

Éléments d'analyse dialectométrique (avec application à l'AIS)

Autor(en): **Goebl, Hans**

Objektyp: **Article**

Zeitschrift: **Revue de linguistique romane**

Band (Jahr): **45 (1981)**

Heft 179-180

PDF erstellt am: **11.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-399711>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

ÉLÉMENTS D'ANALYSE DIALECTOMÉTRIQUE (avec application à l'AIS)⁽¹⁾

0. Géographie linguistique + taxonomie numérique = dialectométrie.

Telle qu'elle se présente à l'heure actuelle, la dialectométrie constitue un amalgame bidisciplinaire réunissant les apports de la géolinguistique et de la taxonomie (ou taxinomie) numérique. Alors que nous porterions de l'eau à la rivière en donnant des explications sur la géolinguistique, il semble par contre indiqué de s'attarder un peu sur la taxonomie numérique. C'est une discipline statistique, qui — par voie numérique — « [...] apporte une réponse au problème suivant : comment décomposer une population donnée d'individus ou d'objets, décrits par un ensemble de caractéristiques, en un certain nombre de sous-groupes homogènes. » (Chandon/Pinson 1981, IX). Ayant cours sous différentes dénominations (en français : analyse typologique, classification automatique, analyse des données, taxométrie, etc. ; en anglais :

(1) Collaboration technique : S. Selberherr (Vienne), W.-D. Rase (Bonn), H. Pudlatz (Münster). Les travaux dont il est question ci-après ont bénéficié de l'aide des formations de recherches et organismes suivants :

- Hochschuljubiläumsstiftung der Stadt Wien (Vienne),
- Jubiläumsfonds der Österreichischen Nationalbank zur Förderung der Forschungs- und Lehraufgaben der Wissenschaft (Vienne),
- Fonds zur Förderung der wissenschaftlichen Forschung in Österreich (Vienne),
- Deutsche Forschungsgemeinschaft (Bonn).

J'exprime ma profonde reconnaissance aux organismes mentionnés ci-dessus, ainsi qu'aux MM. Selberherr (calcul électronique), Rase (cartographie électronique) et Pudlatz (polygones de Thiessen) pour leur collaboration efficace, compétente et précieuse. Je ne veux pas oublier non plus M. L. Karlson (Ratisbonne - Creutzwald) qui a bien voulu se charger de la toilette stylistique de notre texte.

Encore une petite remarque qui, peut-être, n'est pas inutile : le lecteur familier de l'algèbre et/ou de la taxométrie modernes constatera tout de suite que nous avons sacrifié la rigueur du symbolisme mathématique à l'intérêt conceptuel de notre présentation tout en espérant que le public philologique en tirera profit d'un côté, et que les spécialistes avertis ne nous en voudront pas trop de l'autre.

numerical taxonomy, mathematical taxonomy, etc. ; en allemand : Numerische Klassifikation, Numerische Taxonomie, Taxometrie), elle remonte aux débuts des années 60, la première synthèse taxométrique, rédigée par Sokal et Sneath, datant de 1963. Mais ce qui la caractérise tout particulièrement c'est l'universalité de ses applications, d'où une bonne dose de pluridisciplinarité dans laquelle baigne, bon gré mal gré, toute démarche taxométrique : « Cette préoccupation est commune à de nombreuses disciplines : biologie, zoologie, archéologie, botanique, géographie, géologie, agronomie, médecine, psychiatrie, psychologie, sociologie, anthropologie, linguistique, documentation automatique, intelligence artificielle, sciences de la gestion des entreprises. » (Chandon/Pinson 1981, IX).

Pour plus de détails (et pour une introduction générale aux problèmes de la taxométrie) voir les ouvrages suivants ⁽²⁾ :

— en français :

Bertier/Bouroche 1975, Benzécri 1980, Bourouche/Saporta 1980, Chandon/Pinson 1981, Jambu 1978 ;

— en anglais :

Jardine/Sibson 1971, Sneath/Sokal 1973, Sokal/Sneath 1963 ;

— en allemand :

Bock 1974, Opitz 1980, Sodeur 1974, Vogel 1975.

Vu l'affinité apparente entre la géolinguistique et la géographie, et, partant, entre la dialectométrie et la géographie quantitative, il est bon de tenir compte aussi des acquis récents de la « Nouvelle Géographie » (ou géographie quantitative) : cf. les orientations générales de Abler/Adams/Gould 1977, Claval 1977 et French/Racine 1971 ; ainsi que les études méthodologiques de Bahrenberg/Giese 1975, Béguin 1979, Fischer 1980, Haggett 1973 et Racine/Reymond 1973.

Quant à la dialectométrie à proprement parler, elle remonte — de par le nom et l'orientation méthodologique — à J. Séguy (première mention de « dialectométrie » : 1973 a, 1 ; cf. aussi Séguy 1971, 1973 b

(2) Signalons entre parenthèses l'existence de trois Sociétés savantes qui consacrent leurs activités exclusivement (ou presque) au développement et à l'avancement des méthodes taxométriques :

- Classification Society (fondée en 1964, divisée en « European Branch » et en « North American Branch »),
- Société française de classification (fondée en 1976),
- Gesellschaft für Klassifikation (fondée en 1977 ; elle édite la revue « International classification », depuis 1974).

et 1973 c) et aux efforts synthétiques déployés par lui autour des matériaux dialectaux de l'Atlas linguistique de la Gascogne. Il importe cependant de ne pas oublier les contributions non moins fondamentales de H. Guiter (1973, 1974, 1979), qui peut être considéré à juste titre comme le second père fondateur de la dialectométrie, vu aussi l'aide morale et spirituelle qu'il avait prêtée à J. Séguy.

Tout méritoires que soient les contributions de J. Séguy et de l'école de Toulouse après lui (cf. Philips 1975, Fossat 1977 et 1980, etc.) et de H. Guiter, il serait malhonnête de passer sous silence toute une série de travaux antérieurs d'inspiration taxométrique (avant la lettre), où la nécessité d'un groupement numérique de données linguistiques disparates se dessine déjà clairement. Citons à ce propos les travaux de Kroeber/Chrétien 1937, Reed/Spicer 1952, Ellegård 1959, Grimes/Agard 1959, Cowan 1964, Herdan 1964, Muljačić 1967 (et 1980), Houck 1969 et Rensch 1970, parmi lesquels Reed/Spicer, Cowan, Houck et Rensch s'orientent déjà nettement vers l'analyse numérique des données dialectales.

A partir des années 70, le nombre des travaux linguistiques d'allure carrément taxométrique va sans cesse en augmentant : cf. p. ex. Doerfer 1971, Altmann/Lehfeldt 1973, Shaw 1974 et Fronzaroli 1975 (et 1967). Le livre de Altmann/Lehfeldt mérite une mention particulière étant donné le fait qu'il opère la première jonction explicite entre la typologie linguistique et la taxonomie numérique (en se basant sur Sokal/Sneath 1963).

Les débuts de nos propres travaux dialectométriques remontent à 1971 (cf. Goebel/Winterleitner 1971) et débutent par des analyses taxométriques faites sur une matrice de données tirée de l'ALF ; cf. Goebel 1976. Ce n'est qu'à partir de 1975 que nous disposons d'une autre matrice de données (tirée de l' AIS) qui servira de base empirique à nos contributions de 1977 sqq. Signalons encore un rejeton dialectométrique très récent (et greffé sur nos travaux) au sein de la philologie anglaise : Viereck 1980.

Bien que méthodologiquement fort apparentée à maints égards, la dialectométrie (ou dialectologie, voire géolinguistique quantitative) ne doit pas être confondue avec la dialectologie computative, dont les intérêts majeurs se situent moins sur le plan de la synthèse classificatoire que sur celui de l'optimisation du traitement électronique des données géolinguistiques ; cf. le volume-recueil de Putschke 1977, l'étude de Naumann 1976 et l'Atlas linguistique de l'Est du Canada (complètement informatisé) de G. Dulong et de G. Bergeron.

1. De l'AIS à la matrice des données.

Toute procédure taxométrique commence par le choix des *objets* (*individus* ou *éléments*) sur lesquels on veut faire l'analyse, et par le choix des *attributs* (*caractéristiques* ou *variables*) qui permettent de les caractériser. En dialectométrie, ce choix est facile : les objets à classer sont les points d'enquête d'un atlas linguistique, alors que les données dialectales (répertoriées en transcription phonétique sur les planches d'un atlas linguistique à côté de chaque point d'enquête) constituent les attributs respectifs. Dans un second temps, il convient de se demander quelle sera la manière la plus adéquate de mesurer les attributs retenus afin d'obtenir une description codée des phénomènes observés, appelée *matrice* (ou *tableau*) *des données*. La matrice des données est un tableau à double entrée groupant n objets (ou points d'atlas) et p attributs (ou cartes d'atlas).

Matrice des données

		1	2	3	4	5	6
Attributs (cartes d'atlas)	4	j	j	j	j	j	j
	3	i	h	h	×	i	i
	2	e	f	f	e	g	g
	1	a	b	b	c	a	d
		1	2	3	4	5	6
		Objets (points d'atlas)					

× *case vide*

Fig. 1 :

Matrice des données constituée de 6 objets (points d'atlas) et 4 attributs qualitatifs (cartes d'atlas).

En gris : l'équivalent matriciel d'une carte d'atlas.

Vu l'impossibilité de convertir, d'une façon objective et bien réglée, les données dialectales répertoriées sur les planches d'atlas en valeurs numériques, on se contentera d'en retenir les identités et différences respectives tout en les marquant par des symboles conventionnels (p. ex. par des lettres : a, b, c, d, etc. ; voir la Fig. 1). C'est ainsi que l'on peut coder — et, partant, mesurer — selon un critère descriptif préalablement choisi, les données dialectales brutes. Parmi les quatre sortes d'échelles de mesure couramment utilisées, celle que l'on vient d'appliquer est la plus simple, à savoir l'échelle nominale (cf. Chandon/Pinson 1981, 33-36 et Gendre 1977, 9-11). Le codage des données brutes d'un atlas linguistique sur échelle nominale conduit à l'isolement d'un certain nombre d'unités taxatoires, appelées désormais « taxats ». Il convient d'en dresser la liste, le codage des données brutes une fois fait.

Les structures de la matrice de données utilisée dans ce travail sont les suivantes :

- données dialectales brutes : vol. I, II et IV de l' AIS (3)
- objet (points d'atlas) : 247 points AIS + 4 points artificiels :
 - P. 154 : seconde enquête faite à Turin (cf. Jaberg/Jud 1928, 56-57)
 - P. 262 : seconde enquête faite à Milan (cf. Jaberg/Jud 1928, 69)
 - P. 524 : seconde enquête faite à Florence (cf. Jaberg/Jud 1928, 100-101)
 - P. 999 : équivalent de l'italien standard tel qu'il figure dans les en-têtes des planches de l' AIS
- critère discriminatoire pour l'isolement des unités taxatoires (taxats) : différences lexicales et morpho-syntaxiques (4)

(3) Le choix des vol. I, II et IV parmi les huit volumes de l' AIS a été fait par voie aléatoire (tirage au sort).

(4) En fait, le codage descriptif (sur échelle nominale) ressemble de très près à la « typisation » des données dialectales pratiquée couramment par les géolinguistes ; cf. à cela le témoignage suivant de K. Jaberg (1947, 6) :

« Le géographe-linguiste, pour ne pas se perdre dans le détail, se voit obligé de « typiser » les nombreuses variantes d'un mot qui forment le tapis bariolé d'une carte linguistique. Il crée ainsi des mots-types, dont il symbolise les aires par des couleurs ou par des hachures. Cela lui permet de voir en étendue ce que l'étude étymologique voit en profondeur. Le mot-type peut être une base étymologique, un mot littéraire, un mot régional, voire un mot inexistant qu'on a forgé arbitrairement — ce qui importe, c'est qu'il résume les caractères essentiels d'un faisceau de formes similaires. Si l'on

- attributs (cartes analysées) : 696 ⁽⁵⁾
- nombre des unités taxatoires (taxats) consignées dans la matrice des données : 4.836
- nombre de taxats/carte analysée : 6,9483.

Retenons qu'à de petites différences près, cette matrice correspond à celle que nous avons utilisée dans nos études antérieures de 1977 et 1978.

Il n'est pas du tout sans intérêt d'observer de plus près la répartition des taxats consignés dans les lignes de la matrice de données (cartes analysées, attributs). La liste qui en résulte — voir la Fig. 2 — montre que le morcellement onomasiologique des données dialectales considérées (en l'occurrence celles des vol. I, II et IV de l'AIS) revêt — pour disparates que puissent paraître les étendues et la bigarrure des différentes aires lexicales et morpho-syntaxiques de l'AIS — des régularités numériques fort bien accusées.

Il va de soi que l'analyse taxatoire faite par nous est bien loin d'être complètement « objective ». Elle comporte une dose inévitable d'arbitraire dont les conséquences numériques peuvent être enrayées

Fig. 2 : 

Nombre de taxats par carte analysée.

La première ligne de la liste est à interpréter comme suit : « De 696 cartes analysées (colonne N-KART, en bas) 104 (colonne N-KART, en haut), ou 14,96 % (colonne %/0-TOT), ne disposent que d'un seul taxat (colonne N-TYP) et sont donc *mononymes* ».

veut symboliser un ensemble de mots qui diffèrent par leurs éléments morphologiques, on peut se contenter d'en indiquer le radical. On procédera de même quand il s'agira de réunir des mots qui se rapprochent par la similitude de leur phonétisme expressif. »

(5) Il faut distinguer les cartes originales de l'AIS (plan des données brutes) des cartes analysées (plan de la matrice de données). Par exemple, la carte AIS 704 *andava spesso* [dal medico] peut être codée deux fois, le premier codage portant sur l'onomasiologie du concept « andava » (*'andava, anava, giva, allait'*, etc.), le second sur celle du concept « spesso » (*'spesso, parecchie volte, tante volte, ogni tanto, souvent'*, etc.). D'où deux cartes analysées (équivalant à autant d'attributs de la matrice de données) tirées d'une seule carte AIS.

N-TYP	N-KART	%-TOT
1	104	14,94
2	147	21,12
3	75	10,78
4	51	7,33
5	43	6,18
6	45	6,47
7	21	3,02
8	24	3,45
9	19	2,73
10	21	3,02
11	13	1,87
12	10	1,44
13	11	1,58
14	19	2,73
15	15	2,16
16	7	1,01
17	8	1,15
18	10	1,44
19	5	0,72
20	1	0,14
21	7	1,01
22	5	0,72
23	5	0,72
24	1	0,14
25	4	0,57
26	2	0,29
27	1	0,14
28	2	0,29
29	5	0,72
30	3	0,43
31	1	0,14
32	1	0,14
33	2	0,29
40	1	0,14
41	1	0,14
45	1	0,14
46	1	0,14
48	1	0,14
49	1	0,14
50	1	0,14
51	1	0,14
	696	

par un contrôle de fiabilité (angl. reliability, all. Reliabilität) ⁽⁶⁾. Pour pouvoir apprécier correctement la part de l'arbitraire personnel dans notre analyse taxatoire il faudrait la faire refaire par plusieurs dialectométriciens à la fois tout en en comparant les résultats par des mesures appropriées ⁽⁷⁾.

Le fait qu'une analyse de ce genre portera toujours l'empreinte personnelle de celui qui la fait, n'en dévalorise en aucune manière la valeur empirique. Ce phénomène, bien connu en matière de statistique et des sciences empiriques (tant exactes que humaines), montre fort bien que tout travail empirique ne se fait pas directement sur les données brutes, mais bien plutôt sur l'image, voire le modèle que le chercheur en conçoit.

Encore deux remarques relatives à la matrice de données utilisée.

1.1. Pondération des attributs.

Aucune pondération d'attributs n'a été prévue. Tous les attributs sont traités sur un pied d'égalité. Nous avons obéi en cela au principe taxométrique de l'adansonisme (ou de l'isocratie) ⁽⁸⁾.

(6) Le contrôle de fiabilité permet de constater si, parmi les attributs consignés dans la matrice des données, il y en a qui s'écartent considérablement de la variabilité générale et sont donc aptes à biaiser les résultats taxométriques. Pour évaluer l'homogénéité de la variabilité nominale des attributs, nous avons constitué, par voie aléatoire, plusieurs sous-matrices partielles, dont le nombre des attributs va de 25 à 450 en passant par 50, 75, 100, 150, 200 et 300 attributs. L'analyse taxométrique faite sur ces matrices partielles montre que les résultats taxométriques commencent à se stabiliser entre 200 et 300 attributs et qu'au-dessus de ce seuil, « l'amélioration » du bilan taxométrique ne se fait que très lentement. Une expérience similaire a été faite par Guiter (1973, 80).

(7) Un tel contrôle concernerait la validité (ou la standardisation) de l'instrument de mesure utilisé : « Essentially, un instrument de mesure est standardisé lorsque plusieurs personnes différentes et indépendantes, qui utilisent le même instrument, obtiennent des résultats très semblables. La formulation de règles explicites est un aspect majeur de la standardisation. » (Gendre 1977, 9).

(8) Selon le botaniste français Michel Adanson (1727-1806) qui avait proposé le premier le principe de l'évaluation homogène des attributs observés. Quant à l'importance que la recherche taxométrique actuelle attribue à l'adansonisme, cf. les passages suivants : « A priori every character is of equal weight in creating natural taxa. » (Sneath/Sokal 1973, 5), et : « Die Klassifikation soll auf möglichst *vielen* synchron beobachtbaren [...] und dazu *ungewichteten* Sprachmerkmalen beruhen. Es ist also nicht nur nicht notwendig, sondern falsch und illusorisch [...] zu glauben, der Fortschritt

1.2. Cases vides.

Tout atlas linguistique contient des lacunes qui, lors de la mise en matrice (codage) des données brutes de l'atlas, causent l'insertion de cases vides dans les lignes et/ou colonnes respectives de la matrice ; voir la Fig. 1. Les cases vides ont l'inconvénient de perturber quelque peu l'exploitation statistique de la matrice de données et correspondent en ceci aux bruitages parasites (angl. noise) en théorie de l'information. Il faut en tenir compte lors de l'interprétation des résultats taxométriques⁽⁹⁾. Pour d'autres détails cf. 5.1. et 5.2.

2. Mesure de similarité et matrice de similarité.

L'étape suivante de la procédure taxométrique consiste à mesurer la *similarité* (*similitude*, *proximité*, *ressemblance*) réciproque des points d'atlas (comparaison des points d'atlas par paires). Voir la Fig. 3.

Ici encore il se pose le problème du choix d'un indice de similarité approprié. Comme le nombre des indices disponibles est considérable⁽¹⁰⁾ et que d'autre part chacun de ces indices répond à des besoins

der typologischen Sprachforschung hänge von der Möglichkeit ab, die 'saliency' von Sprachmerkmalen und ihre gegenseitigen Beziehungen zu bewerten. Nur wenn wir uns an den Grundsatz halten, die Sprachmerkmale nicht verschieden zu gewichten, können wir hoffen, eine objektive und nachvollziehbare Klassifikationsprozedur zu entwickeln. Erst a posteriori können wir die Merkmale gewichten, und zwar nur numerisch. » (Altmann/Lehfeldt 1973, 32).

(9) Cf. Sneath/Sokal 1973, 178 sqq. ; Sodeur 1974, 48 sqq. ; Vogel 1975, 77 sqq. ; Bock 1974, 75-76.

(10) Pour les indices de similarité cf. Sneath/Sokal 1973, 129-137 ; Chandon/Pinson 1981, 72-79 ; Bock 1974, 48-80.

Ajoutons que, pour d'autres visées classificatoires, par exemple pour le traitement du problème des isoglosses quantitatives, on doit remplacer la mesure de la ressemblance (s) par celle de la distance (d) entre deux points d'atlas ; cf. à cela Séguy 1971 et 1973, Guiter 1973 et 1979, et Goebel 1976, 174-175. Tout indice de similarité (s) peut être transformé en indice de dissimilarité ou de distance (d). Parmi les nombreuses possibilités d'effectuer cette transformation, nous ne signalons ici que la plus simple : $1 = s + d$; cf. Chandon/Pinson 1981, 51-53. Dans Goebel 1976 l'indice de similarité est appelé (provisoirement) « IGI » (Indice Général d'Identité ; 173) alors que l'indice de distance a reçu le nom (également provisoire) de « IGR » (Indice Général de Rendement ; 174). Vu que le maximum de la similarité générale (en valeurs absolues) se situe, pour la matrice de données utilisée dans cet article, à 1468 (173), la relation entre l'IGI et l'IGR serait la suivante : $1468 = \text{IGI} + \text{IGR}$.

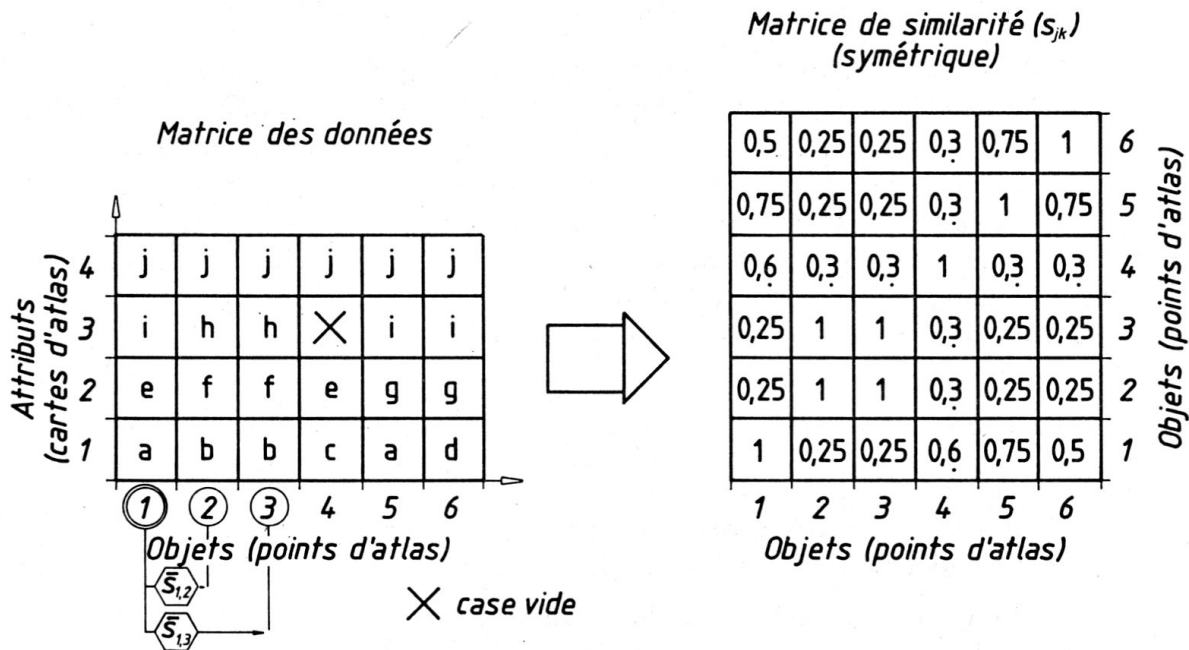


Fig. 3 :

Génération de la matrice de similarité à partir de la matrice des données par l'application d'un indice de similarité \bar{s} .

taxométriques précis, il est donc nécessaire que le dialectométricien règle son choix sur les présuppositions conceptuelles de la géolinguistique traditionnelle. Or, quant au degré de ressemblance de deux dialectes voisins, l'opinion des dialectologues est assez nette, et ceci depuis belle lurette : « Et maintenant, qu'est-ce qui constitue le degré de ressemblance qui rapproche deux langues entre elles, et le degré de dissemblance qui les éloigne l'une de l'autre ? La ressemblance se mesure à la proportion des caractères communs, la dissemblance à la proportion des caractères particuliers. » (Durand 1889, 63).

L'indice de similarité choisi par nous se définit donc comme suit :

$$s = 100 \cdot \frac{\text{nombre des co-identités d'attributs par paire d'objets}}{\text{nombre des co-présences d'attributs par paire d'objets}}$$

Cet indice se situe entre 0 et 100 (%). Ses propriétés algébriques le rapprochent du « simple matching coefficient » de la taxonomie numérique : cf. Sneath/Sokal 1973, 132 ; Bock 1974, 51 ; Chandon/Pinson 1981, 74. Dans nos travaux dialectométriques en langue française nous appelons cet indice « Indice Relatif d'Identité » (IRI), dans ceux de langue allemande « Relativer Identitätswert » (RIW).

