

Investigación de la variación léxica española


Método de la zonificación multivariante y visualización cartográfica

Hiroto Ueda (Universidad de Tokio)

§1. Desde 1993 hemos venido investigando la variación léxica del español moderno en España, Guinea Ecuatorial e Hispanoamérica *in situ* y por correspondencia. Recurrimos a la ayuda de colaboración prestada por los investigadores locales, enviándoles cuatro cuestionarios con dibujos y explicación donde los encuestados de dos sexos y dos generaciones seleccionarían las formas que utilizan usualmente. Hemos subido el resultado de aproximadamente 2000 ítems en la página web (Ueda 1993-).

Explicaciones

1. Subraye la(s) palabra(s) o expresión(es) si usted mismo la(s) utiliza.
Ejemplo:



[A001] JACKET: Prenda de vestir masculina, que forma con el chaleco y los pantalones el traje completo. No es de paño con botones dorados.

(1)americana, (2)capa, (3)chaleco, (4)chaqueta, (5)gabán, (6)leva, (7)paletón, (8)saco, (9)saco de terno, (10)saco de traje, (11)traje, (12)vestón.

&Otro(s) _____, #No se me ocurre.

\$Comentario:

Fig. 1. Cuestionario, parte inicial

En los métodos de la geografía lingüística tradicional de Europa y de la dialectología española, se practicaba la encuesta directa al informante anciano en pueblos pequeños sobre el léxico inmenso residual de la localidad y el investigador transcribía la respuesta en un cuaderno, que posteriormente se copiaba en los mapas de atlas lingüísticos. Nuestro método se difiere del método tradicional en los siguientes puntos: grandes ciudades y pequeños pueblos, múltiples encuestados y único encuestado anciano, encuesta por correspondencia y encuesta directa, encuesta continua y encuesta única, procesamiento informático y trabajo manual. Los dos métodos diferentes en condiciones, objetivos y métodos no son excluyentes

sino complementarios.

Hemos elaborado un sistema de cuantificación y visualización manejable en la página web (Ueda y Moreno Sandoval 2017-). El programa produce la tabla bidimensional de forma y lugar junto con el mapa de distribución, donde se aplican métodos multivariantes para obtener la zonificación geográfica continua que combina las zonas en el esquema y puntos geográficos en el mapa. Para la visualización hemos adoptado el método de red y casco convexo.

§2. En esta ocasión, tomando por ejemplo la variación del ítem número 1 (A001) ‘jacket (for men)’, observaremos su distribución geográfica en los países hispanohablantes y presentaremos nuestro método de zonificación y visualización. Utilizaremos los datos ofrecidos por investigadores locales (datos de tipo 1/0).

En la siguiente tabla, se manifiesta la distribución de cada forma en los países correspondientes:

País	<i>americana</i>	<i>chaqueta</i>	<i>gabán</i>	<i>saco</i>	<i>saco de terno</i>	<i>saco de traje</i>	<i>vestón</i>
1.ES		1					
2.GE	1	1					
3.CU		1		1		1	
4.RD		1		1			
5.PR		1	1	1			
6.MX				1		1	
7.GU				1			
9.HO				1			
9.EL				1			
10.NI				1			
11.CR		1		1			
12.PN				1			
13.CO		1		1			
14.VE		1		1			
15.EC		1			1		
16.PE				1	1		
17.BO				1	1		
18.CH		1					1
19.PA				1			
20.UR				1		1	

Tabla 1. Dstribución de formas en países

(1.ES: España, 2.GE: Guinea Ecuatorial, 3.CU: Cuba, 4.RD: República Dominicana, 5.PR: Puerto Rico, 6.MX: México, 7.GU: Guatemala, 8.HO: Honduras, 9.EL: El Salvador, 10.NI: Nicaragua, 11.CR: Costa Rica, 12.PN: Panamá, 13.CO: Colombia, 14.VE: Venezuela, 15.EC: Ecuador, 16.PE: Perú, 17.BO: Bolivia, 18.CH: Chile, 19.PA: Paraguay, 20.UR: Uruguay, 21.AR: Argentina)

