

Cómo usar el sistema LYNEAL

Letras y Números en Análisis Lingüísticos

Redacción: Hiroto Ueda

Revisión: Inmaculada Martínez

Revisión: Antonio Moreno Sandoval

Revisión: Leyre Martín Aizpuru

ver. 2024-02-24

EXPLICACIÓN

En el sistema LYNEAL se están reuniendo varios proyectos de corpus digitales.

Invitamos a todos los investigadores interesados a formar un grupo de corpus en esta plataforma general, sin hacer distinción de lenguas, de magnitud de materiales, de modo de estructura, etc.

Tenemos dos sitios en Madrid y Tokio para instalar los materiales reunidos y ofrecer la versión más nueva de la herramienta.

De esta manera, podemos presentar y citar los datos con los que se han hecho nuestros estudios, lo que garantiza la fiabilidad y replicabilidad tanto de los datos como de los métodos.

* Para reiniciar el sistema utilice el atajo [Ctrl] + [F5].

Nuestro contacto es:

Hiroto Ueda. Universidad de Tokio: uedahiroto[ARROBA]com.home.ne.jp

Antonio Moreno Sandoval. Universidad Autónoma de Madrid: antonio.msandoval[ARROBA]uam.es

ver. 2020-8-23

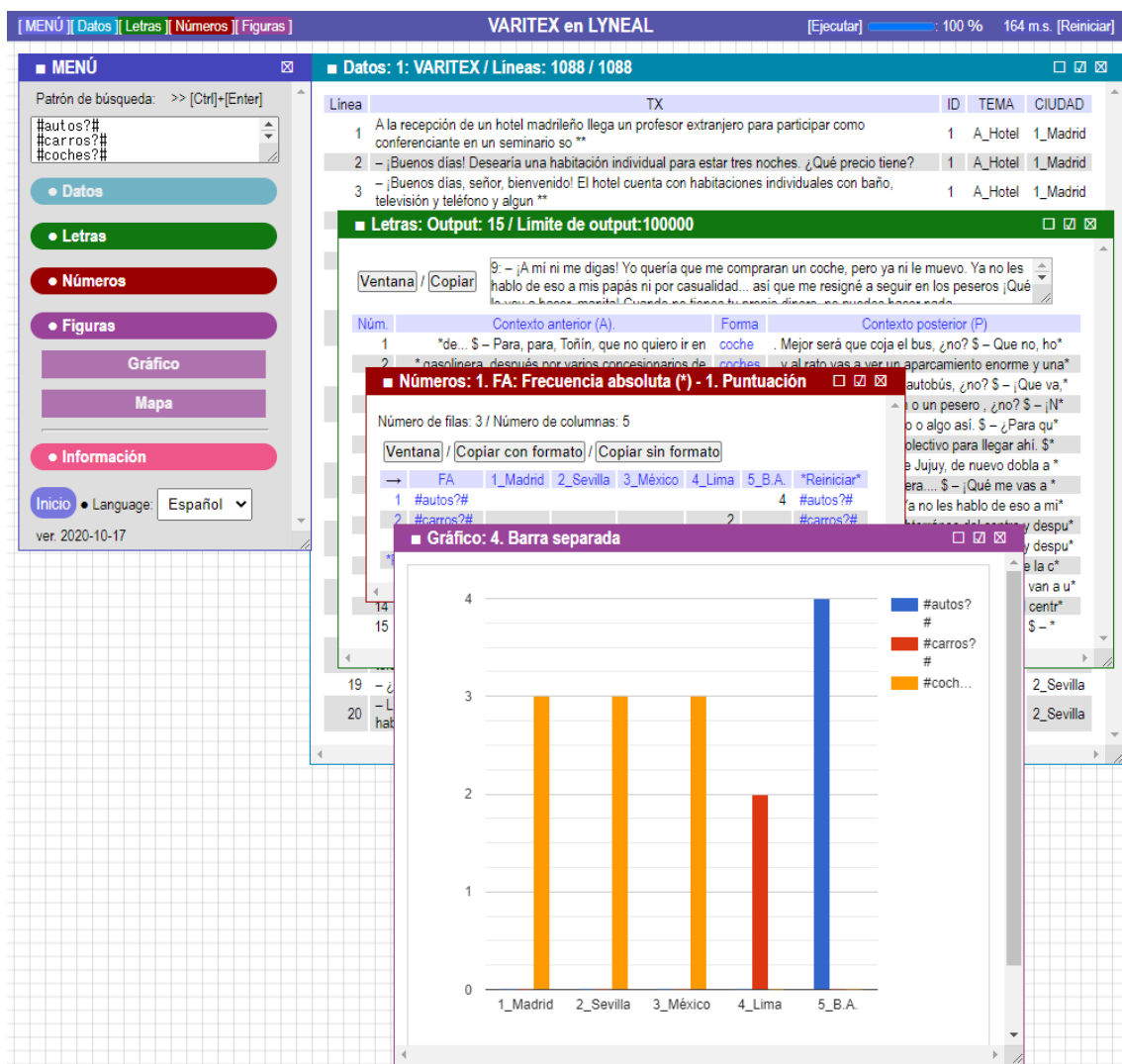
<http://shimoda.lllf.uam.es/ueda/lyneal/>