L'indice \bar{s} de la Fig. 3 ne se distingue de l'IRI que par le manque de mise en pourcentage : $100.\bar{s} = \text{IRI}$. Comme les mesures de proximité se font par paires d'objets (deux à deux), l'on peut calculer, pour n objets, n^2 indices de proximité qui ensuite seront consignés dans une matrice de similarité symétrique (voir la Fig. 3, à droite). Etant donné que l'indice de similarité utilisé est symétrique ($s_{x,y} = s_{y,x}$) et que chacun des n^2 scores sera calculé deux fois et que, finalement, les valeurs inscrites le long de la diagonale sont dépourvues d'intérêt taxométrique, il est possible de transformer la matrice symétrique rectangulaire en une matrice asymétrique triangulaire, sans perte d'informations codées (comparer les Fig. 3 et 4).

Vérification de quelques indices \bar{s} de la matrice de similarité de la Fig. 3.

— Soit la proximité entre les PP. 1 et 2 :

- nombre des co-identités d'attributs : 1 (par le taxat j)
- nombre des co-présences d'attributs : 4 (par les paires de taxats a/b, e/f, i/h, j/j)

$\bar{s}_{1,2}$ est donc de $1/4 = 0,25$.

— Soit la proximité entre les PP. 1 et 3 :

- nombre des co-identités d'attributs : 1 (par le taxat j)
- nombre des co-présences d'attributs : 4 (par les paires de taxats a/b, e/f, i/h, j/j)

$\bar{s}_{1,3}$ est donc de $1/4 = 0,25$.

— Soit la proximité entre les PP. 1 et 4 :

- nombre des co-identités d'attributs : 2 (par les taxats e et j)
- nombre des co-présences d'attributs : 3 (par les paires de taxats a/c, e/e, j/j)
- Remarquer que la case vide dans le vecteur du P. 4 entraîne la disparition d'une co-présence d'attributs.

$\bar{s}_{1,4}$ est donc de $2/3 = 0,6$.

Vérifier ces résultats dans la matrice de similarité de la Fig. 3.

En taxométrie l'ensemble des valeurs de proximité stockées dans la matrice de similarité est appelé « préordonnance » (cf. Chandon/Pinson 1981, 54 ; Bertier/Bouroche 1975, 13). Comme la dialectométrie, de par la contiguïté spatiale des objets à classifier (points d'atlas), est sujette à certaines contraintes conceptuelles, l'on a intérêt à commencer

par faire une analyse partitionnelle de l'information numérique consignée dans les différents vecteurs de la matrice de similarité ; voir la Fig. 4 (surtout les secteurs en gris).

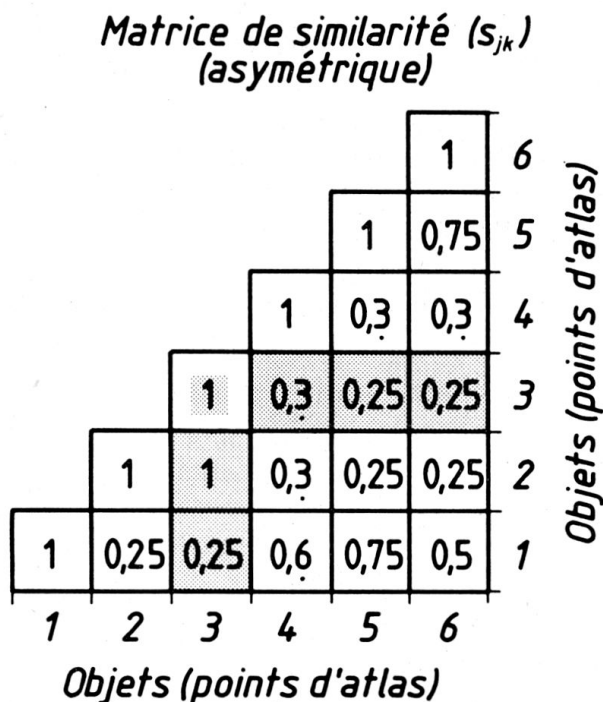


Fig. 4 :

Localisation d'une distribution de similarité (en gris) à l'intérieur de la matrice de similarité.

Chaque vecteur (brisé ou non) contient n valeurs (scores), dont une valeur égale à 1, et $n-1$ valeurs oscillant entre 1 et 0. Ces $n-1$ valeurs seront appelées dorénavant « distribution de similarité ». Chaque distribution de similarité est à interpréter par rapport à la valeur diagonale 1, c'est-à-dire en fonction du point d'atlas auquel elle appartient (point de référence) ⁽¹¹⁾. Dans la Fig. 4, le secteur hachuré montre la distribution de similarité appartenant au point de référence 3. Les cinq valeurs (scores) qui la constituent (0,25/1/0,3/0,25/0,25) représentent

(11) La mise en rapport d'une valeur-étalon avec un certain nombre d'autres valeurs correspond au problème bien connu de la dialectologie classique, qui, le plus souvent, est évoqué dans les termes suivants : « Quelle est la position du dialecte A parmi les dialectes B, C, D, etc. ? ».

les similarités (\bar{s}) entre les points 1, 2, 4, 5 et 6 d'un côté, et le point de référence 3 de l'autre. En outre, l'on constate aisément que le schéma de la Fig. 4 contient en tout six distributions de similarité différentes, donc une pour chaque point d'atlas.

Il en appert que pour 251 points d'atlas (c'est l'étendue de notre réseau-AIS) il y a autant de distributions de similarités différentes, chacune comprenant 250 valeurs au-dessous de 100 % (IRI). Deux de ces 251 distributions de similarité seront présentées par la suite (voir les Fig. 9 - 14). L'analyse des distributions de similarité présuppose la mise au point de procédures taxométriques aptes à en assurer la visualisation adéquate et à en dégager les propriétés statistiques.

3. Notice servant à l'intelligence des cartes choroplèthes et des surfaces statistiques lissées.

Voir les Fig. 9, 12, 17, 20, 23, 27 et 31 (cartes choroplèthes) et les Fig. 10, 13, 18, 21, 25, 28 et 32 (surfaces statistiques lissées).

3.1. Algorithme d'intervallisation MINMWMAX.

Parmi les différentes méthodes cartographiques utilisées pour la visualisation de différences numériques à l'intérieur d'un réseau de données ponctuelles, la carte choroplèthe occupe une place de choix. Cette méthode consiste pour l'essentiel à transposer une variation quasiment continue d'un grand nombre de valeurs numériques en une variation ordonnée (et discrète) d'un nombre fort restreint de paliers visuels (teintes, couleurs, hachures, etc.). La mise en carte des valeurs numériques exige donc l'application d'un algorithme de transposition pour assurer le groupement des valeurs initiales en un nombre limité d'intervalles. Pour ce faire, la cartographie statistique offre tout un arsenal de possibilités classificatoires⁽¹²⁾ : cf. Dickinson 1973, 83 sqq., Jenks 1963, Jenks/Coulson 1963, Jenks/Caspall 1971 et Scriptor 1970.

(12) A ce sujet cf. le passage suivant : « La cartographie constitue un cas spécial de classification lorsqu'il s'agit de constituer des classes pour une variable à cartographier. Les procédés utilisés sont nombreux et il faut choisir le mieux adapté ou en imaginer un qui convienne mieux encore. Il ne saurait être question d'aborder ce vaste thème ici, mais il faut insister sur le fait que les méthodes de classification peuvent aider à le traiter. En effet, des problèmes analogues se posent : celui du nombre de classes, celui de la taille des classes, celui de leur variance interne, celui des discontinuités aux limites de classes. » (Béguin 1979, 246).

L'algorithme utilisé dans cet article (sigle : MINMWMAX) a été mis au point après toute une série d'expériences systématiques en matière de cartographie statistique et s'est avéré, parmi tous les algorithmes expérimentés, comme étant le plus approprié à nos visées classificatoires⁽¹³⁾. Le principe de l'algorithme MINMWMAX est le suivant :

- nombre des intervalles : 6, répartis de par et d'autre de la moyenne arithmétique ;
- limites d'intervalle : les écarts numériques entre le minimum et la moyenne arithmétique et entre le maximum et la moyenne arithmétique sont respectivement divisés par trois. Les valeurs ainsi déterminées constituent les tailles des intervalles situés au-dessous et au-dessus de la moyenne arithmétique.

Exemple : voir les Fig. 9 et 11 :

- moyenne arithmétique : 50,204
- minimum : $\underline{-42,541}$
 $7,663 : 3 = 2,554$
- taille des intervalles au-dessous de la moyenne arithmétique : 2,554 ;
- maximum : 88,071
- moyenne arithmétique : $\underline{-50,204}$
 $37,867 : 3 = 12,622$
- taille des intervalles au-dessus de la moyenne arithmétique : 12,622.

Quant aux limites des intervalles, on peut en dresser la liste suivante (voir aussi les Fig. 9 et 11) :

Intervalle	de	à	Taille de l'intervalle
1	42,541	45,095	2,554
2	> 45,095	47,650	2,554
3	> 47,650	50,204	2,554
4	> 50,204	62,826	12,622
5	> 62,826	75,449	12,622
6	> 75,449	88,071	12,622

(13) Pour les algorithmes MEDMW et MED (pas traités dans cet article) cf. Goebel 1977 b, 345-346.

De tous les avantages qu'offre l'algorithme en question, le plus remarquable réside dans le fait que, par l'affectation de certaines valeurs aux intervalles 6 et 5 et la classification spatiale ainsi établie, il est possible de reconstituer, par voie taxométrique, les « aires dialectales » de la dialectologie classique et de greffer ainsi les résultats de la dialectométrie sur ceux de l'aréologie traditionnelle. En outre ce n'est que par l'application de l'algorithme MINMWMAX que l'on obtient les types iconiques décrits ci-dessous (5.1.2.) et dont la présentation successive en fondu enchaîné constitue le principe de la « cartographie dialectale dynamique » (cf. ci-dessous, 5.1.2.).

Du point de vue taxométrique l'application de tels algorithmes d'intervallisation représente une analyse partitionnelle en classes disjointives.

3.2. Histogrammes.

Voir les Fig. 9, 12, 17, 20, 23, 27 et 31.

Les histogrammes figurant en bas de page sur nos cartes choroplèthes permettent d'analyser rapidement certaines propriétés statistiques des distributions de similarité examinées (position, dispersion, symétrie, voussure, etc.). Leur construction géométrique met à profit les données classificatoires fournies par l'algorithme d'intervallisation MINMWMAX mentionné ci-dessus (3.1.), avec la différence cependant que les écarts entre le maximum, voire le minimum et la moyenne arithmétique sont divisés par six.

Pour la construction des 12 (j) colonnes de l'histogramme on porte

- en abscisse : la taille (d_j) de l'intervalle considéré, et
- en ordonnée : le quotient (p_j/d_j) entre la fréquence relative des points d'atlas affectés à l'intervalle considéré (p_j) et la taille de cet intervalle (d_j).

3.3. Construction des polygones de Thiessen.

Voir les Fig. 9, 12, 17, 20, 23, 27 et 31.

Le fond des cartes choroplèthes a subi un traitement cartographique particulier. Les 251 points d'atlas dont l'espacement sur le fond de carte correspond à la réalité géographique (mis à part quelques généralisations schématiques aux environs de Trieste, P. 369, en Istrie et dans le golfe de Quarnéro) ont été entourés d'un réseau polygonal afin de remplacer les implantations ponctuelles des données du réseau

(représentation discrète) par un agencement disjonctif de surfaces géométriques (représentation quasi-continue). La construction des polygones de Thiessen est d'ailleurs fort simple ; voir la Fig. 5.

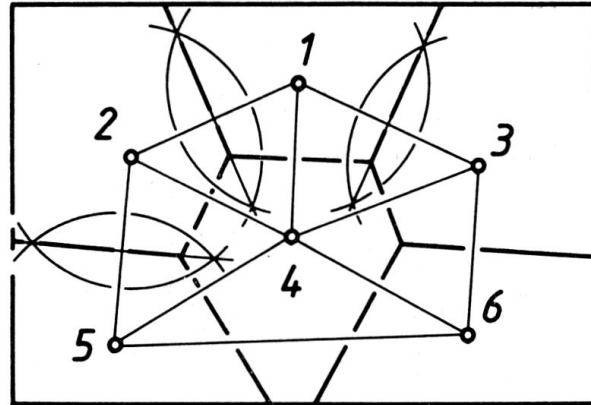


Fig. 5 :

Construction de polygones de Thiessen dans un réseau de 6 points d'atlas.

En voici les étapes constructives :

1. triangulation des points du réseau examiné,
2. construction des médiatrices des segments de droite reliant deux points du réseau,
3. prolongement des médiatrices jusqu'à ce qu'elles se rencontrent,
4. les points de concours des médiatrices constituent les angles des polygones de Thiessen.

Au sujet des polygones de Thiessen cf. Thiessen 1911 et Haggett 1973, 277 ; Brassel/Reif 1979, Pudlatz 1977, Rhynsburger 1973.

3.4. Surfaces statistiques lissées.

Voir les Fig. 10, 13, 18, 21, 25, 28 et 32.

Les cartes choroplèthes sont bien loin d'être les seules manières de visualisation de données spatiales. Grâce aux progrès spectaculaires du calcul électronique, il est désormais possible de calculer les contours d'un relief lissé représentant les proportions des variations numériques d'un certain nombre de valeurs ponctuelles. Bien que l'effet stéréographique de telles surfaces statistiques lissées soit souvent très éloquent

(voir p. ex. les Fig. 10, 18, 25 et 28), il convient de ne pas en surestimer l'importance cartographique. C'est pourquoi leur emploi ne se justifie guère que pour des représentations sommaires où la distribution des masses passe avant l'exactitude des détails. Quant à l'expression des détails, au rendu du modelé et à la représentation exacte des rapports et proportions numériques, les cartes choroplèthes l'emportent de loin sur les surfaces statistiques lissées. Pour plus de détails cf. Monmonier 1978 et Peucker 1977.

4. Notice servant à l'intelligence des cartes numériques.

Voir les Fig. 11, 14, 19, 22, 26, 29 et 33.

Toute carte numérique contient une en-tête (reproduite en haut de page, à gauche et à droite) et des indications numériques relatives aux différents points d'atlas.

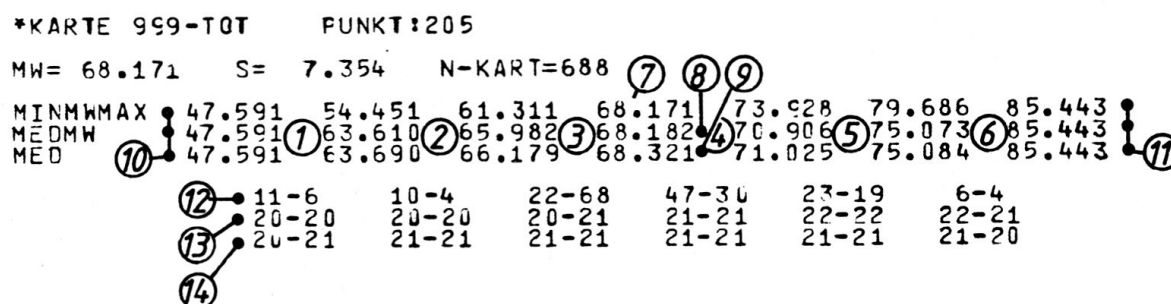


Fig. 6 :

En-tête de la carte numérique de la distribution de similarité relative au point de référence 205.

Voir aussi la Fig. 14.

Les valeurs numériques de l'en-tête doivent être interprétées de la manière suivante :

PUNKT : 205	Point de référence 205.
MW	Moyenne arithmétique (all. Mittelwert).
S	Ecart-type (all. Standardabweichung).
N-KART	Nombre des attributs disponibles (taxats) dans le vecteur du point 205 de la matrice des données. Comme tout vecteur d'un point de la matrice contient 696

cases et que de ces 696 cases 688 seulement détiennent des attributs analysés (taxats), il en ressort que ce vecteur dispose de 8 cases vides.

- 1...6 Intervalles (cf. 3.1).
- 7 Moyenne arithmétique de la distribution de similarité (valeur continue).
- 8 Valeur la plus proche de la moyenne arithmétique (valeur discrète).
- 9 Médian (valeur qui divise la distribution de similarité en deux parties égales).
- 10 Minimum.
- 11 Maximum.
- 12 Indications relatives au nombre des points d'atlas par intervalle selon l'algorithme d'intervallisation MINMWMAX :
- partition en 6 intervalles :
 intervalle 1 : $11 + 6 = 17$ points,
 intervalle 2 : $10 + 4 = 14$ points, etc., cf. 3.1.
 - partition en 12 intervalles :
 intervalle 1 : 11 points,
 intervalle 2 : 6 points, etc., cf. 3.2.
- 13 Indications relatives au nombre des points d'atlas par intervalles selon l'algorithme d'intervallisation MEDMW (pas traité dans cet article) ⁽¹⁴⁾.
- 14 Indications relatives au nombre des points d'atlas par intervalles selon l'algorithme d'intervallisation MED (pas traité dans cet article) ⁽¹⁴⁾.

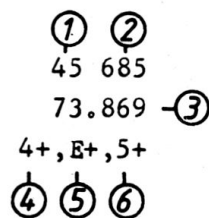


Fig. 7 :

Données numériques relatives à un point d'atlas.

Les valeurs numériques répertoriées à l'emplacement d'un point d'atlas doivent être lues comme suit :

- 1 Indicatif du point d'atlas 45.
- 2 Nombre des attributs disponibles (taxats) dans le vecteur du point 45 de la matrice des données.
- 3 Indice Relatif d'Identité (IRI) entre le point de référence 205 et le point comparé 45.
- 4 Classement par intervalle selon l'algorithme d'intervallisation MINMWMAX :
classement dans l'intervalle 4 selon MINMWMAX à 6 intervalles ;
classement dans l'intervalle 8 (= 4 +) selon MINMWMAX à 12 intervalles.
- 5 Classement par intervalle selon l'algorithme d'intervallisation MEDMW (pas traité dans cet article) ⁽¹⁴⁾.
- 6 Classement par intervalle selon l'algorithme d'intervallisation MED (pas traité dans cet article) ⁽¹⁴⁾.

Pour la bonne intelligence des Fig. 11 et 14 (cartes numériques) contenant chacune une distribution de similarité, il importe de toujours consulter simultanément les indications numériques du point de référence et du point comparé. Voir la Fig. 8.

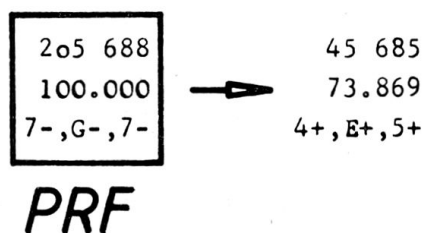


Fig. 8 :

Mesure de la similarité (selon IRI) entre les points 205 et 45.

Voir aussi la Fig. 14.

PRF - point de référence.

(14) Cf. la note 13.

Il s'agit d'une comparaison entre les points 205 et 45. Le nombre des attributs (taxats) disponibles (et, partant, calculables) dans le vecteur du P. 205 est de 688, et de 685 dans celui du P. 45. Entre les vecteurs 205 et 45 il y a donc 685 co-présences d'attributs. De ces 685 attributs (ou taxats) 73,869 % — soit 506 ($= 685 \cdot 0,73869$) — sont co-identiques avec ceux du vecteur 205. Ceci signifie que la similarité entre les PP. 205 et 45 ($IRI_{205,45}$) est de 73,869 %.

5. Deux cartes de similarité.

5.1. Carte de similarité relative au point de référence 15 (Mathon, Grison centraux).

Voir les Fig. 9, 10 et 11. Quant au P. 15 de l'AIS cf. Jaberg/Jud 1928, 41.

5.1.1. Interprétation géolinguistique.

Le relief de la Fig. 9 montre l'agencement spatial des similarités typologiques par rapport au point de référence 15⁽¹⁵⁾. La carte choroplèthe (Fig. 9) révèle que les valeurs maximales (polygones dans l'intervalle 6) se concentrent autour du point de référence et ne dépassent pas le cadre géographique étroit des Grisons occidentaux et centraux. Les polygones de l'Engadine rentrent tous dans l'intervalle 5. Remarquer la déclivité relativement brusque du pavage de similarité entre les Grisons et la Lombardie, ainsi que — vers l'est — l'apparition des affinités intra-rhétoromanes dans le Tyrol du Sud (PP. 312, 314, 313) et — d'une manière déjà fort estompée — dans le Frioul. Les antipodes typologiques (intervalle 1) se situent en Vallée d'Aoste, en Toscane, en Romagne, dans les Marches, en Vénétie et en Istrie. L'IRI maximal (88,071) est au P. 17 dans le voisinage immédiat du point de référence, alors que l'IRI minimal (42,541) se trouve au P. 122 (Vallée d'Aoste). L'on constate en plus qu'un certain nombre de polygones disséminés dans le centre de la carte, répertoriés le plus souvent dans l'intervalle 4 et marqués d'un astérisque (voir la mention de la légende « punti con corpo ridotto ») ne s'intègrent que très mal dans la logique du décalage spatial des valeurs. Ce sont des points d'atlas dont les vecteurs matri-

(15) Le principe des cartes de similarité a déjà été appliqué par Reed/Spicer (1952, 358) et Cowan (1964, 16-17).

ciels contiennent beaucoup de cases vides (cf. 1.2.)⁽¹⁶⁾, d'où de forts biaisements des scores numériques respectifs. Il est bon de les exclure de l'interprétation synoptique des cartes choroplèthes.

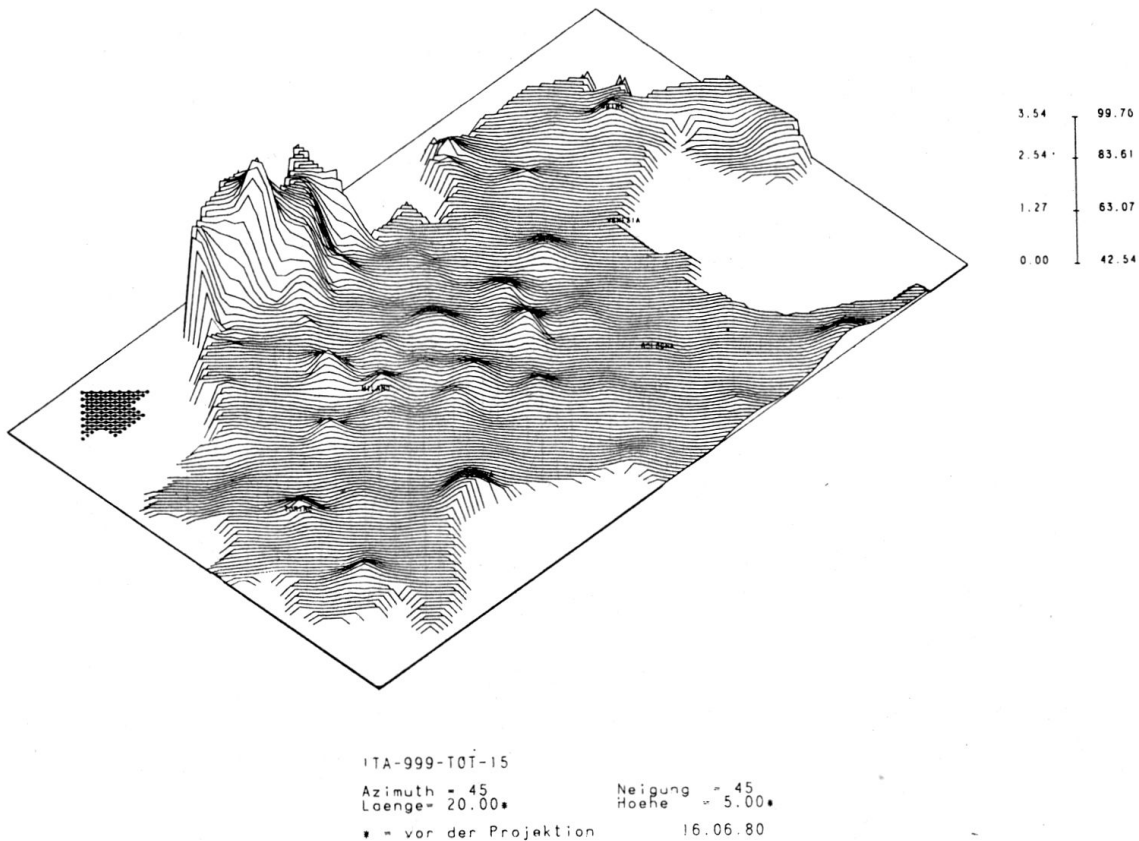


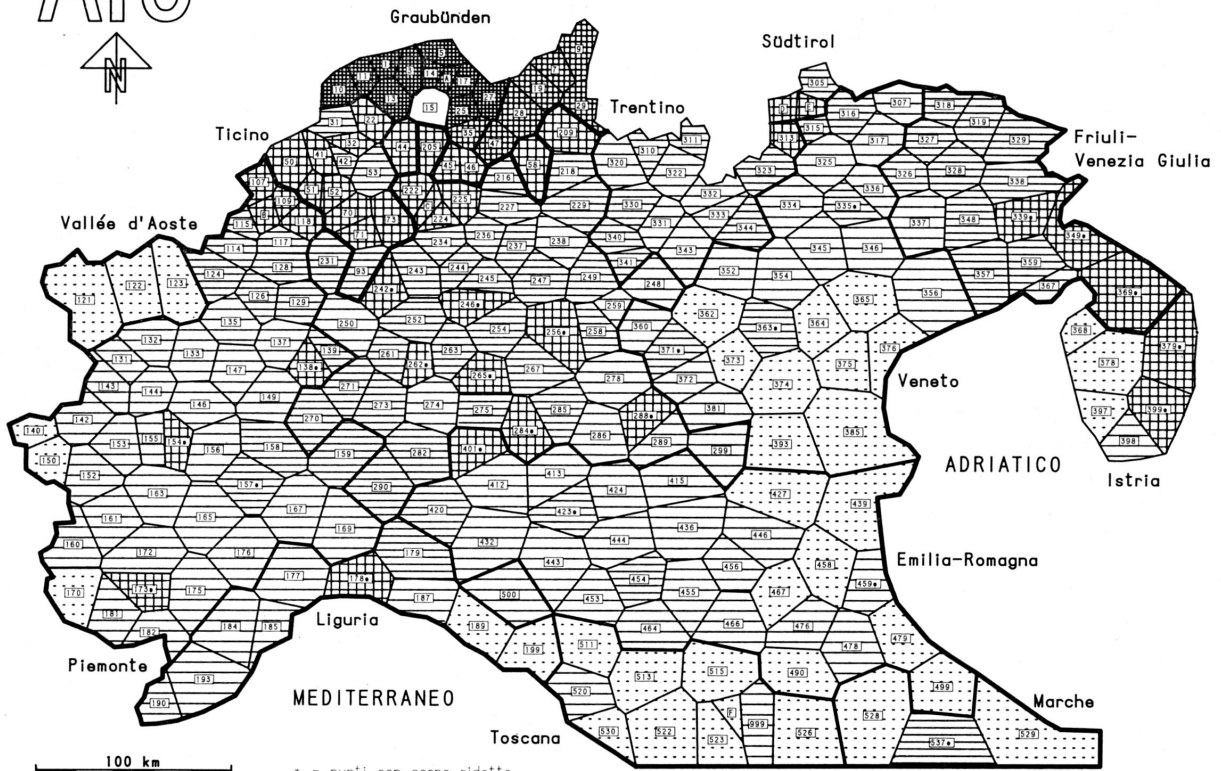
Fig. 10 :

Surface statistique lissée de la distribution de similarité relative au points de référence 15 (Mathon, Grisons centraux).