La siguiente tabla muestra la distribución diagonalizada de las formas frecuentes (*chaqueta, saco, saco de traje, saco de terno*) y países correspondientes, donde hemos utilizado la distancia Minkowski de elevación 3 para obtener los coeficientes verticales (Xn) y los horizontales (Yp):

Dst.	1.s. de terno	2.chaqueta	3.saco	4.s. de traje	Row	Xn
1.EC	1	1			EC	1.651
2.ES.GE.CH		1			ES.GE.CH	2.000
3.PE.BO	1		1		PE.BO	2.410
4.PR.RD.CR.CO.VE		1	1		PR.RD.CR.CO.VE	2.596
5.GU.HO.EL.NI.PN.PA			1		GU.HO.EL.NI.PN.PA	3.000
6.CU		1	1	1	CU	3.208
7.MX.UR.AR			1	1	MX.UT.AR	3.570

Column	s. de terno	chaqueta	saco	s. de traje
Yp	2.410	4.165	5.372	6.538

Tabla 2. Dstribución diagonalizada

La ditribución diagonalizada que indica la flecha oblicua explica la correlación continua que hay entre las formas y países. Para obtener la zonificación separada, en primer lugar, preparamos la matriz simétrica de distancia (Xn) por medio del valor absoluto de la diferencia. Por ejemplo, la distancia entre EC y [ES.GE.CH] es: $|1.651 - 2.000| = .349$.

Dis.	1.EC	2.ES.GE.CH	3.PE.BO	4.PR.RD.CR.CO.VE	5.GU.HO.EL.NI.PN.PA	6.CU	7.MX.UR.AR
1.EC	.000	.349	.759	.945	1.349	1.557	1.919
2.ES.GE.CH	.349	.000	.410	.596	1.000	1.208	1.570
3.PE.BO	.759	.410	.000	.186	.590	.797	1.160
4.PR.RD.CR.CO.VE	.945	.596	.186	.000	.404	.611	.974

5.GU.HO.EL.NI.PN.PA	1.349	1.000	.590	.404	.000	.208	.570
6.CU	1.557	1.208	.797	.611	.208	.000	.362
7.MX.UT.AR	1.919	1.570	1.160	.974	.570	.362	.000

Tabla 3. Matriz simétrica de distancia

A partir de la misma matriz, por medio del análisis de conglomerados, se obtiene el siguiente dendrograma:

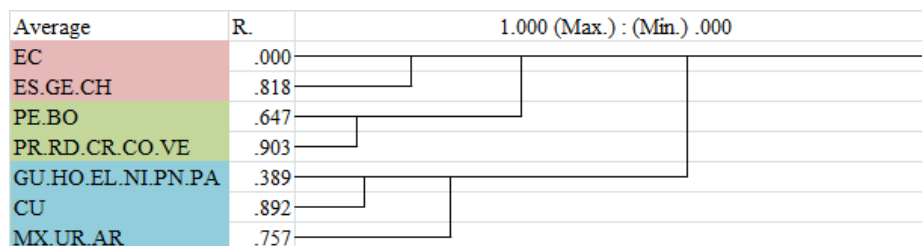


Fig. 2. Dendrograma de países

De la misma manera, realizamos el análisis de conglomerados de formas léxicas (Yp):

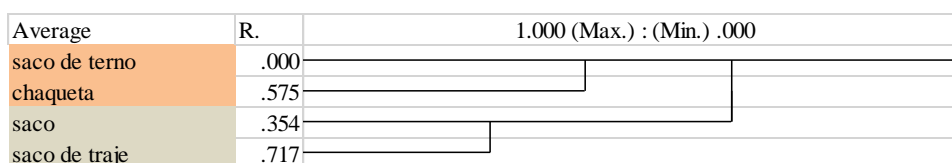


Fig. 3. Dendrograma de formas

Reflejamos el resultado del análisis de conglomerados en la tabla diagonalizada anterior:

Dst.a	saco de terno	chaqueta	saco	saco de traje
EC	1	1		
ES.GE.CH		1		
PE.BO	1		1	
PR.RD.CR.CO.VE		1	1	
GU.HO.EL.NI.PN.PA			1	
CU		1	1	1
MX.UT.AR			1	1

Fig. 4. Distribución diagonalizada conglomerada

§3. A partir de esta tabla, vamos a considerar la distribución de *saco* : *chaqueta*.