<https://h-ueda.sakura.ne.jp/lyneal/>

LYNEAL (*Letras y Números en Análisis Lingüísticos*) es un sistema de análisis de textos que facilita los procesamientos de datos textuales tanto de los archivos almacenados en el servidor como los propios del usuario. En la elaboración del programa, hemos buscado la comodidad de uso y la rapidez de distintas funciones al mismo tiempo. El método se ordena en forma lineal siguiendo la dirección de:

● Datos → ● Letras → ● Números → ● Figuras

El manejo es interactivo, fácil e intuitivo, de manera que el usuario puede comprobar las funciones inmediatamente al hacer clic en los botones correspondientes. Solo hacen falta algunas instrucciones sobre los patrones con los que buscamos las formas lingüísticas y sobre las estructuras del archivo, con las que construimos y filtramos los parámetros y elaboramos las tablas de frecuencias.



ÍNDICE

1. Datos	13
1.1. Archivo	13
1.2. Atributos y filtros	14
1.2.1. Atributos	14
1.2.2. Filtros	14
1.2.3. Agrupamiento	15
1.3. Datos de usuario	15
1.3.1. Caja de texto.....	15
1.3.2. Archivo de usuario	17
1.4. Datos numéricos	18
2. Letras	20
2.1. Patrón de búsqueda	20
2.2. Opción	21
2.2.1. Elementos en output	21
2.2.2. Procesamiento.....	23
2.2.3. Límite de output.....	26
2.3. Contexto	26
2.4. Casos y atributos	28
2.5. Dos variables.....	29
2.5.1. Separar variables.....	30
2.5.2. Unir variables	30
2.5.3. Variable por variable (total).....	31
2.5.4. Variable por variable (ratio: $X/(X+Y)$)	31
2.5.5. Variable por variable (contraste: $(X-Y)/(X+Y)$)	31
2.6. Intervalo numérico.....	32
3. Números	33

3.1. Método.....	33
3.1.1. Matriz	33
3.1.2. Análisis	41
3.2. Opción	50
4. Figuras	54
4.1. Gráfico.....	54
4.2. Mapa.....	61
4.2.1. Procesos generales	61
4.2.2. Tipo de mapa	66
4.2.3. Marca.....	66
4.2.4. Línea.....	67
4.2.5. Posición	72
4.2.6. Longitud / Latitud	74
4.2.7. Ancho y alto del mapa	74
5. Información	75
6. Apéndice-1: Patrones.....	76
6.1. Patrones simples	76
6.1.1. Almohadilla y porcentaje : #, %.....	76
6.1.2. Signo de et: &.....	78
6.1.3. Signo de arroba: @.....	78
6.1.4. Signo de punto:	78
6.1.5. Signo de raya inversa más la letra "d": \d	79
6.1.6. Corchetes: [...].....	79
6.1.7. Paréntesis y raya vertical: (... ...)	79
6.2. Patrones especiales	79
6.2.1. Cantidad: ?, *, +, {...}	80
6.2.2. Raya reversa: \.....	80

6.2.3. Contexto: ...<<...>>.....	80
6.2.4. Reproducción: (...) \1	81
6.2.5. Redefinición: ...==.....	81
6.2.6. Reemplazo: ...=>.....	81
6.2.7. Exclusión {-}.....	82
6.2.8. Escape {E}	82
6.2.9. Patrones propios de español	83
7. Apéndice-2: Tratamiento de texto	84
7.1. Combinar / dividir texto	84
7.1.1. Combinar texto	84
7.1.2. Dividir texto	85
7.2. Comparar palabras	85
7.3. Lematizar y anotar texto.....	86
7.4. Mapa lingüístico	87
7.5. Palabras españolas de poca frecuencia.....	88
7.6. Expresión regular.....	88
7.7. Repetir celda	88
7.8. Correspondencia léxica	89
7.9. Comparación de frecuencia de palabras	90