Voir aussi les Fig. 9 et 11.

(16) Rappelons que le réseau de l'AIS comprend un certain nombre de points d'enquête où les explorateurs n'ont utilisé qu'un questionnaire réduit (cf. Jaberg/Jud 1928, 175-176). Il va de soi que les vecteurs de ces points de la matrice de données ne peuvent contenir qu'un nombre réduit de taxats par rapport aux autres qui, eux, reposent sur l'analyse du questionnaire complet.

AI5



100 km

* = punti con corpo d'acqua

LEX - ITALIEN

999-TOT-15

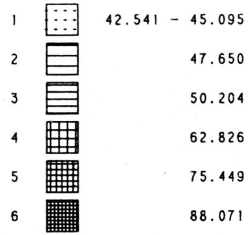
Punkte mit Buchstaben:

- A - 16
- B - 116*
- C - 223*
- D - 312
- E - 314*
- F - 524



Kartographie: W.-D. Rase

Choroplethenkarte
MINWMAX 6-fach



Häufigkeitsverteilung
MINWMAX 12-fach

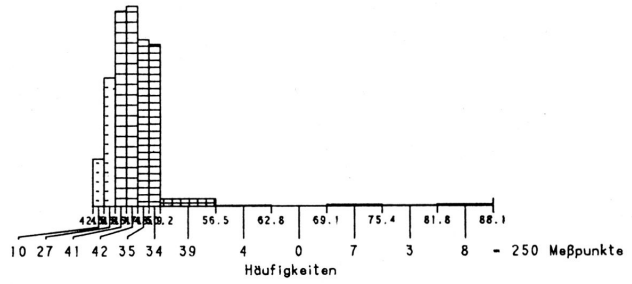


Fig. 9:

Carte choroplèthe de la distribution de similarité relative au point de référence 15 (Mathon, Grisons centraux).
Voir aussi les Fig. 10 et 11.

*KARTE 509-101 PUNKT: 15
NW= 50.204 S= 8.851 N-KERT=097

MINMAX 42.541 45.095 47.650 50.204 52.758 55.312 57.866 60.420 62.974 65.528
MEIN 42.541 45.095 47.650 50.204 52.758 55.312 57.866 60.420 62.974 65.528
MEI 42.541 45.095 47.650 50.204 52.758 55.312 57.866 60.420 62.974 65.528

STREIFEN 1

SWIZZERA
RUMANTSCHA

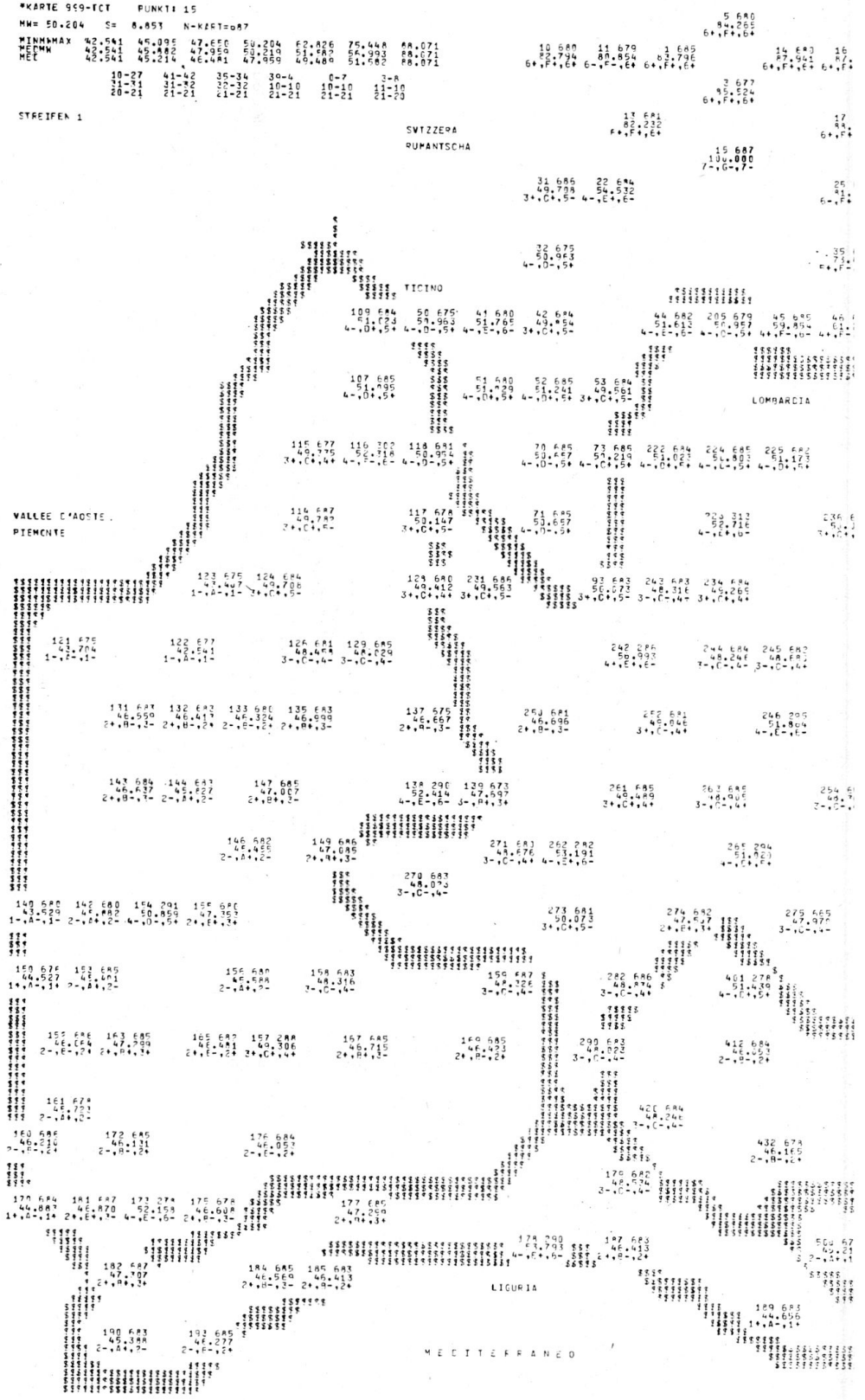
TICINO

LOMBARDIA

VALLEE D'AOSTE
PIEMONTE

LIGURIA

MEDITERRANEO



L'indice de similarité utilisé (IRI) est d'un abord interprétatif facile. Soient les deux scores suivants :

$$\text{IRI}_{15,3} = 85,524 \text{ et}$$

$$\text{IRI}_{15,522} = 42,815.$$

Vu la logique algébrique de l'IRI, l'on peut donc dire tout simplement que l'affinité linguistique entre les points 15 et 3 est à peu près le double de l'affinité linguistique entre les points 15 et 522. Pour d'autres indices de proximité dont la facture algébrique est plus compliquée, de telles évaluations interprétatives des résultats numériques sont beaucoup plus délicates.

La logique iconique de la stratification spatiale des indices de proximité ressort également fort bien de la Fig. 10 (surface statistique lissée).

5.1.2. Dialectométrie et cartographie dynamique.

Il importe de souligner que lors d'un déplacement progressif du point de référence à l'intérieur d'un domaine dialectal typologiquement cohérent (en l'occurrence dans les Grisons), les différentes distributions de similarité restent tant soit peu invariées. Ce n'est qu'en franchissant une barrière linguistique d'une certaine importance (p. ex. entre les Grisons et la Lombardie septentrionale) que l'on obtient des profils de similarité sensiblement différents. Voir p. ex. la Fig. 12 qui montre un profil typiquement lombard (point de référence 205) et dont la syntaxe iconique est complètement différente de celle de la Fig. 9.

Par la présentation consécutive de plusieurs profils de proximité dont les points de référence sont alignés le long d'un parcours préalablement choisi, on obtient un effet cinématographique très éloquent. Cet effet cinématographique réussit particulièrement bien quand la projection consécutive des profils de proximité est effectuée en fondu enchaîné à l'aide de deux projecteurs de diaporama. Nous avons présenté de tels montages diaporamiques à plusieurs reprises avec toujours des parcours d'épreuve différents⁽¹⁷⁾. Les changements iconiques qui interviennent d'un profil à l'autre, se font en fonction de la proximité

(17) Ces projections avaient lieu en 1977 à Vienne (cf. Goebel 1978 b), Coire et Marburg (cf. Goebel 1977 b, 335), en 1978 à Brixen (Tyrol du Sud), en 1979 à Sarrebruck, en 1980 à Salzbourg, et en 1981 à Caen et à Klagenfurt (Carinthie).

taxatoire des points de référence des distributions de similarité converties en images. Ils permettent d'obtenir une vision très claire de l'enchevêtrement typologique du réseau dialectal examiné.

5.1.3. Interprétation géographique.

En géographie, c'est la notion d'*interaction* qui répond à la notion linguistique de l'*affinité* ou *parenté* (phonétique, morpho-syntaxique, lexicale, étymologique ou autre). En regardant la Fig. 9 on peut donc dire que les flux (ou flots) d'interaction (de trafic, de transport ou de communication) sont beaucoup plus importants en direction ouest-est que dans la direction nord-sud d'une part, et que de l'autre leur capacité interactive n'est vraiment bonne que dans un espace fort limité. D'une comparaison des Fig. 9 et 12 (point de référence 205) il ressort que la centralité du point de référence 205 est de loin supérieure à celle du point de référence 15.

Une considération serrée des problèmes posés par les flots circulant à l'intérieur d'un réseau interactif mène directement à la théorie des graphes, à la théorie des réseaux de transport, à la théorie des circuits électriques et à la théorie des jeux⁽¹⁸⁾. C'est ici que la dialectométrie gagne rapidement en profondeur théorique et en intérêt pluridisciplinaire. Comme il est bien évident que de tels approfondissements dépasseraient largement le cadre de cet article, nous nous passons ici d'autres remarques, quitte à reprendre la trace indiquée ailleurs et une autre fois.

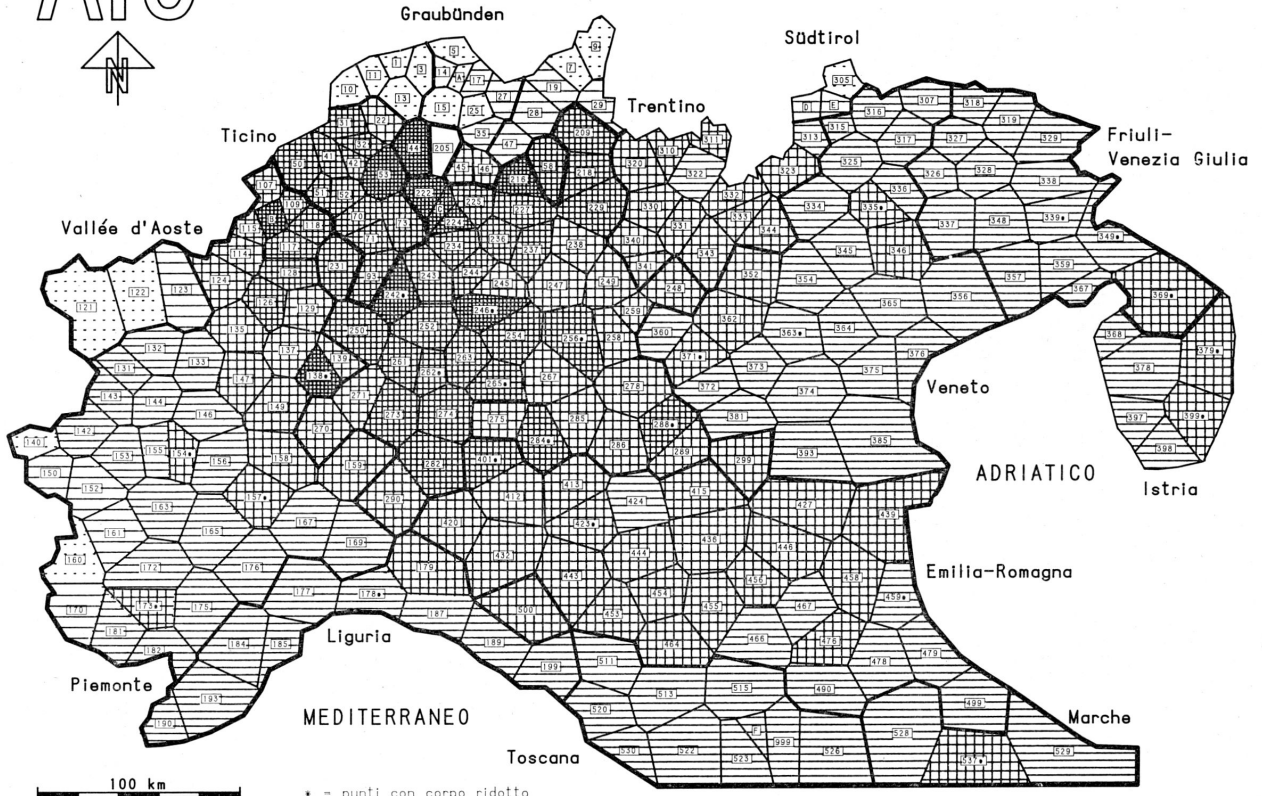
Avant de conclure ce paragraphe nous aimerions souligner l'intérêt qu'offrent les cartes de similarité pour l'étude géographique de la diffusion, de l'évolution de l'habitat et des groupements de population, ainsi que pour celle de la théorie des places centrales (cf. Haggett 1973, 135-153).

5.2. Carte de similarité relative au point de référence 205 (Cam-podolcino, Lombardie septentrionale).

Voir les Fig. 12, 13 et 14. Quant au P. 205 de l' AIS cf. Jaberg/Jud 1928, 63.

(18) Cf. les ouvrages de Berge 1967, Berge/Ghouila-Houri 1962, Flament 1968 et Sache 1974, où le lecteur intéressé trouvera des démonstrations bien équilibrées entre la théorie et l'application.

AI S



100 km

* = punti con corpo ridotto

LEX - ITALIEN

999-T0T-205

Punkte mit Buchstaben:

- A - 16
- B - 116*
- C - 223*
- D - 312
- E - 314*
- F - 524



Kartographie: W.-D. Rose

Choroplethenkarte
MINMWMAX 6-fach

1		47.591 - 54.451
2		61.311
3		68.171
4		73.928
5		79.686
6		85.443

Häufigkeitsverteilung
MINMWMAX 12-fach

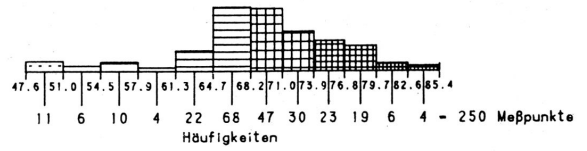


Fig. 12 :

Carte choroplèthe de la distribution de similarité relative au point de référence 205
(Campodolcino, Lombardie septentrionale).

Voir aussi les Fig. 13 et 14.

*KARTE 959-TCT PUNKT1205

MW= 68.171 S= 7.354 N-KART=688

MINMAX 47.991 54.451 61.311 68.171 73.928 79.686 85.443
MEDMM 47.991 63.610 65.982 68.182 70.966 75.073 85.443
MED 47.991 63.690 66.179 68.321 71.025 75.084 85.443

5 688
49.555
1-A,-1-
10 681 11 680 1 685
49.226 48.824 47.591
1-A,-1- 1-A,-1- 1-A,-1-
14 678 16
69.709 72.
1-A,-1- 1-A,-1-
3 677
48.445
1-A,-1-
13 683
48.023
1-A,-1-
17
54.
2-A,-1-
15 679
1-A,-1-
25
52.
1-A,-1-
31 685 22 685
77.843 69.635
5-F,-6- 4-D,-6-
32 678
78.991
5-F,-6-
35
68.
2-A,-1-
109 684 90 677 41 682 42 685
74.415 78.139 77.495 77.226
5-E,+5+ 5-F,+6+ 5-F,-6- 5-F,-6-
44 683 205 688 45 685 46
6-F,+6+ 7-D,-7- 4-E,+5+ 4-E,-
107 686 51 680 52 685 53 685
73.032 77.794 77.810 79.854
4-F,+5+ 5-F,-6- 5-F,-6- 6-F,+6+
115 679 116 302 118 692 \$ 70 685 73 686 222 686 224 685 225 682
70.987 80.132 75.073 78.102 78.863 82.653 83.212 78.446
4-E,-4+ 6-F,+6+ 5-F,-5+ 5-F,+6+ 5-F,+6+ 6-F,+6+ 6-F,+6+ 5-F,+6+

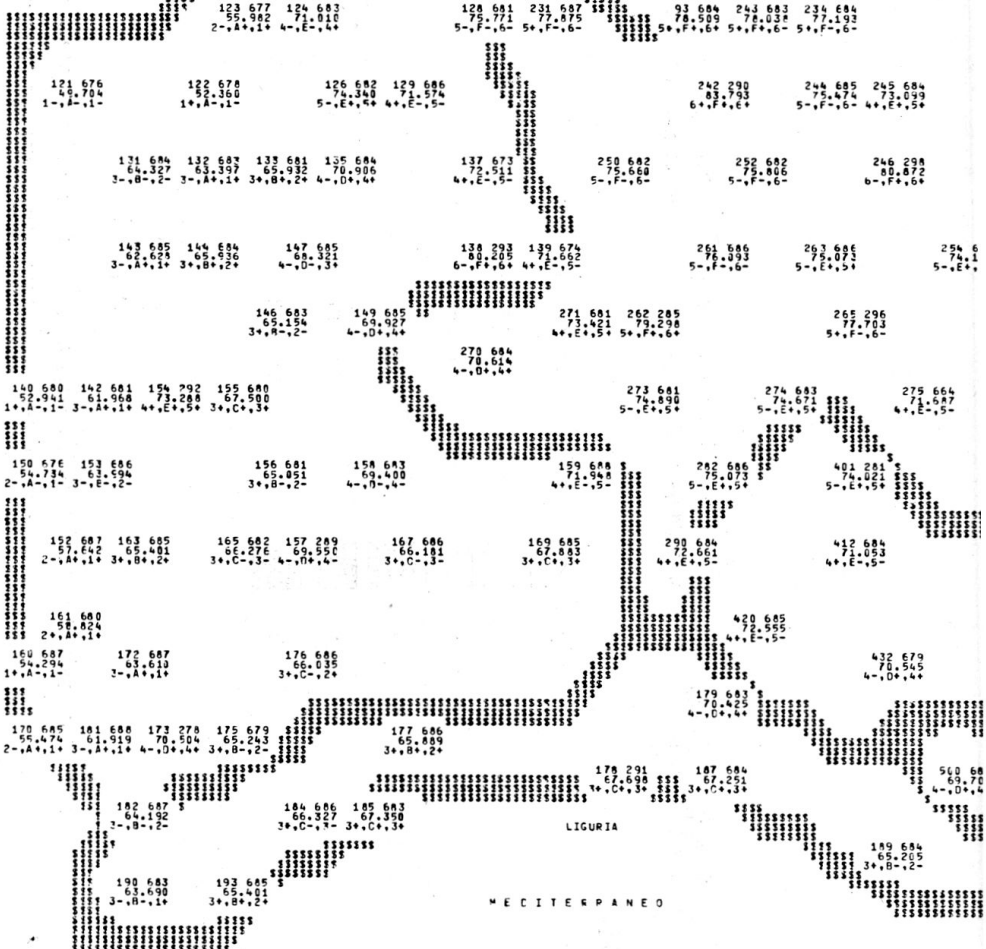
STREIFEN 1

SVIZZERA
ROMANTSCHA

VALLEE D'AOSTE
PIEMONTE

TICINO

LOMBARDIA



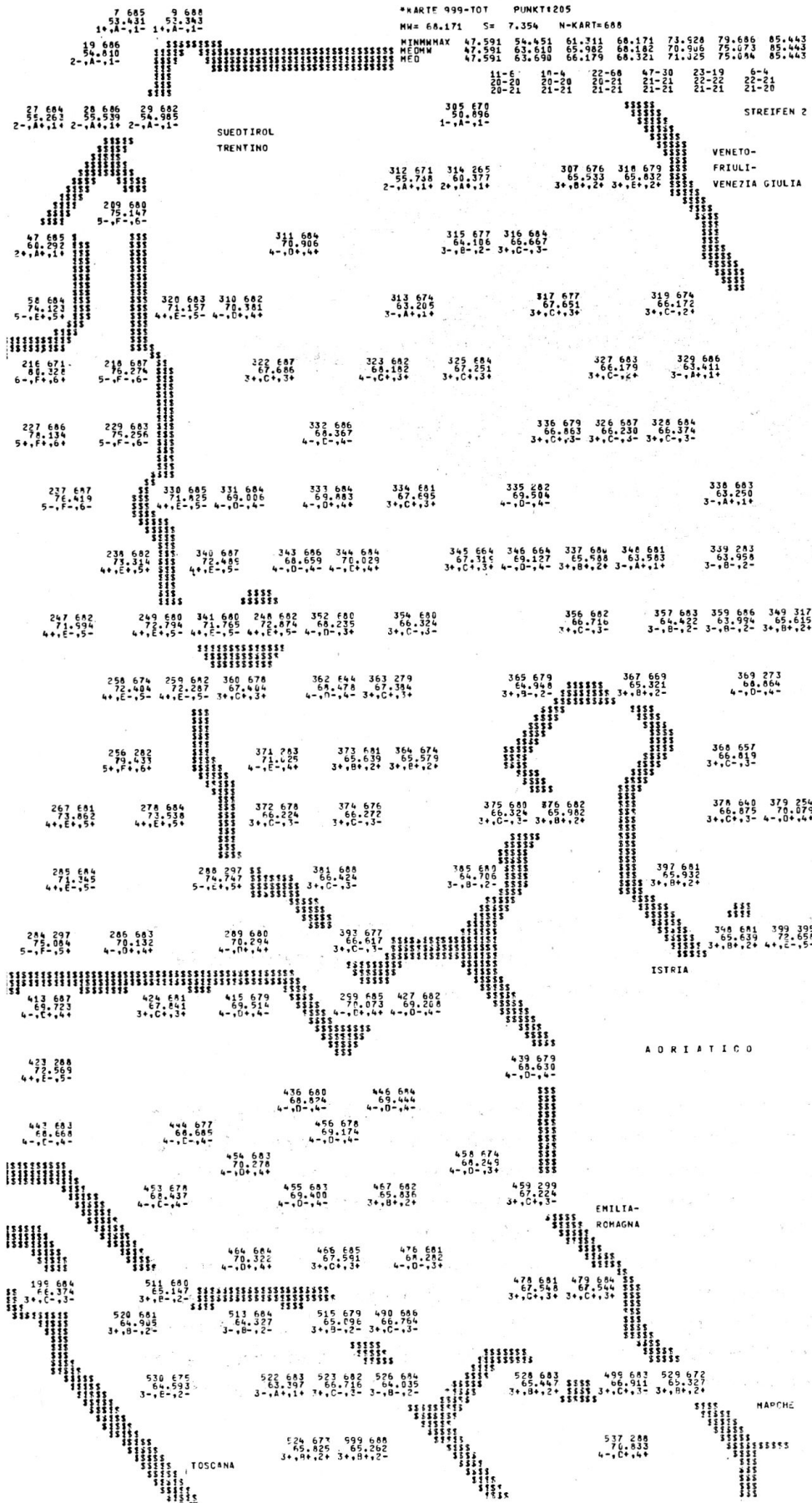
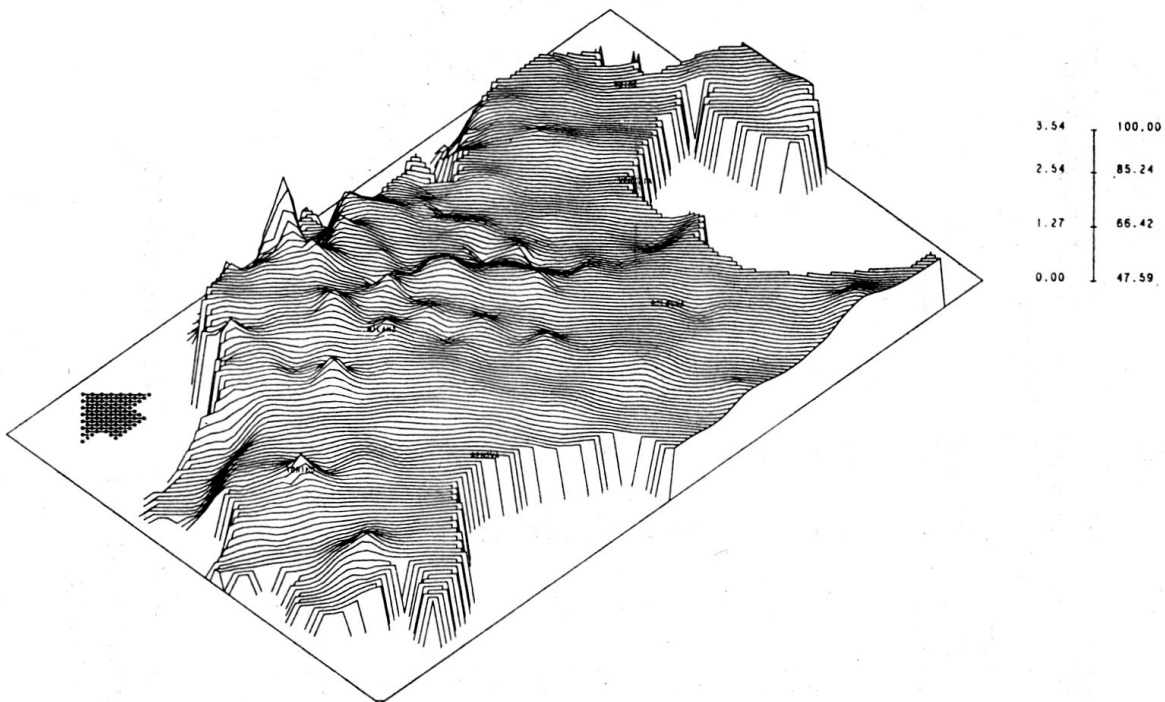