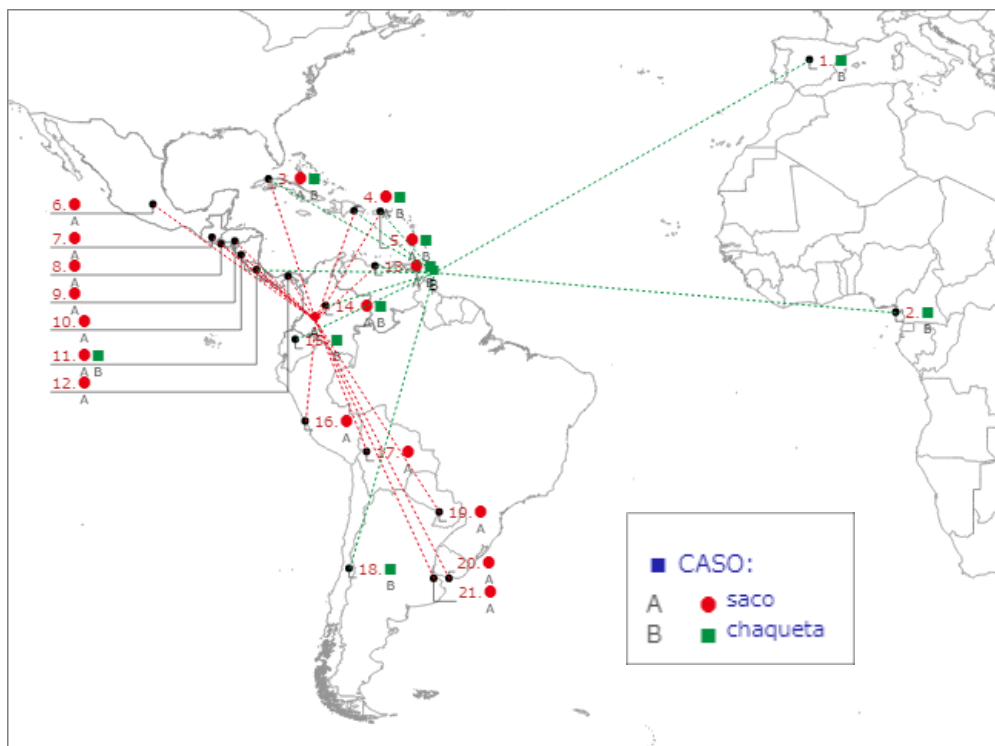


Fig. 5. Distribución geográfica de *saco* y *chaqueta*

Al observar este mapa, comprobamos la distribución de A: *saco* en México y América Central, que corresponde a la zona occidental, la de B: *chaqueta* en España, Guinea Ecuatorial y Chile, que constituyen la zona oriental, y la de A: *saco* + B: *chaqueta* en los países de Caribe, es decir, zona central. De modo que comprobamos la continuidad geográfica en forma de A : A+B : B. En general, el cambio cronológico normal de la misma distribución es $A \rightarrow A+B > B$ o $B \rightarrow A+B \rightarrow A$, mientras que es difícil de pensar en $*A \rightarrow B \rightarrow A+B$ ($*B \rightarrow A, A+B$), $*A+B \rightarrow A \rightarrow B$ ($*A+B \rightarrow B \rightarrow A$), $*A+B \rightarrow A=B$ (A, B simultáneos).

Según Corominas y Pascual (1983, s.v. *saco*, *chaqueta*), *saco* de origen latino (<lat. SACCUS) se utilizaba como nombre de prenda de vestir en el siglo XIV y a principios del siglo XIX se introdujo la forma francesa *jaquette* en forma de *chaqueta*. Actualmente en España, Guinea Ecuatorial y Chile, la palabra *saco* con el significado de ‘chaqueta’ no se utiliza, de modo que se ha comprobado el cambio histórico de A: *saco* > A+B: *saco+chaqueta* > B: *chaqueta*.