El sistema LYNEAL está instalado en dos sitios, en Madrid (Universidad Autónoma de Madrid) y en el servidor Sakura:

<http://shimoda.llf.uam.es/ueda/lyneal/>

<https://h-ueda.sakura.ne.jp/lyneal/>

Son iguales de contenido. Al dar acceso a cualquiera de los dos, aparece la primera página:

LYNEAL: Letras y Números en Análisis Lingüísticos

Language: Español

PROYECTO

- * ALDICAM.m
- * Biblia Medieval
- * Bolivia (HLCB, ILSC)
- * CICA
- * CODCAR-NORM
- * CODEA.m
- * CORHEN.m
- * COSER.m
- * DCVB
- * Documentos Bolivianos.m
- * GREC-UNED
- * LEMI
- * LLI-UAM
- * PRESEEA.m
- * PROGRAMES
- * TESIS
- * VARIGRAMA-ES.m
- * VARIGRAMA-LA.m
- * VARILEX.m
- * VARITEX.m
- * VARIOS
- * レクシコ

(.m: + mapa lingüístico)

EXPLICACIÓN

En el sistema LYNEAL se están reuniendo varios proyectos de corpus digitales.

Invitamos a todos los investigadores interesados para formar un grupo de corpus en esta plataforma general, sin hacer distinción de lenguas, de magnitud de materiales, de modo de estructura, etc.

Tenemos dos sitios en Madrid y Tokio para instalar los materiales reunidos y ofrecer la versión más nueva de la herramienta.

De esta manera, podemos presentar y citar nuestros datos con los que se han hecho nuestros estudios, lo que garantiza la fiabilidad y replicabilidad tanto de los datos como de los métodos.

* Para reiniciar el sistema utilice el atajo [Ctrl] + [F5].

Nuestro contacto es:

Hiroto Ueda. Universidad de Tokio: uedahiroto@jcom.home.ne.jp

Antonio Moreno Sandoval. Universidad Autónoma de Madrid: antonio.msandoval@uam.es

ver. 2020-3-11

En la parte izquierda se encuentra el índice de proyectos de distintos corpus que colaboran con nosotros. Está aumentando progresivamente el número de grupos internacionales de investigación en el proyecto. En la primera página hemos anotado nuestros contactos en Madrid y Tokio para recibir comentarios, sugerencias, preguntas y propuestas de colaboración. A continuación, reproducimos la información que ofrecemos en la Explicación:

En el sistema LYNEAL se están reuniendo varios proyectos de corpus

digitales.

Invitamos a todos los investigadores interesados para formar un grupo de corpus en esta plataforma general, sin hacer distinción de lenguas, de magnitud de materiales, de modo de estructura, etc.

Tenemos dos sitios en Madrid y Tokio para instalar los materiales reunidos y ofrecer la versión más nueva de la herramienta.

De esta manera, podemos presentar y citar nuestros datos con los que se han hecho nuestros estudios, lo que garantiza la fiabilidad y replicabilidad tanto de los datos como de los métodos.

* Para reiniciar el sistema utilice el atajo [Ctrl] + [F5].

Nuestro contacto es:

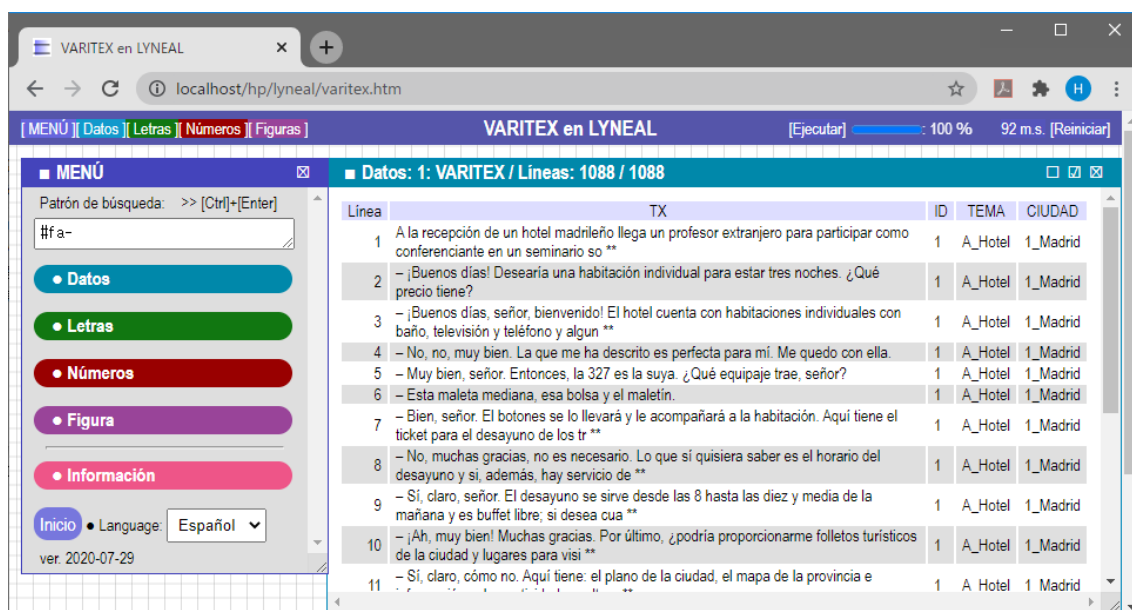
- Hiroto Ueda. Universidad de Tokio:

uedahiroto[at]jcom.home.ne.jp

- Antonio Moreno Sandoval. Universidad Autónoma de Madrid:

antonio.msandoval[at]uam.es

Al hacer clic en uno de los proyectos que figuran en la lista de [Corpus] llegamos directamente a la página de herramientas. Para familiarizarnos con la interfaz del sistema, vamos a utilizar, a modo de ejemplo, el sitio de VARITEX (Variación Textual de Español). Se trata de textos paralelos de distintas ciudades hispanohablantes: Madrid, Sevilla, Ciudad de México, Lima y Buenos Aires:

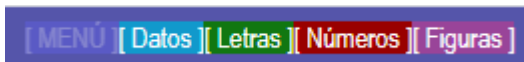


La página está dividida en tres partes: encabezamiento en la parte superior, [Menú] en la parte inferior izquierda y [Output] en la parte inferior derecha.

En el encabezamiento, a la izquierda se encuentran cinco botones: [MENÚ], [Datos], [Letras], [Números], [Figuras].

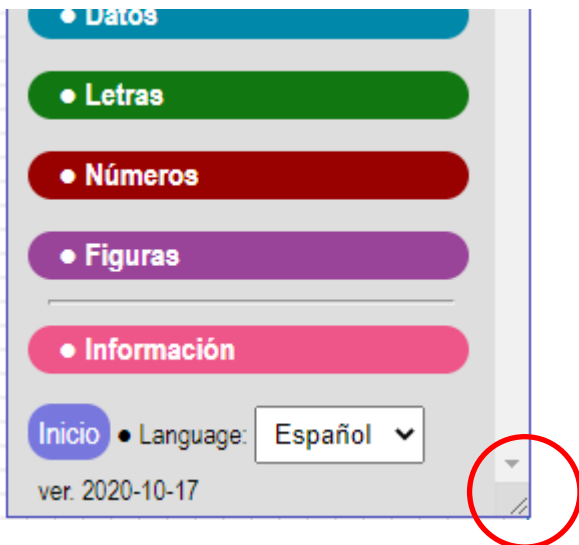


Al hacer clic sobre [MENÚ] aparece y desaparece alternativamente el [* MENÚ] en la parte inferior izquierda. Al hacer clic sobre una de las opciones [Datos], [Letras], [Números], [Figuras], aparecen y desaparecen alternativamente las divisiones correspondientes tanto en [■ MENÚ], en la parte inferior izquierda como en la parte inferior derecha. Al desaparecer la división, el botón se hace borroso, con media opacidad, por ejemplo, en el botón de [MENÚ].

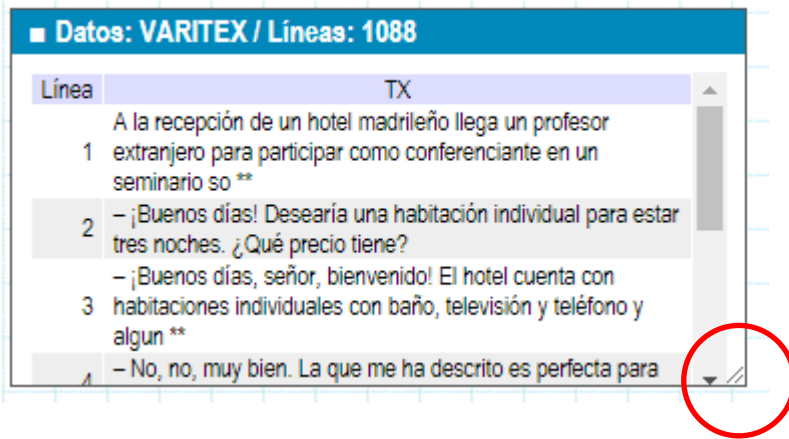


Al sobreponer el ratón en uno de estos botones, aparece la división correspondiente en primer plano y al mover el ratón del botón, vuelve al plano anterior. Al hacerle un clic, la división aparece siempre en el primer plano.

