

Patternized configuration of geolinguistic data:

Application to Spanish lexical variation

Methods XI at Joensuu University (Finland)

Hiroto Ueda (University of Tokyo)

Geolinguistic data, composed of linguistic forms and geographic points (places), may be analyzed by various methods, such as cluster analysis, multidimensional analysis, factor analysis and principal component analysis, among others. All these are variants on a common statistical procedure, i.e. multivariate calculation, in which linguistic data and/or geographic points are considered as multiple variables. Each of these methods has its own characteristics, which are so different from each other that it is recommended to use them in accordance with the purpose of the study, data size and, naturally, research resources such as the available computer applications, human resources and time. The methods presented above are relatively complicated and, as a result, researchers in the humanities may tend to refrain from using them. On this occasion I wish to present a method for patternizing configuration which is much more simple and easy to understand.

1. Procedure

1.1. Average distance from zero point

Before I proceed to explain the operation, it will be convenient to reconsider the ‘paradox of any taxonomical study’. Supposing that we intend to classify the distributional data (n multiplied by m), consisting of n individuals and m attributes, such as n elicited word forms and m places, presented as follows:

Table 1: Before patternization

Word	L-1	L-2	L-3	L-4
W-1	+	+		
W-2			+	
W-3		+		
W-4			+	+
W-5	+	+	+	

In this table we have 5 cases represented in terms of 5 words (W-1, W-2, ... W-5); and 4 attributes, which are 4 geographic places (L-1, L-2, L-3 and L-4). This is a simplified form of the usual descriptive presentation of the result of any field work. We would say that W-1 is

found in L-1 and L-2, W-2 in L-3, and so on. In practice we would have hundreds or thousands of words and the number of places would be a variable according to the scale of investigation, but the manner of composition would be the same. With this method, however, we would not be able to classify either the words or the places because of the lack of any criterion. Or, to put it in another way, the words and the places are so interdependent that we could not classify *a priori* the words according to the places, nor the places according to the words. This is what we call the ‘taxonomical paradox’. For example, assuming that we have 5 individuals and 4 books selected by these individuals and we want to divide these individuals into two groups according to their particular reading taste. One could say that Group A consists of those who select the books in Group Z. Then we could ask the question: How is Group Z of the books defined? The answer would be that Group Z consists of those books which the people in Group A selected. We notice that our argument seems to be circular.

Now, our interest is to reconsider the order of the words presented in the figure above. Except in the case in which the words are arranged by means of some predetermined external criterion, such as alphabetical order or semantic feature, we would not be able to justify the order of W-1 to W-5. Then proceeding to the rearrangement of the words so as to obtain an order in the form of a gradual transition in a diagonal line, without changing the absolute points of reaction, we could present the result as follows:

Table 2: Patternization (External criterion: L)

[PATTERN]	L-1	L-2	L-3	L-4
W-1	+	+		
W-3		+		
W-5	+	+	+	
W-2			+	
W-4			+	+

The method which we use to obtain such a patternized configuration is relatively simple and needs no complicated mathematical operation. It is derived from the deviation in terms of average distance from zero point (a.d.z.p.), which is formulated:

$$\text{a.d.z.p.} = \sqrt{\frac{\sum x_i^2}{n}}$$

where x_i is the distance of each reaction point and n , the number of reactions.

For example in Table 1 we observe in the case of W-1 reaction points in L-1 and L-2. Consequently the a.d.z.p. will be calculated as follows:

$$\text{a.d.z.p. (W-1)} = \sqrt{\frac{1^2 + 2^2}{2}} = 1.581$$

Likewise in the case of W-2, we have only one reaction in location 3, so the a.d.z.p. result is:

$$\text{a.d.z.p. (W-2)} = \sqrt{\frac{3^2}{1}} = 3$$

The cases of the other Words may be calculated in the same manner:

$$\text{a.d.z.p. (W-3)} = \sqrt{\frac{2^2}{1}} = 2$$

$$\text{a.d.z.p. (W-4)} = \sqrt{\frac{3^2 + 4^2}{2}} = 3.535$$

$$\text{a.d.z.p. (W-5)} = \sqrt{\frac{1^2 + 2^2 + 3^2}{2}} = 2.160$$

Once we obtain the results of all the Words, we proceed to rearrange them in accordance with the a.d.z.p. and we obtain the order: W-1 (1.581), W-3 (2), W-5 (2.160), W-2 (3), W-4 (3.535), which is the patterned configuration represented in Table 2.

1.2. Internal criterion

The patternized configuration acquired by means of the rearrangement of Words (W-1, W-2, ..., W-n) is useful for discovering the relationship of the Words in question between themselves in terms of the reactions in each Place. The order of Places, L-1, L-2, ..., L-m, however, may also be reconfigured if we take into consideration the reaction points in terms of the Words. Effectively by the rearrangement of Places we obtain a more neatly patternized configuration, as Table 3 shows:

Table 3: Patternization (Internal criterion)

[PATTERN]	L-2	L-1	L-3	L-4
W-3	+			
W-1	+	+		
W-5	+	+	+	
W-2			+	
W-4			+	+

In this table we observe the rearrangement of the order not only of the Words (W-3, W-1, W-5, W-2, W-4) but also of the Places (L-2, L-1, L-3, L-4). Now we can affirm that the five words may be divided into two groups: the three Words, W-3, W-1 and W-5 are in a group which is characterized by the selection of Places L-2 and L-1, on the one hand, while W-2 and W-4 constitute a group corresponding to the Places L-3 and L-4, on the other.

In order to arrive at such a neatly patternized configuration, we need a reiteration of the same calculation as the one we performed in the previous section. Departing from the result of the rearrangement of the Words, we proceed to calculate the a.d.z.p. of the Places in terms of the selection of Words. In Table 2, we have the L-1 with two reaction points W-1 and W-5 which occupy the points 1 and 3, starting from a zero point, respectively. So the a.d.z.p. of the L-1 will be calculated as follows.

$$\text{a.d.z.p. (L-1)} = \sqrt{\frac{1^2 + 3^2}{2}} = 2.236$$

The other Places may be calculated successively:

$$\text{a.d.z.p. (L-2)} = \sqrt{\frac{1^2 + 2^2 + 3^2}{3}} = 2.160$$

$$\text{a.d.z.p. (L-3)} = \sqrt{\frac{3^2 + 4^2 + 5^2}{3}} = 4.082$$

$$\text{a.d.z.p. (L-4)} = \sqrt{\frac{5^2}{1}} = 5$$

Now we notice that the order of L-1 (a.d.z.p.=2.236) and L-2 (2.160) must be inverted and we obtain the configuration:

Table 4

[PATTERN]	L-2	L-1	L-3	L-4
W-1	+	+		
W-3	+			
W-5	+	+	+	
W-2			+	
W-4			+	+

At this stage, we calculate once more the a.d.z.p. of Words in respect of the selection of places:

$$\text{a.d.z.p. (W-1)} = \sqrt{\frac{1^2 + 2^2}{2}} = 1.581$$

$$\text{a.d.z.p. (W-3)} = \sqrt{\frac{1^2}{1}} = 1$$

$$\text{a.d.z.p. (W-5)} = \sqrt{\frac{1^2 + 2^2 + 3^2}{3}} = 2.160$$

$$\text{a.d.z.p. (W-2)} = \sqrt{\frac{3^2}{1}} = 3$$

$$\text{a.d.z.p. (W-4)} = \sqrt{\frac{3^2 + 4^2}{2}} = 3.535$$

Then the order of W-1 (1.581) and W-3 (1.000) must be inverted. By rearranging this order, we obtain Table 3. Now we calculate the a.d.z.p. once more and we can confirm that we do not need any more rearrangement, since the order is correct:

$$\text{a.d.z.p. (L-2)} = \sqrt{\frac{1^2 + 2^2 + 3^2}{3}} = 2.160$$

$$\text{a.d.z.p. (L-1)} = \sqrt{\frac{2^2 + 3^2}{2}} = 2.550$$

$$\text{a.d.z.p. (L-3)} = \sqrt{\frac{3^2 + 4^2 + 5^2}{3}} = 4.082$$

$$\text{a.d.z.p. (L-4)} = \sqrt{\frac{5^2}{1}} = 5$$

In practice, the distribution pattern of a certain scale composed of, for example, 500 individuals by 500 attributes will converge on a patternized configuration after fewer than 20 reiterations.

1.3. Differentiation of the score

We would like to point out that there may be cases in which different distribution patterns may present the same a.d.z.p. value; for example:

Word-a: L-4, L-6, L-11

Word-b: L-3, L-8, L-10

We obtain the result as follows:

$$\text{a.d.z.p. (Word-a)} = \sqrt{\frac{4^2 + 6^2 + 11^2}{3}} = 7.5939$$

$$\text{a.d.z.p. (Word-b)} = \sqrt{\frac{3^2 + 8^2 + 10^2}{3}} = 7.5939$$

In order to differentiate the a.d.z.p. score, we add a slight value to a.d.z.p.:

$$\text{dif. score} = \frac{(n + m)}{(c \times 100)}$$

With this formula, we obtain the dif. score of each case:

$$\text{dif. score (Word-a)} = \frac{(3 + 6)}{(11 \times 100)} = 0.0127$$

$$\text{dif. score (Word-b)} = \frac{(3 + 10)}{(11 \times 100)} = 0.0118$$

This is an extremely slight value, used only for differentiating words with the same a.d.z.p. but with a different distribution pattern. If we added a larger value than this, the result would be so high that it would destroy the order of the rearrangement itself. The a.d.z.p. with dif. score of each case now becomes:

$$\text{a.d.z.p. + dif. score (Word-a)} = 7.5939 + 0.0127 = 7.6066$$

$$\text{a.d.z.p. + dif. score (Word-b)} = 7.5939 + 0.0118 = 7.6057$$

1. 4. Interpretation of the two axes

The a.d.z.p. indicates the distance with respect to the zero point. The next table shows the result of patternization according to the horizontal criterion (Place):

Table 5

W	a.d.z.p.	L	a.d.z.p.
W-1	1.581	L-1	2.236
W-3	1.732	L-2	2.191
W-5	1.958	L-3	3.041
W-2	2.138	L-4	3.317
W-4	2.517		

This table may be represented by means of a graph:

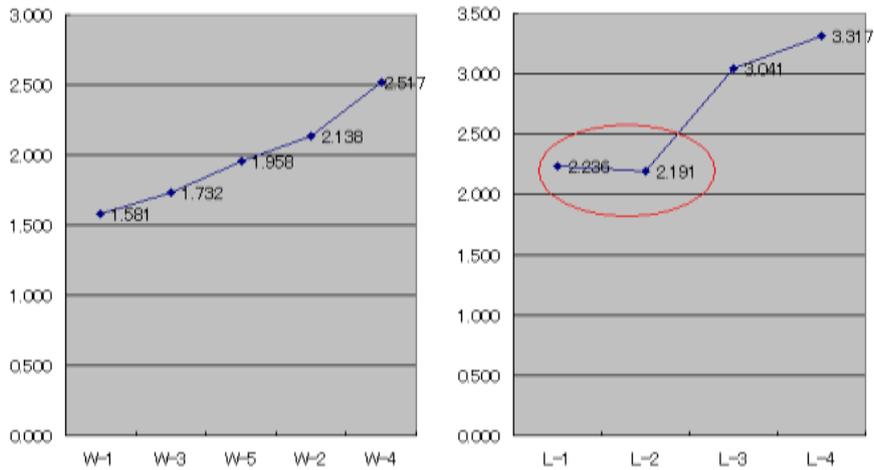


Figure 1: Rearrangement of W

We observe in the first of these graphs that the arrangement is correct in the sense that the line is ascending gradually (from W-1 to W-4), while that of the geographic points shows no such graduation in the transition between L-1 and L-2.

On the other hand we have the results of the patternization according to the internal criterion, which show the graduation in the order of geographic points as well as in that of words.

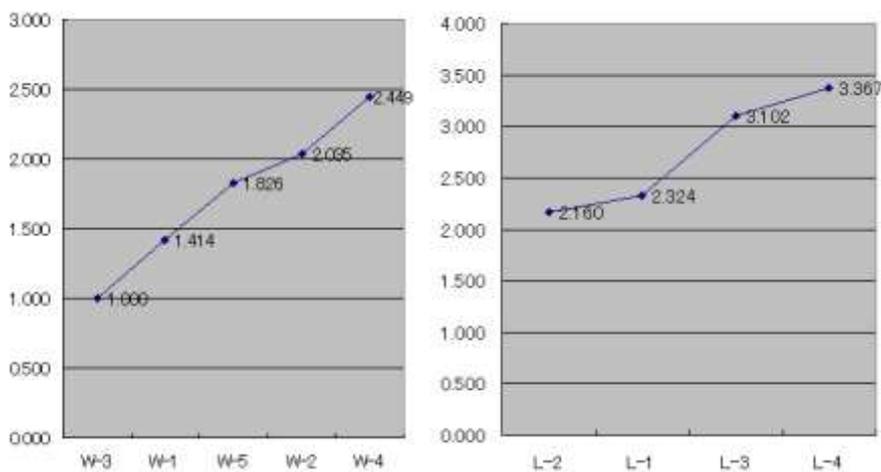


Figure-2: Rearrangement of W and L

2. Application

2.1. World-wide Spanish-speaking cities

Since 1993, I have been participating in an investigation of Spanish lexical variation along with researchers in Spanish speaking countries¹. We have been able to investigate, so far, 55 cities in the Iberian Peninsula and in the Spanish-speaking parts of the Americas and the Caribbean, as well as some points in Africa and the United States. The following is a list of the cities we have investigated, together with the abbreviation code used in the patternization analysis:

1 ES-COR: La Coruña, SPAIN; 2 ES-SCO Santiago de Compostela, SPAIN; 3 ES-OVI Oviedo, SPAIN; 4 ES-STD Santander, SPAIN; 5 ES-ZAR Zaragoza, SPAIN; 6 ES-BAR Barcelona, SPAIN; 7 ES-GDL Guadalajara, SPAIN; 8 ES-MAD Madrid, SPAIN; 9 ES-VAL Valencia, SPAIN; 10 ES-GRA Granada, SPAIN; 11 ES-MLG Málaga, SPAIN; 12 ES-ALM Almería, SPAIN; 13 ES-MUR Murcia, SPAIN; 14 ES-TEN Santa Cruz de Tenerife, SPAIN ; 15 ES-PAL Las Palmas de Gran Canaria, SPAIN; 16 ES-SLM Salamanca, SPAIN; 17 GE-MAL Malabo, ECUATORIAL GUINEA; 18 CU-HAB La Habana, CUBA; 19 CU-SCU Santiago de Cuba, CUBA; 20 RD-STI Santiago, DOMINICAN REPUBLIC; 21 RD-SDO Santo Domingo, DOMINICAN REPUBLIC; 22 PR-SJU San Juan, PUERTO RICO; 23 PR-DOR Dorado, PUERTO RICO; 24 PR-MAY Mayagüez, PUERTO RICO; 25 EU-NVO Nuevo Orleans, UNITED STATES; 26 MX-MON Monterrey, MEXICO; 27 MX-AGS Aguas Calientes, MEXICO; 28 MX-MEX Ciudad de México, MEXICO; 29 MX-MRD Mérida, MEXICO; 30 GU-GUA Guatemala, GUATEMALA; 31 EL-SSV San Salvador, EL SALVADOR; 32 HO-TEG Tegucigalpa, HONDURAS; 33 HO-NAC Nacaome, HONDURAS; 34 NI-LEO León, NICARAGUA; 35 NI-MAN Managua, NICARAGUA; 36 CR-LMN Puerto Limón, COSTA RICA; 37 CR-SJO San José, COSTA RICA; 38 PN-PAN Panamá, PANAMA; 39 CO-MED Medellín, COLOMBIA; 40 CO-BOG Santafé de Bogotá, COLOMBIA; 41 VE-CAR Caracas, VENEZUELA; 42 VE-MER Mérida, VENEZUELA; 43 EC-QUI Quito, ECUADOR; 44 PE-LIM Lima, PERU; 45 PE-ARE Arequipa, PERU; 46 BO-PAZ La Paz, BOLIVIA; 47 CH-ARI Arica, CHILE; 48 CH-SCH Santiago, CHILE; 49 CH-CON Concepción, CHILE; 50 CH-TEM Temuco, CHILE; 51 PA-ASU Asunción, PARAGUAY; 52 UR-MTV Montevideo, URUGUAY; 53 AR-SAL Salta, ARGENTINA; 54 AR-SJN San Juan, ARGENTINA; 55 AR-BUE Buenos Aires, ARGENTINA.

2.2. Spanish variable lexicon: Parts of the automobile

The study focuses on the Spanish variable lexicon concerning the modern urban life and on this occasion I will extract from our data base the semantic field of the parts of an automobile². The concept code number (V) corresponds to the identifier used in the Varilex

¹ See: Takagaki (1993), Ueda (1994, 1995a, 1995b, 1996a, 1996b, 1996c, 2000). The project is referred to by López Morales (1996; p.23). The homepage of the project is:

<http://gamp.c.u-tokyo.ac.jp/ueda>

² For previous studies on the Spanish variable lexicon of the concept of transportation, see: Quilis (1982), Kovacci (1988) and Ueda (2000).

project data base. In each city four informants have been asked what they call the part in question. They were shown a list of terms and asked to mark the one(s) they used. If they found that the form they used was not included in the list they added their form(s) in the space prepared for “Others”³. Positive answers were considered to be confirmed if more than two persons chose the forms in question. Consequently cases of hapax (only one registration) are excluded. In the following example, we have seven different forms corresponding to the same signified object: “steering wheel”:

V142. STEERING WHEEL

dirección CR-SJO EC-QUI.

guía RD-STI RD-SDO PR-SJU PR-DOR PR-MAY MX-MRD.

manivela CR-SJO.

manubrio CH-ARI CH-SCH CH-CON.

timón CU-HAB CU-SCU EU-NVO GU-GUA EL-SSV HO-TEG HO-NAC NI-LEO NI-MAN PN-PAN
CO-BOG PE-LIM PE-ARE.

volante ES-COR ES-SCO ES-OVI ES-STD ES-ZAR ES-BAR ES-GDL ES-MAD ES-VAL ES-GRA
ES-MLG ES-ALM ES-MUR ES-TEN ES-PAL ES-SLM GE-MAL CU-HAB RD-STI PR-SJU
MX-MON MX-AGS MX-MEX MX-MRD EL-SSV HO-TEG NI-LEO NI-MAN CR-SJO PN-PAN
CO-MED VE-CAR VE-MER EC-QUI BO-PAZ CH-ARI CH-CON CH-TEM PA-ASU UR-MTV
AR-SAL AR-SJN AR-BUE.

All the variable words treated in this study are presented in a two-dimensional matrix in Figure 3, in which we collocate the city code, presented in 2.2., in the vertical axis; and the linguistic forms in the horizontal one. The forms treated in this study are:

V142: STEERING_WHEEL: dirección; guía; manivela; manubrio; timón; volante

V143: NUMBER_PLATE: chapa; matrícula; patente; placa; tablilla

V145: HEADLIGHT: farolas; faros; faros delanteros; focos; luces; cristal; cristal delantero; luna delantera.

V146: WINDSCREEN: panorámico; parabrisa; parabrisas; vidrio; vidrio delantero; windshield

V147: BONNET: bonete; capó; capota; cofre; tapa; tapa del motor

V152: TRUNK: baúl; cajuela; joroba; maleta; maletera; maletero; portabultos; portaequipajes; portamaletas; valija

V154: FENDER: guardabarro; guardabarros; guardafango; guardalodo; guardapolvo; polvera;

³ This method is recommended for investigations which gather data by means of postal questionnaires and is very useful for later analyses by computer. Wood (1990) points out that “a printed vocabulary checklist will serve computer users in four ways at least: (1) The checklist text can easily be entered into computer storage; (2) The text itself becomes a stored, fixed data base; (3) These personal responses can be quickly matched against the synonyms in the fixed data base; (4) The stored matches can be sorted and displayed for immediate study and for comparison with printed lists of regional vocabularies (...).”

salpicadera

In Figure 3, we observe a relatively high degree of convergence in neighboring countries. If we go on to an analysis of average distance from zero point (a.d.z.p.), however, we are confronted with abrupt zigzag curves (Figure 4 and Figure 5).

