode : cela consiste à dresser simultanément un atlas des pratiques dites linguistiques, et un atlas des pratiques dites ethnographiques ; sur ce dernier point, nous sommes très fortement influencés par les travaux des ethnographes de la tradition hongroise. L'idée nous est naturellement venue d'appeler ethnodialectométrie, la nouvelle discipline, en voie de création.

L'objet du présent travail est de présenter un historique de la genèse des méthodes dialectométriques, pour le domaine méridional, à partir de l'examen critique des thèses célèbres de l'abbé Lalanne pour le gascon maritime, et de J. Cremona, pour le gascon de la vallée d'Aure, et de la vallée du Louron, dans les Hautes Pyrénées. Nous exposerons successivement les données selon l'enquête classique menée par J. Cremona, puis les mesures dialectométriques effectuées sur cette base ; l'interprétation personnelle que nous proposons est entièrement fondée sur les cartes microdialectométriques réalisées d'après les maquettes de J. Cremona ; nous avons seulement dû renoncer à la représentation cartographique sémiotisée par poteaux frontières ; l'objectif de ce compte-rendu vingt ans après, est de dire pourquoi la dialectologie a mis si longtemps à accéder au rang de science humaine.

CARTE N° 1

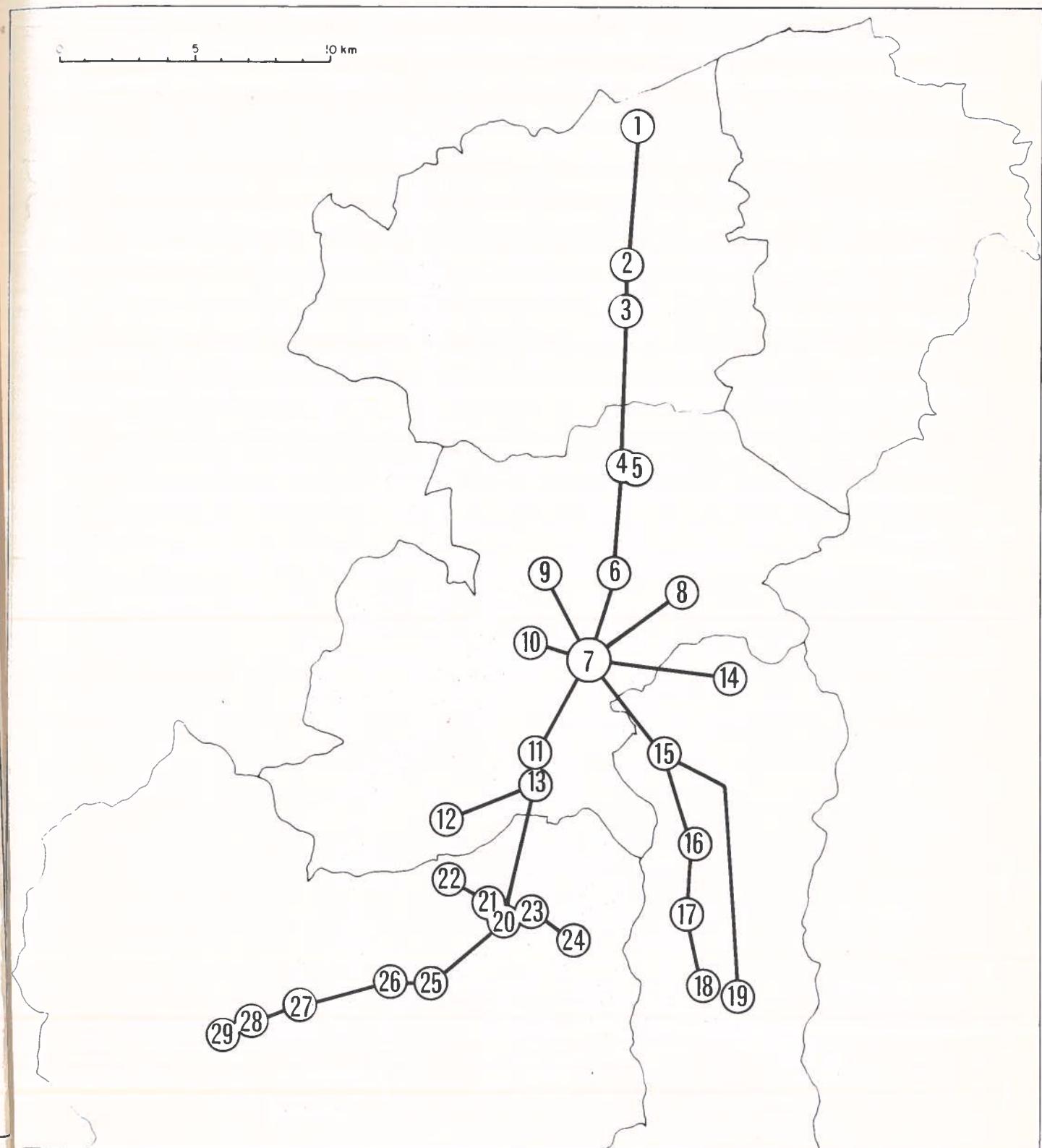
MICRODIALECTOLOGIE ET MICRODIALECTOMETRIE DE DEUX VALLEES DES PYRENEES CENTRALES:  
VALLEE D'AURE ET VALLEE DU LOURON



dessiné par R.ESPARBES Institut de Géographie, d'après une maquette de J.CREMONA (1956), pour les travaux de l'E.R.A 352 C.N.R.S Atlas pastoral pyrénéen

CARTE N° 2

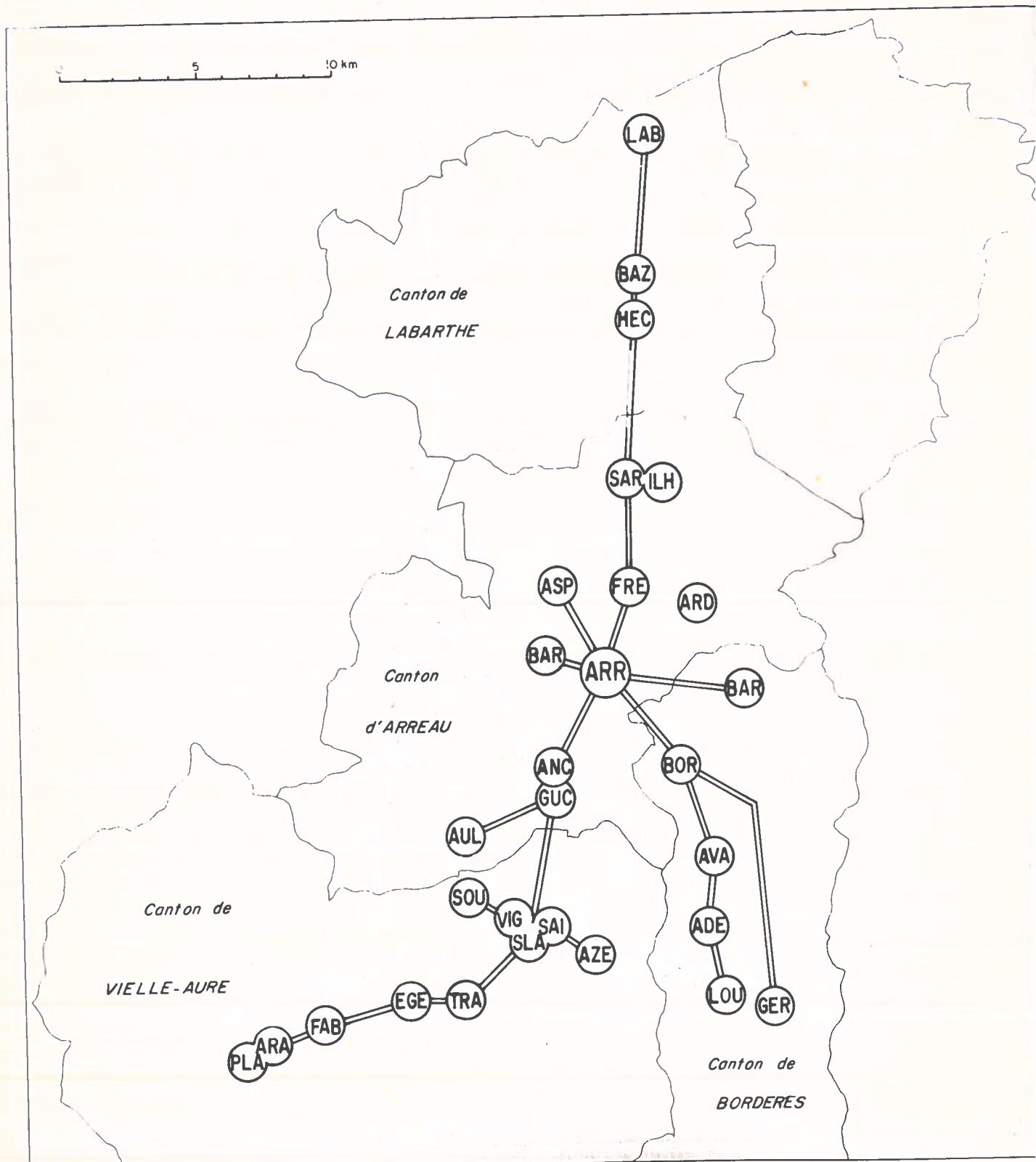
PREMIERE REPRESENTATION DU RESEAU D'ANALYSE MICRODIALECTOMETRIQUE  
VALLEES D'AURE ET DU LOURON



dessiné par R ESPARBES Institut de Géographie, d'après une maquette de J CREMONA (1956), pour les travaux de I E R A 352 C.N.R.S Atlas pastoral pyrénéen

CARTE N° 3

REPRESENTATION DU RESEAU DES COMMUNICATIONS NATURELLES POUR LES  
VALLEES D'AURE ET DU LOURON



CARTE N° 4

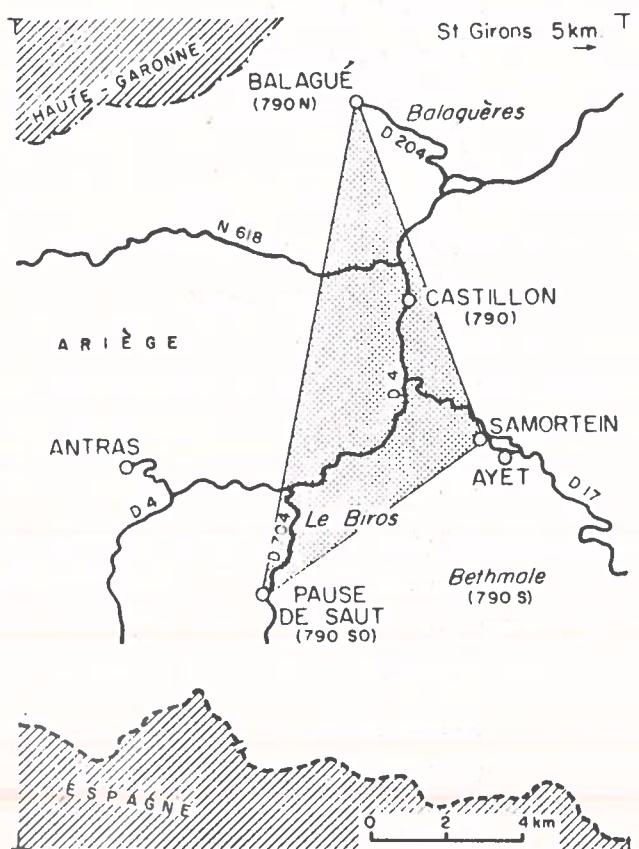
CADRE GENERAL DE L'OBSERVATION MICRODIALECTOMETRIQUE DANS LES PYRENEES  
CENTRALES: BALAGUERES-BETHMALE-BIROS .

Dans le cadre des travaux de microdialectométrie menés par les chercheurs de l'E.R.A. 352 C.N.R.S., Dennis Philps, publie, avec le concours du service des publications du Centre National de la Recherche Scientifique, ses travaux dialectométriques sur la zone castillonnaise des Pyrénées Ariégeoises. D. Philps a publié en 1976 dans le cadre des travaux de l'E.R.A. 352 l'ensemble des mesures qu'il a pratiquées manuellement, pour la zone des Pyrénées Centrales, qui constituent son champ d'observation actuel.

Ce second travail est publié en même temps que le manuscrit Saudinos, qui représente un document de première main important pour les études microtoponymiques des Pyrénées Centrales (Vallées de Larboust et d'Oueil).

La carte type établie par le service cartographique de l'Institut de géographie de l'Université de Toulouse Le Mirail, a valeur de prototype pour l'ensemble des observations à venir; c'est sur ce schéma que sera établi l'Atlas pastoral des Pyrénées Centrales, en voie de réalisation par l'E.R.A. 352.

D. Philps présente ici au grand public l'état actuel des études dialectométriques; le texte de présentation, rédigé en langue anglaise, a été présenté en communication lors de la Seconde conférence internationale de dialectologie de l'Université du Prince Edouard, à Charlottetown (Prince Edouard's Island, Canada) en juillet 1975.



Légende des unités ponctuelles du  
complexe Aure-Louron (65)

Sigle	N°	Nom de lieu	Sigle	N°	NL
ADE	17	Adervielle	TUZ	32	Tuzaguet
ANC	11	Ancizan	VIG	21	Vignec
ANE	33	Anères			
ARA	28	Aragnouet			Cette liste est un état des
ARD	8	Ardengost			enquêtes ponctuelles de J.
ARR	7	Arreau			Cremona(1956).
ASP	9	Aspin			Une enquête de l'E. R. A.
AUL	12	Aulon			352 menée en 1976 reprend
AVA	16	Avajan			systématiquement le même
AVE	30	Avezac			quadrillage.
AVEN	37	Aventignan			
AZE	24	Azet			
BAR	14	Bareilles			
BARR	10	Barrancoueou			
BAZ	2	Bazus-Neste			
BOR	15	Bordères-Louron			
BOU	40	Bourg-Bigorre			
EGE	26	Eget			
ESP	39	Esparros			
FAB	27	Fabian			
FRE	6	Fréchet-Aure			
GER	19	Germ Point ALG			
GUC	13	Guchen			
HEC	3	Héchettes			
ILH	5	Ilhet			
LAB	1	Labarthe-de-Neste			
LAU	34	St-Laurent-de-Neste			
LOU	18	Loudenvielle			
LOU 2		Loudervielle			
MON	36	Montégut			
NES	35	Nestier			
NIS	38	Nistos			
PLA	29	Le Plan			
PRA	31	Prat			
SAI	23	Sailhan			
SAR	4	Sarrancolin			
SLA	20	St-Lary			
SOU	22	Soulan			
TRA	25	Tramezaygues			

Distance phonétique : légende des items

Colonne N°		Paramètre N°	Convention adoptée
5		1	If b è d é =0 ; b e j é=1 ; b e d y é=2
6		2	Ky- =0 ; K-= 1 kyèbé;kèbé
7		3	-tr- =0 ; -t = 1 ; h u r t i g o,-t r i g o
8		4	NT =0 ; YT = 1 mentat;meitat
9		5	gw- = 0 ; w- = 1 guelh;uelh
10		6	br = 0 ; pr = 1 ombra;ompra
11		7	ng = 0 ; nk = 1 hanga;lenga;unpla
12		8	g r o y l o =0 ; g r a g u l'o =1
13		9	desbrelharse=0;desvelharse =1
14		10	Aradilho =0 ; rederilho =1
15		11	posieura =0;porsingla=1
16	2	12	Atonique +N : a conservé=0; o =1
17	2	13	-a final produit -e =0 ; -o =1
18		14	-as produit -as = 0;produit -es =1
19		15	-d- =1 ; -z- = 0
20	2	16	-s- produit /z/=0 ; /d/=1 gleisa
21	2	17	K+e produit /z/=0 ; /d/=1 cosina
22		18	k o r =1 ; k o = 0
23		19	k o r s = 0 ; k o s = 1
24		20	h e = 0 ; h y e = 1
25		21	hienudero =0 ; hienedura = 1

Commentaire : le total des items décomptés est de 21; les items 12,13,etc... sont comptés avec coefficient 1; les items 16,17,12 et 13 sont comptés avec coefficient 2.