Sin ejecutar, no aparece más que el [MENÚ] y [Datos]. El tamaño del [MENÚ] es graduable verticalmente con el tirador que se encuentra en la parte inferior derecha.

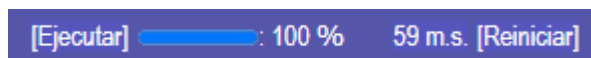


El tamaño de la división de [Datos] es graduable tanto vertical como horizontalmente con el mismo tirador:



La misma operación es aplicable a otras divisiones, [Letras], [Números], [Gráficos], que aparecen a la hora de la ejecución.

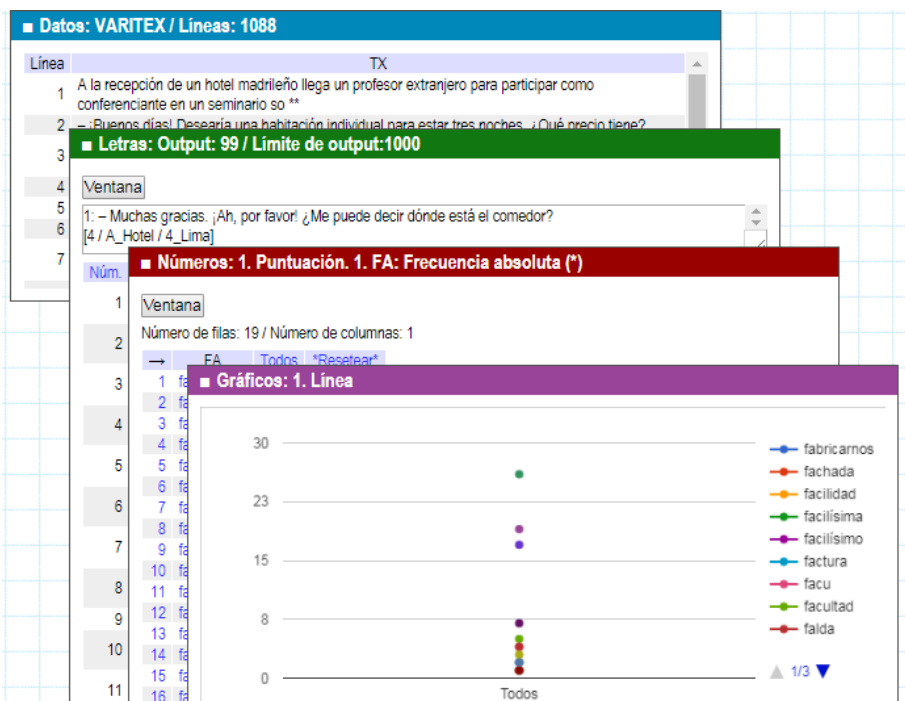
En el encabezamiento, a la derecha del título de la página, se encuentran dos botones: [Ejecutar], al principio, y [Reiniciar] al final:



En lugar de estos botones, podemos utilizar las teclas de atajo: para [Ejecutar], [Ctrl]+[Enter]; para [Reiniciar], [Ctrl]+R.

Entre [Ejecutar] y [Reiniciar] se encuentran la barra de progreso, porcentaje (%) de progreso y tiempo de ejecución en milisegundo (m.s.).

Al hacer un clic sobre [Ejecutar] (o [Ctrl]+[Enter]), aparecen tres divisiones más, en total cuatro, en forma sobrepuesta:



Al sobreponer el ratón en alguna parte de la barra de título, aparece toda la división en el primer plano, por ejemplo, la división de [Letras]:

■ Datos: VARITEX / Líneas: 1088

Línea TX

1 A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario so **

2 - ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio tiene?

3 ■ Letras: Output: 99 / Limite de output:1000

4 Ventana

5 1: - Muchas gracias. ¡Ah, por favor! ¿Me puede decir dónde está el comedor?