2.3. Patternization according to the internal criterion

If we are not so certain in selecting either of the two axes as an external criterion, the usual way to analyze the distribution is to proceed directly to the patternization according to the internal criterion. The result may be seen in Figure 6. In the order of countries we observe the gradual ascending of the value a.d.z.p. in Figure 7. Concretely the order we obtained is:

16_ES-SLM; 03_ES-OVI; 10_ES-GRA; 07_ES-GDL; 13_ES-MUR; 02_ES-SCO; 05_ES-ZAR;
06_ES-BAR; 11_ES-MLG; 01_ES-COR; 04_ES-STD; 12_ES-ALM; 08_ES-MAD; 49_CH-CON;
14_ES-TEN; 17_GE-MAL; 15_ES-PAL; 48_CH-SCH; 09_ES-VAL; 50_CH-TEM; 47_CH-ARI;
41_VE-CAR; 51_PA-ASU; 54_AR-SJN; 55_AR-BUE; 52_UR-MTV; 53_AR-SAL; 42_VE-MER;
27_MX-AGS; 28_MX-MEX; 37_CR-SJO; 44_PE-LIM; 36_CR-LMN; 43_EC-QUI; 18_CU-HAB;
19_CU-SCU; 46_BO-PAZ; 31_EL-SSV; 39_CO-MED; 26_MX-MON; 45_PE-ARE; 38_PN-PAN;
29_MX-MRD; 25_EU-NVO; 30_GU-GUA; 35_NI-MAN; 32_HO-TEG; 33_HO-NAC; 40_CO-BOG;
34_NI-LEO; 24_PR-MAY; 20_RD-STI; 22_PR-SJU; 21_RD-SDO

The order of the linguistic forms is:

V146: WINDSREEN: luna_delantera; V152: TRUNK: portamaletas; V154: FENDER: guardabarros;
V146: WINDSREEN: cristal; V152: TRUNK: portabultos; V142: STEERING_WHEEL: manubrio;
V154: FENDER: tapabarro; V152: TRUNK: portaequipajes; V143: NUMBER_PLATE: matrícula;
V152: TRUNK: maletero; V143: NUMBER_PLATE: patente; V145: HEADLIGHT: faros; V147:
BONNET: capó; V145: HEADLIGHT: faros_delanteros; V146: WINDSREEN: parabrisas; V152:
TRUNK: maleta; V142: STEERING_WHEEL: volante; V154: FENDER: guardapolvo; V143:
NUMBER_PLATE: chapa; V142: STEERING_WHEEL: manivela; V152: TRUNK: joroba; V146:
WINDSREEN: parabrisa; V142: STEERING_WHEEL: dirección; V154: FENDER: guardabarro;
V152: TRUNK: maletera; V154: FENDER: polvera; V154: FENDER: salpicadera; V147: BONNET:
cofre; V152: TRUNK: cajuela; V145: HEADLIGHT: luces; V146: WINDSREEN: cristal_delantero;
V143: NUMBER_PLATE: placa; V154: FENDER: guardafango; V147: BONNET: tapa_del_motor;
V145: HEADLIGHT: focos; V146: WINDSREEN: windshield; V152: TRUNK: baúl; V142:
STEERING_WHEEL: timón; V145: HEADLIGHT: farolas; V147: BONNET: tapa; V147: BONNET:
capota; V146: WINDSREEN: panorámico; V152: TRUNK: valija; V146: WINDSREEN:
vidrio_delantero; V154: FENDER: guardalodo; V142: STEERING_WHEEL: guía; V146:
WINDSREEN: vidrio; V147: BONNET: bonete; V143: NUMBER_PLATE: tablilla.

All these data, rearranged according to the internal criterion, indicate that the places and the linguistic forms presented in the relatively anterior part of the graphs correspond to each other. Naturally the same is true in the middle groups and final groups.

2.4. Patternization according to an external criterion

It is well known in Spanish dialectology that the entire Spanish-speaking world may be divided into the following 6 groups: (1) Spain and Africa; (2) the Caribbean countries; (3) Mexico and Central America; (4) Colombia and Venezuela; (5) the Andean countries and (6) Chile and the La Plata countries⁴. Taking into consideration this grouping and the result of the previous analysis (patternization according to internal criterion), the vertical axis of the cities was rearranged manually as follows:

16_ES-SLM; 03_ES-OVI; 10_ES-GRA; 07_ES-GDL; 13_ES-MUR; 02_ES-SCO; 05_ES-ZAR;
06_ES-BAR; 11_ES-MLG; 01_ES-COR; 04_ES-STD; 12_ES-ALM; 08_ES-MAD; 14_ES-TEN;
15_ES-PAL; 09_ES-VAL; 17_GE-MAL; 49_CH-CON; 48_CH-SCH; 50_CH-TEM; 47_CH-ARI;
41_VE-CAR; 42_VE-MER; 51_PA-ASU; 52_UR-MTV; 54_AR-SJN; 55_AR-BUE; 53_AR-SAL;
27_MX-AGS; 28_MX-MEX; 26_MX-MON; 29_MX-MRD; 37_CR-SJO; 36_CR-LMN; 44_PE-LIM;
45_PE-ARE; 46_BO-PAZ; 43_EC-QUI; 18_CU-HAB; 19_CU-SCU; 39_CO-MED; 40_CO-BOG;
31_EL-SSV; 38_PN-PAN; 34_NI-LEO; 25_EU-NVO; 30_GU-GUA; 35_NI-MAN; 32_HO-TEG;
33_HO-NAC; 24_PR-MAY; 20_RD-STI; 22_PR-SJU; 21_RD-SDO

Determining this axis as an external criterion, the order of linguistic forms was rearranged by the method of patternization (with external criterion = vertical). Now we obtain the Figure 9, in which we have an order of linguistic forms as follows:

V146: WINDSREEN: luna_delantera; V146: WINDSREEN: cristal; V154: FENDER: guardabarros;
V152: TRUNK: portabultos; V152: TRUNK: portamaletas; V142: STEERING_WHEEL: manubrio;
V154: FENDER: tapabarro; V143: NUMBER_PLATE: matrícula; V152: TRUNK: maletero; V152:
TRUNK: portaequipajes; V145: HEADLIGHT: faros; V143: NUMBER_PLATE: patente; V147:
BONNET: capó; V145: HEADLIGHT: faros_delanteros; V146: WINDSREEN: parabrisas; V152:
TRUNK: maleta; V142: STEERING_WHEEL: volante; V154: FENDER: guardapolvo; V154:
FENDER: salpicadera; V154: FENDER: polvera; V147: BONNET: cofre; V143: NUMBER_PLATE:
chapa; V146: WINDSREEN: parabrisa; V154: FENDER: guardabarro; V152: TRUNK: maletera;
V142: STEERING_WHEEL: manivela; V152: TRUNK: joroba; V142: STEERING_WHEEL:
dirección; V152: TRUNK: cajuela; V146: WINDSREEN: panorámico; V145: HEADLIGHT: luces;
V146: WINDSREEN: cristal_delantero; V143: NUMBER_PLATE: placa; V145: HEADLIGHT:

⁴ See, for example, Cahuzac (1980), Canfield (1976), Henriquez Ureña (1921-1931), Honsa (1975), Resnick (1969, 1975) and Ueda (1995b).

focos; V147: BONNET: tapa_del_motor; V154: FENDER: guardafango; V145: HEADLIGHT: farolas; V152: TRUNK: baúl; V142: STEERING_WHEEL: timón; V146: WINDSREEN: windshield; V147: BONNET: tapa; V146: WINDSREEN: vidrio_delantero; V147: BONNET: capota; V154: FENDER: guardalodo; V142: STEERING_WHEEL: guía; V152: TRUNK: valija; V146: WINDSREEN: vidrio; V147: BONNET: bonete

Comparing the Figures of a.d.z.p. of places as well as of linguistic forms (Figures 7, 8 and 10, 11, respectively), we notice no significant difference. So if we respect the geographical order of Figure 9, we are able to affirm that the corresponding group of linguistic forms are in correct order.

3. Conclusion

In the field of geolinguistics it is a common practice to analyze the data not in order to reach a general conclusion but rather to obtain a tentative conclusion which will allow us to reformulate a hypothesis in order to go on later to further studies. I believe that this method of patternization is adequate for such a purpose in the sense that it permits us to reformulate the order of the horizontal and/or vertical variables, according to the requirements of the investigation.

Finally I would like to emphasize the importance of letting the data tell the facts. Thus, the external criterion used in this method would be valid only if we are sure about the selection of that criterion. Furthermore the figures obtained in these analyses should not be interpreted as an absolute ones. The computer and its algorithm are certainly useful, but they cannot offer firm conclusions. The final decision should be taken by researchers who know the concrete linguistic situation, something which is impossible for any machine like a computer.

References

- Cahuzac, Philippe. 1980. "La división del español de América en zonas dialectales: Solución etnolingüística o semántico-dialectal", *Lingüística Española Actual*, 2, pp.385-461.
- Canfield, Lincoln. 1976. "La identificación de dialectos del español americano a base de rasgos distintivos", *Homenaje a Fernando Antonio Martínez*, Bogotá: Instituto Caro y Cuervo.
- Henríquez Ureña, Pedro. 1921, 1930, 1931. "Observaciones sobre el español de América", *Revista de Filología Española*, 7, pp.357-390; 17, pp.277-284; 18, pp.120-148.
- Honsa, Vladimir. 1975. "Clasificación de los dialectos españoles de América y la estructura de los dialectos de Colombia", *Actas del III Congreso de Asociación Lingüística y Filología de la América Latina*, México: Asociación de Lingüística y Enseñanza de Idiomas.

- Kovacci, Ofelia. (1988) "El léxico relacionado con el automóvil en Buenos Aires", *Anuario de Letras*, 26, pp.239-246.
- López Morales, Humberto. "Rasgos Generales" in Manuel Alvar (dir.) *Manual de dialectología hispánica. El Español de América*. Barcelona: Ariel, pp.19-27..
- Moreno Fernández, Francisco. 1993b. "Las áreas dialectales del español americano. Historia de un problema", en Moreno Fernández, F. (ed.) *La división dialectal del español de América*. Alcalá de Henares: Univ. de Alcalá de Henares, pp.10-38.
- Quilis, Antonio. (1982) "Léxico relacionado con el automóvil en Hispanoamérica y en España", *Anuario de Letras*, 22, pp.115-144.
- Resnick, Melvyn C. 1969. "Dialect zones and automatic dialect identification in Latin American Spanish", *Hispania*, 52, pp.553-568.
- Resnick, Melvyn C. 1975. *Phonological variants and dialect identification in Latin American Spanish*. The Hague.
- Takagaki, Toshihiro. 1993. "Hacia la descripción del español contemporáneo de las grandes ciudades del mundo hispánico", *Lingüística Hispánica (Japón)*, 16, pp.65-86.
- Ueda, Hiroto. 1994. "Banco de datos léxico del español. Un proyecto internacional de investigación", *Verba* (Univ. de Santiago de Compostela), 21, pp.397-416.
- . 1995a. "Variación fonológica en el léxico urbano español", *Lingüística Hispánica (Japón)*, 18, pp.105-125.
- . 1995b. "Zonificación del español. Palabras y cosas de la vida urbana", *Lingüística (ALFAL)*, 7, pp.43-86.
- . 1996a. "Análisis demolingüístico del léxico variable español", *Lingüística Hispánica*, vol. 19, pp.63-98.
- . 1996b. "Estudio de la variación léxica del español. Métodos de investigación", *Homenaje al profesor Makoto Hara. Trabajos reunidos con motivo de la jubilación universitaria*. Tokio, pp.341-375.
- . 1996c. "Variación léxica del español urbano. Distribución: Vestuario y equipo (2)", *The Proceedings of the Foreign Language Sections, Graduate School of Arts and Sciences, The University of Tokyo*, vol.1, pp.41-83.
- . 2000. "Distribución de las palabras variables en España y en América. Léxico del transporte", *Estudios de Lingüística Hispánica. Homenaje a María Vaquero*, pp.637-655.
- Wood, Gordon R. (1990) "Using a printed vocabulary checklist". In *Computer Methods in Dialectology*. (ed. by William A. Kretzchmar, Jr. Edgar W. Schneider and Ellen Johnson), *Journal of English Linguistics* 22.1. A special issue, University of Georgia, Athens, Georgia, pp. 1-5.