L'enquête telle qu'elle est établie par Cremona, en 1956, fait mal le départ entre les paramètres phonétique et phonologique; la seconde phase de l'enquête qui sera microdialectométrique reprendra les items arrêtés par Jean Séguy dans l'Atlas linguistique de la Gascogne (volume VI).

Le coefficient de péréquation adopté permet de majorer les faits relatifs au système phonologique aurois.

Après péréquation, le nombre total d'items est de 25 pour le paramètre phonétique décrit par l'enquête de Cremona en 1956; notre mesure dialectométrique se fonde sur cette base.

Légende des items de la distance morphologique .

Colonne N°	paramètre N°	Identification et code
5	1	Imparfait : 0 = a y è w i ; 1 = a w y è i 2 = a w è i
6	2	Article 0 = u a ; 1 = w o
7	3	Article 0 = ed + V° ; 1 = er + V°
8,9	4	Prétérit 0 cantei,-eres 1 canteri 2 cantei,cantes
+0	5	Imparfait 2° conj. 0 = è w i ; 1 = y è w i
11	6	NIGRU 0 = néré ; 1 = nér

Commentaire de la distance morphologique.

Dans les résultats dialectométriques obtenus après dépouillement de l'enquête de J. Cremona(1956) la distance morphologique est stabilisée à 6 items dont 3 concernent la morphologie verbale.

Il s'agit en fait d'un choix des différences perçues,et non d'une description complète des traits de ressemblance et de différence extraits d'une observation de la totalité des phénomènes de morphologie nominale et verbale.

A seconde phase de l'enquête microdialectométrique,l'ALG VI servira de base tant pour la morphologie nominale que pour la morphologie verbale.

La première lecture du document dialectométrique établit que la distance morphologique est,pour l'ensemble de ces 6 paramètres,tout à fait stabilisée à trois paliers: 100 % ; 50 % ; 0 % .

Le lecteur se réfèrera en premier lieu,pour les opérations,à l'établissement du procès verbal des distances;en second lieu à la table du total des distances morphologiques;enfin à la table de distance morphologique appréciée en pourcentage.

Pour comparer la distance morphologique en pourcentage et la distance phonétique appréciée dans les mêmes conditions, on pratique une règle de peréquation ou de pondération;on se contentera ici de pratiquer la pondération sur les moyennes générales phonétique et morphologique.On procèdera de même pour la distance lexicale en pourcentage.

Légende des items de la distance lexicale pour les vallées d'Aure et de Louron (Pyrénées Centrales) d'après l'enquête de J. Cremona(1956).

Colonne N°	item N°	vocables	code	champ lexical
5	1	emporte passadera escorredou	0 1 2	champ primaire=habitat champ secondaire=vestibule de la maison
6	2	galatra soler	0 1	champ primaire=habitat champ secondaire = grenier
7	3	arrepos trepadér	0 1	habitat palier
8	4	bocau bocalér bocalhér bocalh	0 1 2 3	habitat animal/annexes entrée du fenil
9	5	gariero porsinglo	0 1	habitat animal : porcherie
10	6	cubat chilato sonto	0 1 2	ustensiles seau à traire en fer
11	7	sier,escher leitou	0 1	Champ 1 = lait champ 2 = 2ème petit lait
12	8	serou brulh gurrop 0	0 1 2 3	champ 1 = lait champ 2=recuit du petit-lait = pas de terme = 3
13	9	pon sarralho	0 1	habitat serrure
14	10	mentat d'arrat mentat d'audet tirahust	0 1	oiseaux chauve-souris

Colonne N°	item N°	vocables	code	champ lexical
15	11	bonsou houladou	0 1	Animaux domestiques: pathologie/nosographie "coup de sang au ventre des ruminants"
16	12	cauhopanso bridér	0 1	habitat humain centre-coeur du foyer
17	14	emballes per arrè ta plo	0 1 2	Champ 1=opposition = en vain coefficient 2
18	15	cap not	0 1	champ 1 =négation animée aucun, personne
19	16	assieta's sede's	0 1	opérations s'asseoir
20	17	brau tauré	0 1	Animaux d'élevage taureau
21	18	umbrè OPACU Amarè héredè 0	0 1 2 3 4	Ecosystème "pente donnant vers le Nord" = pas de terme.
22	19	trot trou troat	0 1 2	habitat fenil
23	20	trento,trenco eschat	0 1	ustensiles petite pioche
24	21	dalho dalh	0 1	ustensiles agraires faux
25	22	agalè agau	0 1	irrigation/agrosystème rigole

colonne N°	item N°	vocables	code	champ lexical
26	23	babol rosa	0 1	ethnobotanique coquelicot
27	24	papabelho piharonho	0 1	ethnozootechnie mésange
28	25	corrau cortelh	0 1	habitat = champ primaire cour de la ferme
29	26	cledo brono	0 1	objet pastoral claie de bercail
30	27	parech coupas	0 1	auge du porc
31	28	arrehierba agor	0 1	pratiques agraires regain
32	29	andutou budet	0 1	alimentation carnée saucisson de porc
33	30	harri sapou tolou	0 1 2	ethnozootechnie crapaud
34	31	artigoto legonho ticoino	0 1 2	ethnobotanique pissenlit 2
35	32	solo arralho cara solelh	0 1 2	Ecosystème "pente donnant vers le Sud"
36	33	debara baisha	0 1	opérations descendre
37	34	arrastelher arrastilher arrastet	0 1 2	ustensiles agraires

La ventilation des vocables se fait selon les champs ethnographiques; c'est pourquoi la distance lexicale intéresse principalement l'ethnolinguiste.

COURBES DES DISTANCES GLOBALES APPRECIÉES EN POURCENTAGE POUR LES VALLEES  
D'AURE ET DE LOURON

Ligne N°	Total des distances	%	Ligne N°	Total des distances	%
L 1	1451	40,3	28	544	60,44
L 2	1395	39,8	29	541	67,6
L 3	1393	40,9	30	0	0
L 4	1045	31,6	31	0	0
L 5	1030	32,18	32	0	0
L 6	912	29,4	33	0	0
L 7	896	29,8	34	0	0
L 8	979	33,75	35	0	0
L 9	877	31,3	36	0	0
L 10	814	30,14			
L 11	845	32,5			
L 12	821	32,8			
L 13	827	34,45			
L 14	1148	49,9			
L 15	1134	51,5			
L 16	1117	53,19			
L 17	1099	54,95			
L 18	1117	58,78			
L 19	1120	58,94			
L 20	523	30,7			
L 21	526	32,8			
L 22	523	34,8			
L 23	542	38,7			
L 24	550	42,6			
L 25	514	42,8			
L 26	522	47,45			
L 27	526	52,6			

Moyenne générale de la distance globale ramenée en pourcentage = 38,05 % .

Commentaire : A ce niveau, aucune pondération des différents paramètres ne peut être opérée; la pondération ne peut se faire qu'à partir des moyennes paramétriques; on se réfèrera au dépouillement de D. Philps, pour l'ensemble des Pyrénées Centrales, Toulouse, 1976.



DISTANCE LEXICALE

### TOTALS DE LA DISTANCE PHONETIQUE

## TOTALS DE LA DISTANCE MORPHOLOGIQUE

## TOTALS DE LA DISTANCE LEXICALE

DISTANCE CUMULEE (Phon. + Morphologie + lexique) =65 items

## MOYENNE PARAMETRIQUE DE LA DISTANCE PHONIQUE

MOYENNE PARAMETRIQUE DE LA DISTANCE MORPHOLOGIQUE

## MOYENNE PARAMETRIQUE DE LA DISTANCE LEXICALE

DISTANCE MOYENNE GÉNÉRALE

### Le cadre théorique des recherches ethnodiaclectométriques

L'intitulé implique une double pratique ; pratique des mesures ethnographiques ; pratique des mesures linguistiques ; la première pratique est conditionnée par la réalisation, puis l'exploitation d'un atlas des pratiques par des ethnologues professionnels ; ceux qui recherchent des sujets de stages pratiques pour les DEA de 3<sup>ème</sup> cycle collectifs, auraient là un beau sujet à réflexion, puis matière à action ; il suffirait de convaincre deux ou 3 formations et/ou laboratoires de s'associer pour pratiquer une politique documentaire cohérente et coordonnée.

La seconde pratique concerne la mesure du fait dialectal proprement dit.

La dialectologie naïve ou réaliste s'est fixé comme objectif la détermination des limites objectives des variations spatiales ; ainsi, de manière plus ou moins crédible, un atlas linguistique permet de dresser le bilan chiffré des différences pour les unités phonématisques dans une zone donnée.

En ce qui touche au domaine gascon, l'abbé Lalanne est un des tout premiers chercheurs à avoir orienté la dialectologie vers la dialectométrie ; je ne reviens pas sur ce fait souligné dans l'avant-propos de l'étude de D. Philps, Balaguères-Bethmale, Biros, Etude dialectométrique, Toulouse, 1976, Travaux de l'ERA 352 CNRS, publiée avec le concours du Centre National de la Recherche Scientifique.

Dans son analyse, Lalanne donne une carte composite des aires phonétiques, dont les chiffres permettent une appréciation quantitative des éléments constitutifs de la phonétique du parler gascon maritime landais.

Dans son travail monographique sur le parler gascon de la Vallée d'Aure, Cremona procède de façon similaire, pour opposer la variété dialectale auroise aux variétés de gascon du triangle aquitain.

La procédure est alors la suivante, dans l'état actuel des choses.

Tout d'abord, nous disposons à présent, avec l'ALG VI d'un outil de mesures générales, qui permet de mesurer, point à point, pour le domaine gascon, les distances appréciées en chiffres ramenés ou non à des moyennes paramétriques ; D. Philps, pour le domaine gascon pyrénéen, a pratiqué l'ensemble de ces mesures et établi que la distance moyenne pyrénéenne pour les Pyrénées Centrales était de 26,6 %.

Le travail présenté par J. Cremona, en 1956 prétendait déduire d'une observation des caractéristiques, une définition théorique du dialecte aurois.

Pour cela, il convenait de déterminer objectivement - et non de choisir, les traits caractéristiques qui opposent le gascon dans sa variété auroise aux parlers voisins ; l'ensemble des traits communs de ressemblance typologique, n'est donc pas pris en considération par l'analyse de Cremona ; on néglige par principe tous les faits communs : chute de -n- intervocalique etc...

La notion de fait commun aquitain s'accorde très bien d'une représentation quantitative par recours aux méthodes de cartographie de densité ; sur ce point on

renvoie aux analyses de H. Goebel (cr. de P. Bec, les interférences linguistiques entre gascon et languedocien) in Zeitschrift für Rom. Phil. 1973, 89/4-6, 557-588 ; on se référera aux cartes dressées par H. Goebel, d'après les dépouillements de l'ALG I-IV. Dans la mesure où l'analyse comporte des chiffres, on peut la considérer comme pratique dialectométrique, à sa manière.

En microdialectométrie, à présent, l'objet est d'étudier, à micro-échelle, l'homogénéité, pour un espace naturel donné, de phénomènes qui caractérisent la zone mise à l'étude. La question est de savoir quelle est l'importance de ces caractéristiques. On pourrait la dire grande ou petite ; l'objet de la mesure que nous avons pratiquée sur la base des enquêtes de J. Cremona, était de substituer des chiffres, et plus précisément des moyennes paramétriques aux adjectifs grand et petit, qui n'ont pas de contenu stable.

On peut admettre, pour ce qui est du cadre théorique de l'enquête, que Cremona a pratiqué un questionnaire comportant assez de vocabulaire pour faire apparaître de façon crédible, la présence de traits, qui, sans créer un état anarchique, démarquent systématiquement les gascons aurois de leurs voisins gascons de Comminges.

La mesure donnée n'est donc pas la mesure des traits caractéristiques du gascon, tel qu'il est considéré par un chercheur anglo-saxon qui veut en avoir quelque idée vague ; c'est la mesure des traits qui constituent l'idiolecte aurois ; j'appelle idiolecte aurois la langue parlée par l'individu collectif de la zone considérée ; l'analyse de Cremona porte sur 37 localités.

L'observation des faits est apparemment décevante ; la lecture des tables de moyennes paramétriques, et plus spécialement la lecture des plages de distances, faite à partir des matrices, établit qu'il n'existe pas un dialecte aurois ; il n'existe pas d'aire dialectale auroise ; il existe une aire de bilinguisme ou de trilinguisme, dont la distance globale moyenne est de 38 % ; en d'autres termes, l'enquêteur qui, à Labarthe-de-Neste, pénètre en vallée d'Aure, rencontre, au cours de son voyage ethnologique au moins deux langues différentes, sinon 3, et des zones de relative sécurité/insécurité linguistique.

Pour la dialectologie, le travail est terminé de manière crédible lorsque le dialectologue a planté les poteaux-frontières du dialecte I et du dialecte II.

Nous disposons, avec le travail de Cremona, d'une première mesure de ce type réalisée en 1956.

L'ethnodialectométrie, pour nous, est une pratique qui consiste à apprécier et mesurer le changement social et culturel intervenu, en répétant 20 ans après la même mesure, et dans des conditions identiques. Le changement observé peut se ramener sous l'aspect théorique, aux contacts qui sont intervenus entre plusieurs dialectes contigus, en état de diglossie/bilinguisme ; le dialecte de la basse vallée d'Aure a pu tendre à assimiler le dialecte de la Haute Vallée d'Aure, et modifier ainsi le profil des frontières linguistiques et/ou culturelles.

Quelles hypothèses peuvent se présenter en cours d'analyse, doit être examiné. En première hypothèse, les traits généraux de la langue gasconne (en morpho-syntaxe, et sur le plan des unités phonématisques) seraient altérés ; le gradient de -n- intervocalique présent serait passé de 25 % à 75 % etc...; ceci ne peut être connu de façon rigoureuse que par la pratique d'une méthode qui échantillonnerait les locuteurs de façon à faire apparaître de manière crédible leur caractère de locuteur collectif, notion qui paraît répugner à certains, et qui pourtant peut seule fonder la notion d'aire dialectale, dans notre perspective. Le nombre de ces traits est fourni par l'ALG ; l'enquête complémentaire d'ethnodialectométrie doit donc reprendre pas à pas les mesures paramétriques de l'ALG VI.

En seconde hypothèse, ces traits généraux persistent avec une plus ou moins grande stabilité, à apprécier en chiffres.

En troisième hypothèse, les traits idiolectaux aurois, au nombre de 65 selon l'enquête de Cremona menée en 1956, sont stabilisés ; ces traits idiolectaux par conséquent continuent à opposer le groupe culturel aurois aux groupes culturels voisins de Bigorre ou de Comminges. On pourra parler, en de telles conditions, de persistance d'une identité culturelle auroise.

Dans le cas où ces traits idiolectaux seraient gravement altérés, on pourrait dire - et ceci est une hypothèse à examiner - qu'il y a déperdition de personnalité collective ; un locuteur qui perd plus de N % de ces 65 traits idiolectaux, cesse d'appartenir à l'échantillon représentatif du groupe culturel aurois ; on doit ainsi admettre, en zone d'insécurité linguistique, créée par le contact de deux variétés dialectales, qu'un individu collectif peut voir se diluer sa personnalité ; la théorie linguistique pour nous s'identifiant à la pratique, selon une doctrine qui nous est chère, on constate que, selon le bilan de l'enquête de Cremona, la tolérance des écarts, c'est-à-dire des personnalités collectives originales est très grande dans un groupe culturel homogène ; la lecture des plages de distances permet de remplacer grand/petit/moyen, par des chiffres.