Fig. 14 :
 Carte numérique de la distribution de similarité relative au point de référence 205
 (Campodolcino, Lombardie septentrionale).
 Voir aussi les Fig. 12 et 13.



ITA-999-101-205
 Azimuth = 45 Neigung = 45
 Loenge = 20.00 Hoehe = 5.00
 * = vor der Projektion 16.06.80

Fig. 13 :

Surface statistique lissée de la distribution de similarité relative au point de référence 205 (Campodolcino, Lombardie septentrionale).

Voir aussi les Fig. 12 et 14.

Au premier coup d'œil on note que par rapport à la Fig. 9, l'agencement des plages choroplèthes a été complètement renversé. Face au bloc central (constitué par les polygones classés dans les intervalles 5 et 6) et qui recouvre grosso modo l'espace de la Lombardie politique (avec des empiètements sur le Piémont, en province de Novare), on remarque de vastes zones intermédiaires (plages dans l'intervalle 3) groupant le Piémont central, la Ligurie, la Toscane, la Vénétie et le Frioul, alors que les antipodes typologiques du point de référence considéré occupent les extrémités occidentales (vallées vaudoises, Vallée d'Aoste), septentrionales (Grisons) et nord-orientales (Tyrol du Sud).

L'IRI maximal (85,443) est au point 223 (dont le vecteur matriciel contient plus de 350 cases vides), l'IRI minimal (47,591) au point 1 (Grisons occidentaux). La différence fondamentale entre les types iconiques des Fig. 9 et 12 s'explique par le fait qu'entre les points de référence 15 et 205 se situe la frontière linguistique lombardo-rhétoromane et que, partant, les vecteurs d'attributs des PP. 15 et 205 disposent de charges taxatoires carrément différentes.

Voir aussi la Fig. 13 qui montre fort bien le renforcement de la position centrale de l'entourage du point de référence. On remarquera en outre la différence intrinsèque entre les silhouettes des histogrammes respectifs (voir les Fig. 9 et 12, à droite, en bas).

6. Au-delà de la mesure de similarité.

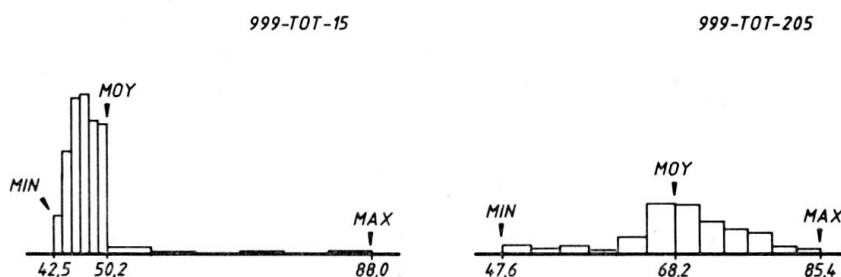


Fig. 15 :

Silhouettes de deux distributions de similarité (relatives aux points de référence 15 et 205).

Voir aussi les Fig. 9 et 12.

MIN — minimum
 MOY — moyenne arithmétique
 MAX — maximum

Pour pousser plus avant nos connaissances des données examinées et pour avancer donc dans l'exploration taxométrique de la matrice de similarité, il convient de se pencher sur une analyse synoptique des paramètres caractéristiques (minimum, moyenne arithmétique, maximum, etc.) des distributions de similarité calculées. La Fig. 15 montre non seulement qu'à deux types bien accusés des distributions de similarité correspondent deux silhouettes bien distinctes des histogrammes respectifs, mais aussi que la taille de trois des paramètres caractéristiques (minimum, moyenne, maximum) est sujette à certaines variations

numériques dont on pourrait présumer qu'elles seraient utilisables à des fins taxométriques. On est donc amené à envisager une présentation synoptique de ces paramètres caractéristiques.

*Matrice de similarité (s_{jk})
(asymétrique)*

					1	6
				1	0,75	5
			1	0,3	0,3	4
		1	0,3	0,25	0,25	3
	1	1	0,3	0,25	0,25	2
1	0,25	0,25	0,6	0,75	0,5	1
1	2	3	4	5	6	

Objets (points d'atlas)

*Matrice de similarité (s_{jk})
(asymétrique)*

					1	6
				1	0,75	5
			1	0,3	0,3	4
		1	0,3	0,25	0,25	3
	1	1	0,3	0,25	0,25	2
1	0,25	0,25	0,6	0,75	0,5	1
1	2	3	4	5	6	

Objets (points d'atlas)

Fig. 16 :

En gris : Secteurs de la matrice de similarité concernés
(à gauche) : lors de la mesure de similarité (voir les Fig. 9 - 11), et
(à droite) : lors de la synopse des paramètres caractéristiques des distributions de similarité (voir les Fig. 17 - 33).

Il est bon toutefois de se rendre compte des différences taxométriques existant entre l'analyse de vecteurs isolés et la considération synoptique des valeurs consignées dans la matrice de similarité ; voir cela à la Fig. 16.

6.1. Synopse des maxima des distributions de similarité.

Voir les Fig. 17, 18 et 19.

Comme nous l'avons démontré plus haut, toute distribution de similarité dispose d'une valeur maximale située le plus souvent dans le voisinage immédiat du point de référence. Pour chacune des 251 distributions de similarité stockées dans la matrice de similarité, cette

valeur sera transportée à l'endroit du point de référence afin d'obtenir une série de 251 valeurs différentes, qui, elle, sera soumise au traitement cartographique habituel. Les résultats en sont très éloquentes ; voir les Fig. 17 et 18. En effet, ce procédé synoptique permet de dégager ce qu'on appelle communément, en dialectologie classique, des « noyaux dialectaux ». C'est que la quasi-totalité des dialectes italo-, gallo- et rhétoromans traditionnellement répertoriés dans les manuels de linguistique romane comme figurant en Italie septentrionale et dans la partie méridionale de la Suisse, peut être repérée dans les plis et replis des reliefs fort accidentés des Fig. 17 et 18.

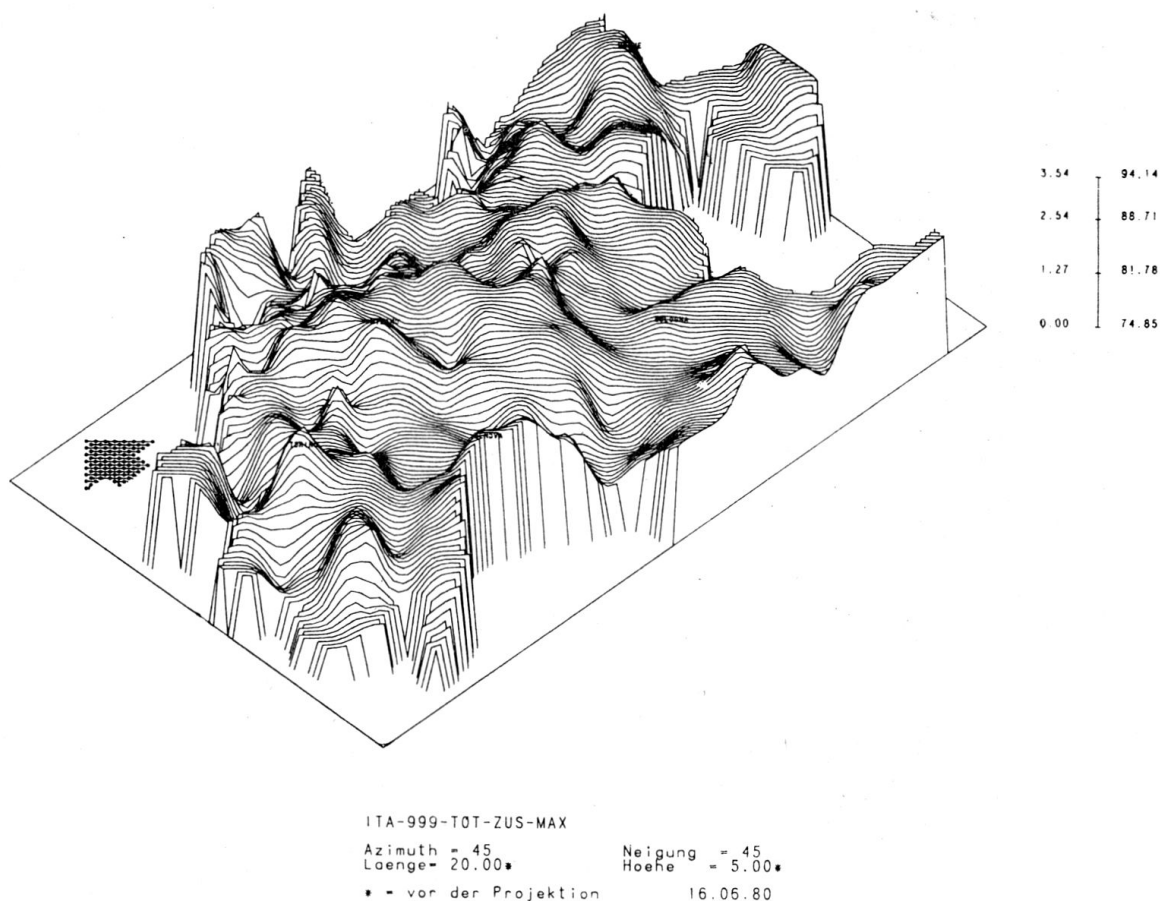
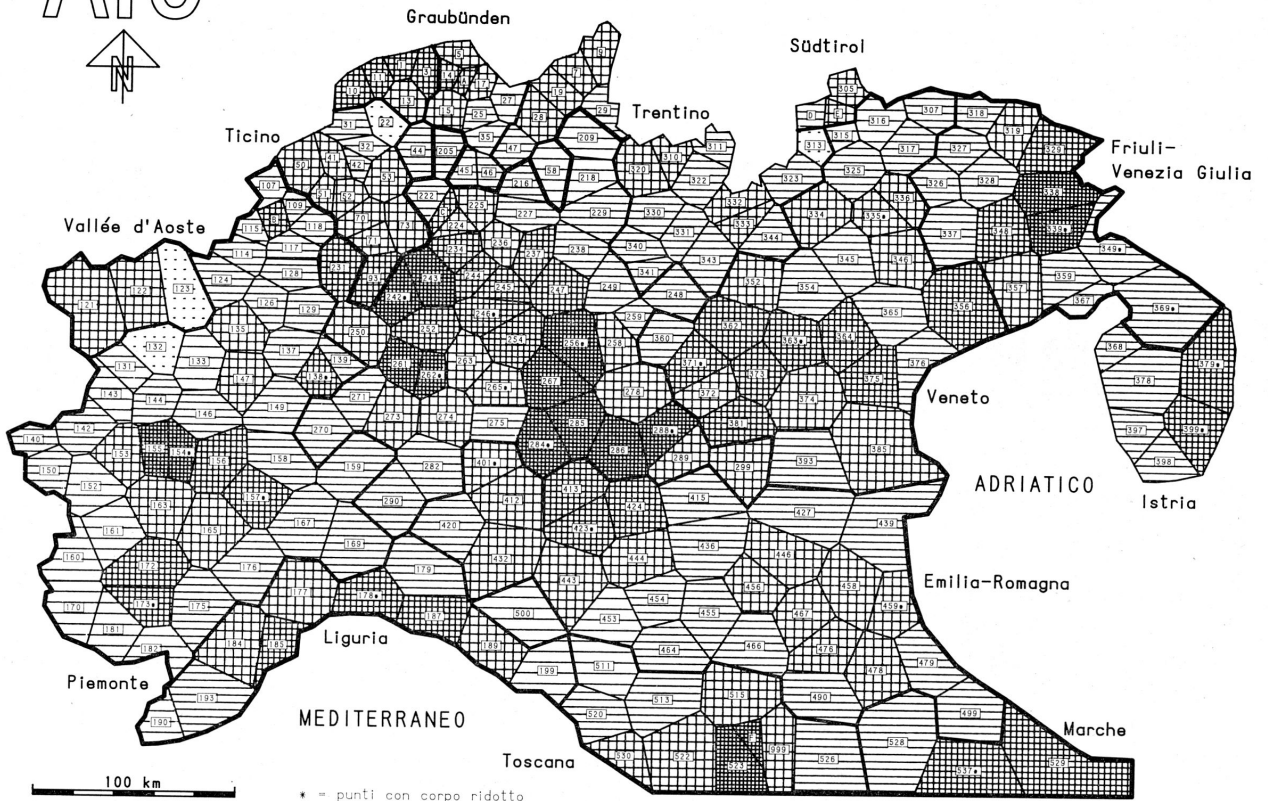


Fig. 18 :

Surface statistique lissée de la synopse des maxima des 251 distributions de similarité.

Voir aussi les Fig. 17 et 19.

AIS



100 km

* = punti con corpo ridotto

LEX - ITALIEN

999-TOT-ZUS-MAX

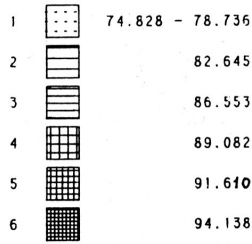
Punkte mit Buchstaben:

- A - 16
- B - 116*
- C - 223*
- D - 312
- E - 314*
- F - 524



Kartographie: W.-D. Raab

Choroplethenkarte
MINWMAX 6-fach



Häufigkeitsverteilung
MINWMAX 12-fach

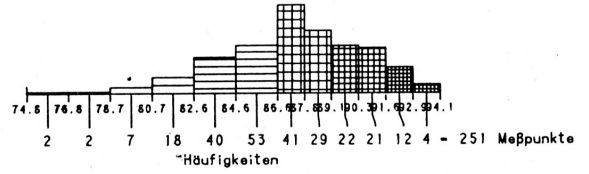


Fig. 17:

Carte choroplèthe de la synopse des maxima des 251 distributions de similarité.
Voir aussi les Fig. 18 et 19.

*KARTE 559-ZUS MAX

MW= 26.553 S= 3.417

MINMAX 74.828 74.736 82.645 86.553 89.081 91.516 94.138
NEHMK 74.828 83.556 85.239 85.569 87.833 90.000 94.138
NEC 74.828 83.571 85.366 84.618 87.960 90.044 94.138

0-3 7-18 8-1 9-1 10-1 11-1 12-4
0-20 0-1 0-0 0-0 0-0 0-0 0-1
0-1 0-1 0-1 0-1 0-1 0-1 0-1

STREIFEN 1

SVIZZERA
PUNANTSCHA

13 88.954
4+E+5-

17 86.
4+C-

31 84.515 22 74.828
3+D+2+ 1+D+1-

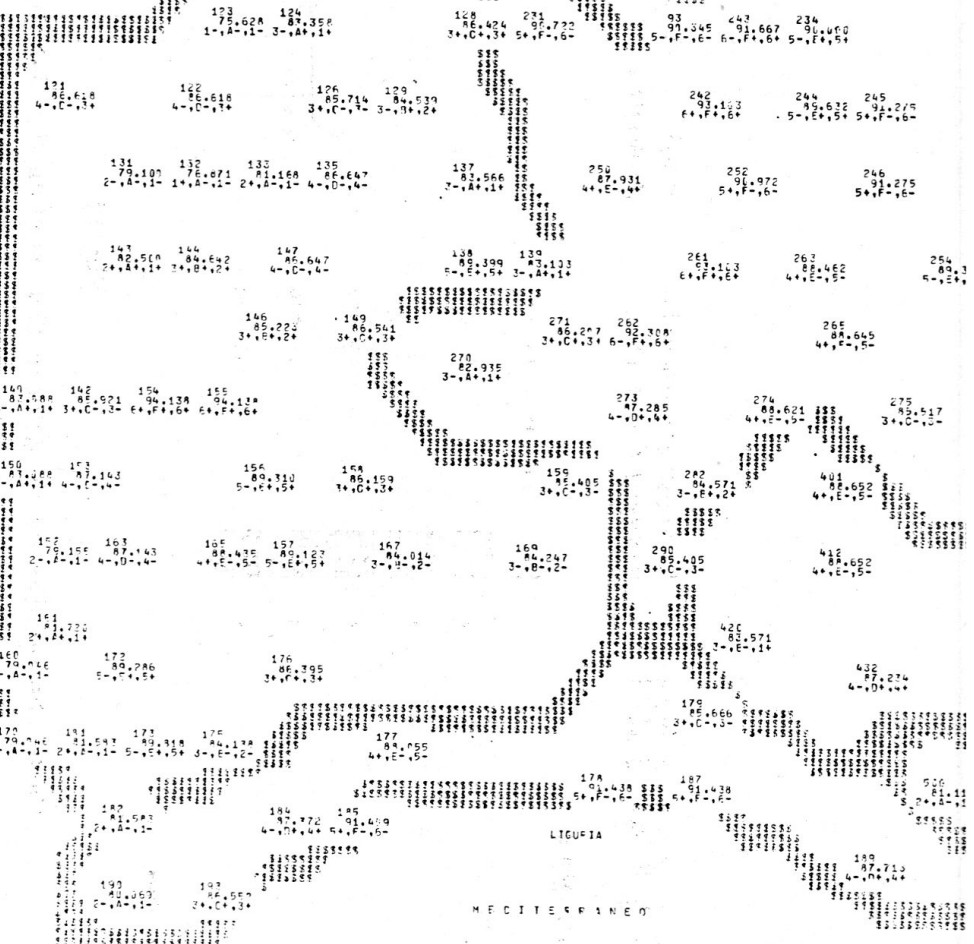
33 83.378
2+A+1-

25 84.
3+D+

TICINO

LOMBARDIA

VALLÉE D'AOSTE
PIEMONTE



MEDITERRANEO

*KARTE 999-705 MAX
 MW= 86,553 S= 7,417
 MINMMAX 74,228 76,736 82,645 86,553 89,081 91,610 94,138
 MECHM 74,228 81,560 85,239 86,553 87,003 91,003 91,610
 MECC 74,228 81,571 85,166 86,553 86,610 91,003 91,610

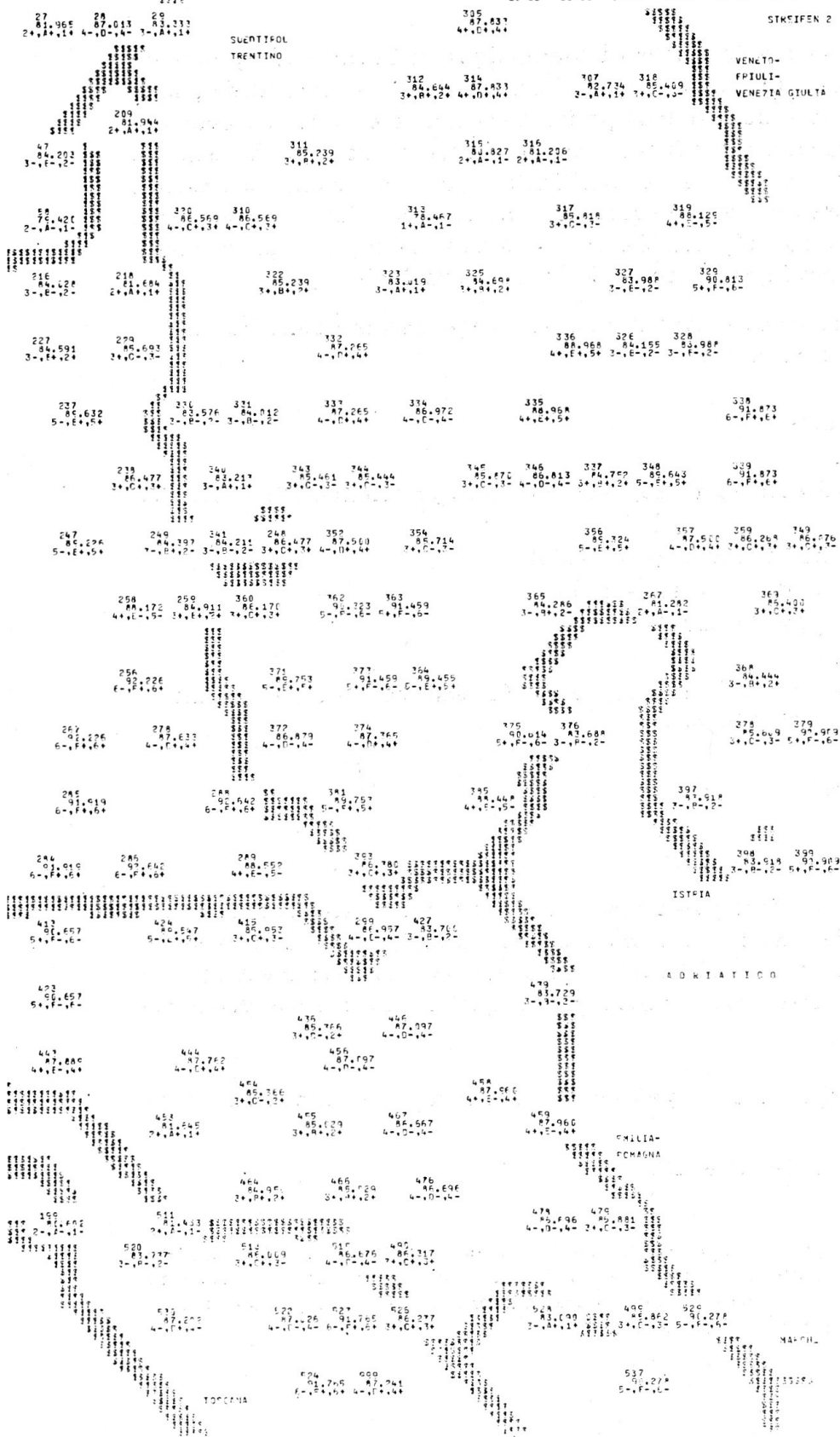


Fig. 19 :
 Carte numérique de la synopse des maxima des 251 distributions de similarité.
 Voir aussi les Fig. 17 et 18.