§4. En la tabla diagonalizada anterior, las distribuciones de *saco de terno* (Ecuador, Perú y Bolivia) y *saco de traje* (Cuba, México, Uruguay y Argentina) no están sobrepuestas, lo que se visualiza de la siguiente manera.

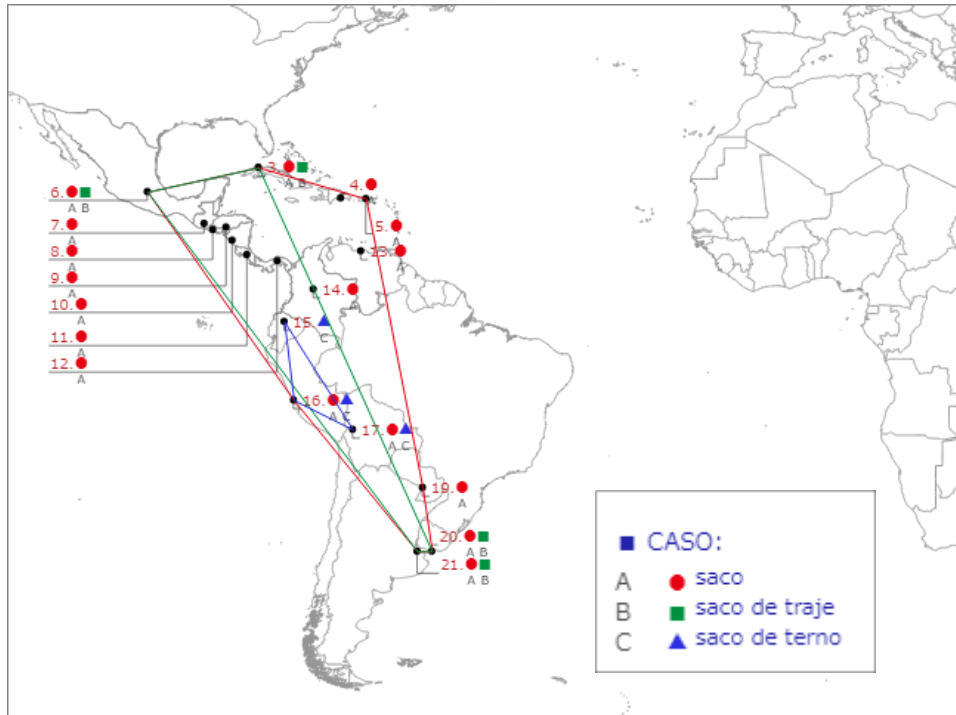


Fig. 6. Distribución geográfica de *saco*, *saco de traje*, *saco de terno*

Se trata del gráfico de casco convexo que encubre el conjunto de los puntos reactivos. La desventaja de este método está en que puede incluir los puntos negativos. No obstante, aquí es conveniente por construir relaciones envolventes: A: *saco* \supset B: *saco de traje* \supset C: *saco de terno*. Generalmente cuando estamos ante la distribución envolvente A-B-A, suponemos el cambio cronológico de A \rightarrow B a exclusión de A = B (simultáneo) y B \rightarrow A. A pesar de que no encontramos información al respecto en los estudios anteriores ni en el corpus históricos, podemos suponer el doble cambio de A: *saco* \rightarrow B: *saco de traje* \rightarrow C: *saco de terno*.

Referencia citada:

Corominas, J. / Pascual, J. A. (1983) *Diccionario crítico etimológico castellano e hispánico*. Madrid, Gredos.

RAE. CORDE. Real Academia Española. *Corpus Diacrónico del Español (CORDE)*
<http://corpus.rae.es/> [2019/9/29]

Ueda, H. (1993-) VARILEX, Variación léxica del español en el Mundo
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/varilex/index.html>

Ueda, H. y Moreno Sandoval, A. (2017-) LYNEAL, Letras y números en análisis lingüísticos [2019/9/29]

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/> [2019/9/29]

Ueda, H. y Moreno Fernandez, Francisco. (2016) VARILEX-R.

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/varilex-r/> [2019/9/29]