Il peut se faire que pour des raisons historiques précises, les personnalités originales disparaissent ou soient de plus en plus diluées dans l'unité d'une norme gasconne ; le résultat admirable serait qu'il n'y aurait plus de différence entre un individu collectif aurois et un individu collectif gédrois ; on se ferait sans doute à juste titre en 1976 traiter de mauvais plaisant en affirmant qu'il n'y a plus de distance entre Gèdre et Germ ; la distance moyenne linguistique est nécessairement de 26,6 % ; le seuil de 38 % (d'après Cremona) s'explique en partie par le privilège accordé à la distance ethnographique, par le biais du vocabulaire.

Une bonne partie de nos analyses précédentes permet de conclure que par une procédure du type de celle de Cremona, il est impossible d'aboutir à déterminer les structures phonétiques et morphosyntaxiques du gascon aurois, puisqu'on n'en retient que les éléments apparemment exotiques.

Ceci appelle la réflexion suivante ; lorsque Cremona détermine 6 localismes ou exotismes aurois pour la phonétique en 1956, il étudie la stabilité de ces 6 traits caractéristiques ou particuliers, qui assurent, dans un voisinage donné, une communauté de pratique ; vélariser /a/ au contact de N est une pratique au même titre que tondre les bourrecs de telle ou telle manière est une pratique ; la carte qui dresse les poteaux frontières du fait linguistique est donc une carte qui fait partie intégrante d'un atlas des pratiques, tel que nous le définissons ; il est extrêmement fâcheux, pour l'avenir des études ethnodialectométriques que l'on ait pris la déplorable habitude de considérer comme activité et comme pratique mineure celle qui consistait à décrire comment dans un groupe de voisinage, sinon dans un milieu unifié par une pratique collective ou collectiviste, on tond les moutons ; la carte de cette pratique est au moins d'importance égale à la carte de la pratique phonétique, en principe ; si une pondération devait intervenir, rien ne dit qu'il ne faut pas privilégier le coefficient à donner aux pratiques du second type ; étant entendu qu'il n'existe pas de différence théorique entre nommer borrec un "mouton de 1 an" et pratiquer sur lui tel type de tonte (cul tonedura).

La notion de particularisme, d'exotisme, telle qu'elle se déduit aisément des analyses des dialectologues est suspecte ; la notion de communisme pastoral peut l'être tout autant ; pourtant, cette pratique collectiviste est prouvée historiquement par l'analyse des documents ethnographiques établis par l'enquête de Saudinos, pour le groupe culturel Larboust-Oueil (cf. Toponymie du canton de Bagnères-de-Luchon, manuscrit Saudinos, publié et commenté par J. L. Fossat et D. Philps, Travaux de l'ERA 352 Cnrs, 1976) ; une étude des structures familiales et de la démographie jusqu'au XIXème s. viendrait heureusement conforter nos positions et hypothèses.

Il s'agit de savoir si les limites des particularismes sont ou non des limites de nature linguistique ou non ; un particularisme pan-gascon peut recouvrir un fait ethnique ; ceci serait de nature à encourager les travaux sur l'hypothèse substratiste au sein du Laboratoire d'Etudes Méridionales. Un particularisme limité à une aire donnée, renvoie à une classe d'hommes réunis autour de pratiques extralinguistiques identiques ; une somme de particularismes est nécessaire pour garantir la stabilité de la personnalité collective du groupe culturel.

## EXAMEN DES TABLES DE DISTANCE

La lecture des matrices fait apparaître la présence de zones de relative stabilité.

Il suffit, pour exploiter en ce sens les matrices de masquer une à une les zones de concentration.

D'une part, on a reporté les distances paramétriques, appréciées en totaux, sur les cartes ; la cartographie est un premier moyen de lecture des zones de stabilité de la distance.

Le second moyen de lecture, s'établit de manière très artisanale, par masquage sur la matrice elle-même.

### DISTANCE PHONETIQUE.

On peut envisager, pour une meilleure exploitation, une représentation par cartographie de densité, à partir de la matrice elle-même.

Soit une zone caractérisée par 0 traits de distance phonique ; une zone représentée par 18 traits de distance phonique, sur un total de 25 items de la matrice phonétique dans les conditions de l'enquête ; le premier chiffre représente un minimum ; le second chiffre un maximum ; on voit ainsi se dessiner des zones de sécurité, ou d'insécurité minimale, maximale et moyenne ; ce concept d'insécurité linguistique est fondamental pour les études d'attitude et les analyses de contenu pratiquées par différentes écoles sociolinguistiques ; son application en microdiialectologie pourrait donner de bons résultats.

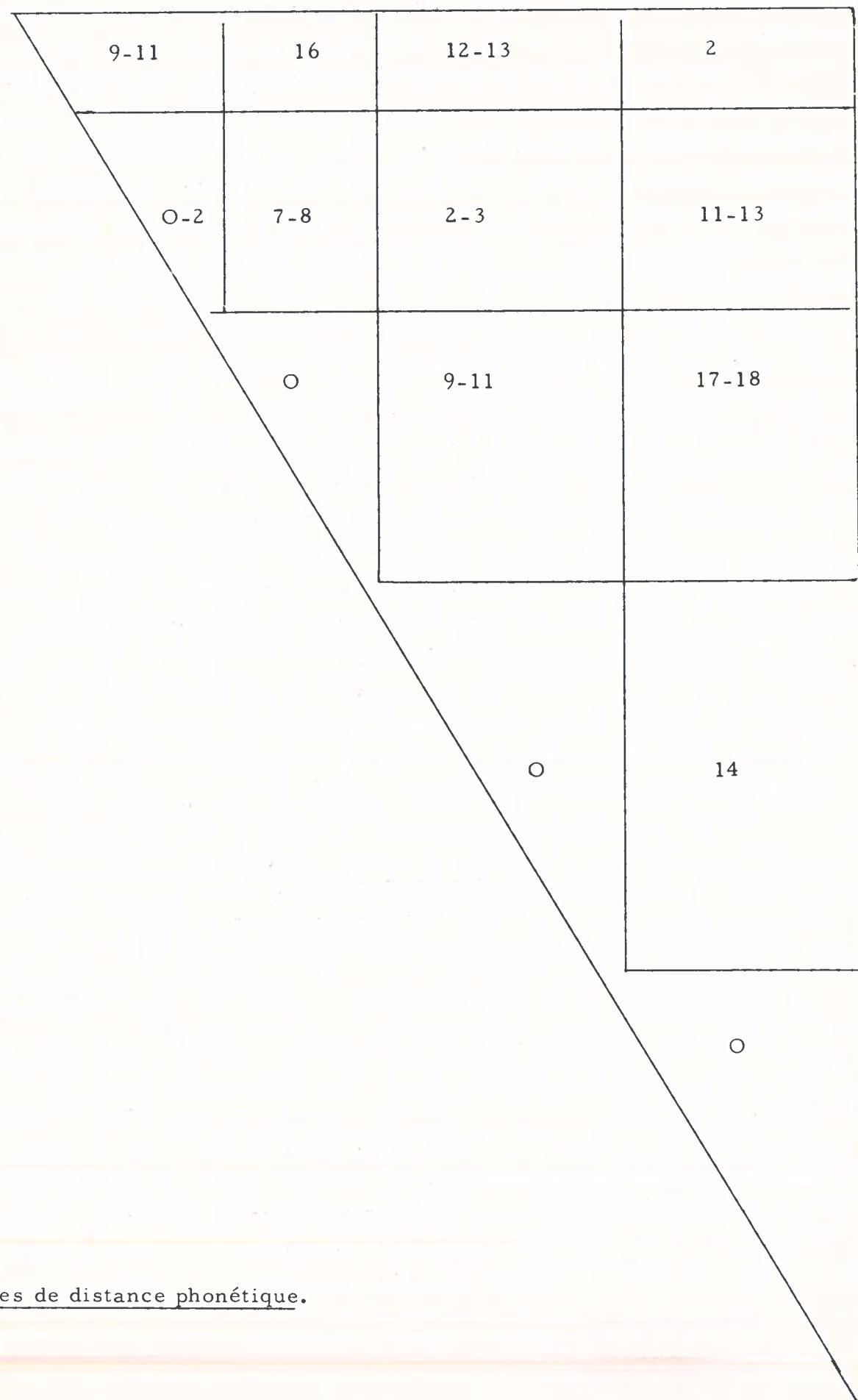
Pour ce qui est de la distance phonétique, le bilan est arrêté à 8 zones représentables en gradient (cf. ALG VI, gradient de la gasconité).

La zone de distance phonétique maxima le correspond à ce que Jean Séguy a appelé zone du plus fort gradient de gasconité.

Zone N°	Nombre de traits différentiels	Sécurité
1	0	
2	0 à 3	maximale
3	7 à 8	
4	9 à 11	moyenne
5	12 à 13	
6	14	
7	15 à 16	
8	17 à 18	minimale

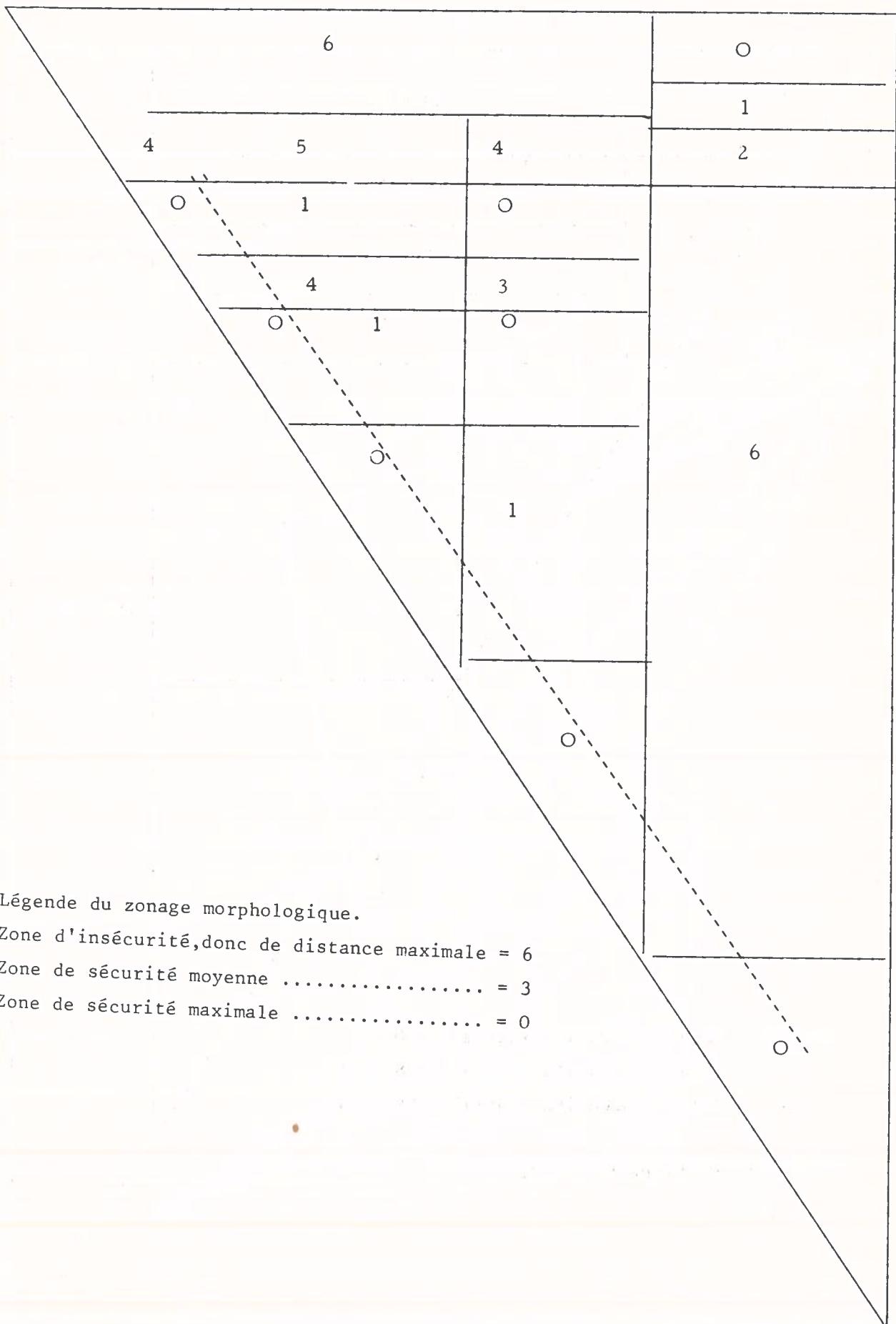
La lecture des distances paramétriques ainsi pratiquée est un des moyens les plus efficaces de déabstraction de la table de la distance globale cumulée.

## Représentation du zonage sur la matrice



Plages de distance phonétique.

Observation de la matrice de la distance morphologique



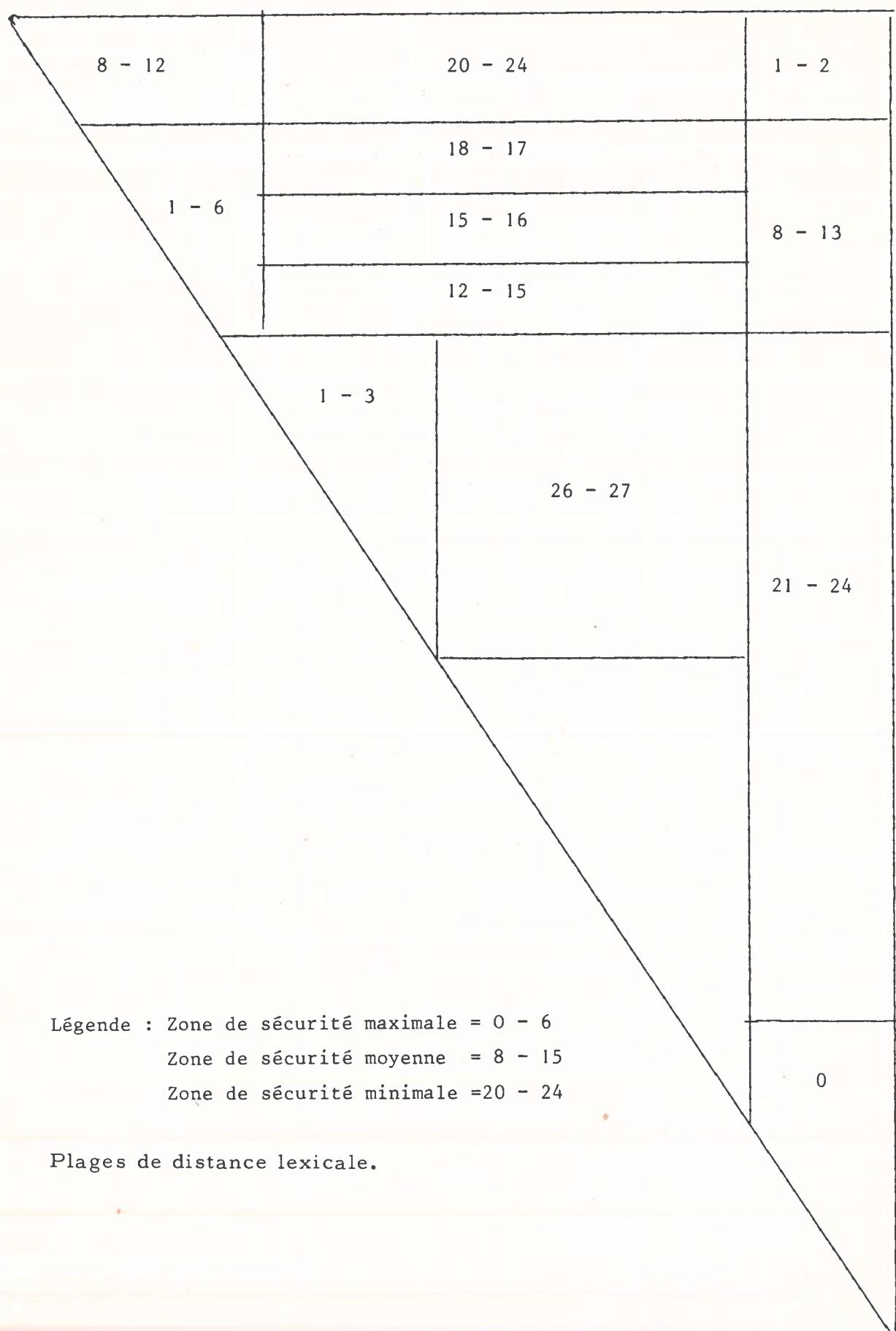
Légende du zonage morphologique.