Remarquer que la valeur cartographique de la carte synoptique en question repose surtout sur le jeu alternatif de dépressions et de points hauts, alternance qui est tout particulièrement bien mise en évidence par la surface statistique lissée (Fig. 18). Ce n'est qu'à l'intersection des domaines lombard, émilien et vénitien que les dépressions font parfois défaut, si bien que l'on est forcé de chercher les lieux d'origine des valeurs maximales répertoriées sur la carte synoptique, pour constater l'orientation des lignes de force des points considérés ⁽¹⁹⁾.

La Fig. 17 montre donc des noyaux agglomératifs dans les domaines dialectaux suivants :

- ligure,
- piémontais,
- valdotain,
- lombard occidental (PP. 243, 261, etc.) et oriental (PP. 267, 285, etc.),
- romanche occidental (sursilvain),
- romanche oriental (engadinois),
- ladin (Tyrol du Sud),
- frioulan,
- vénitien (blocs occidental : PP. 362, 363, 364, etc. ; et oriental : PP. 346, 335, 356, etc.),
- émilien,
- romagnol,
- toscan,
- dialectes des Marches.

Les résultats fort patents de cette analyse confirment non seulement la validité du procédé taxométrique utilisé, mais aussi celle du réseau d'enquête de l' AIS dont les mailles sont suffisamment serrées pour donner une bonne vision modélisée de la réalité dialectale observée.

6.2. Synopse des moyennes arithmétiques des distributions de similarité.

Voir les Fig. 20, 21 et 22.

(19) Le lieu d'origine d'une valeur maximale ne peut être trouvé que par l'inspection de la carte numérique de la distribution de similarité relative au point d'atlas concerné. Si cette carte numérique n'est pas publiée, il faut recourir aux archives de l'auteur.

Dans une perspective communicative (cf. 5.1.3.), la moyenne arithmétique d'une distribution de similarité peut être utilisée à en évaluer numériquement la position centrale au sein du réseau examiné. En effet, les Fig. 20 et 21 montrent un décalage bien accusé entre régions centrales (intervalles 6, 5 et 4) et régions périphériques (intervalles 1, 2 et 3). On remarquera que ce décalage se fait le plus souvent le long de frontières linguistiques bien connues : entre le romanche et le lombard, entre le valdotain (francoprovençal) et le piémontais, entre le ladin et le vénitien, voire le trentin, entre le frioulan et le vénitien, entre le toscan et le romagnol, voire l'émilien, etc.

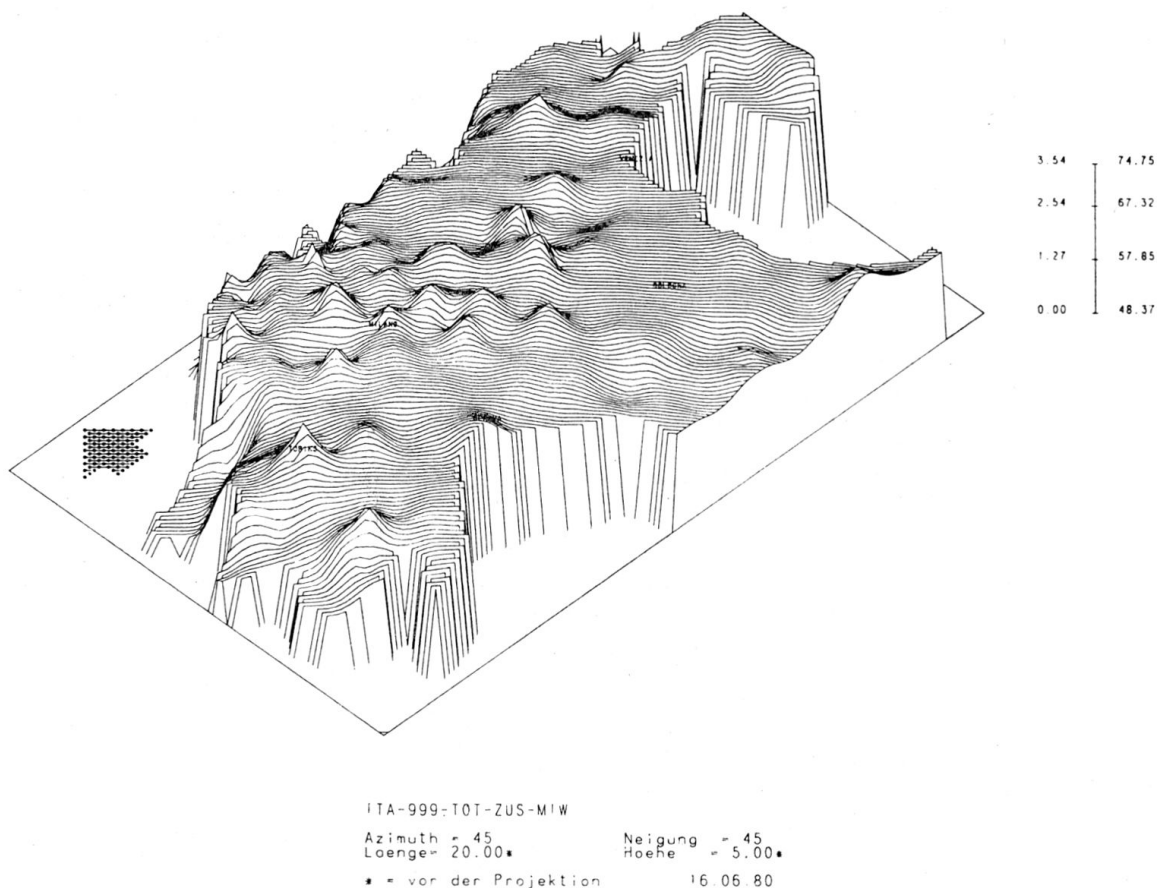
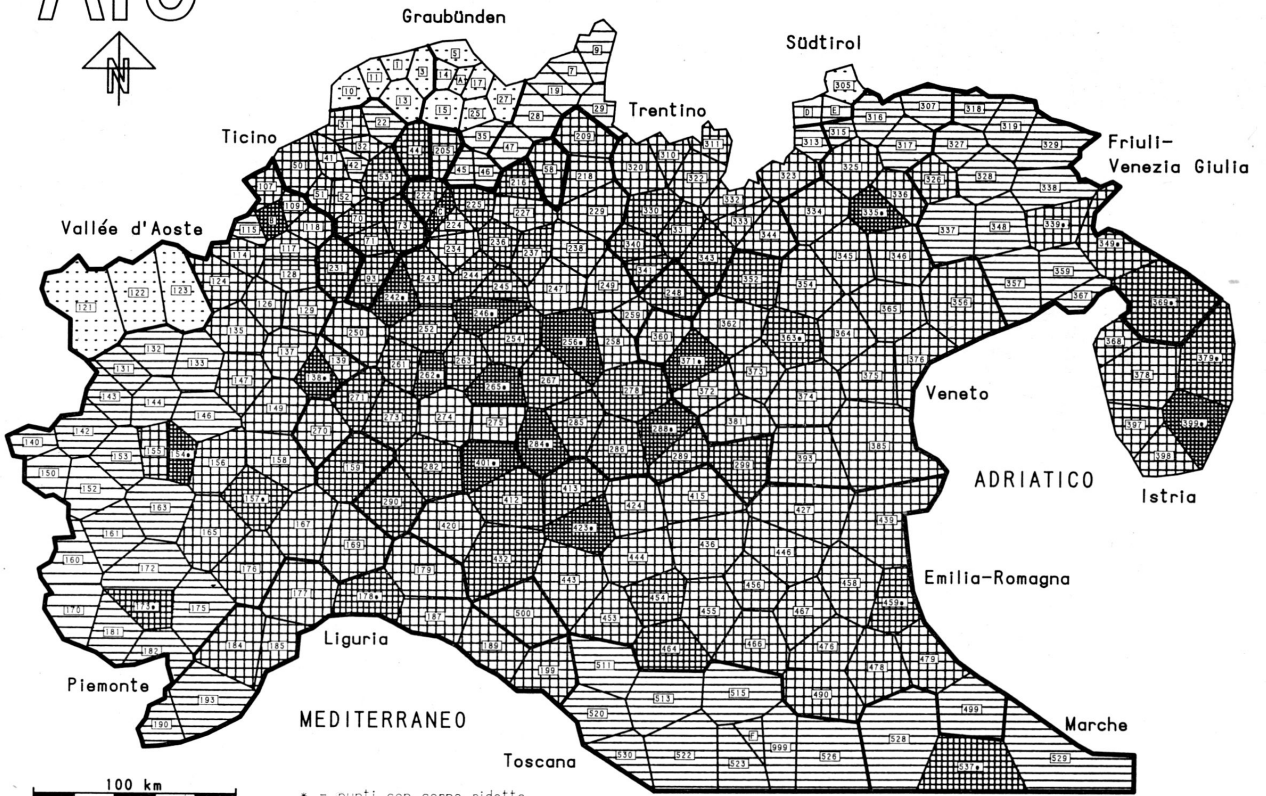


Fig. 21 :

Surface statistique lissée de la synopsis des moyennes arithmétiques des 251 distributions de similarité.

Voir aussi les Fig. 20 et 22.

AIS



100 km

* = punti con corpo ridotto

LEX - ITALIEN

999-TOT-ZUS-MIW

Punkte mit Buchstaben:

- A - 16
- B - 116*
- C - 223*
- D - 312
- E - 314*
- F - 524



Choroplethenkarte
MINWMAX 6-fach

1		48.372 - 54.354
2		60.335
3		66.317
4		69.129
5		71.941
6		74.753

Häufigkeitsverteilung
MINWMAX 12-fach

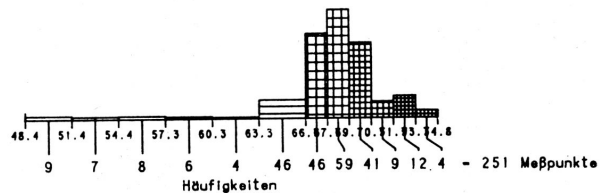


Fig. 20 :

Carte choroplèthe de la synopsis des moyennes arithmétiques des 251 distributions de similarité.
Voir aussi les Fig. 21 et 22.

*KARTE 909-ZUS MIN

MM: FG.216 S: 5.269

MIN	48.272	54.351	60.135	66.316	69.129	71.941	74.753
MAX	48.375	57.795	67.176	77.515	88.129	89.390	74.753
MED	48.375	54.282	66.641	77.706	88.678	89.741	74.753

13	11	1
49.847	49.348	48.372
1-A,-1-	1-A,-1-	1-A,-1-

5

1-A,-1-

14	16
58.099	68.149
1-A,-1-	1-A,-1-

9-7	8-6	4-46	46-59	41-9	12-4
11-11	13-13	14-14	18-28	29-29	23-28
20-21	21-21	21-21	21-21	21-21	21-21

STÄLLEN 1

SWIZZERA
ROMANTSCHA

13	13
49.173	49.173
1-A,-1-	1-A,-1-

17

1-A,-1-

22	22
64.343	64.343
3-A,+2-	3-A,+2-

25

1-A,-1-

32	32
67.733	67.733
4-D,-3-	4-D,-3-

35

2-A,+4-

TICINO

100	59	61	62
67.524	67.342	61.373	62.949
4-D,-3-	4-D,-3-	4-D,-3-	3-A,+2-

44	205	45	46
65.138	68.171	68.333	68.333
5-E,+5-	4-A,+1-	3-A,+2-	3-A,+2-

107	51	52	53
65.647	68.320	67.315	69.494
4-D,-3-	4-D,-3-	4-D,-3-	5-F,+5+

LOMBARDIA

115	116	118	70	73	222	244	245
65.746	67.272	67.328	70.157	69.552	69.796	68.675	69.326
3-A,+2+	6-F,+6+	4-D,-3-	5-F,+6-	5-F,+5+	5-F,+6-	4-A,+1+	5-E,+5+

VALLEE D'AOSTE
PIEMONTE

114	117	71	223	235
65.043	68.124	69.513	72.541	73.141
4-A,+1-	4-A,+1-	5-F,+5+	6-F,+6+	5-F,+6+

121	122	126	129	242	244	245
52.108	63.692	69.590	67.118	74.278	77.657	68.704
4-A,+1-	4-A,+1-	4-E,+5-	4-D,-3-	6-F,+6+	5-F,+6-	5-F,+6-

131	132	133	135	137	250	252	256
63.409	61.988	64.286	68.848	65.075	67.470	66.396	72.141
3-A,+1+	3-E,+4-	3-A,+2-	4-E,+5-	4-E,+4-	4-D,-3-	5-F,+5+	6-F,+6+

143	144	147	138	139	261	263	254
63.884	65.128	66.916	72.662	68.824	74.179	68.988	69.513
3-A,+1+	3-A,+2-	4-D,-3-	6-F,+6+	4-D,-3-	5-F,+6-	5-F,+6-	5-F,+6+

146	149	271	262	265
65.597	67.050	68.750	72.750	73.589
3-A,+2+	4-A,+4-	5-E,+5-	6-A,+6+	6-A,+6+

150	142	154	155	273	274	275
55.178	64.464	72.452	68.693	71.108	68.860	68.403
2-A,+1-	3-A,+2-	6-F,+6+	4-A,+5-	5-F,+6+	4-E,+5-	4-A,+1+

153	163	150	154	159	282	401
57.534	65.167	65.515	68.004	68.154	69.171	72.933
2-A,+1-	3-A,+2-	4-A,+2+	4-A,+4-	5-F,+6-	5-E,+5-	6-F,+6+

161	163	165	167	169	290	413
61.448	65.873	68.929	70.171	69.700	69.872	66.256
2-A,+1-	3-A,+2+	4-E,+5-	5-F,+6-	4-D,-3-	5-F,+6-	5-E,+5+

161	173	175	177	179	420	430
61.578	65.177	68.872	67.001	68.824	68.507	69.741
3-E,+4-	3-A,+2-	4-D,-3-	4-D,-3-	4-D,-3-	4-A,+1-	5-F,+5+

170	184	173	175	177	178	187
57.833	60.564	70.165	65.650	67.001	70.925	67.570
2-A,+1-	3-A,+2-	5-F,+6-	3-A,+2+	4-D,-3-	5-A,+0	5-A,+0

182	184	185	177	178	187	189
64.124	67.548	67.881	67.001	70.925	67.570	68.824
3-A,+1+	4-D,-3-	4-D,-3-	4-D,-3-	5-A,+0	5-A,+0	5-A,+0

190	193	194	195	196	197	198
64.100	65.037	67.548	67.881	68.824	69.872	70.925
3-A,+1-	3-A,+2+	4-D,-3-	4-D,-3-	4-D,-3-	4-D,-3-	4-D,-3-

MEDITERRANEO

199	200	201	202	203	204	205
64.100	65.037	67.548	67.881	68.824	69.872	70.925
3-A,+1-	3-A,+2+	4-D,-3-	4-D,-3-	4-D,-3-	4-D,-3-	4-D,-3-



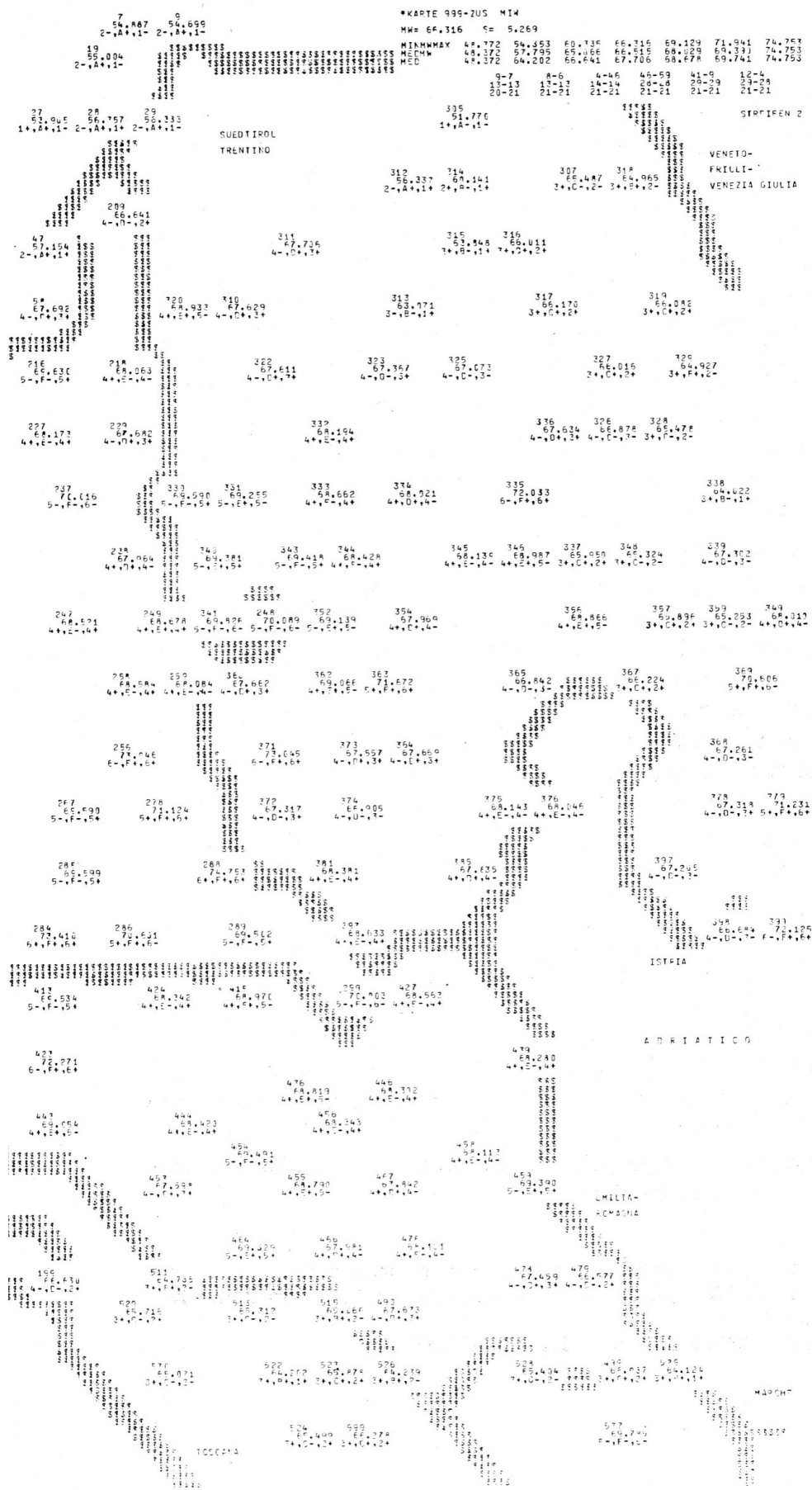


Fig. 22 :

Carte numérique de la synopse des moyennes arithmétiques des 251 distributions de similarité.
 Voir aussi les Fig. 20 et 21.

Il est d'ailleurs intéressant de noter que les valeurs maximales de la distribution visualisée (isoplèthes dans l'intervalle 6 : Fig. 20 ; voir les pics et les sommets bien visibles du relief de la Fig. 21) coïncident sans exception aucune avec des points d'atlas à corpus réduit (« punti con corpo ridotto »), c'est-à-dire dont le vecteur d'attributs contient beaucoup de cases vides (cf. 1.2.)⁽²⁰⁾. Vu la dégradation des résultats taxométriques causée par la présence d'un nombre trop élevé de cases vides dans un vecteur ponctuel de la matrice de données, il semble indiqué de ne pas surestimer le fait que les plages choroplèthes affectées à l'intervalle 6 correspondent à des centres urbains majeurs.

6.3. Synopse des coefficients de symétrie (de Fisher) des distributions de similarité.

Voir les Fig. 23, 24, 25 et 26.

Ainsi que le montre la Fig. 15, une des particularités les plus saillantes des différentes distributions de similarité réside dans la symétrie, voire l'asymétrie de leurs histogrammes respectifs. Il est donc nécessaire de recourir à un indice approprié pour mesurer le degré de symétrie de chacune des distributions de similarité et de réunir ensuite les coefficients calculés en synopse. La meilleure mesure de symétrie est le troisième moment de la distribution divisé par le cube de l'écart-type. En voici la formule⁽²¹⁾ :

$$g = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n \cdot s^3},$$

- où g est le coefficient de symétrie (de Fisher),
 i est l'indicatif (allant de 1...250) des différents scores de similarité,
 n est le total (= 250) des scores de similarité,
 x_i est un des 250 scores d'une distribution de similarité,
 \bar{x} est la moyenne arithmétique d'une distribution de similarité,
 s est l'écart-type (all. Standardabweichung) d'une distribution de similarité.

(20) Les points concernés correspondent à des centres urbains de quelque importance (p. ex. P. 242 - Côte, P. 401 - Plaisance, P. 284 - Crémone, etc.) ; cf. aussi la note 16.

(21) Cf. Bahrenberg/Giese 1975, 54 ; Gendre 1977, 38.