Addenda. VBA modules for EXCEL

1. Patternization

```
Sub pattern()
```

```
Dim mt As Variant
```

```
Dim sel, i, j, m, n, fh, fv, rt, h(250), v(250) As Integer
```

```
Dim mh(250), mv(500), sm As Double
```

```
sel = Val(InputBox("Rearrangement : vertical =1 / Horizontal =2 / Both =[3]", "Patternization, 3))
```

```
If sel = 0 Or sel > 3 Then End
```

```
mt = ActiveSheet.UsedRange
```

```
m = ActiveSheet.UsedRange.Columns.Count - 1: n = ActiveSheet.UsedRange.Rows.Count - 1
```

```
For j = 1 To m: h(j) = j: Next j
```

```
For i = 1 To n: v(i) = i: Next i
```

```
Do
```

```
For i = 1 To n
```

```
sm = 0: t = 0: mx = 0
```

```
For j = 1 To m
```

```
If mt(v(i) + 1, h(j) + 1) = "+" Then
```

```
sm = sm + j ^ 2: t = t + 1
```

```
If j > mx Then mx = j
```

```
End If
```

```
Next j
```

```
mh(v(i)) = 0
```

```
If t > 0 Then mh(v(i)) = Sqr(sm / t) + (t + mx) / (n * 100)
```

```
Next i
```

```
If sel = 1 Or sel = 3 Then
```

```
fh = 0
```

```
For i = 1 To n - 1: For j = i + 1 To n
```

```
If mh(v(j)) < mh(v(i)) Then fh = 1: k = v(i): v(i) = v(j): v(j) = k
```

```
Next: Next
```

```
End If
```

```
For i = 1 To m
```

```
sm = 0: t = 0: mx = 0
```

```
For j = 1 To n
```

```
If mt(v(j) + 1, h(i) + 1) = "+" Then
```

```
sm = sm + j ^ 2: t = t + 1
```

```
If j > mx Then mx = j
```

```
End If
```

```
Next j
```

```
mv(h(i)) = 0
```

```

    If t > 0 Then mv(h(i)) = Sqr(sm / t) + (t + mx) / (m * 100)
Next i
If sel = 2 Or sel = 3 Then
    fv = 0
    For i = 1 To m - 1: For j = i + 1 To m
        If mv(h(j)) < mv(h(i)) Then fv = 1: k = h(i): h(i) = h(j): h(j) = k
    Next: Next
End If
rt = rt + 1: Application.StatusBar = "Reiteration: " & rt & String(rt, "*")
Loop While fh + fv > 0 And rt <= 50
Application.StatusBar = False
Worksheets.Add after:=ActiveSheet
Cells(1, 1) = "[PATTERN] sel=" & sel
For j = 1 To m: Cells(1, j + 1) = mt(1, h(j) + 1): Cells(1, j + 1).ColumnWidth = 3: Next j
For i = 1 To n
    Cells(i + 1, 1) = mt(v(i) + 1, 1)
    For j = 1 To m: Cells(i + 1, j + 1) = mt(v(i) + 1, h(j) + 1): Next j
Next i
End Sub

```

A.2. Axes

```

Sub axes()
Dim mt As Variant
Dim sel, i, j, m, n, th, tv As Integer
Dim mh(250), mv(2000), sh, sv As Double
mt = ActiveSheet.UsedRange
m = ActiveSheet.UsedRange.Columns.Count - 1: n = ActiveSheet.UsedRange.Rows.Count - 1
For i = 1 To n
    For j = 1 To m
        If mt(i + 1, j + 1) = "+" Then sh = sh + j ^ 2: th = th + 1
    Next j
    If th > 0 Then mh(i) = Sqr(sh / th)
Next i
For i = 1 To m
    For j = 1 To n
        If mt(j + 1, i + 1) = "+" Then sv = sv + j ^ 2: tv = tv + 1
    Next j
    If tv > 0 Then mv(i) = Sqr(sv / tv)
Next i

```

```
Worksheets.Add after:=ActiveSheet
Cells(1, 1) = "VER.": Cells(1, 2) = "Dist. (v)"
For i = 1 To n
    Cells(i + 1, 1) = mt(i + 1, 1)
    Cells(i + 1, 2) = mh(i): Cells(i + 1, 2).NumberFormatLocal = "0.000"
Next
Cells(1, 3) = "HOR.": Cells(1, 4) = "Dist. (h)"
For i = 1 To m
    Cells(i + 1, 3) = mt(1, i + 1)
    Cells(i + 1, 4) = mv(i): Cells(i + 1, 4).NumberFormatLocal = "0.000"
Next
End Sub
```