Zone d'insécurité, donc de distance maximale = 6

Zone de sécurité moyenne ..... = 3

Zone de sécurité maximale ..... = 0

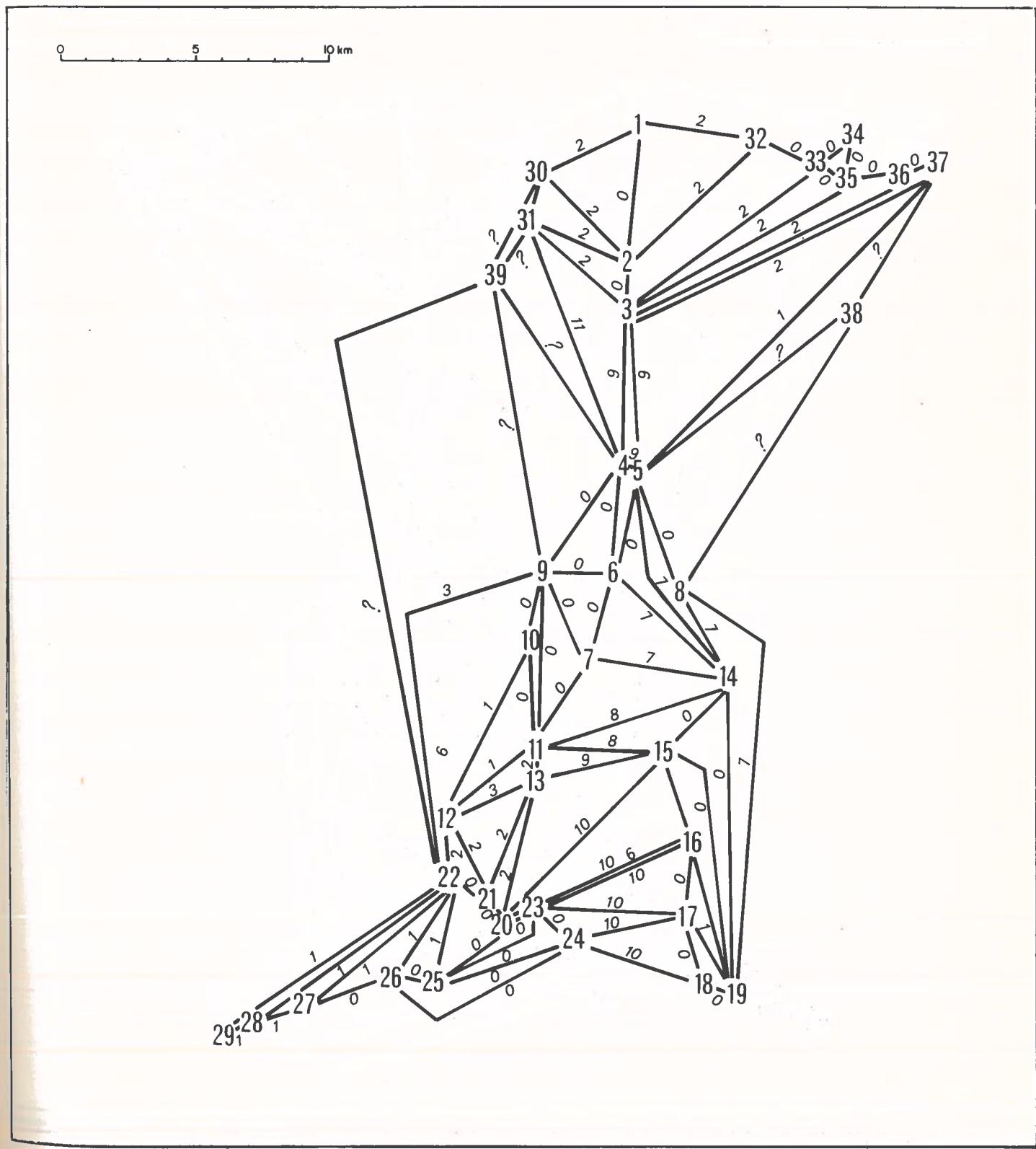
Observation de la table des distances lexicales sur 34 items.



CARTE N° 5

DISTANCE PHONETIQUE APPRECIÉE EN CHIFFRES

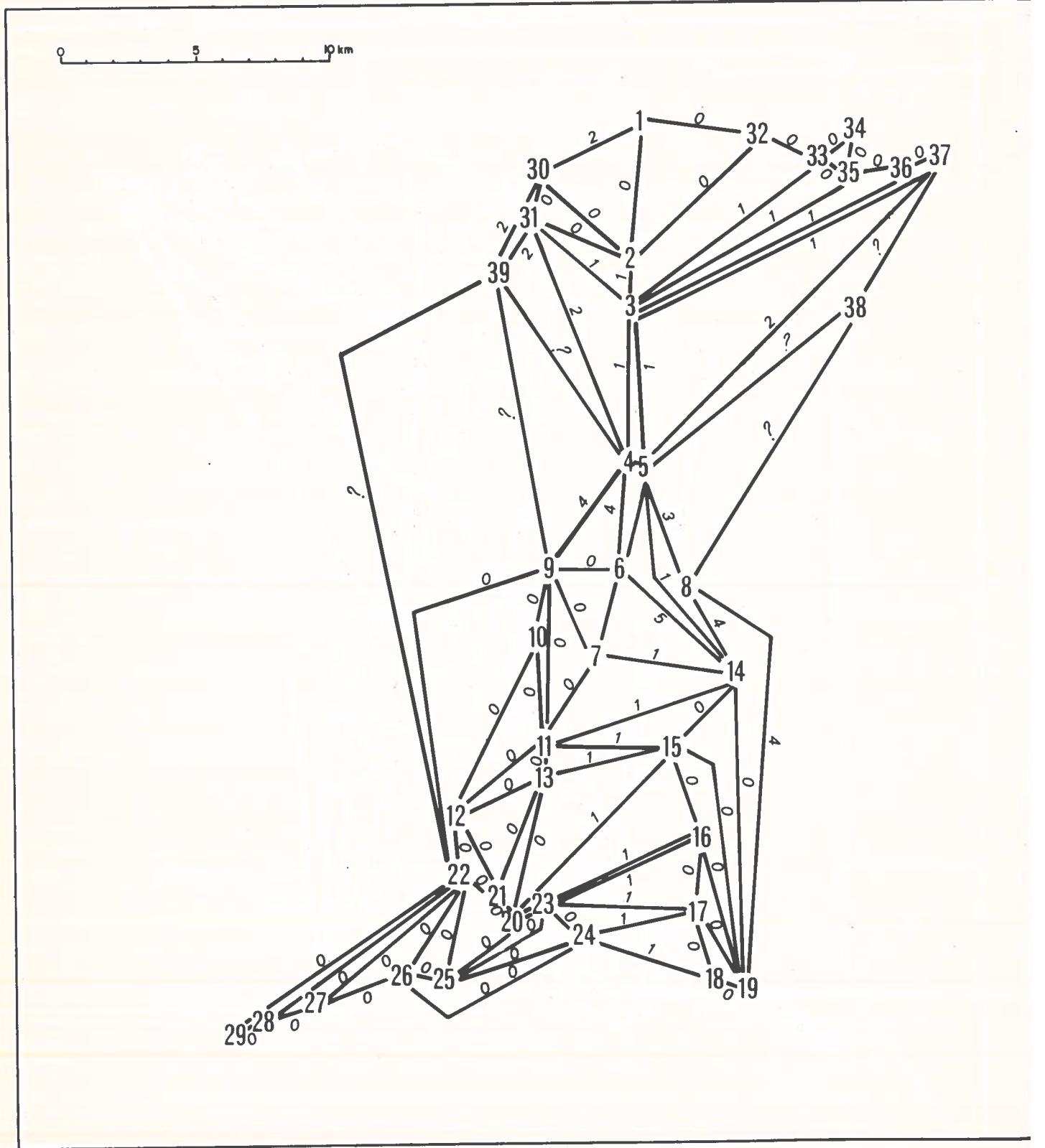
Stabilité et sécurité linguistique en vallées d'Aure et du Louron  
d'après les données de J. Cremona(1956)



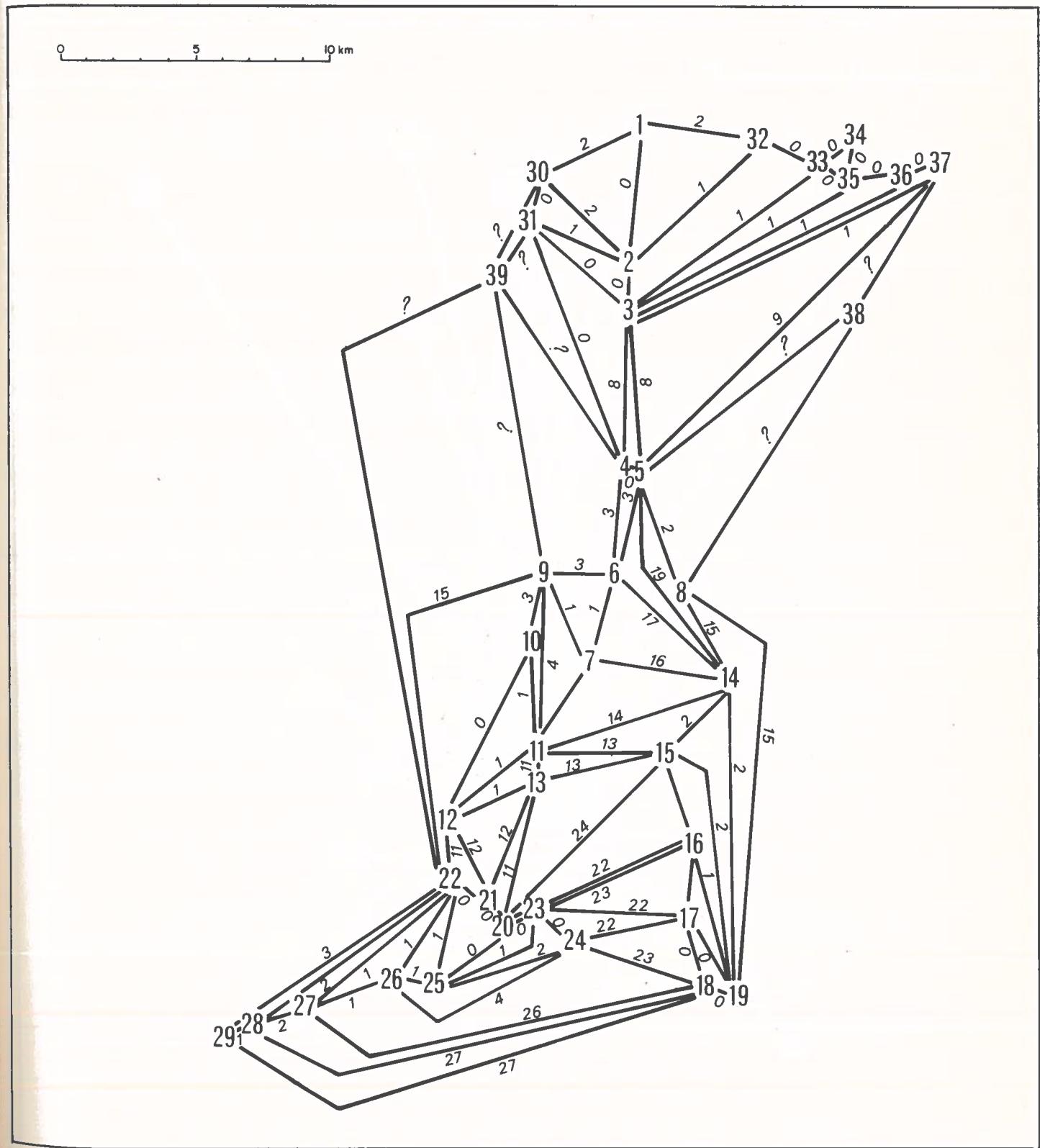
CARTE N° 6

DISTANCE MORPHOLOGIQUE APPRECIÉE EN CHIFFRES

Stabilité et sécurité morphologiques en vallées d'Aure et du Louron  
d'après les données de l'enquête de J. Cremona(1956).



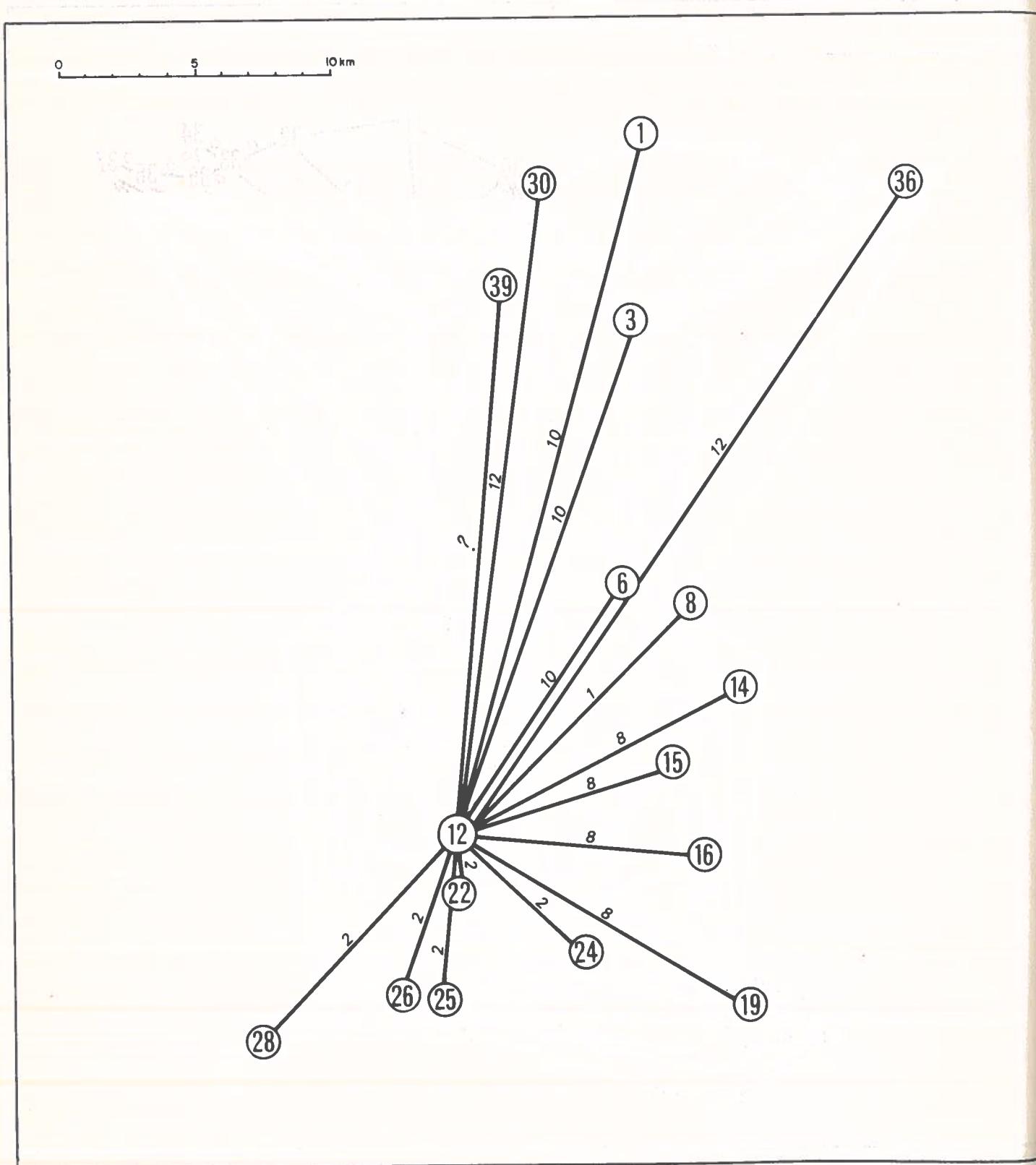
CARTE N° 7  
DISTANCE LEXICALE APPRECIÉE EN CHIFFRES  
Stabilité et sécurité lexicale en vallées d'Aure et du Louron  
d'après les données de l'enquête de J. Cremona (1956).