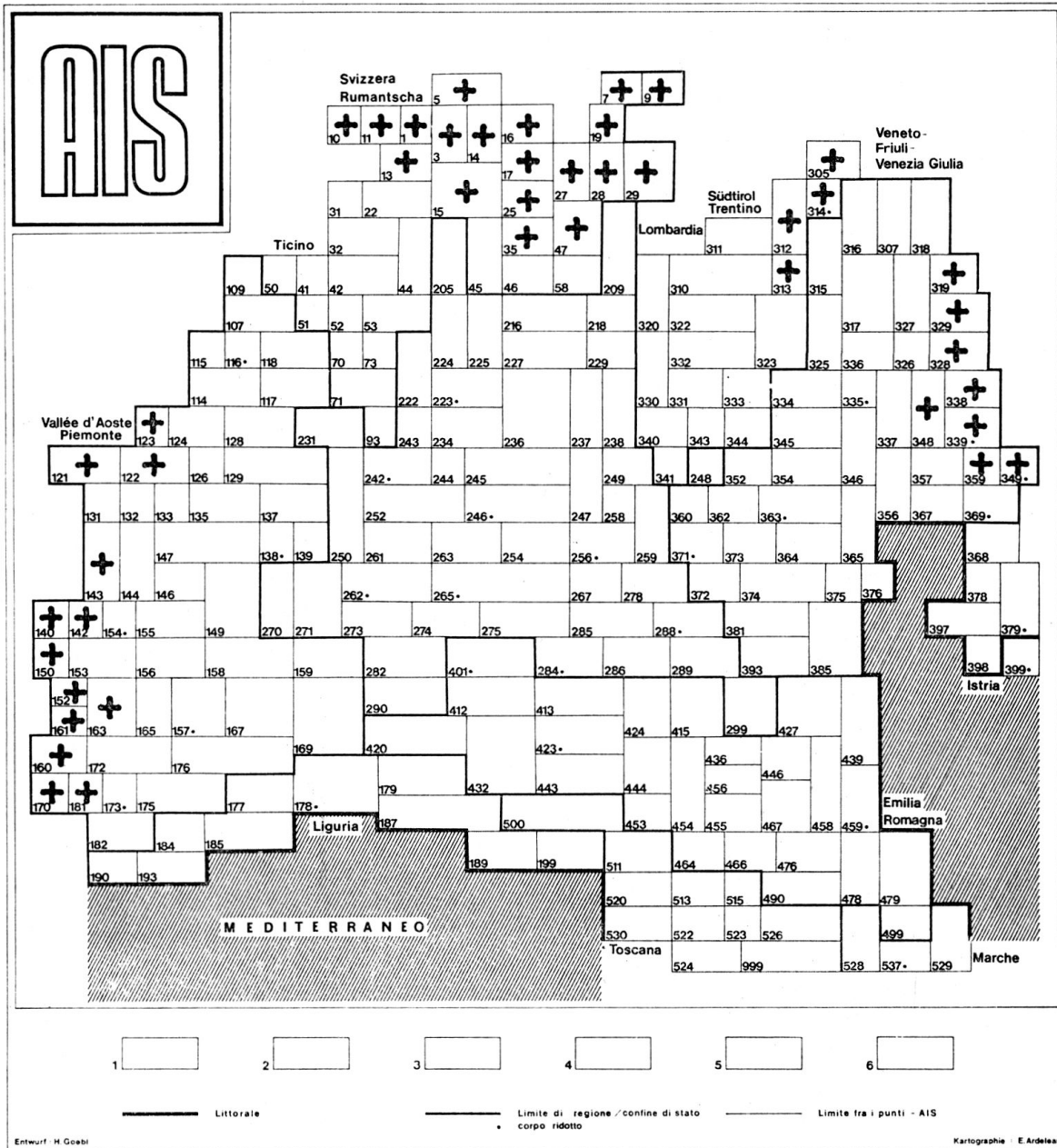
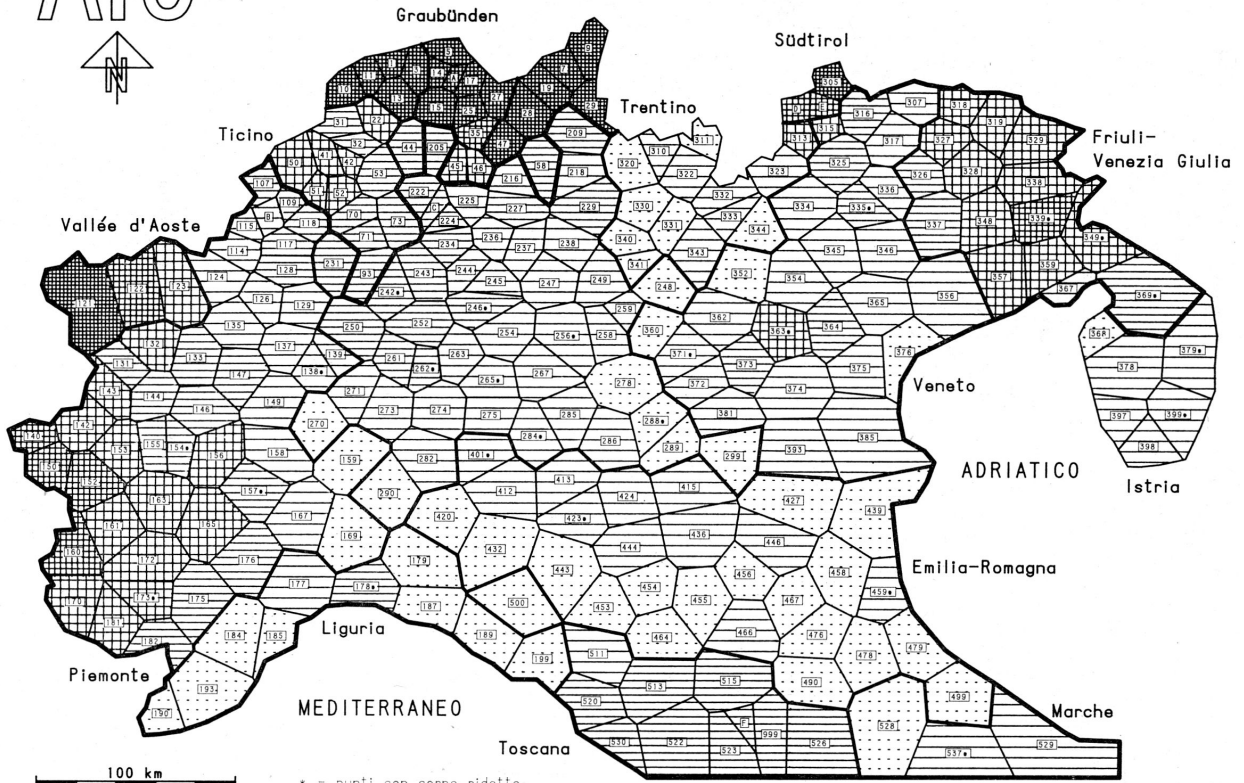


Fig. 24 :

Localisation des coefficients de symétrie (de Fisher) positifs.

Voir aussi les Fig. 23, 25 et 26.

AI S



* = punti con corpo ridotto

LEX - ITALIEN

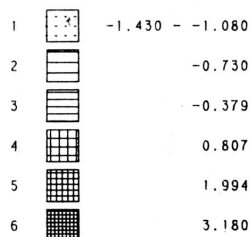
999-TOT-ZUS-SCHIEFE

Punkte mit Buchstaben:

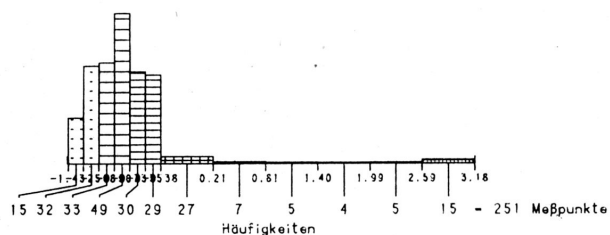
- A = 16
- B = 116
- C = 223
- D = 312
- E = 314
- F = 524



Choroplethenkarte
MINWMAX 6-fach



Häufigkeitsverteilung
MINWMAX 12-fach



Kartographie: W.-D. Rase

Fig. 23 :

Carte choroplèthe de la synopse des coefficients de symétrie (de Fisher) des 251 distributions de similarité.
Voir aussi les Fig. 24, 25 et 26.

LEX-ITALIEN
KARTE 999-TOT-ZUS SCHIEFE

MW= -0.378 S= 1.492

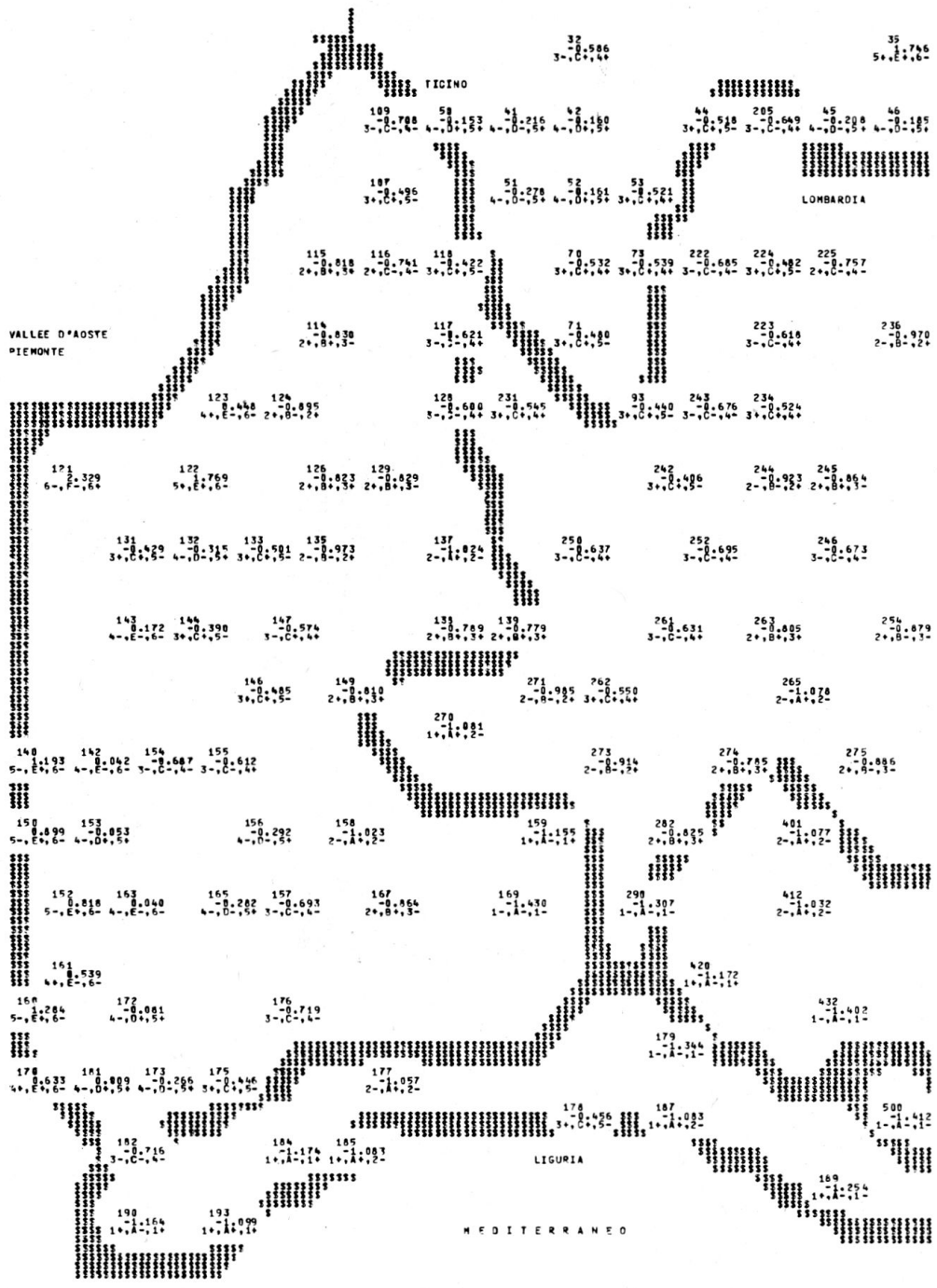
MINMAX	-1.429	-1.879	-0.729	-0.378	0.807	1.994	3.130
MEHMW	-1.429	-1.999	-0.789	-0.318	0.926	1.926	3.188
MEB	-1.429	-1.899	-0.692	-0.796	0.939	1.948	3.188

18	11	1	14	16
6+,F+,6+	6+,F+,17	6+,F+,18	6+,F+,93	6+,F+,83

STREIFEN 1

SVIZZERA
ROMANTSCHA

13	17
6+,F+,872	6+,F+,710
15	25
6+,F+,910	6+,F+,784
21	22
3+,C+,78	4-,B-,272



VALLEE D'AOSTE
PIEMONTE

TIGINO

LOMBARDIA

LIGURIA

MEDITERRANEO

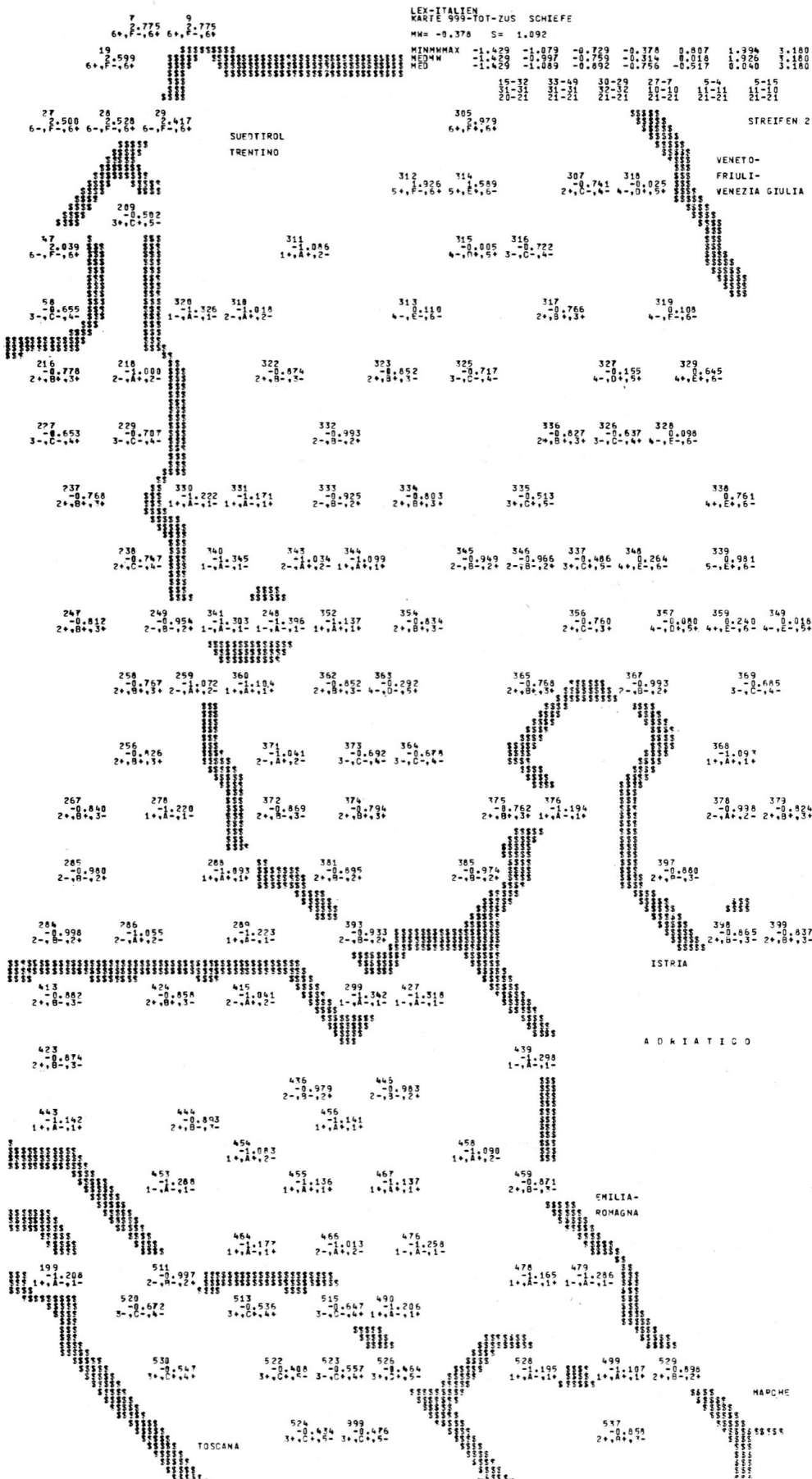
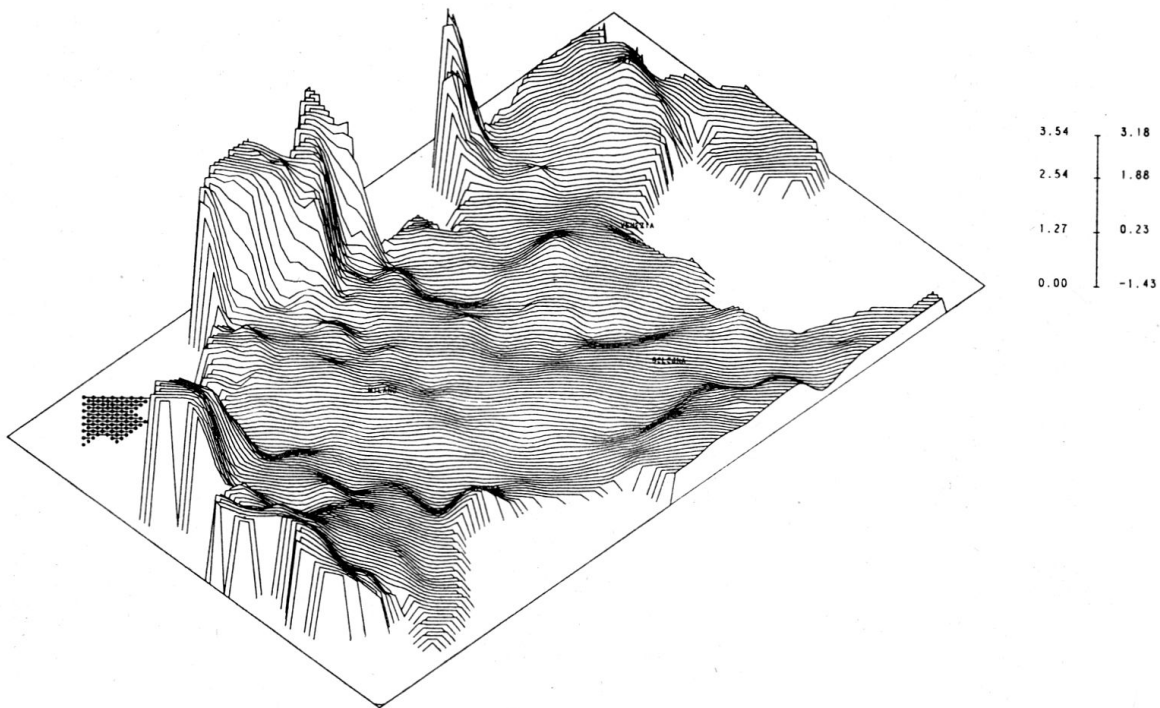


Fig. 26 :

Carte numérique de la synopsis des coefficients de symétrie (de Fisher) des 251 distributions de similarité.
 Voir aussi les Fig. 23, 24 et 25.



ITA-999-TOT-ZUS-SCHIEFE
 Azimuth = 45 Neigung = 45
 Loenge = 20.00 Hoehe = 5.00
 * = vor der Projektion 16.06.80

Fig. 25 :

Surface statistique lissée de la synopse des coefficients de symétrie (de Fisher) des 251 distributions de similarité.

Voir aussi les Fig. 23, 24 et 26.

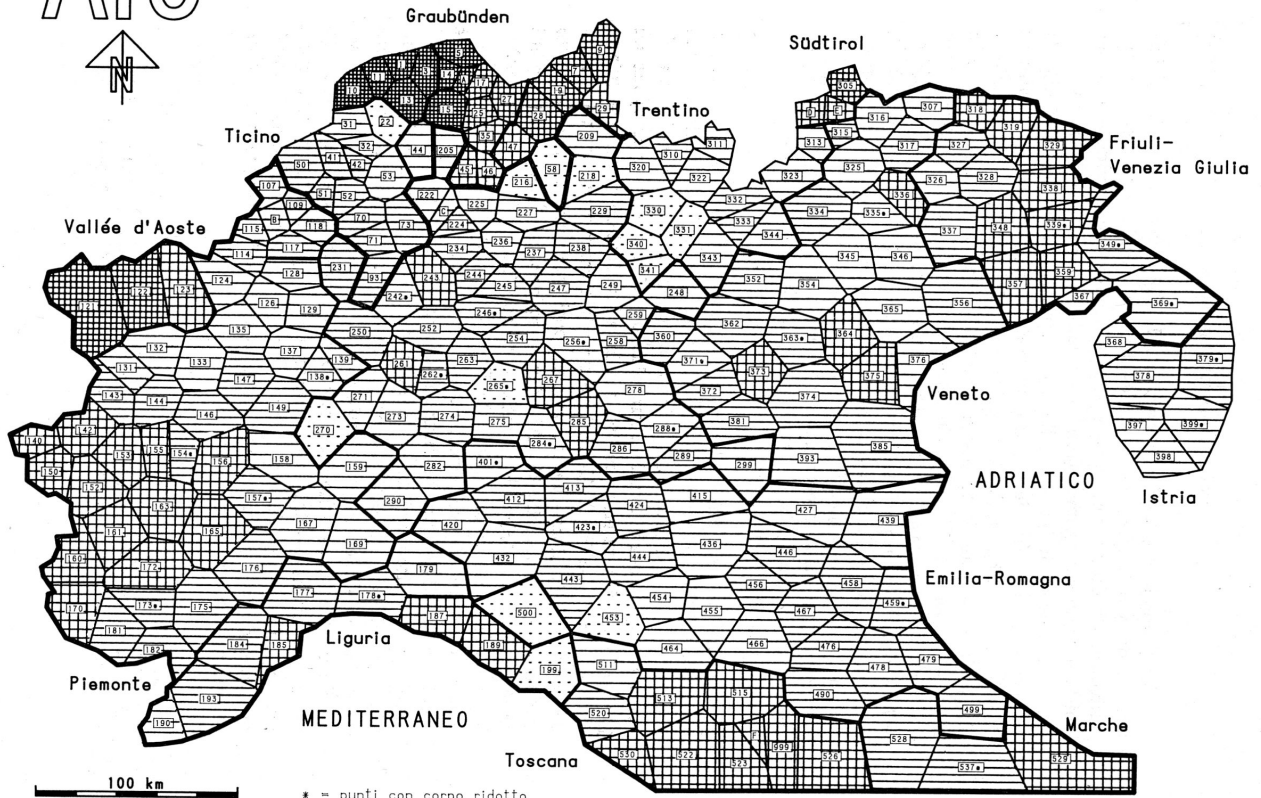
Le coefficient ainsi calculé (proposé par le statisticien R. A. Fisher, 1890-1962) n'a pas de nom particulier en français ; il s'appelle *skewness* en anglais et *Schiefe* en allemand. Il prend la valeur 0 quand la distribution est symétrique, des valeurs positives quand la distribution est dissymétrique avec accumulation sur la gauche (voir p. ex. la Fig. 15), et des valeurs négatives quand la distribution est dissymétrique avec accumulation sur la droite. La duplicité des tiroirs numériques du coefficient de symétrie — variations de la taille du coefficient, variation du signe opérationnel — lui confère une importance taxométrique accrue. Aussi est-il bon de prévoir deux analyses cartographiques : voir les Fig. 23 et 24.

Alors que la Fig. 24 fait apparaître, de par la répartition des distributions dissymétriques positives, des domaines dialectaux bien connus, à savoir ceux du provençal alpin, du francoprovençal valdotain et des trois volets du rhétoroman (romanche, ladin, frioulan) tout en les détachant nettement de leurs corollaires italo-romans (voir aussi la Fig. 25 en ceci non moins explicite), le pavage de la Fig. 23 prend des allures franchement esthétiques. Les polygones classés dans l'intervalle 1 (équivalents à une forte dissymétrie négative) se situent — grosso modo — le long de la chaîne de l'Apennin et du cours de l'Adige tout en revêtant ainsi l'aspect d'une charpente cunéiforme entourée de ses solives (plages choroplèthes dans les intervalles 2 sqq.). De toute façon, la logique spatiale du pavage choroplèthe est fort cohérente et suggestive à la fois, étant donné qu'elle cadre assez bien avec nos connaissances relatives à la genèse des grands domaines dialectaux de l'Italie septentrionale (romanisation de la plaine padane d'abord, et du « Venetorum angulus », situé au-delà de l'Adige, ensuite, etc.).

La raison en est qu'avec la mesure de l'asymétrie d'une distribution de similarité l'on aborde une des propriétés les plus importantes de tous réseaux interactifs, à savoir celle de la connectivité d'un point (ou nœud) au sein d'un réseau donné. Pour plus de détails nous renvoyons au paragraphe 6.5. tout en nous limitant ici à l'indication sommaire qu'aux valeurs négatives du coefficient de symétrie correspond un maximum de connectivité, alors que les valeurs positives vont de pair avec un minimum de connectivité. Ajoutons rapidement que l'agencement cunéiforme des extrémités négatives du coefficient de symétrie (polygones dans l'intervalle 1 ; voir la Fig. 23) est bien loin d'être dû au hasard. C'est qu'il correspond de très près aux amphizones situées entre les domaines dialectaux les plus importants du réseau AIS examiné, à savoir entre le lombard d'un côté et le vénitien, le toscan et le piémontais de l'autre (22). Il semble donc bien que les zones de contact — dans lesquelles la dialectologie classique n'a vu le plus souvent que des agrégats régionaux sillonnés par des bourrelets-frontières plus ou moins importants — deviennent, dans une perspective méthodiquement plus rigoureuse (en l'espèce dans notre perspective dialectométrique), des aires de mouvement dont les points constitutifs accusent une optimisation de leurs facultés d'interaction.

(22) La valeur minimale (— 1,430) peut être trouvée au point 169 situé à l'intersection des domaines dialectaux lombard, piémontais, émilien et ligure.

AI S



LEX - ITALIEN

999-TOT-ZUS-MAX/MIW

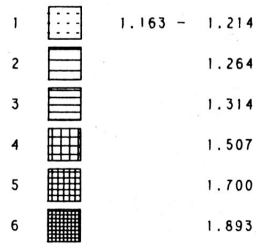
Punkte mit Buchstaben:

- A - 16
- B - 116*
- C - 223*
- D - 312.
- E - 314*
- F - 524



Kartographie: W.-D. Rase

Choroplethenkarte
MINWMAX 6-fach



Häufigkeitsverteilung
MINWMAX 12-fach

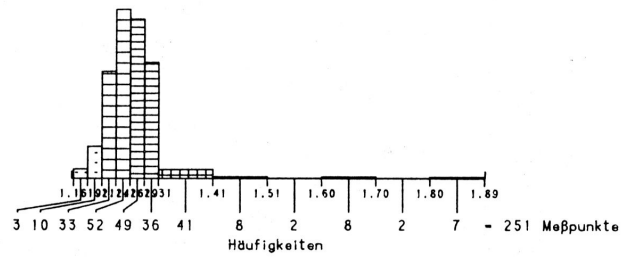


Fig. 27 :

Carte choroplèthe de la synopse des quotients entre le maximum et la moyenne arithmétique des 251 distributions de similarité.

Voir aussi les Fig. 28 et 29.