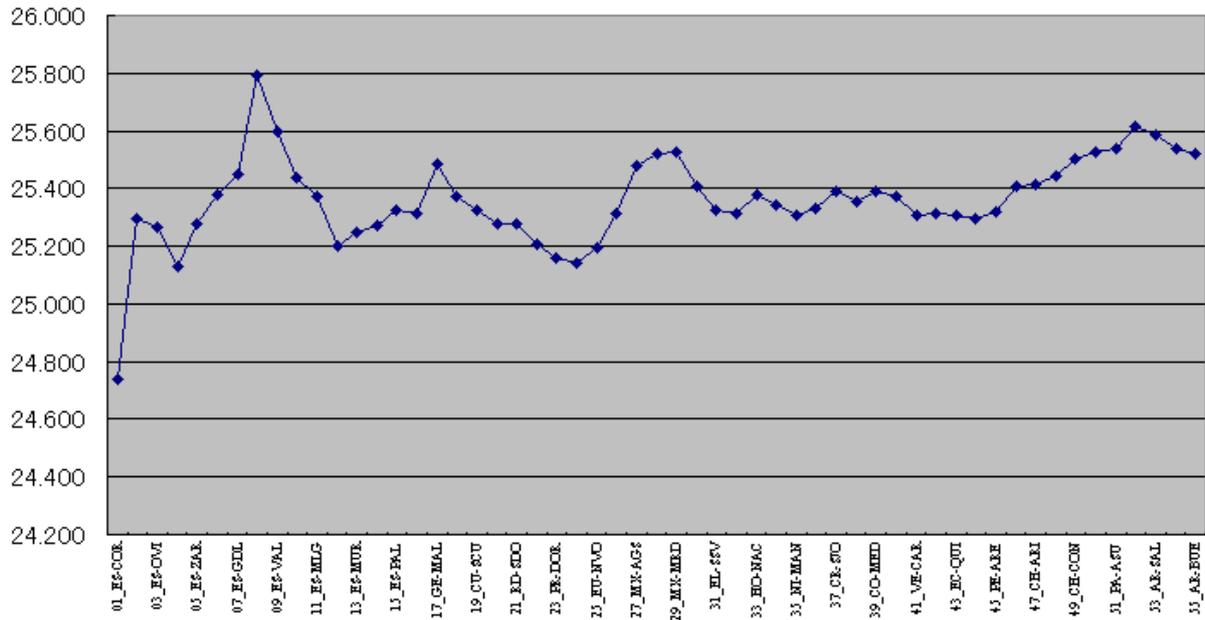



Figure 4: A.d.z.p. of cities. Before patternization

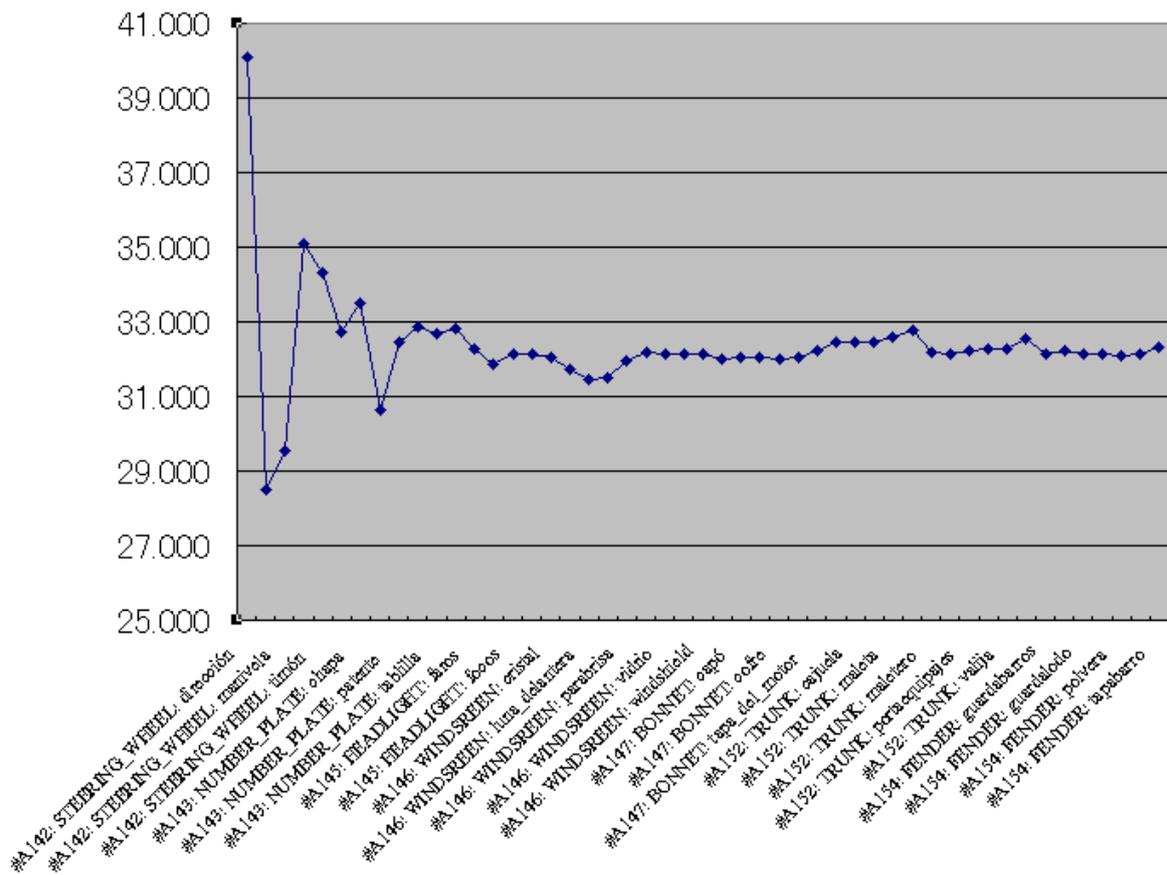


Figure 5: A.d.z.p. of linguistic forms. Before patternization

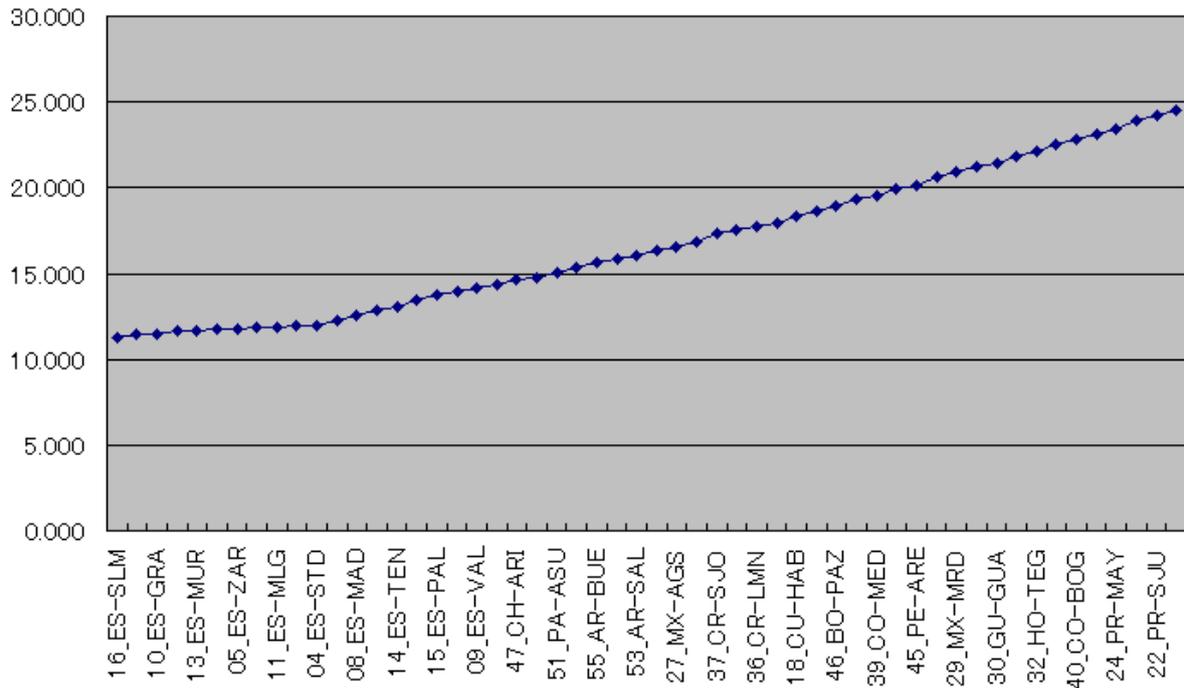


Figure 7: A.d.z.p. of cities. Internal criterion