CARTE N° 8

DISTANCE PHONETIQUE APPRECIÉE EN CHIFFRES

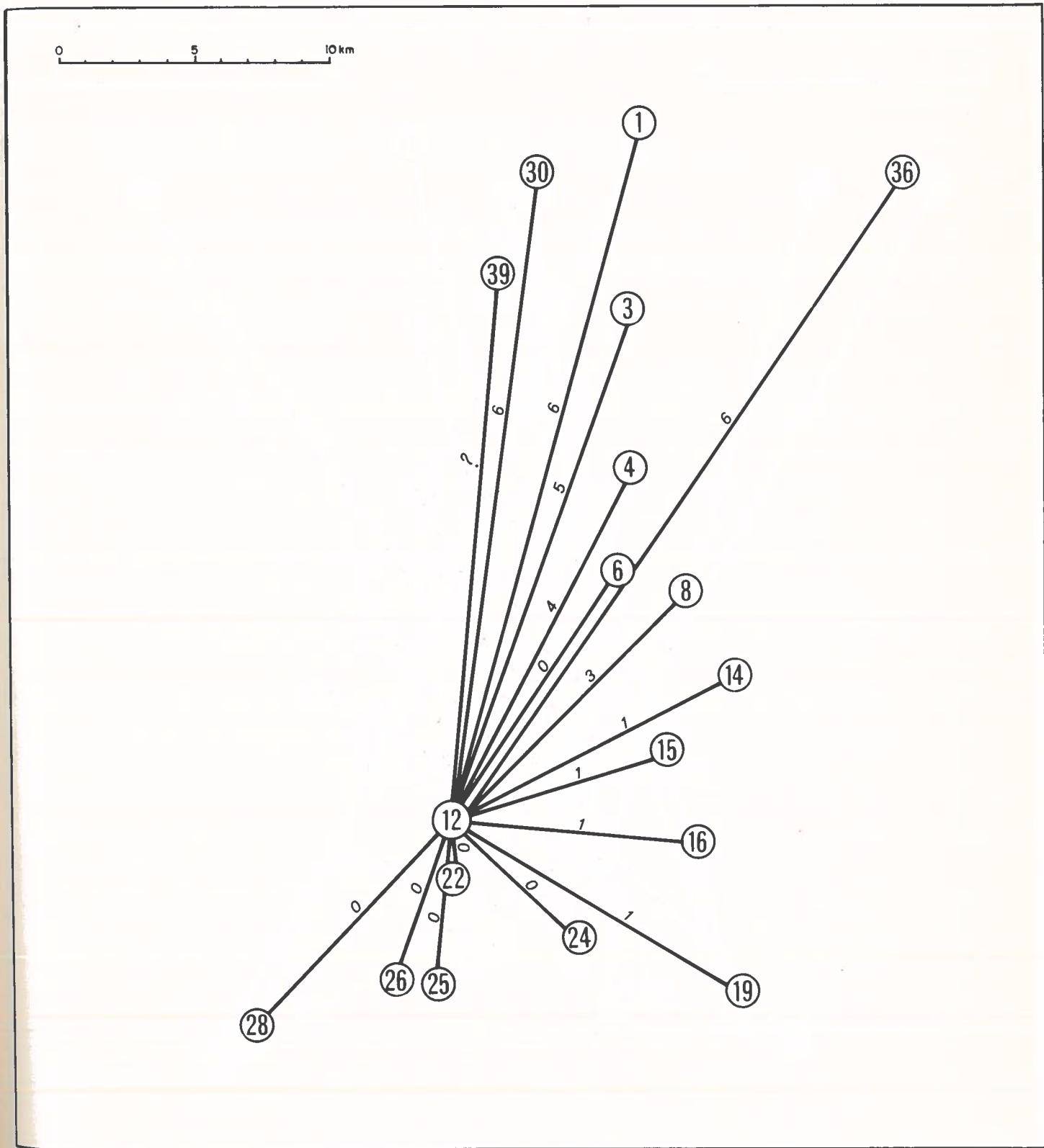
Appréciation quantitative de l'insécurité linguistique en vallées  
d'Aure et du Louron d'après les données de J. Cremona (1956).



CARTE N° 9

DISTANCE MORPHOLOGIQUE APPRECIÉE EN CHIFFRES

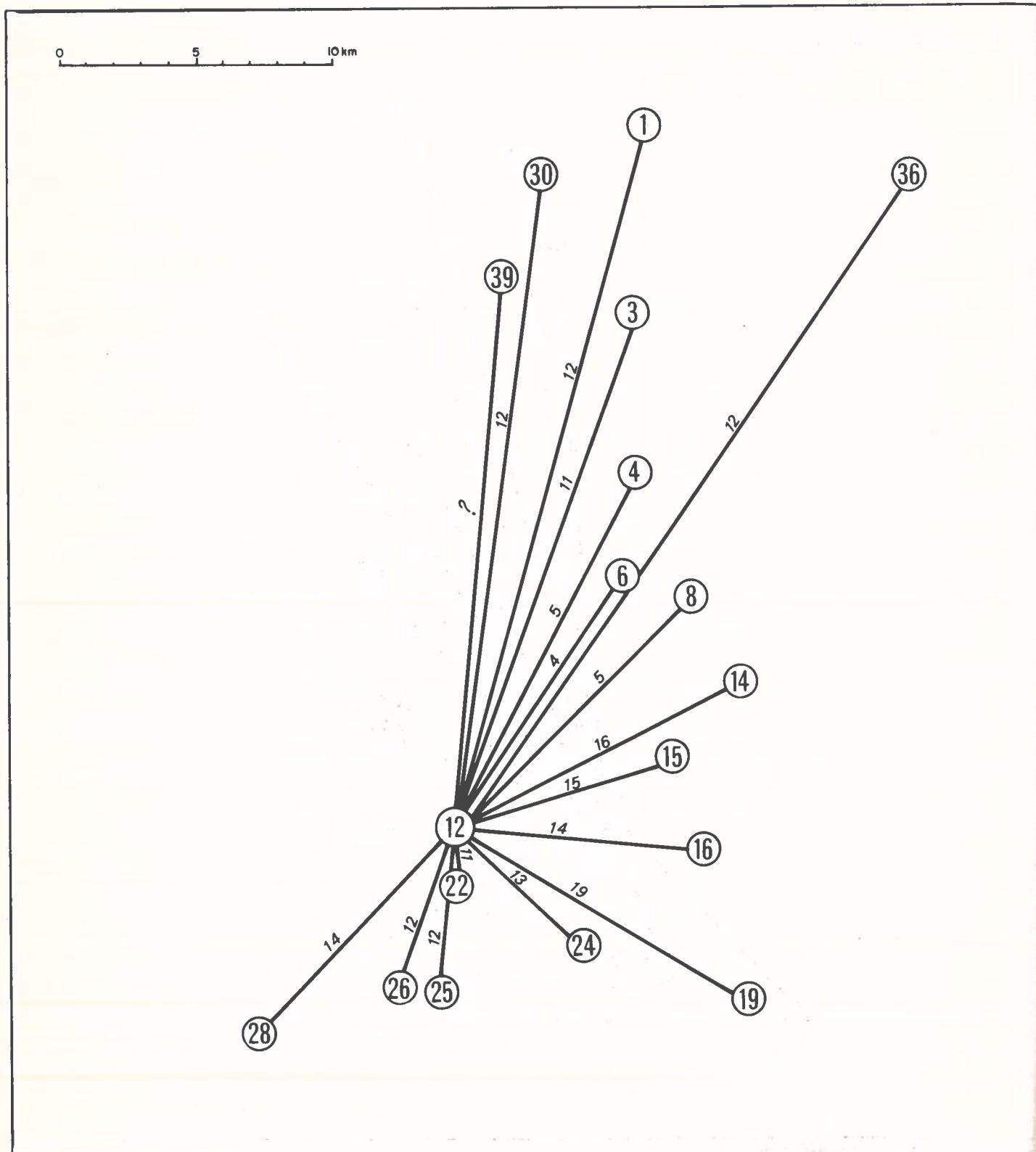
Appréciation quantitative de l'insécurité morphologique en vallées d'Aure et du Louron, d'après les données de J. Cremona(1956).



CARTE N° 10

DISTANCE LEXICALE APPRECIÉE EN CHIFFRES

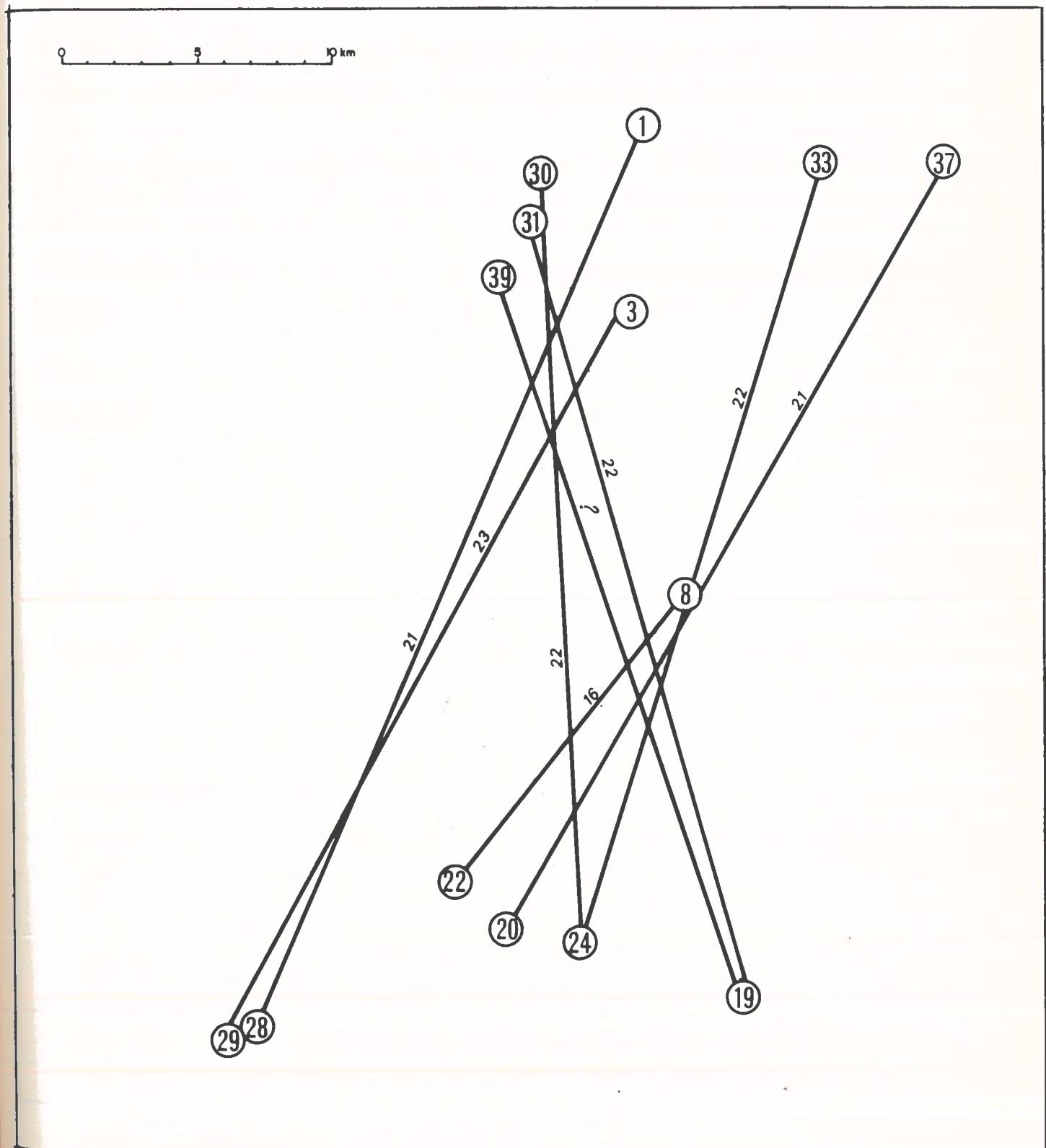
Appréciation quantitative de l'insécurité lexicale dans les vallées d'Aure et du Louron d'après les données de J. Cremona (1956)



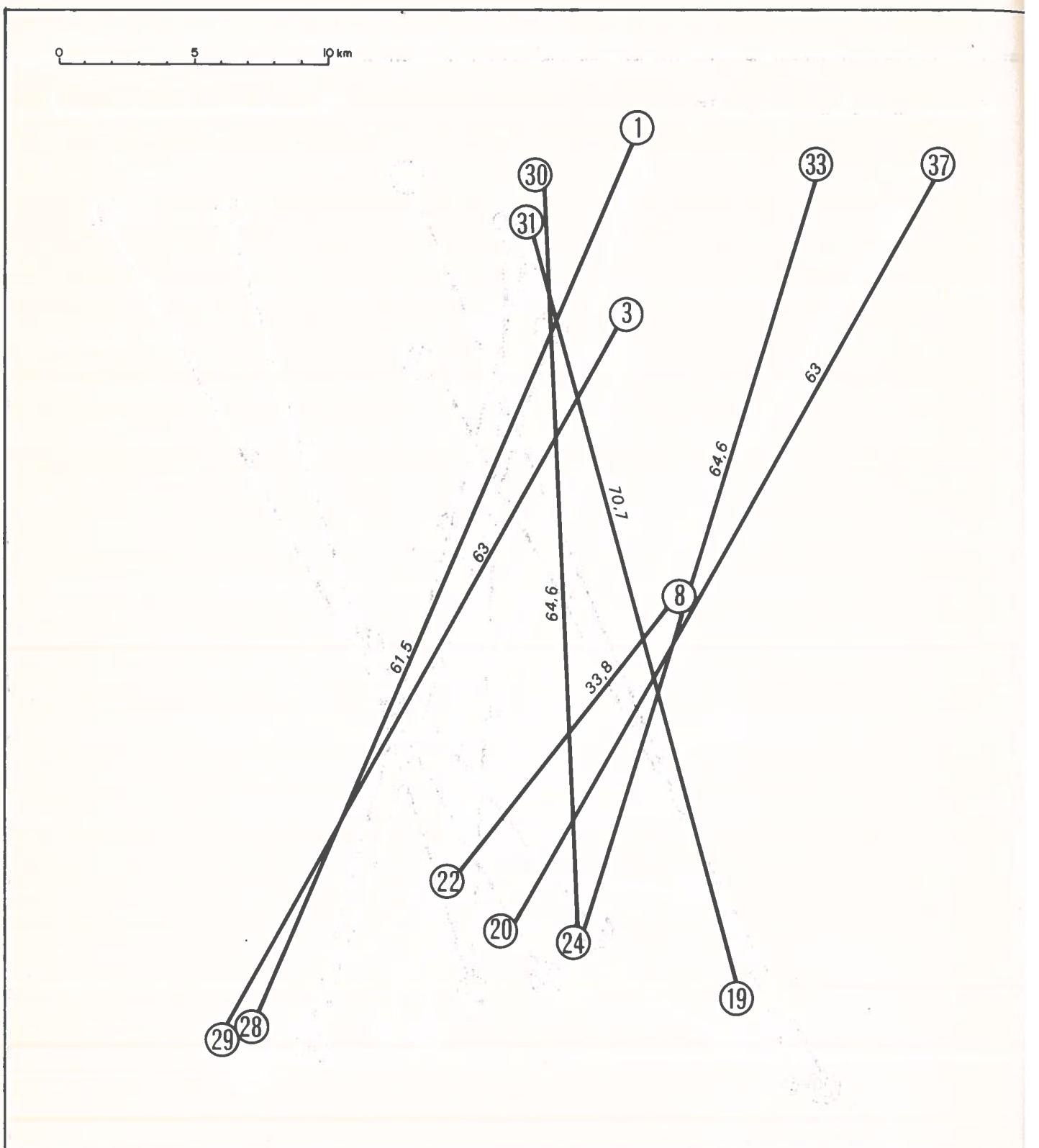
CARTE N° 11

CARTOGRAPHIE DE LE DISTANCE LEXICALE MAXIMALE

Appréciation quantitative de l'insécurité lexicale maximale  
d'après les données de l'enquête dialectologique de J. Cremona(1956).