KARTE 999-TOT-ZUS MAX/PIW

MW= 1.314 S= 0.130

MINMAX	1.163	1.213	1.264	1.314	1.367	1.420	1.473	1.526
MDMW	1.163	1.247	1.331	1.415	1.500	1.584	1.668	1.753
PDW	1.163	1.235	1.258	1.274	1.303	1.347	1.393	1.440

3-18	33-52	49-36	41-8	2-8	2-7
70-30	30-31	71-31	11-11	11-12	10-11
20-71	21-21	21-21	21-21	21-21	21-21

STREIFEN 1

10	11	1	14	16
1.837	1.834	1.893	1.826	1.819
5+,F+,6+	5+,F+,6+	5+,F+,6+	6+,F+,6+	6+,F+,6+

13	17
1.809	1.819
6+,F+,6+	6+,F+,6+

31	22	25
1.244	1.163	1.163
2+,A+,2-	1-,A-,1-	5+,F+

72	35
1.219	1.219
2-,A-,1-	4+,E

109	50	41	42	44	205	45	46
1.241	1.104	1.291	1.274	1.238	1.253	1.324	1.324
2+,A+,2-	3+,C+,5-	3+,C+,4+	3-,C-,1+	2-,A+,2-	2+,A+,2-	4-,D+,5+	4-,D

107	51	52	53	70	73	222	224	225
1.240	1.277	1.275	1.253	1.268	1.265	1.235	1.295	1.258
2+,A+,2-	3-,C-,4-	3+,C+,5-	2+,B-,2+	3-,B+,3+	3-,B+,3-	2-,A+,2-	3+,C+,4+	2+,B-,2+

114	116	118	71	223	236
1.250	1.235	1.256	1.266	1.252	1.252
2+,B-,2+	2-,A+,1+	2+,B-,2+	3-,B+,1-	2+,B-,2+	2-,A-

VALLEE D'AOSTE
PIEMONTE

121	122	124	123	128	231	93	247	234
1.662	1.614	1.241	1.260	1.257	1.298	1.295	1.322	1.385
5+,F+,6+	5+,F+,6+	2+,A+,2-	2+,B-,3-	2+,B-,2+	3+,C+,4+	3+,C+,4+	4-,C-,5+	3+,C+,5-

131	132	133	135	137	250	252	246	265
1.246	1.240	1.267	1.259	1.228	1.303	1.311	1.266	1.265
2+,A+,2-	2+,A+,2-	2+,B+,3-	2+,B-,3-	2-,A-,1+	3+,C+,4+	3+,C+,4-	3-,B+,3+	3-,B+,3-

143	144	147	146	149	138	139	261	263	264	254
1.291	1.300	1.295	1.297	1.273	1.240	1.244	1.327	1.264	1.264	1.264
3+,C-,4+	3+,C+,4+	3+,C+,4+	3+,C+,4+	3-,B+,1+	2-,A+,1+	2+,A+,2-	4-,B+,5+	3-,B+,3-	3-,B+,3-	3-,C-

140	142	154	155	158	159	273	274	275
1.506	1.333	1.299	1.370	1.256	1.217	1.224	1.287	1.249
4+,F-,6+	4-,D+,5+	3+,C+,4+	4-,E+,6-	2+,B-,2+	2-,A-,1-	2-,A-,1+	7-,C-,4-	2+,B-,2+

150	151	154	158	159	282	401
1.444	1.377	1.347	1.256	1.217	1.223	1.216
4+,C+,5-	4-,E+,5+	4-,E-,5+	2+,B-,2+	2-,A-,1-	2-,A-,1-	2-,A-,1-

152	163	165	157	167	169	290	412
1.332	1.324	1.304	1.270	1.254	1.222	1.222	1.280
4-,D+,5+	4-,D+,5+	4-,A-,3+	3-,B+,3+	2+,A-,2+	2-,A-,1-	2-,A-,1-	3-,C-,4-

161	172	176	177	179	400	432
1.327	1.370	1.292	1.310	1.246	1.218	1.251
4-,D+,5+	4-,E-,6+	3+,C-,4+	3+,C+,5-	2+,A+,2-	2-,A-,1-	2+,B-,2+

170	181	174	175	177	174	187	500
1.367	1.304	1.280	1.282	1.310	1.289	1.153	1.153
4-,F-,6+	3+,C+,5-	2-,D-,4-	3-,C-,4-	3+,C+,5-	3+,C-,4+	1-,A-,1-	1-,A-,1-

180	184	185	187	189	190	191	189
1.272	1.297	1.267	1.267	1.249	1.313	1.313	1.249
2-,B+,3+	3+,C+,4+	4-,E-,5+	4-,E-,5+	2+,B-,2+	3+,C+,5-	3+,C+,5-	4-,C+,5+

SVIZZERA
RUMANTSCHA

TICINO

LOMBARDIA

LIGURIA

MERIDIONALE

KARTE. 999-TOT-ZUS MAX/MTW
 MW= 1.314 S= 0.130
 MINMAX 1.161 1.213 1.264 1.314 1.507 1.700 1.897
 MENNA 1.163 1.247 1.273 1.354 1.357 1.473 1.803
 MED 1.163 1.235 1.268 1.274 1.303 1.347 1.893

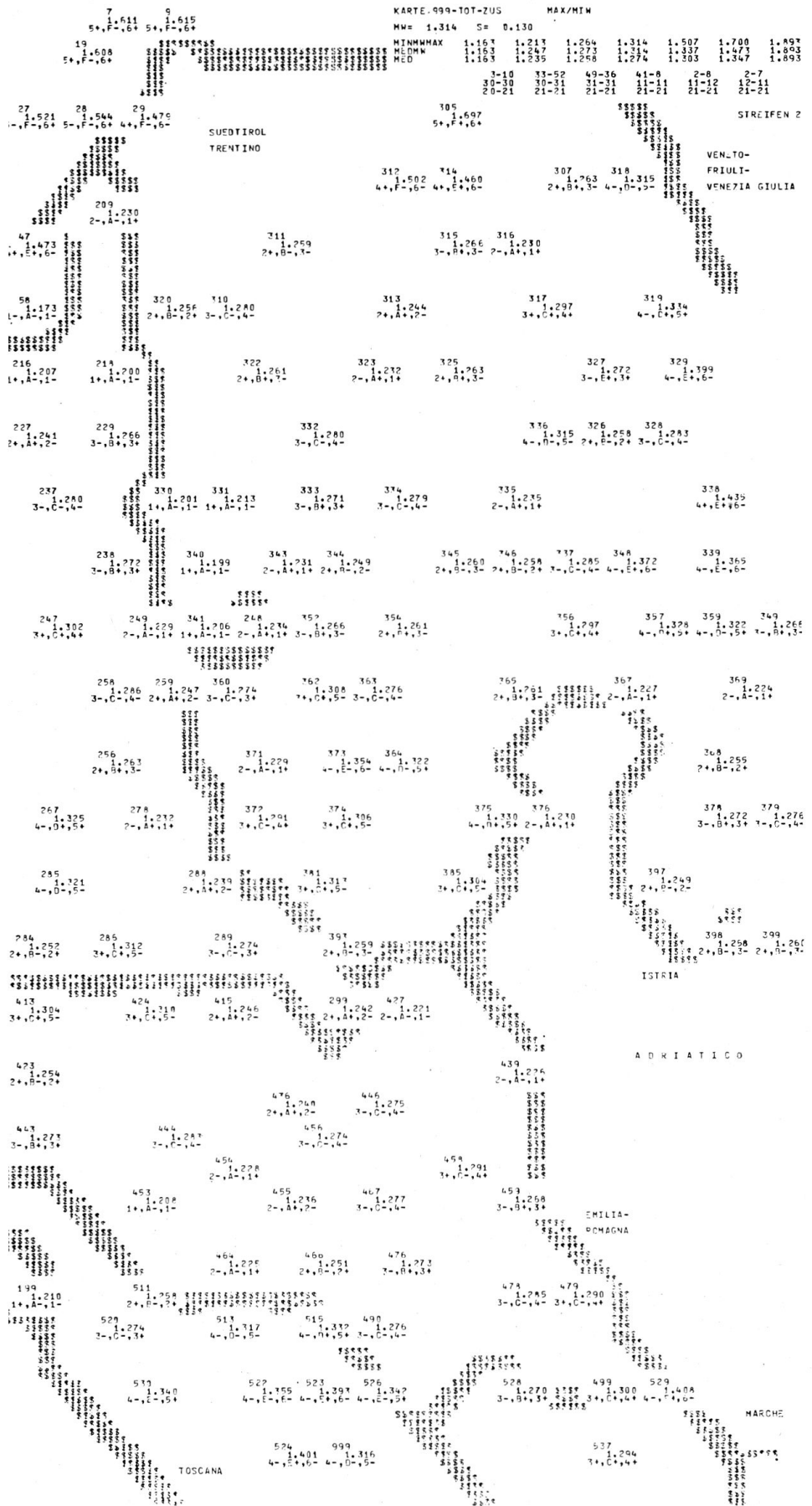


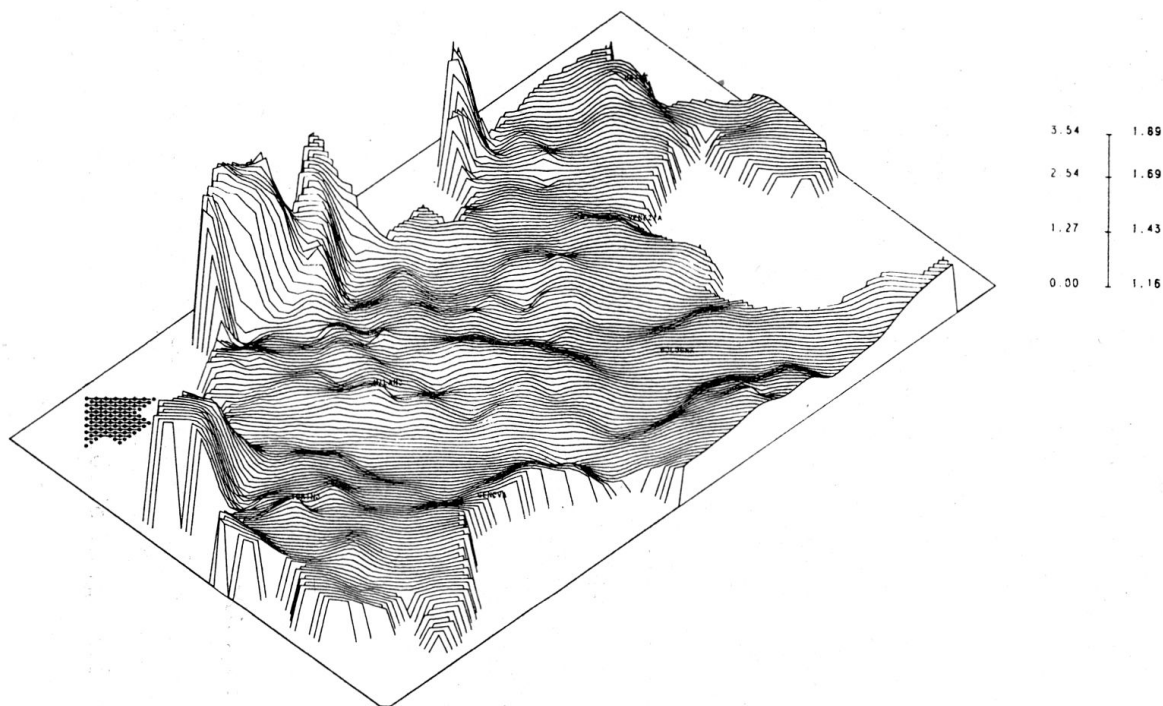
Fig. 29 :

Carte numérique de la synopse des quotients entre le maximum et la moyenne arithmétique des 251 distributions de similarité.

Voir aussi les Fig. 27 et 28.

6.4. Synopse des quotients entre le maximum et la moyenne arithmétique des distributions de similarité.

Voir les Fig. 27, 28 et 29.



ITA-999-TOT-ZUS-MAX/MIW

Azimuth = 45 Neigung = 45
 Loenge = 20.00* Hoehe = 5.00*

* * vor der Projektion 16.06.80

Fig. 28 :

Surface statistique lissée de la synopse des quotients entre le maximum et la moyenne arithmétique des 251 distributions de similarité.

Voir aussi les Fig. 27 et 29.

Comme le calcul des coefficients de symétrie est très fastidieux (et électroniquement onéreux), on peut recourir à un indice de symétrie simplifié représenté par le quotient entre le maximum et la moyenne arithmétique d'une distribution de similarité donnée. Voir les Fig. 27 et 29 qui démontrent clairement que les résultats taxométriques sont très proches de ceux obtenus par le calcul du coefficient de Fisher. C'est

surtout d'une comparaison visuelle des Fig. 25 et 28 (surfaces statistiques lissées) que cet état de choses ressort très nettement.

Il est d'ailleurs facile de contrôler les valeurs de la Fig. 29 par la confrontation des bilans numériques des Fig. 19 et 22. Soit la valeur du quotient MAX/MOY pour le point 5 : 1,765 (voir la Fig. 29). Sur les Fig. 19 (synopse des maxima) et 22 (synopse des moyennes arithmétiques) l'ont peut prélever au même endroit (P. 5), les valeurs suivantes :

$$\text{MAX}_5 = 87,172 \text{ (voir la Fig. 19),}$$

$$\text{MOY}_5 = 49,389 \text{ (voir la Fig. 22).}$$

La division entre 87,172 et 49,389 donne bien la valeur du quotient mentionnée ci-dessus (1,765).

6.5. Synopse des coefficients d'interaction des distributions de similarité.

Voir les Fig. 30, 31, 32 et 33.

Pour expliquer les propriétés taxométriques du coefficient d'interaction (pour la formule, cf. infra) on doit considérer de plus près les relations numériques à l'intérieur d'une distribution de similarité.

*KARTE 999-TCT	PUNKT:205						
MW= 68.171	S= 7.354	N-KART=688					
MINMWMAX	47.591	54.451	61.311	68.171	73.928	79.686	85.443
MEDMW	47.591	63.610	65.982	68.182	70.906	75.073	85.443
MED	47.591	63.690	66.179	68.321	71.025	75.084	85.443
	11-6	10-4	22-68	(47+30)	(23+19)	(6+4)	= $\sum P_{sm_x}$
	20-20	20-20	20-21	21-21	22-22	22-21	
	20-21	21-21	21-21	21-21	21-21	21-20	

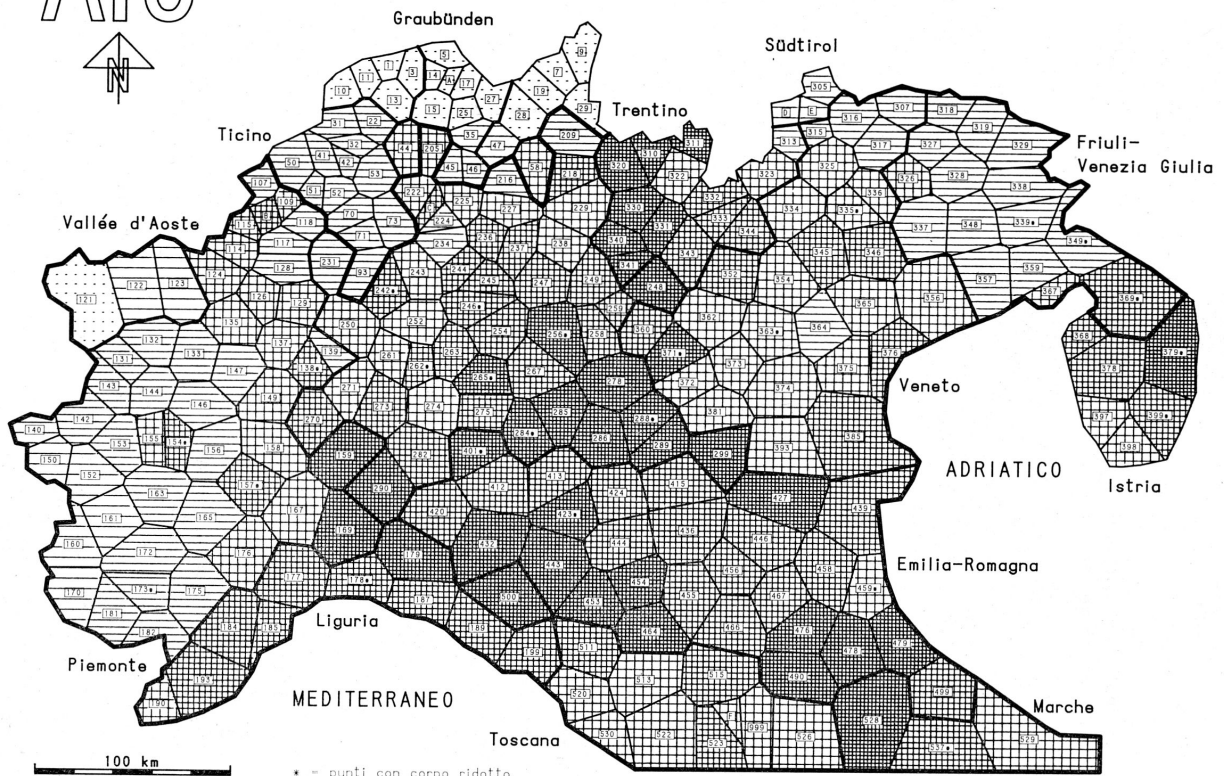
Fig. 30 :

Calcul de la valeur $\sum P_{sm_x}$.

L'intervallisation selon l'algorithme MINMWMAX (voir la Fig. 30 et le paragraphe 3.1.) permet de déterminer le nombre des points d'atlas situés de par et d'autre de la moyenne arithmétique de la distribution de similarité. Pour la distribution de similarité relative au point de référence 205, le bilan de ce partage est le suivant (voir la Fig. 30) :

- au-dessous de la moyenne arithmétique :
 $11 + 6 + 10 + 4 + 22 + 68 = 121$ points d'atlas,
- au-dessus de la moyenne arithmétique :
 $47 + 30 + 23 + 19 + 6 + 4 = 129$ points d'atlas (= $\sum P_{sm_{205}}$).

AI S



100 km

* = punti con corpo ridotto

LEX - ITALIEN

999-TOT-ZUS-INTERAKTION

Punkte mit Buchstaben:

- A = 16
- B = 116
- C = 223
- D = 312
- E = 314
- F = 524



Kartographie: W.-D. Raab

Choroplethenkarte
MINWMAX 6-fach

1		9.481 - 17.885
2		26.290
3		34.694
4		38.269
5		41.843
6		45.418

Häufigkeitsverteilung
MINWMAX 12-fach

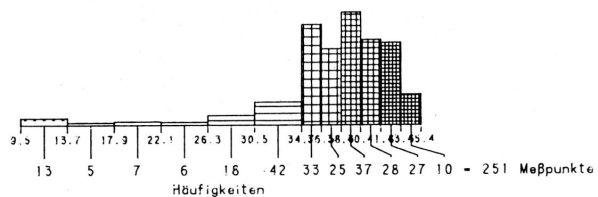


Fig. 31 :

Carte choroplèthe de la synopsis des coefficients d'interaction des 251 distributions de similarité.
Voir aussi les Fig. 32 et 33.

LEY-ITALIEN
KARTE 969-101-7US INTERAKTION

MM# 34.694 S= A.201

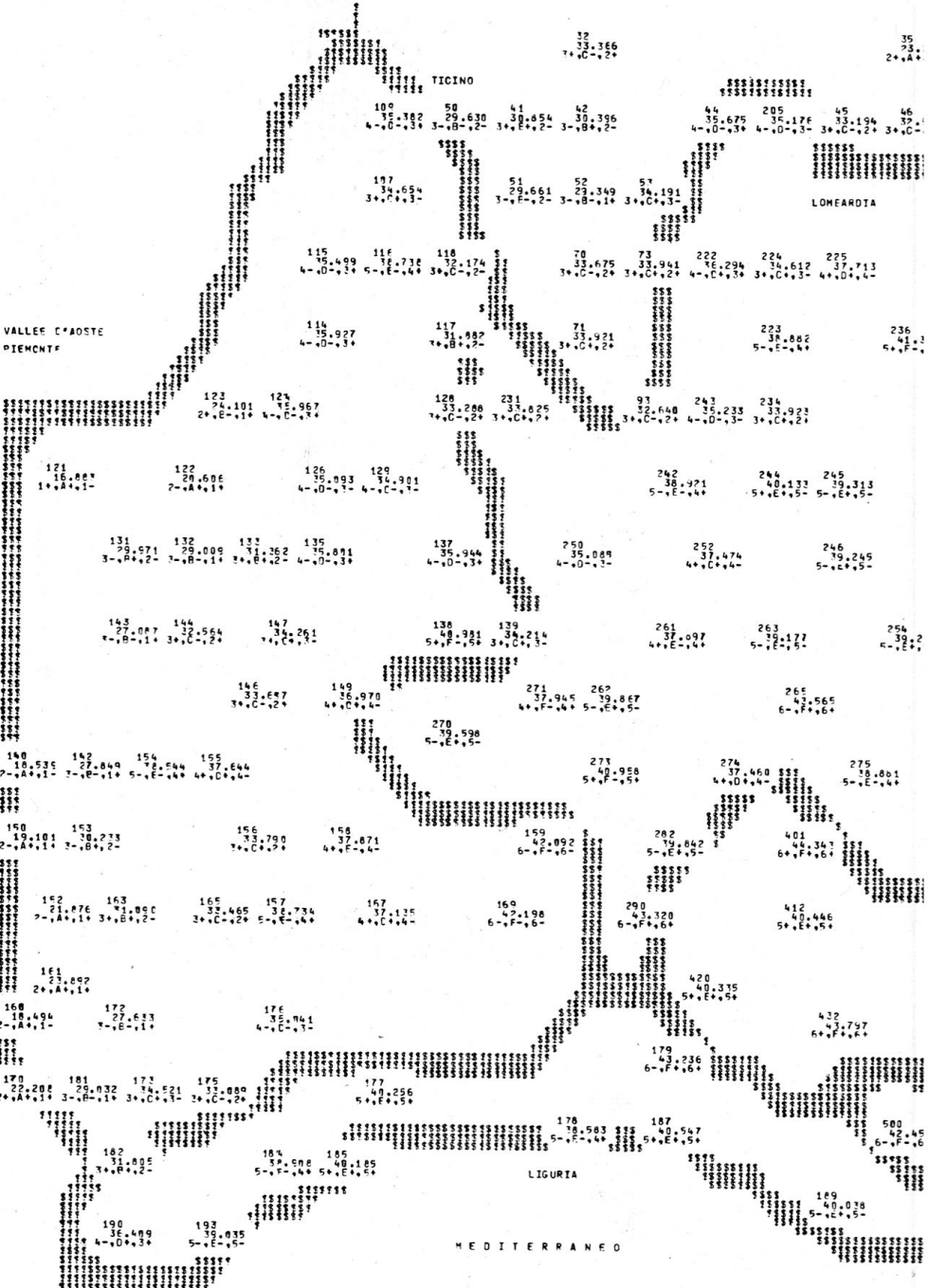
MIN	0.481	17.8A5	26.298	34.694	38.269	41.843	45.418
MAX	0.481	24.101	32.174	34.756	37.748	40.640	45.418
MED	0.481	29.030	34.191	36.555	38.035	41.595	45.418

10	11	1
10.767	9.670	9.441
1-A-1-1	1-A-1-1	1-A-1-1

STREIFEN 1

SVIZZERA
RUMANTSCHA

13	17
10.425	15
1-A-1-1	1-A-1-1



MEDITERRANEO

LEX-ITALIEN
KÄRTE 999-TOT-ZUS INTERAKTION
M= 34.694 S= 8.201
MINMAX 9.481 17.885 26.290 34.694 36.269 41.843 45.418
MEDMM 9.481 24.101 32.174 34.759 37.748 40.644 45.418
MED 9.481 29.030 34.191 36.956 39.835 44.395 45.418