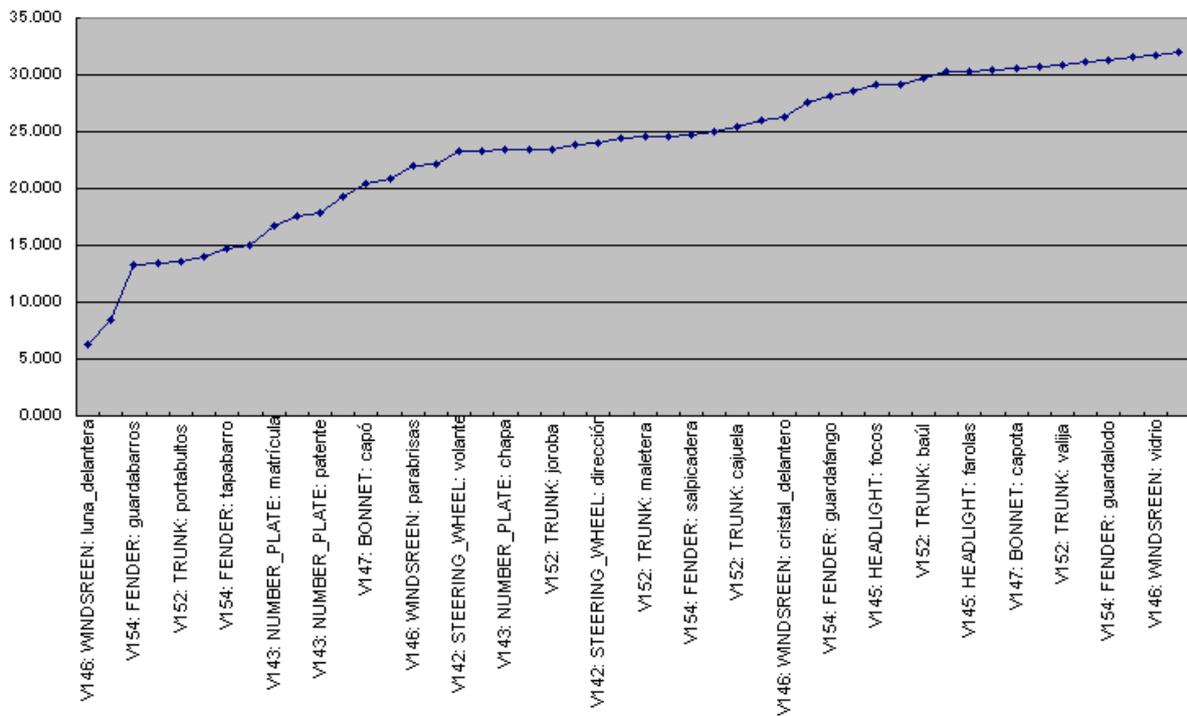


Figure-8: A.d.z.p. of linguistic forms. Internal criterion

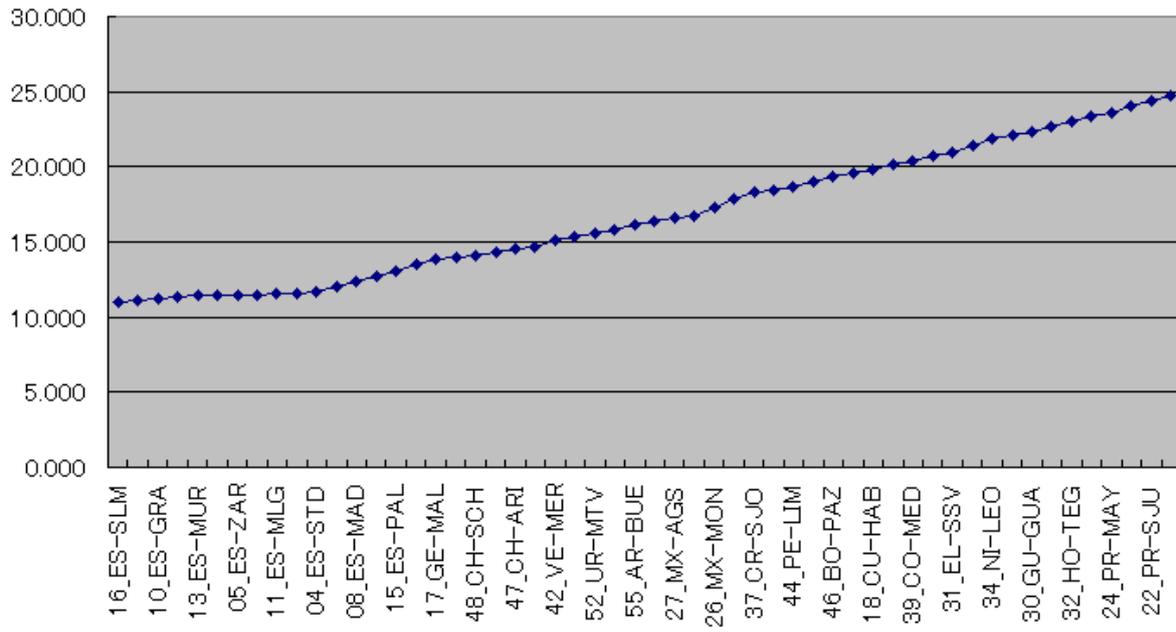


Figure-10: A.d.z.p. of cities. External criterion. Vertical

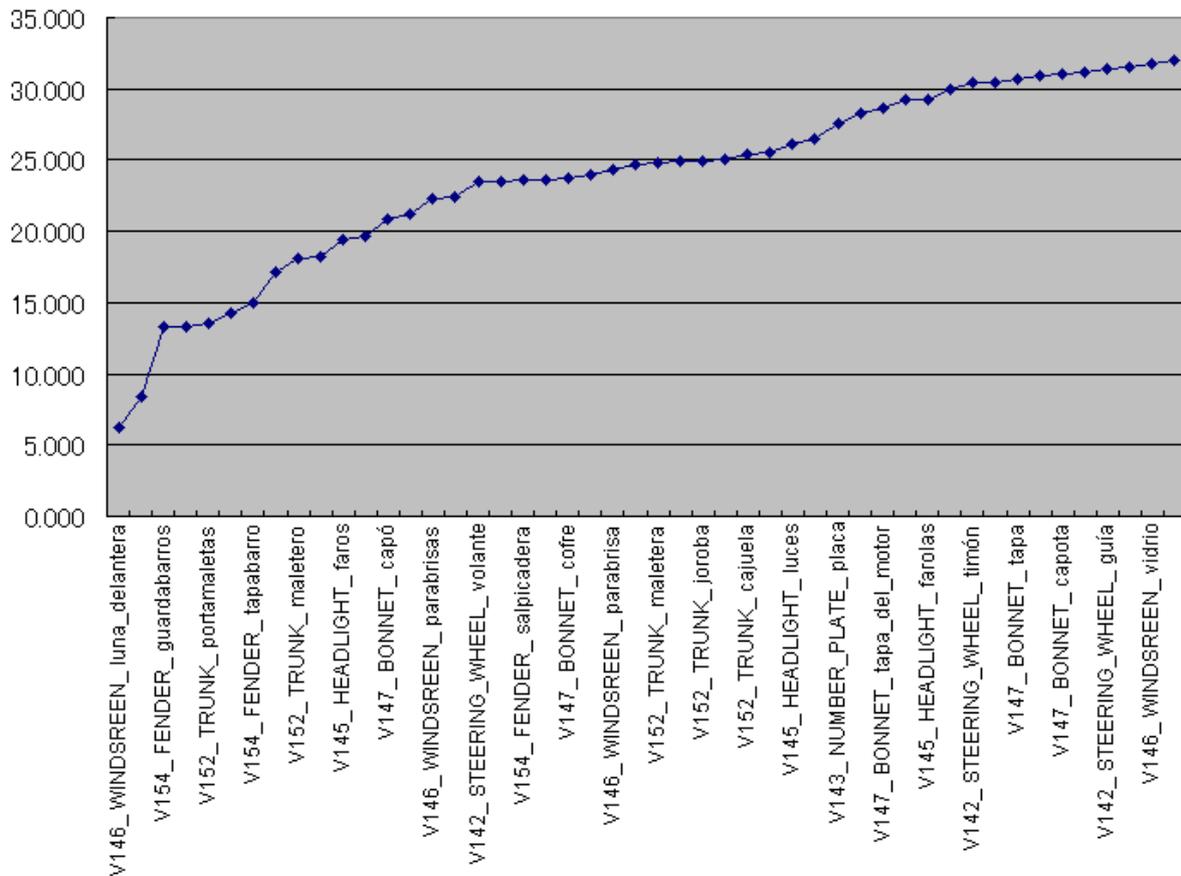


Figure-11: A.d.z.p. of linguistic forms. External criterion. Vertical