CARTE N° 12  
Distance globale en pourcentage



CARTE N° 12

DISTANCE GLOBALE MAXIMALE APPRECIÉE EN POURCENTAGE

Appréciation de l'insécurité linguistique maximale en vallées d'Aure et du Louron,d'après les données de l'enquête de J. Cremona(1956).

Ceci signifie que sur un ensemble de traits différentiels perçus, arrêté arbitrairement à 65 par J.Cremona, les unités géographiques 1 et 29 se démarquent à 61,5% et ne signifie rien d'autre; la distance moyenne calculée sur cette base est de 38,05%; comme la distance moyenne intrapyrénéenne est de 26,6% selon les calculs de J. Séguys, vérifiés par Dennis Philps(cf. Toponymie et microdialectométrie des Pyrénées Centrales,ERA 352 CNRS,1976) on peut estimer à 11,45% en moyenne le coefficient de ressemblance générale non évalué par les résultats de l'enquête de J. Cremona(1956). On comparera ces résultats à la carte du gradient de gasconité établie par J. Séguy pour la vallée d'Aure et celle du Louron.

Cette enquête pose de façon nette aux chercheurs plusieurs problèmes.

Le premier concerne la procédure de questionnaire microdialectométrique à mettre en place pour les travaux à effectuer dans les Pyrénées en général,d'est en ouest de la chaîne; il faudrait imaginer un questionnaire pratique qui fasse apparaître avec certitude l'ensemble des items tels que les a calculés J. Séguy,pour l'établissement des moyennes paramétriques(ALG VI,volume annexe introduction).

Le second problème, qui n'a pas trouvé sa solution, reste le caractère arbitraire de la mesure de la distance lexicale ouverte, comparée aux distances fermées de la phonétique, de la phonologie et de la morphologie.

Certains chercheurs novateurs voudraient inclure au calcul des distances un calcul des différences ethnographiques(appréciation quantitative des différences relatives aux pratiques; les croyances n'étant qu'une pratique parmi d'autres); un des objets de cette étude est de proposer une réflexion collective pour l'élaboration du cadre théorique d'une telle entreprise,pour la France Méridionale.

Inclure à nos calculs un coefficient de distance tonométrique couronnerait l'entreprise,mais ceci nous paraît encore relever du futurisme,et nous ne voyons pas bien comment cette mesure pourrait se constituer en paramètre autonome.

Au stade de l'interprétation,reste un sérieux problème; il nous faudra un jour confronter les mesures proprement dialectométriques et des mesures pratiquées par d'autres disciplines(démographie ,économie politique,économie,histoire des religions et des mentalités etc...); on sait l'intérêt de telles confrontations telles que les pratiquent certains chercheurs; une telle confrontation mériterait une concertation plus systématique des entreprises scientifiques trop souvent individuelles.

Enfin,l'ensemble de nos analyses actuelles milite fortement en faveur d'un traitement automatique des données en vue de l'analyse dialectométrique.

TRAVAUX DE L'INSTITUT D'ETUDES MERIDIONALES

E. R. A. 352 C. N. R. S.

Dialectometry: The Scientific Measurement  
of Linguistic Distance .

by

Dennis PHILPS (Université de Londres)

## Dialectometry : The Scientific Measurement of Linguistic Distance.

As is widely known, the National Scientific Research Centre in France (C.N.R.S.), has for many years been promoting the publication of regional linguistic atlases covering the whole of metropolitan France. The present state of each atlas varies according to region. For example, work on the Linguistic Atlas of Gascony (S.W. France), covering an area of approximately 26,000 square miles, has come to a close upon the publication of the sixth volume (A.L.G.6), whereas the neighbouring region, Western Languedoc, has yet to see the appearance, due shortly, of volume one (A.L.L.O.).

The present writer, based at the University of Toulouse, is working in the field of Romance linguistics, and more specifically in that of Gascon dialectology. Our special concern is therefore with the Linguistic Atlas of Gascony (A.L.G.), since the author of this atlas, the late Jean Séguy, directed research operations from his laboratory in Toulouse. His untimely death, in March 1973, brought to an end the present phase of dialect investigation in Gascony, which culminated in the publication of the A.L.G.6 (Paris, C.N.R.S., 1973). The paradox inherent in this event is that the volume in question does not constitute a terminal stage, but a take-off point.

The characteristic that sets the A.L.G.6 apart from any other French linguistic atlas is that not only does it contain linguistic maps of the traditional kind (onomasiological), similar to those in most other atlases, but also maps that outline a new discipline, which J. Séguy baptised "dialectometry". Dialectometry may be defined as the mathematical measurement of dialect variation in geographical space.

It has been said that every researcher invents his own "-ism". In this case, however, such a remark is inappropriate. Jean Séguy's discipline is not a "particularism". The author of the A.L.G. has designed a scientific instrument capable of measuring any linguistic surface, whether it be Pyrenean Gascon, or dialect variations of Quechua, the second official language of Peru. The universal nature of dialectometry is essentially the factor which has prompted the team of

.../...

researchers who have collaborated with Jean Séguy, to publicise their work outside the traditional sphere of Romance linguistics research activity. At present, at least three of the C.N.R.S. laboratories in Toulouse are actively engaged in dialectometric studies: 1) The Laboratoire d'Etudes Méridionales, directed by Philippe Wolff and J. Allières; 2) Anthropologie et écologie pyrénéennes (RCP 323, responsable J. Guilaine, I.P.E.A.)

3) Lexique de l'économie des viandes : boucherie - charcuterie - nécce (E.R.A. 352, Jean-Louis Fossat). This concentration gives some indication of the importance which Toulouse attaches to the development of dialectometric research, both nationally and internationally. The E.R.A. 352, for example, is at present organising a dialectometric study of a region in Northern Spain (Vega de Pas), whilst other researchers are carrying out similar work in the French Pyrenees.

In order to understand Séguyan dialectometry, it is evidently necessary to refer to the A.L.G.6, in which the dialectometric approach is divided into two sections. The first one deals with the statistical operations involved in the mathematical treatment to which the linguistic data contained in A.L.G.s 1-5 is subjected. The second section deals with linguistic distanciation and the existence of dialect boundaries.

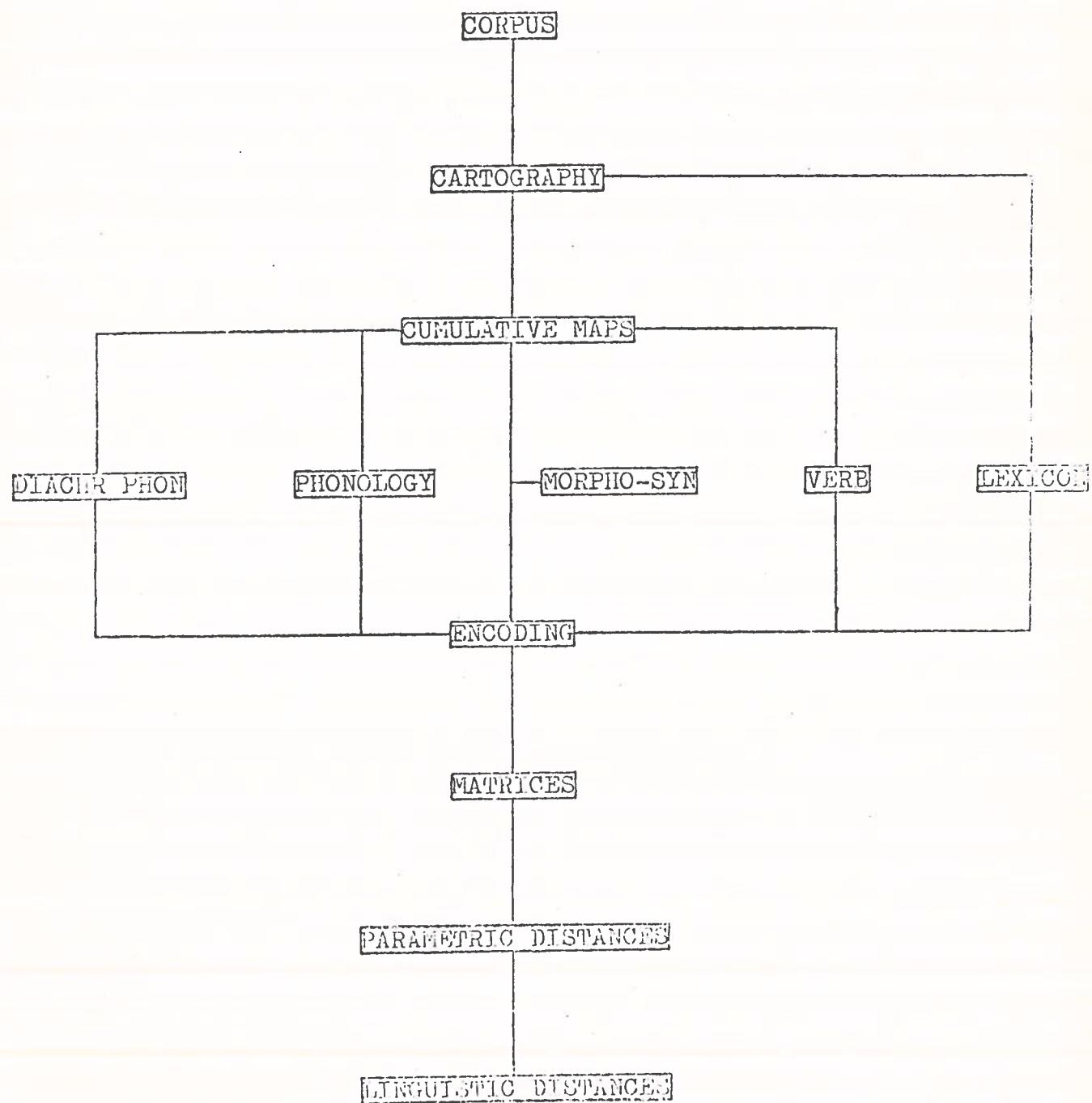
Part one of the A.L.G.6 contains series of maps grouped under four headings : diachronic phonetics, phonology, pronoun morpho-syntax and verb morphology. These parameters, with one other, constitute the basis of dialectometric analysis. Each series of maps referred to, assembles statistically the data scattered throughout the previous volumes, under the four headings mentioned above. Remarkably, all calculations were made manually; at no time were the services of a computer required.

However, it is the second section of the A.L.G.6 which we are particularly interested in. This section is devoted to the measurement of the linguistic distances between each of the 154 points explored in the atlas. These distances are measured by employing a mathematical method of analysis known as the "Hamming distance". The technique involved is described in a little book called "Mathématiques nouvelles pour le recyclage des parents", by A. Kaufmann & G. Cullmann. The authors state that "the Hamming distance is one method of evaluating a certain degree of resemblance between two objects defined in a

.../...

"grill structure". This notion appears to possess four principle characteristics, the first, and most important of which, we have already mentioned : its universal application. It is also an integral method, in that as many parameters as are found in the field may be incorporated in terminal calculations. Other characteristics are its objectivity, i.e. each distance is coded identically : 0-1, and the fact that it disassociates the concept of distanciation from that of differentiation.

### The Dialectometric Process



The above diagram illustrates the complete dialectometric operation. The input is, of course, the corpus. The recorded data is transcribed phonetically, and mapped. Then begins the statistical phase, which produces cumulative maps. These maps bear numerical calculations taking into account all occurrences of a given phoneme or other linguistic unit. Data is then separated into relevant parameters. The matrix coding and decoding process will thereupon generate all linguistic distances, either parametrically, or in overall form. These terminal percentages constitute the global linguistic distance between two points. It should be noted in this connection that although one can calculate the global linguistic distance between the points A-B, B-C, and C-D, it is not possible to establish any relation between, for example, A-D, passing through B and C. In other words, the notion of binary resemblance does not embrace that of transitivity.

Linguistic distances are measured according to a certain number of parameters, in theory, as many as present themselves in the field. For the A.L.G., Jean Séguy selected five, chosen as being representative of the linguistic spectrum as a whole : these are, as previously mentioned, diachronic phonetics, phonology, pronoun morphology, verb morphology, and a lexical component.

### 1. Diachronic Phonetics

The 67 criteria (or "items"), were chosen in order to represent the phonetic evolution of the diverse entities which constitute differential measurements in Gascon. The following list contains an indicative selection of the type of criteria used to measure the phonetic evolution of dialect features in geographical space, along with coding information :

- 1) Consonant assimilation (e.g. [rok > rɔts]); 0=no; 1=yes; 2=both
- 5) -D- : 0>z; 1>d 7) FR- : 1>fr; 0=no 8) FR- : 1>hr; 0=no
- 18) -MB- : 1>mb; 0>m 33) AU pretonic : 1>o; 0>au
- 35) -ARTU : 1>ei; 2>je; 0>e 36) -ARIA : 1>ei; 2>je; 0>e
- 42) Open O diphthongised (conditioned by u) : 1>(w)eu; 0=no
- 51) Suffix -ORIU : 1>ei; 0=no 52) 1>oei; 0=no
- 62) Open E + nasal cons. : 1>ɛ; 0>e
- 67) Open O before implosive nasal cons. : 1>o; 0>u

.../...

Only cases of phonetic evolution attested over the whole territory can be accepted as criteria. The area comprising positive occurrences may nonetheless be very small, as concerning tonic Q followed by a nasal consonant. At only five points in the Pyrenees does [ɔ] remain unchanged. Elsewhere, it has evolved to [u].

## 2. Phonology

This parameter consists of 77 phonological criteria. The phonemic character of these criteria will be discussed later.

- 1) /p/ word initial and intervocalic    2) /p/ word final
- 16) /ə/ in all three positions    20) /dz/ wd. initial
- 30) /w/ intervocalic and wd. final    42) /ʌ/ intervocalic
- 48) /r:/ wd. final    50) /h/ wd. initial and intervocalic
- 51) Tonic vowels : /i/    52) /y/ etc.    61) /all nasal vowels/
- 62) Pretonic vowels : /i/    63) /y/ etc.
- 70) Posttonic vowels : /i/    71) /y/ etc.....    77) /u/

The phonological parameter is perhaps the most succinct of the five, in that all coding is strictly binary. Each criterion is a phoneme attested in modern Gascon. For example, localities where /ə/ is to be found are coded 1 in the appropriate column. Localities not possessing this phoneme in their phonological system are coded 0. The data on which this parameter is based originates from the relevant cumulative maps in the first part of the A.L.G.6. The implicit phonological analysis is of a strictly distributional kind. Consonants are divided into three positions : word initial, intervocalic, and word final. Vowels are separated into three categories too : pretonic, tonic, and posttonic.