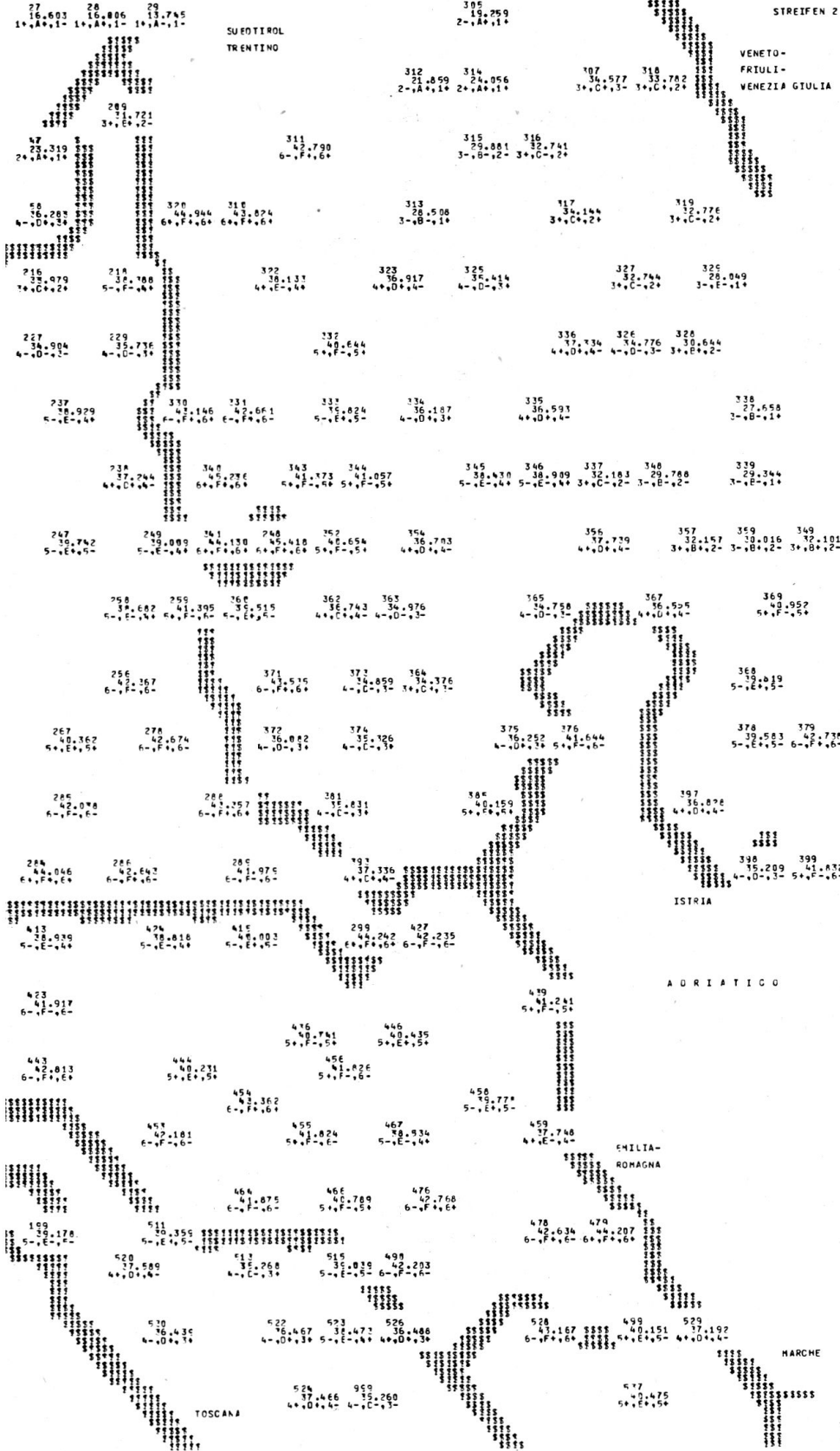
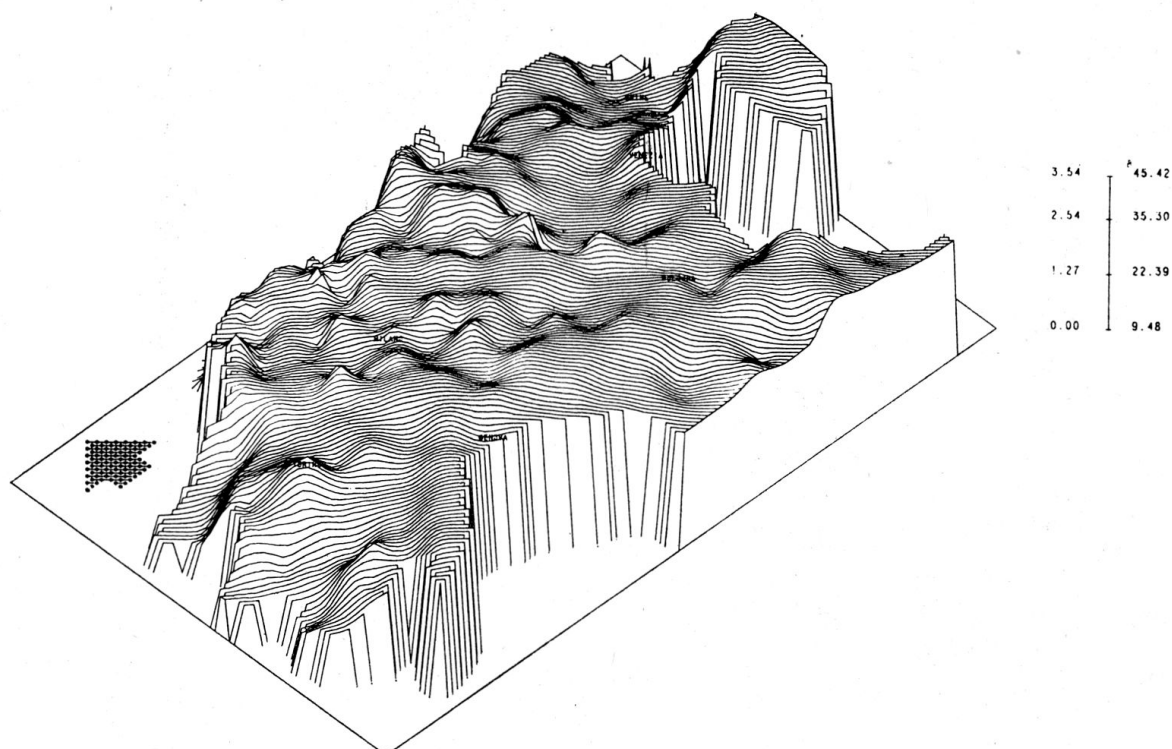


Fig. 33 :

Carte numérique de la synopse des coefficients d'interaction des 251 distributions de similarité.
Voir aussi les Fig. 31 et 32.



ITA-999-TOT-ZUS-INTERAKTION

Azimuth = 45 Neigung = 45
 Laenge = 20.00* Hoehe = 5.00*

* = vor der Projektion 16.06.80

Fig. 32 :

Surface statistique lissée de la synopse des coefficients d'interaction des 251 distributions de similarité.

Voir aussi les Fig. 31 et 33.

Des 250 points d'atlas mis en contact avec le point de référence 205 par l'intermédiaire d'une mesure de proximité (IRI), 121 (ou 48,4 %) en sont typologiquement plutôt éloignés, c'est-à-dire numériquement au-dessous de la moyenne arithmétique) alors que la majorité, à savoir 129 points (ou 51,6 %) en sont typologiquement très proches (c'est-à-dire situés au-dessus de la moyenne arithmétique).

Le même calcul effectué sur les données de la Fig. 11 (carte numérique du point de référence 15) se présente comme suit (voir l'en-tête de la Fig. 11) :

- au-dessous de la moyenne arithmétique :
 $10 + 27 + 41 + 42 + 35 + 34 = 189$ points d'atlas,
- au-dessus de la moyenne arithmétique :
 $39 + 4 + 0 + 7 + 3 + 8 = 61$ points d'atlas ($= \sum P_{sm_{15}}$).

La majorité des points concernés, à savoir 189 sur 250 (ou 75,6 %), se trouvent donc dans une situation fort peu privilégiée par rapport au point de référence 15, étant donnée leur position numérique au-dessous de la moyenne arithmétique. En revanche, la valeur $\sum P_{sm_{15}}$ n'est que de 61 (ou 24,4 %), résultat fort modeste par rapport à ce que nous avons pu constater pour le point de référence 205. Il en appert qu'en dernière analyse, la valeur $\sum P_{sm_x}$ constitue, à l'instar du coefficient de Fisher, un coefficient de symétrie.

Comme toute mesure de proximité — et qui en plus se fait toujours par paires — peut être comparée, par une métaphore ludique, à une interaction sous la forme d'un jeu entre deux partenaires (la règle du jeu étant représentée par les propriétés de l'indice de proximité utilisé), nous pouvons reformuler le problème de la mesure des proximités entre les différents points (nœuds) d'un réseau d'atlas dans les termes suivants :

- Soient n individus ayant chacun p propriétés nominales, prêts à jouer une seule fois contre tout autre individu.
- Soit en plus une règle de jeu représentée par un indice de proximité (en l'espèce l'IRI).
- Le tournoi comprendra donc n^2 jeux.
- Chaque jeu donnera lieu à une évaluation numérique exprimant la réussite des joueurs.

Bien que pour l'évaluation numérique des jeux on puisse se contenter de l'utilisation de la valeur $\sum P_{sm_x}$ décrite ci-dessus, il semble préférable d'intégrer dans la formule définitive aussi la moyenne arithmétique dont la vertu classificatoire a déjà été éprouvée (cf. 6.2.). La formule du coefficient d'interaction se présente donc comme suit :

$$INT_x = \frac{\sum P_{sm_x} \cdot MOY_x}{250},$$

- où INT_x est le coefficient d'interaction de la distribution de similarité relative au point de référence x ,
- ΣP_{sm_x} est le nombre des scores (P) situés au-dessus (sm) de la moyenne arithmétique d'une distribution de similarité relative au point de référence x ,
- MOY_x est la moyenne arithmétique de la distribution de similarité relative au point de référence x .

Les Fig. 31, 32 et 33 montrent — pour nous en tenir à la métaphore ludique — les résultats des n^2 jeux mentionnés ci-dessus (23). Au vu de ces graphiques on note aisément que l'agencement spatial des coefficients d'interaction rappelle celui de la Fig. 23, à la différence près qu'aux valeurs minimales de la Fig. 23 correspondent les valeurs maximales de la Fig. 31.

Le minimum des coefficients d'interaction (9,481) se trouve au point 1 (Grisons occidentaux), alors que le maximum (45,418) peut être repéré au point 248 (Lombardie orientale), là où se rencontrent les domaines dialectaux lombard, trentin et vénitien. Nous sommes donc à même de dire que, parmi les joueurs qui ont participé à ce tournoi, c'est le joueur n° 248 (i.e. le point 248) qui, à force d'avoir obtenu le meilleur résultat, en est le gagnant, alors que le joueur n° 1 (i.e. le point 1), complètement isolé et occupant une position stratégique désavantageuse, est le perdant du tournoi.

A l'appui des Fig. 11 et 14 il est possible de vérifier deux des scores répertoriés dans la Fig. 33, à savoir INT_{15} et INT_{205} .

Pour INT_{15} voir la Fig. 11 :

- $\Sigma P_{sm_{15}} = 61$
- $MOY_{15} = 50,204$
- $INT_{15} = \frac{61 \cdot 50,240}{250} = 12,250.$

Repérer cette valeur dans la Fig. 33 à l'endroit du point 15.

(23) En réalité le nombre des jeux différents est de $n/2 \cdot (n-1)$ étant donné le fait qu'un jeu entre les joueurs A et B correspond exactement à celui entre les joueurs B et A (symétrie des indices de proximité) et qu'un jeu entre A et A est dépourvu de sens (réflexivité des indices de proximité).

Pour INT_{205} voir la Fig. 14 :

- $\sum P_{sm_{205}} = 129$
- $MOY_{205} = 68,174$
- $INT_{205} = \frac{129 \cdot 68,171}{250} = 35,173.$

Repérer cette valeur dans la Fig. 33 à l'endroit du point 205.

7. Epilogue et perspectives interdisciplinaires.

Nous tenons à souligner encore une fois que les réflexions amorcées ci-dessus débouchent sur la théorie des graphes, la théorie des réseaux de transport et des circuits électriques (angl. network theory ; all. Netzwerkktheorie) et la théorie des jeux ⁽²⁴⁾. Bien que ces disciplines mathématiques soient d'un abord difficile pour la dialectologie classique, vu le peu d'envie que ressentent beaucoup de représentants des sciences humaines d'utiliser des méthodes censées appartenir exclusivement aux sciences exactes, il semble néanmoins évident que nos connaissances sur la nature des réseaux dialectaux, et, partant, sur tout ce qui a trait à l'enchevêtrement des registres et niveaux linguistiques, profiteraient énormément d'une ouverture méthodologique pratiquée dans un esprit de pluridisciplinarité. La dialectométrie, telle qu'elle a été présentée ici, n'en constitue qu'une première et modeste ébauche.

8. Abréviations et glossaire (termes techniques et allemands).

Choroplethenkarte	carte choroplèthe (voir les cartes choroplèthes).
Graubünden	Grisons (voir les cartes choroplèthes).
Häufigkeiten	fréquences (absolues) (voir les cartes choroplèthes).
Häufigkeitsverteilung	distribution de fréquence ou de similarité (voir les cartes choroplèthes).
Hoehe	hauteur (voir les surfaces statistiques lissées).
IRI	Indice Relatif d'Identité (cf. 2.).
Laenge	longueur (voir les surfaces statistiques lissées).
MAX	valeur maximale d'une distribution de similarité (cf. 6.1., les cartes choroplèthes et les cartes numériques).

(24) Cf. la note 18.

Meßpunkte	points d'atlas (voir les cartes choroplèthes).
MED, MEDMW	algorithmes d'intervallisation (pas traités dans cet article) (voir les cartes numériques et la note 13).
MINMWMAX (6-fach, 12-fach)	algorithme d'intervallisation (à 6 intervalles, à 12 intervalles) (cf. 3.1., les cartes choroplèthes et les cartes numériques).
MW, MIW	moyenne arithmétique (all. Mittelwert) (voir les cartes choroplèthes et les cartes numériques).
Neigung	inclinaison (voir les surfaces statistiques lissées).
N-KART	nombre des attributus disponibles dans un vecteur ponctuel de la matrice des données (cf. 4.).
PRF	point de référence (voir la Fig. 8).
Punkte mit Buchstaben	points d'atlas marqués par des lettres (voir les cartes choroplèthes).
RIW	Indice Relatif d'Identité (all. Relativer Identitätswert) (cf. 2.).
S	écart-type (all. Standardabweichung) (voir les cartes numériques).
Schiefe	coefficient de symétrie (de Fisher) (cf. 6.3., les Fig. 23, 25 et 26).
Südtirol, Suedtirol	Tyrol du Sud (voir les cartes choroplèthes et les cartes numériques).
vor der Projektion	avant la projection (voir les surfaces statistiques lissées).

9. Bibliographie.

Il nous a paru utile de fournir aux lecteurs des indications bibliographiques qui, pour internationales qu'elles soient dans leur ensemble, donnent une certaine préférence aux travaux en langue française.

Abler, R./J. S. Adams/P. Gould : Spatial organization. The Geographer's view of the world, Londres 1977.

AIS : Sprach- und Sachatlas Italiens und der Südschweiz, ed. K. Jaberg/J. Jud, vol. I-VIII, Zofingen 1928-1940.

Altmann, G./W. Lehfeldt : Allgemeine Sprachtypologie. Prinzipien und Meßverfahren, Munich 1973.

Bahrenberg, G./E. Giese : Statistische Methoden und ihre Anwendung in der Geographie, Stuttgart 1975.

Béguin, H. : Méthodes d'analyse géographique quantitative, Paris 1979.

Benzécri, J.-P./F. Benzécri et alii : Pratique de l'analyse des données. 1. Analyse des correspondances. Exposé élémentaire, Paris 1980.

Berge, Cl. : Théorie des graphes et ses applications, Paris 1967².

- Berge, Cl./A. Ghouila-Houri : Programmes, jeux et réseaux de transport, Paris 1962.
- Bertier, P./J.-M. Bouroche : Analyse des données multidimensionnelles, Paris, 1975.
- Bock, H. H. : Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse), Göttingen 1974.
- Bouroche, J.-M./G. Saporta : L'analyse des données, Paris 1980.
- Brassel, K. E./D. Reif : A procedure to generate Thiessen polygons, in : Geographical Analysis 11 (1979), 289-303.
- Chandon, J./L.-S. Pinson : Analyse typologique. Théories et applications, Paris, New York, Barcelone, Milan 1981.
- Claval, P. : La Nouvelle Géographie, Paris 1977.
- Cowan, H. K. J. : Old Limburgian and the quantitative classification of West germanic dialects, in : Leuvense Bijdragen 53 (1964), 1-21.
- Dickinson, G. C. : Statistical mapping and the presentation of statistics, Londres 1973².
- Doerfer, G. : Bemerkungen zur linguistischen Klassifikation, in : Indogermanische Forschungen 76 (1971), 1-14.
- Durand, J.-P. : Notes de philologie rouergate (suite), in : Revue des langues romanes 33 (1889), 47-84.
- Ellegård, A. : Statistical measurement of linguistic relationship, in : Language 35 (1959), 131-156.
- Fischer, M. M. : Regional taxonomy. A comparison of some hierarchic and non-hierarchic strategies, in : Regional Science and Urban Economics 10 (1980), 503-537.
- Flament, Cl. : Théorie des graphes et structures sociales, Paris 1968.
- Fossat, J.-L. : Vers un traitement automatique des données dialectologiques, en dialectométrie, in : Putschke 1977, 314-334.
- Fossat, J.-L. : Variation et théorie des isoglosses quantitatives, in : Cahiers du Centre interdisciplinaire des sciences du langage. Sections de linguistique générale et appliquée (Université de Toulouse - Le Mirail) 2 (1980), 155-164.
- French, H. M./J.-B. Racine : Quantitative and qualitative geography. La nécessité d'un dialogue, Ottawa 1971.
- Fronzaroli, P. : Metodi statistici nella linguistica diacronica, in : Il Verri 24 (1967), 100-115.
- Fronzaroli, P. : Problemi di classificazione delle lingue su base quantitativa, in : Colloquio sul tema : Le tecniche di classificazione e loro applicazione linguistica (Firenze 1972), Accademia Nazionale dei Lincei, anno 372/1975, Rome 1975, 123-141.
- Gendre, F. : L'analyse statistique univariée. Introduction à son utilisation pratique, Genève 1977.

- Goebel, H. : La dialectométrie appliquée à l'ALF (Normandie), in : Congresso internazionale di linguistica e filologia romanza. Atti, ed. A. Vàrvaro, Naples, Amsterdam 196, vol. II, 165-195.
- Goebel, H. : Rätoromanisch versus Hochitalienisch versus Oberitalienisch. Dialektometrische Beobachtungen innerhalb eines Diasystems, in : Ladinia 1 (1977 a), 39-71.
- Goebel, H. : Zu Methoden und Problemen einiger dialektometrischer Meßverfahren, in : Putschke 1977 b, 335-365.
- Goebel, H. : Analyse dialectométrique de quelques points de l'AIS (italien standard, valdotain, provençal alpin, turinois, milanais), in : Lingue e dialetti nell'arco alpino occidentale. Atti del Convegno internazionale di Torino (1976), ed. G. P. Clivio/G. Gasca Queirazza, Turin 1978 a, 282-294 (cartes).
- Goebel, H. : Dialektometrie - Methoden und Probleme einer empirischen Betrachtung von Sprache im Raum (Tondiaschau), in : Proceedings of the twelfth international Congress of Linguists (Vienna 1977), ed. W. U. Dressler/W. Meid, Vienne 1978 b, 611-613.
- Goebel, H. : Dialektgeographie + Numerische Taxonomie = Dialektometrie. Anhand rätoromanischer und oberitalienischer Dialektmaterialien (AIS), in : Ladinia 4 (1980), 31-95.
- Goebel, H./N. Winterleitner : Projekt einer sprachstatistischen Auswertung von in Sprachatlanten gespeicherter linguistischer Information mit Hilfe elektronischer Rechenanlagen, in : Linguistische Berichte 14 (1971), 60-61.
- Grimes, J. E./F. B. Agard : Linguistic divergence in Romance, in : Language 35 (1969), 588-604.
- Guiter, H. : Atlas et frontières linguistiques, in : Les dialectes romans de France à la lumière des atlas régionaux (Colloque de Strasbourg 1971), éd. G. Straka/P. Gardette, Paris 1973, 61-109.
- Guiter, H. : Une vérification de loi linguistique par corrélation, in : Revue de linguistique romane 38 (1974), 253-264.
- Guiter, H. : Critique et limites d'une méthode, in : Mélanges à la mémoire de L. Michel, Montpellier 1979, 261-272.
- Haggett, P. : L'analyse spatiale en géographie humaine, Paris 1973.
- Herdan, G. : Mathematics of genealogical relationship between languages, in : Proceedings of the ninth international Congress of Linguists (Cambridge 1962), Londres, La Haye, Paris 1964, 51-60.
- Houck, Ch. L. : A statistical and computerized methodology for analysing dialect material, Diss. Iowa 1969.
- Jaberg, K. : Géographie linguistique et expressivisme phonétique : les noms de la balançoire en portugais, in : Revista portuguesa de filologia 1 (1947), 1-44 (aussi in : K. Jaberg, Sprachwissenschaftliche Forschungen und Erlebnisse, N. F., éd. S. Heinemann, Berne 1965, 63-90).
- Jaberg, K./J. Jud : Der Sprachatlas als Forschungsinstrument. Kritische Grundlegung und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz, Halle 1928.

- Jambu, M. : Classification automatique pour l'analyse des données. 1. Méthodes et algorithmes, Paris 1978.
- Jardine, N./R. Sibson : Mathematical taxonomy, Londres 1971.
- Jenks, G. F. : Generalization in statistical mapping, in : Annals of the Association of American Geographers 53 (1963), 15-26.
- Jenks, G. F./F. C. Caspall : Error on choroplethic maps. Definition, measurement, reduction, in : Annals of the Association of American Geographers 61 (1971), 217-244.
- Jenks, G. F./M. R. C. Coulson : Class intervalls for statistical maps, in : Internationales Jahrbuch für Kartographie 3 (1963), 119-134.
- Kroeber, A. L./C. D. Chrétien : Quantitative classification of Indo-european languages, in : Language 13 (1937), 83-103.
- Monmonier, M. S. : Viewing azimuth and map clarity, in : Annals of the Association of American Geographers 68 (1978), 180-195.
- Muljačić, Ž. : Die Klassifikation der romanischen Sprachen, in : Romanistisches Jahrbuch 18 (1967), 23-37.
- Muljačić, Ž. : Per una classificazione 'standardologica' delle lingue romanze, in : Italic and Romance (Linguistic studies in honor of E. Pulgram), éd. H. J. Izzo, Amsterdam 1980, 85-96.
- Naumann, C. L. : Grundzüge der Sprachkartographie und ihrer Automatisierung, in : Germanistische Linguistik 1-3/76 [1976], 1-285.
- Opitz, O. : Numerische Taxonomie, Stuttgart, New York 1980.
- Peucker, T. K. : Die Strukturen der digitalen Erfassung und Darstellung topographischer Oberflächen, in : Beiträge zur Theoretischen Kartographie, éd. I. Kretschmer, Vienne 1977, 125-140.
- Philps, D. : Balaguère - Bethmale - Biros. Etude dialectométrique. Thèse de doctorat de III^e cycle, Toulouse 1975.
- Pudlatz, H. : Automatische Erzeugung von Isoglossen auf dem Plotter mit Hilfe von Thiessen-Polygonen, in : Putschke 1977, 245-258.
- Putschke, W. éd. : Automatische Sprachkartographie. Vorträge des Internationalen Kolloquiums zur Automatischen Sprachkartographie (Marburg 1977), in : Germanistische Linguistik 3-4/77 [1977], 1-395.
- Racine, J.-B./H. Raymond : L'analyse quantitative en géographie, Paris 1973.
- Reed, D. W./J. L. Spicer : Correlation methods of comparing idiolects in a transition area, in : Language 28 (1952), 348-359.
- Rensch, K. H. : Die dia-synchronische Verwandtschaft von Phonemsystemen. Ein Beitrag zur strukturellen Dialektologie, in : Orbis 19 (1970), 94-101.
- Rhynsburger, D. : Analytic delineation of Thiessen polygons, in : Geographical Analysis 5 (1973), 133-144.
- Sache, H. : La théorie des graphes, Paris 1974.

- Scripter, M. W. : Nested-means map classes for statistical maps, in : *Annals of the Association of American Geographers* 60 (1970), 385-393.
- Séguy, J. : La relation entre la distance spatiale et la distance lexicale, in : *Revue de linguistique romane* 35 (1971), 335-357.
- Séguy, J. : La dialectométrie dans l'Atlas linguistique de la Gascogne, in : *Revue de linguistique romane* 37 (1973 a), 1-24.
- Séguy, J. : Atlas linguistique et ethnographique de la Gascogne, vol. V-VI/1+2, Paris 1973 b.
- Shaw, D. : Statistical analysis of dialectal boundaries, in : *Computers and the Humanities* 8/3 (1974), 173-177.
- Sneath, P. H. A./R. R. Sokal : Numerical taxonomy. The principles and practice of numerical classification, San Francisco 1973.
- Sodeur, W. : Empirische Verfahren zur Klassifikation, Stuttgart 1974.
- Sokal, R. R./P. H. A. Sneath : Principles of numerical taxonomy, San Francisco 1963.
- Thiessen, A. H. : Precipitation averages for large areas, in : *Monthly weather Review* 39 (1911), 1082-1084.
- Viereck, W. : Dialektometrie und englische Dialektologie, in : *Grazer linguistische Studien* 11/12 (1980), 335-356.
- Vogel F. : Probleme und Verfahren der numerischen Klassifikation unter besonderer Berücksichtigung von Alternativmerkmalen, Göttingen 1975.

Ratisbonne.

Hans GOEBL