## 3. Pronoun Morpho-syntax and Verb Morphology

These two parameters are dealt with together, as they were conceived on the same principle. The morpho-syntactic component comprises essentially paradigmatic criteria, such as appear in morphological systems, or form grammatical rules. It is widely realised that syntactic information is the least accessible linguistic data of all, largely due to the fact that spontaneous conversation is an elusive

..../....

phenomenon which proves extremely difficult to record without transgressing certain unwritten rules of etiquette. Nevertheless, the traditional method of using a questionnaire sufficed to provide the necessary corpus. The questionnaire was created and exploited by X. Ravier, chief researcher for the A.L.G., and author of the forthcoming Linguistic Atlas of Western Languedoc (A.L.L.O.). Examples :

Item 28 - Existence or not of the Gascon énonciatif : que.

Items 52,53 - Existence or not of the definite article eth, era.

As has been mentioned, the principles described above apply equally well to the parameter involving verb morphology. Furthermore, lexicalised criteria such as the conjugations of the verbs "to be", "to go" etc., have been omitted in order to avoid saturation of the lexical component. The 44 criteria contained in this parameter (compared to 68 in pronoun morpho-syntax), are based entirely on the A.I.G.5, by Jacques Allières. Examples :

Item 4 - Existence or not of the futur<sup>e</sup> imperfect.

Item 31 - Accentuation : (Stresse<sup>d</sup>); 0=teleotonic; 1=rhizotonic.

#### 4. Lexicon

The lexical parameter is based on 170 maps chosen at random from A.L.G.s 1-4. This figure equals approximately 10% of the total number of maps. With regard to an alphabetical atlas, it would suffice to select the first batch of maps coming to hand, as semantic diversity is guaranteed by the hazards of French spelling. However, in the case of the A.L.G. and other such atlases based on a logical classification, it is necessary to effect a random selection. Each of the 170 criteria designates a referend, which is also the heading of the map concerned. Examples :

1) (Rabbit's) burrow	91) Nephew	151) Uvula
2) Hantén	92) Hairs	152) Left
3) Skunk	93) Hare	153) Sleep in eyes
6) Boar	94) Nail	154) Brother
8) Wagtail	95) Little finger	155) Cream
9) Green lizard	96) Arse	170) Too much

.../...

Each referend engenders a certain number of signifiants, in practice from one (mononyms) to fifty-two different terms. Hence there are as many code numbers as there are terms. The formal criterion, in 89% of all cases, is etymological. If the etymon is different, then the code number is also different. A small number of differences are coded according to phonetic considerations, for example [halazɛtʃ], [flau] (flail) <FLAGELLU. Synonymous terms grouped under one referend are given separate numerical codes. If, when comparing two points, there is even one identical number, the lexical distance is cancelled out. The principle of binary coding can scarcely be applied to the lexical parameter, in which as many as 52 different terms can be contrasted in Gascon for one referend. However, from a mathematical point of view, the measurement is identical. Whether 17-6 or 0-1, the result is always the same : one lexical distance.

The total number of items, taking into account each of the five parameters, comes to 426. It will have been noted that the encoding process attributes every item with a distinct numerical code.

The next step in the operation is to design a suitable matrix. This is accomplished by grouping each locality's code number in a vertical column, and distributing that of each criterion (or item), on a horizontal plane. Such a lay-out facilitates code number comparison, and hence the calculation of linguistic distances. Ideally, each matrix is constructed using a binary system, within the limitations mentioned above. The actual process of measuring individual and parametric distances is no more difficult than consulting a railway timetable. One merely juxtaposes the two horizontal lines that contain the items to be compared. Whenever the code numbers agree, no distance is counted; if, however, they are not identical, one distance will be marked down. This operation is repeated for every item; the total obtained for each parameter constitutes the parametric distance. If the five parametric distances are added together, the overall linguistic distance is arrived at. Below appears a specific example, taken from our own thesis, entitled "Balaguères - Bethmale - Biros : A Dialectometric Study" (Toulouse, 1975). The localities in question are to be found within a triangulation zone in the Central Pyrenees (09-Ariège), which was researched in 1974.

.../...

a. Diachronic phonetics (67 items)

Letter code :                    BAL = Balagué                    BET = Bethmale  
                                    BIR = Biros                      CAS = Castillon

BAL-BET	3.0% (two distances)	BET-BIR	1.5% (one distance)
BAL-BIR	1.5% (one distance)	BET-CAS	3.0% (two distances)
BAL-CAS	3.0% (two distances)	BIR-CAS	1.5% (one distance)

b. Phonology (77 items)

BAL-BET	9.0% (seven distances)	BET-BIR	4.0% (three distances)
BAL-BIR	5.0% (four         "         )	BET-CAS	4.0% (         "         )
BAL-CAS	7.8% (six         "         )	BIR-CAS	2.6% (two         "         )

c. Morpho-syntax/Verb morphology (112 items)

All interpoints (estimated except BET-CAS) : 17% (nineteen distances)

d. Lexicon (170 items)

BAL-BET	23% (39 distances)	BET-BIR	17% (29 distances)
BAL-BIR	19% (32         "         )	BET-CAS	21% (35         "         )
BAL-CAS	18% (31         "         )	BIR-CAS	15% (26         "         )

e. Global linguistic distance (426 items)

BAL-BET	16% (67 distances)	BET-EIR	12% (52 distances)
BAL-BIR	13% (56         "         )	BET-CAS	14% (59         "         )
BAL-CAS	14% (58         "         )	BIR-CAS	11% (48         "         )

The equation used to reduce the rather unwieldy distance totals to percentage form is :

$$D = \frac{d \times 100}{c}$$

where D = Hamming distance (Global linguistic distance)

d = sum of individual distances

and c = number of criteria

.../...

### The Significance of Linguistic Distances

Linguistic distanciation plays an important part in the notion of dialect typology. The scientific identification of dialects has long been one of the main preoccupations of the dialectologist. However, although it is recognized that the concept of dialect areas exists in the naïve (non scientific) judgement of the native speaker, analysis of the same linguistic data in geographical space does not appear to confirm the existence of dialects which can be identified according to a set of precisely defined criteria. The dialectometric approach reveals that, despite the non existence of "discrete" dialect areas, the formulation of judgements of resemblance or dissemblance, constitutes a type of analysis conducive to a mathematical definition of such areas. It is this observation which brings us to the second part of our article, during the course of which, we shall discuss our personal attempts to explore the dialectometric method of analysis.

The location researched forms a triangulation zone in the Central Pyrenees, within the compass of the A.L.G. (Atlantic to the West, Pyrenees to the South, River Garonne to the North and (roughly) to the East).

Apart from Jean Séguy, few people in France have as yet concerned themselves with dialect measurement. The best known is perhaps H. Guiter, author of the Linguistic Atlas of the Eastern Pyrenees (Catalan), published in 1965, and of an important paper given at the 1971 Strasbourg Congress, entitled "Atlas et frontières linguistiques". In this paper, H. Guiter outlines a second means of dialectometric analysis, based on a concept similar to that which characterises the Séguyan method : the exploitation of a number (100) of lexical maps. Methodologically, both forms of analysis (to which we shall refer respectively as G and S), differ little, up to a certain point. Both are objective, in that all distances are coded identically; structured, in that measurements are made similarly between two contiguous points; and integral, in that all types of linguistic information can be incorporated in terminal calculations. Henceforth, G and S diverge. The 100 lexical maps in G serve to measure each linguistic parameter - phonetic, syntactic, morphological, etc., as well as the lexical component. For example, let us take the refrend barrette (hair pin), and measure the

.../...

interpoints BAL-BET and BAL-BIR. For BAL, we recorded the significant [waf:cto] ; for BET, we have [yggladəro], and for BIR, [bar:cto]. Between BAL-BET, BET-BIR, we therefore count one lexical distance, based on etymological criteria. Between BAL-BIR, there is no lexical distance, but two phonological distances : 1) alternation b/w; 2) existence of a geo-allophone, namely dorso-uvular [f:] at BAL, whereas BIR attests the universal apico-alveolar articulation [r:].

Jean Séguy, being aware of the existence of a second method of dialect measurement, undertook to apply it to the data contained in the A.L.G. The results he obtained amount to a scientific verification of both methods : though individual figures rarely corresponded, neither did they vary significantly. Values were then plotted on graphs, and compared. The curves produced proved totally compatible : peaks with peaks, and troughs with troughs. The advantage of method G is immediately apparent, in that it dispenses at a stroke with the laborious calculations necessary to establish the statistical data of the cumulative maps.

Less extensive applications of dialectometric analysis have been made by X. Ravier, chief field researcher for the A.L.G., in his article "Le discours sur les isolats pyrénéens au pays de Barèges", and by Jean-Louis Fossat, Director of the E.R.A. 352 (C.N.R.S.), in "Mesures écolinguistiques sur l'axe des marchés transaronnais".

It is in this context that we came to Toulouse from Reading University (G.B.), late in 1973, at the invitation of Jean Séguy, and later Jean-Louis Fossat, to carry out initial research into dialectometric methods of linguistic analysis. This work, lasting eighteen months, culminated in the successful defence of our doctoral thesis (see above), which is to be published early in 1976 (C.R.D.P. - University of Toulouse). In so far as this thesis represents, to our knowledge, the first such study of its kind in France, we shall venture to describe some of its principle aims and conclusions, modest though they may be in comparison to J. Séguy's work.

A short period of pre-research allowed us to define the objectives which seemed relevant to the type of study envisaged. These were as follows :

..../....

1) Employ the dialectometric method of analysis evolved by Jean Séguy.

2) Increase the density of the A.L.G. survey network within a given area.

3) Explore certain Pyrenean localities whose linguistic systems were little known or unknown.

4) Integrate the mathematical data provided by the localities researched into the A.L.G. matrices.

5) Determine the possible existence of dialect boundaries in relation to the area investigated.

6) Define further applications, if any, of the methods of dialectometric analysis.

The choice of the triangulation zone BAL-BET-BIR satisfied objectives two and three. Very little linguistic information was available regarding this location. In fact, the commune of Balaquères had apparently been systematically ignored by geographical, linguistic, and ethnological studies of the region, if not by British Television (B.B.C.), who filmed a nationally networked documentary in the area in September 1973. The three points chosen are located in a well defined geolinguistic zone called Le Castillonnais. Two of these three points had never been investigated (BAL, BIR). The third, BET, appears in the A.L.G. as 790S. The survey density in this area (230 square miles), previously represented by only two localities, was therefore doubled. The incorporation of BET into our research also enabled us to satisfy objective four.

Perhaps the most interesting results of our dialectometric study are those relating to the notion of geolinguistic areas, that is to say, dialect surfaces whose geographical and linguistic boundaries coincide, whatever the morphological composition of such boundaries. These results can be conveniently grouped under two joint headings :

1) Geolinguistic zone projections/Mathematical isoglosses.

2) The concept of geolinguistic centres/Prediction of survey points in a representative atlas.

.../...

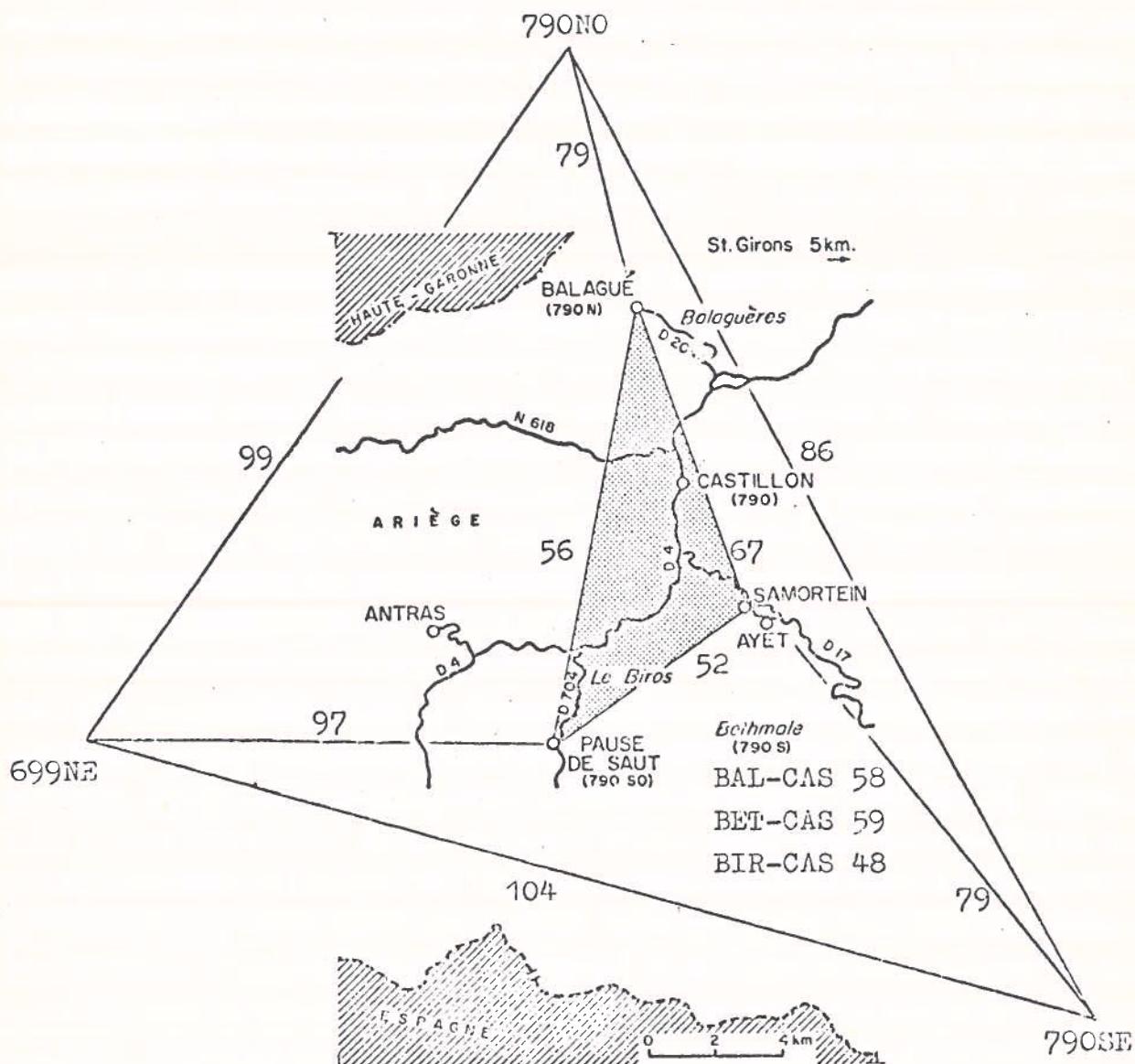
1) The area called Le Castillonnais has traditionally been regarded as an independent entity, both geographically and ethnologically. Geomorphological and hydrographic criteria confirm the territorial independence of this zone, bordered to the south by the lofty mountain peaks announcing the Spanish frontier, to the north by a wide expanse of forest, to the west by an administrative boundary (département), and to the east by a separate hydrographic configuration. We therefore attempted to characterise the degree and type of linguistic independence enjoyed by the area in question. To accomplish this aim, we first used the traditional method of tracing isoglosses for each linguistic parameter. It was found that several isoglosses coincided entirely or partially with the perimeter of Le Castillonnais. However, from the methodological point of view, this approach proved unsatisfactory, in that there merely emerged a mass of inextricable lines. Each line was, of course, significant in its own right, but allowed of no analytic interpretation with regard to the composition of the differential linguistic features of the area, in relation to the external linguistic environment.

Having calculated the overall and parametric distances within Le Castillonnais, we projected the three vertices of our triangulation zone (which coincide with the northern, south-western, and south-eastern geographical limits of this area). The vertices of the external, or projected triangle, coincide with three peripheral localities, 790NO Saleich, 790SE Couflens, and 699NE Melles. These points, though outside the confines of Le Castillonnais, are nevertheless contiguous, in the sense that no intervening community is to be found on the three projected axes in question.

Within the triangulation zone, it will be noted that the six overall (or global) distances measured (five of which do not appear in the A.L.G.6, the exception being BET-CAS), vary between 11% and 16%. Furthermore, there exists no fixed mathematical ratio between the linguistic distances and corresponding geographical distances, contrary to Jean Séguy's findings on a macroscopic level. The internal distances overlap in such a way that no micro-arcs can be distinguished inside our triangulation zone. However, the projection demonstrates that once the geographical limits of Le Castillonnais are crossed, the range of overall linguistic distances increases to between 18.5% and 24%. In other words, there is a mathematically significant dialect boundary

.../...

which coincides with the perimeter of Le Castillonnais. Jean Séguy has in fact established an experimental graduation for the Gascon domain that defines mathematically significant percentage ranges :  $\leq 13\%$  (zero degree - no linguistically relevant dialect boundary),  $13\% - 17\%$ ,  $18\% - 22\%$ , and  $\geq 23\%$ , in ascending order of importance. These calculations are based on the distances overlapping the isogloss which separates the area speaking "gascon noir" from that speaking "gascon clair". The map below illustrates the projection operation just described.



The external distances are measured from 790 Castillon.

... / ...

It is possible, and indeed highly interesting, to carry out this projection using not only the overall distance percentages, but also those concerning each parameter. Such an operation constitutes a first stage in the "deabstraction" of the mathematical calculations under consideration, and a return to concrete linguistic data.

With regard to the lexical parameter, the same phenomenon can be observed : coinciding with the geographical limits of Le Castillonnais, the range of distance percentages increases abruptly. The internal/external ratio of lexical distances is 15%-23%/31%-35%. The lexical component is in fact responsible for a statistically massive percentage of linguistic distances (an average of 56% in our triangulation zone). In this connection, X. Ravier's article "L'incidence maximale du fait dialectal", given as a paper at the 1971 Strasbourg Congress, is particularly illuminating, on a macroscopic scale.

The contour of our dialect boundary is confirmed by a study of each of the other linguistic parameters, excepting phonology, where the reverse is true. In other words, the range of internal phonological distances (2.6%-9%), is higher than that of the external, or projected distances (2.6%-7.8%). In our thesis, we offer a possible explanation for this phenomenon.

We may therefore characterise the type of geolinguistic independence enjoyed by Le Castillonnais, as being essentially global, except that no phonological boundary coincides with the geographical limits defined above. There now intervenes a second stage in the process of "deabstracting" this mathematical data. We took the lexical component and probed into the endemic linguistic units which constituted differential measurements. To our surprise, it happened that of all the lexical distances measured, only five could be imputed to the existence of lexemes enjoying some degree of territorial independence. In other words, the traditional linguistic homogeneity attributed to Le Castillonnais is due to a mathematical independence rather than qualitative lexical features. An attempt to characterise the specific relationship between the lexical features of this area and the external lexical environment, seems to point to the existence of blurred linguistic structures which function precisely as such. Research is

.../...

at present taking place in this direction. The utility of geolinguistic zone projections resides in the fact that they enable us to trace what we call "mathematical isoglosses", that is to say, isoglosses formed by the occurrence, in relation to a given geographical configuration, of a statistically significant number of linguistic distances (or differences).

The notion of a mathematical isogloss is clearly different to that of a traditional isogloss. Although they are both analytic, in the sense that they divide the dialect surface into discrete areas, the traditional isogloss is valid only for one linguistic fact (e.g. latin FR->hr/fr), whereas the mathematical isogloss is a concept which synthesises a whole number of individual elements. It is therefore, in the broadest sense of the term, a structural concept, and not a purely descriptive one.

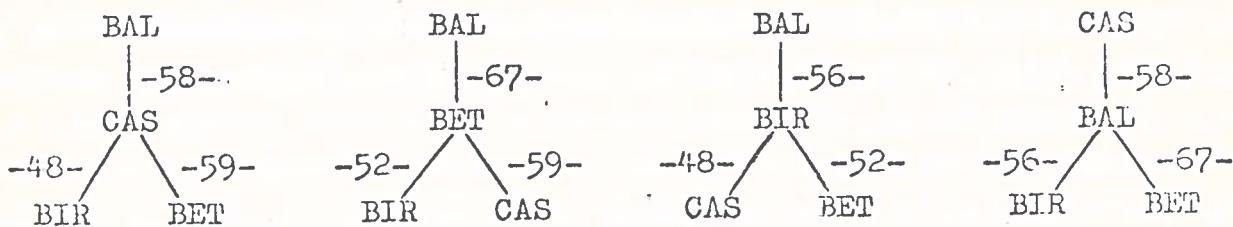
The implications of such an observation are clear : whilst many refute the existence of discrete dialect areas, we do not reject the possibility of structuring the dialect surface meaningfully, using quantitative and not qualitative methods of analysis. For the present, our research has been limited to a triangulation zone in the Central Pyrenees. However, this territory is to be extended over the next few years to cover the whole of the Central Pyrenees, for it is evident that such mountainous areas provide ideal testing grounds for dialectometric projects.

2) We shall now discuss the concept of geolinguistic centres, as well as the prediction of survey points for a representative linguistic atlas...In France, at the present time, there exist essentially two different kinds of linguistic atlas; the most common type is the representative atlas, i.e. in which the dialect surface is represented by the choice of a certain number of survey points. For instance, the A.L.G. has a survey point density of 0.36 per square kilometre (total area 42,000 square kilometres). The rarer type is the exhaustive atlas, such as that created by H. Guiter (Linguistic Atlas of the Eastern Pyrenees), each commune within a given geographical area being researched. Our concern here is not with the latter type of atlas, but with the representative variety. How are the survey points chosen? Are they indeed representative of the dialect they purport to reflect? With

.../...

regard to the A.L.G., we can affirm that the choice was scarcely more than arbitrary, and rarely justified by linguistic reasoning. The geographical factor, in particular, seems to have been a predominant consideration. There would appear to be good grounds for basing the choice of survey points on more scientific criteria.

In the case of our own triangulation zone, we have shown in our thesis that the two points chosen are far from being the most representative of the linguistic systems of Le Castillonnais. Bethmale is, in fact, one of the least representative points, though its inclusion can be justified on other grounds. How, then, is it possible to determine objectively the most representative point in a predefined geolinguistic environment? Our own method constitutes an application of the notion of geolinguistic centres, allied to the less clearly defined concept of parametric covariance, that is to say, the apparently stable mathematical relation existing between the different linguistic parameters in a given geolinguistic system. The parametric approach used in dialectometry is retained for the needs of our analysis. In order to establish the phonological, lexical, morphological, etc., centre of a given area, in other words, the point in relation to which the linguistic systems of all other points within this area differ minimally, one simply compares, centrifugally, the number of linguistic distances emanating from each point researched. The diagram below illustrates the type of comparison envisaged. The figures refer to the sum of all linguistic distances for each interpoint, and correspond to those appearing in tabular form earlier in this article.



In the case of our triangulation zone, the results show the following order of "representativity", taking into account the notion of parametric composition. With regard to the morphological parameter (morpho-syntax and verb combined for the purposes of presentation), the figures are all estimated, except for Bethmale.

.../...

Locality	D. Phon	Phonology	Morphology	Lexicon	Global
Le Biros	3	8	57	87	155
Castillon	5	9	57	92	163
Bethmale	5	10	57	103	175
Balagué	5	17	57	103	182

Le Biros therefore proves to be the geolinguistic centre of our triangulation zone, in other words the locality possessing the most representative linguistic features of Le Castillonnais. Clearly, the success of this method will depend largely on the carrying-out of the pre-research necessary to provide sufficient data for the measurement of linguistic distances. Since the number of points to be pre-researched must be relatively high (five or six for Le Castillonnais), it is imperative for the field-worker to have at his disposal a questionnaire of sufficiently short duration to enable him to investigate a maximum number of points in a minimum amount of time. This requirement may be fulfilled if our conclusions relative to the concept of parametric covariance prove to be valid. Our results (see above table), show that the lexical parameter reflects most unequivocally the percentage gap separating the number of linguistic distances measured between a given series of localities in a predefined area. The same results show that the other parameters in question imitate the tendency manifested by the lexical component, though to a narrower degree.

If this hypothesis can be successfully verified over a greater geographical expanse, it would suffice for the field-worker to use a simple lexical questionnaire covering perhaps 150-200 lexemes. Such a questionnaire would be exhausted in two-three hours. In conditions such as exist in the Pyrenees, this kind of schedule would mean that an area similar to Le Castillonnais could be researched in roughly one week, providing that the field-work was consecutive (one investigation per day), and prepared in advance.

The type of method we have just described has the additional advantage of pinpointing not only the most representative linguistic features of the area concerned, but also the existence of important geo-variants. The presence of a geo-variant /ə/ at Bethmale, for example, justifies the inclusion of this locality in the A.L.G. net-

.../...

work. La Cantillonneuse is the last surviving location of this phoneme in Gascon. It originates from 1) latin C (+e,i), e.g. [ɛq̪de] < CICERO; 2) the Romance group ky ( CY, CCY), e.g. [k̪eo] < COCIA; 3) the voiceless Romance affricate 'ts' ( TSY), in other words from the following latin groups : a) ty; e.g. [kaneu] < CANTIONE; b) tty; e.g. [b̪ee] < BETTIU; c) -cty- et -pty- e.g. [kaea] < CAPTIARE. Our own research lead to the discovery, at 790N Balagué, of a new phenomenon in Gascon phonology, namely the existence of /r:/, whose dorso-uvular articulation, at times similar to the well-known French [r̪], can vary, depending on phonetic environment and position, between [x:] and [h:]. The typical phoneme resulting from latin R- and -RR- in Gascon is /r:/. This phenomenon may also be considered as geo-allophonic, depending on the type of analysis being conducted. The occurrence of the phoneme /r:/ poses a methodological problem concerning the integration of the phonological distances thus created into the appropriate parameter, due to the fact that the latter tends to identify the concept of phonological differentiation with phonemic considerations. In order to codify the existence of /r:/, whether it be considered as a phoneme /r:/, or as a geo-allophone [r̪], we propose three solutions in our thesis, all presenting certain drawbacks. The only satisfactory solution involves restructuring the phonological parameter with a view to the incorporation of differential phonetic criteria.

#### Applications of Dialectometry

Through working in a research team closely connected with ethnolinguistic studies, we have been made especially aware of the implications our research holds for this discipline. It may be noted that the full title of the A.L.G., and of all French regional linguistic atlases, is "Linguistic and Ethnographic Atlas of"...the area concerned.

We were particularly struck by the fact that Jean Séguy, having applied dialectometric methods of analysis to the linguistic content of his atlases, had made no theoretical or practical attempt to subject the ethnolinguistic material contained therein, to the same process. We envisaged the possibility of a Pastoral Atlas of the Pyrenees being founded, not on a purely descriptive basis, but on differential criteria, materialised by measurements of ethnolinguistic

.../...

distanciation. The neologism "ethnodialectometry" seems to be a suitable term to characterise such a discipline, and has already been adopted by one scientific publication in the C.N.R.S. framework. The following example will help to clarify our idea.

In Le Castillonnais, three lexemes are obtained for the single referend "flail" : [flau], [halazɛtʃ], and [lato]. A strict dialectometric analysis would in this case reveal only two lexical distances, since the first two lexemes both originate from FLAGELLU. At 790S Bethmale, the [lato] is still in use. It consists of a length of flexible wood, which does not harm the wheat/straw. In the form of thatch, the latter can be used to roof houses. At 790SO Le Biros, our informers used to carry out the same operation by hand; the significant recorded there is based on modern French "fléau". The flail is no longer used at 790N Balagué, though the informer was able to name each part. In this case, we are in the presence of an ethnodialectometric type of distanciation, dependent not only on linguistic criteria, but equally on the existence of a cultural phenomenon, whose evolution is spaced geographically and temporally over the whole length and breadth of the Pyrenees. A large part of the research programme the E.R.A. 352 is presently engaged in, concerns the creation of an atlas embracing the Pyrenean cultural phenomenon. The ethnodialectometric approach sketched above constitutes one of the pillars of this programme, conceived and directed by Jean-Louis Fossat.

We should like to conclude this short article by insisting on the fact that the discipline Jean Séguy baptised "dialectometry", constitutes a first attempt to make use of the mass of dormant dialect data contained in modern linguistic atlases. To date, only one atlas, the A.L.G.6, comprises a precise application of dialectometric analysis. The publication of this atlas inaugurates a whole new field of research in dialectology. Like other recently evolved scientific methods of analysis, dialectometry creates more problems than it resolves. Nevertheless, these problems are decisive ones, relevant in particular to the understanding of the dialect concept, and of the interaction of linguistic dynamics and geographical space. Furthermore, the application of mathematical models of analysis to linguistic phenomena would seem the logical corollary to the act of calling linguistics a "science".

.../...

If this article has helped to publicise the type of dialect research being conducted at present in the field of Romance Linguistics and more especially in that of Séguyan dialectometry, by the Toulouse School of dialectology, it will have fulfilled its principle aim.

Toulouse, July 1975, Dennis Philips.

+++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++

### Indicative Bibliography

Bec P. - Les interférences linguistiques entre gascon et languedocien dans les parlers du Comminges et du Couserans; P.U.F., Paris, 1968.

Fossat J.-L. - Mesures écolinguistiques sur l'axe des marchés transgaronnais; colloque d'Auvillar, 1973; to be published .(1975-1976)

Guiter H. - Atlas et frontières linguistiques; actes du colloque de Strasbourg, 1971.

Guiter H. - Atlas linguistique et ethnographique des Pyrénées orientales; 1965.

Philps D. - Balaguères - Bethmale - Biros : Etude dialectométrique; C.R.D.P. - University of Toulouse; to be published early 1976.

Ravier X. - Le discours sur les isolats pyrénéens au pays de Barèges; Revue de linguistique romane, tome 38, pp. 420-428, 1974.

Ravier X. - L'incidence maximale du fait dialectal; C.N.R.S., actes du colloque de Strasbourg, May 1971.

Schöntaler W. - Die Mundart der Vallée de Bethmale; Tubingen, 1933.

Séguy J. - Atlas linguistique et ethnographique de la Gascogne; C.N.R.S. Vols. I-VI, Paris, 1954, 1956, 1958, 1966, 1972, 1973.

Séguy J. - La dialectométrie dans l'Atlas linguistique de la Gascogne; Revue de linguistique romane, 1973 (n° 1), pp. 1-25.

Séguy J. - La relation entre la distance spatiale et la distance lexicale; Revue de linguistique romane, tome 35, pp. 335-357, 1971.

### SOMMAIRE

FOSSAT (Jean-Louis) La dialectométrie:petite ..... pp. 1-6  
introduction méthodologique

FOSSAT (Jean-Louis) La microdialectométrie ..... pp. 7-46  
Problèmes et méthodes;microdialectologie  
et microdialectométrie pour deux vallées  
pyrénéennes(Aure et Louron) d'après les  
données de J. Cremona(1956).

PHILPS (Dennis) Dialectometry:The Scientific ..... pp. 47-67  
measurement of Linguistic Distance.

### BIBLIOGRAPHIE

KURATH (Hans) Studies in Area Linguistics, Indiana University  
Press, Bloomington and London, Indiana University Studies  
in the History and Theory of Linguistics, Second Reprint  
1974, 202 pp.;bibliography pp. 185-202.

Dépot légal 2<sup>e</sup> trimestre 1976

---

SERVICE DE REPROGRAPHIE  
ACADEMIE DE TOULOUSE  
3, RUE ROQUELAINE  
31069 TOULOUSE CEDEX

---