6 [4 / A_Hotel / 4_Lima]

7

Núm.	Contexto anterior (A)	Forma	Contexto posterior (P)
1	- Muchas gracias. ¡Ah, por	favor	! ¿Me puede decir dónde está el comedor?
2	laro, tío! Allí tienes de todo: por ejemplo, hay	faldas	guapas para tu mujer o chándal para los crios. ¡T
3	ianistas, que se conoce enseguida porque tiene la	fachada	de ladrillo. Y de ahí al "Campeón", nada, dos o t
4	vios para la limpieza... en fin, todo lo que haga	falta	en la casa para la semana. Así que cuando me sobr
5	laro, mano! Allí tienes de todo: por ejemplo, hay	faldas	padrísimas pa' tu mujer o ropa deportiva
6	- Y... compadre ¿Cómo estás? Oye, un	favor	¿Sabes dónde hay un supermercado cerca?
7	rdoname que te moleste, pero ¿me podés decir, por	favor	, dónde queda el Supermercado "El Disco"?
8	me puedo quejar. Bueno, la semana pasada estuve	fatal	, con un gripazo impresionante. Me dieron la baja
9	- Ya, pero hoy, con la	falta	de trabajo que hay...
10	s así! Sería una gozada venir todos los días a la	facu	...
11	me chifla este curro, porque, así, el venir a la	facu	se me hace más llevadero. A mi madre no le hace n
12	clases son la mar de entretenidas y	falta	emponillar tanto

19 fa

- blue arrow fabricarnos
- orange arrow fachada
- yellow arrow facilidad
- green arrow facilisima
- purple arrow facilisimo
- cyan arrow factura
- pink arrow facu
- light green arrow facultad
- red arrow falda

1/3 ▼

Al mover el ratón fuera de la barra de título, vuelve al sitio original. Para moverla definitivamente al primer plano, hay que darle un clic:

■ Datos: VARITEX / Líneas: 1088

Línea TX

1 A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario so **

2 - ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio tiene?

3 ■ Letras: Output: 99 / Limite de output:1000

4 Ventana

5 1: - Muchas gracias. ¡Ah, por favor! ¿Me puede decir dónde está el comedor?

6 [4 / A_Hotel / 4_Lima]

7

Núm.	Contexto anterior (A).	Forma	Contexto posterior (P)
1	- Muchas gracias. ¡Ah, por	favor	! ¿Me puede decir dónde está el comedor?
2	laro, tío ! Allí tienes de todo: por ejemplo, hay	faldas	guapas para tu mujer o chándal para los crios. ¡T
3	ianistas, que se conoce enseguida porque tiene la	fachada	de ladrillo. Y de ahí al "Campeón", nada, dos o t
4	víos para la limpieza... en fin, todo lo que haga	falta	en la casa para la semana. Así que cuando me sobr
5	laro, mano! Allí tienes de todo: por ejemplo, hay	faldas	padrísimas pa' tu mujer o ropa deportiva 2 sudade
6	- Y... compadre ¿Cómo estás? Oye, un	favor	¿Sabes dónde hay un supermercado cerca?
7	rdoname que te moleste, pero ¿me podés decir, por	favor	, dónde queda el Supermercado "El Disco"?
8	me puedo quejar. Bueno , la semana pasada estuve	fatal	, con un gripazo impresionante. Me dieron la baja
9	- Ya, pero hoy, con la	falta	de trabajo que hay...
10	s así! Sería una gozada venir todos los días a la	facu	...
11	me chifla este curro, porque, así, el venir a la	facu	se me hace más llevadero. A mi madre no le hace n
12	clases son la mar de entretenidas y	falta	emñillar tanto

1/3 ▼

- fabricamos
- fachada
- facilidad
- facilísima
- facilísimo
- factura
- facu
- facultad
- falda

Podemos cambiar el sitio de la división arrastrando la barra de título en color, por ejemplo:

[MENÚ] Datos: Letras | Números | Gráficos VARITEX en LYNEAL [Ejecutar] 100 % 211 m.s. [Resets]

■ Letras: Output: 99 / Limite de output:1000

Ventana

5 - ¡Pues, claro, mano! Allí tienes de todo: por ejemplo, hay guapas para tu mujer o chándal para los crios. ¡Te gustan los chicos de la universidad?

6 - ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio tiene?

7 ■ Números: 1. Puntuación. 1. FA: Frecuencia absoluta (*)

Ventana

Número de filas: 19 / Número de columnas: 1

FA	Todos	Resetear
1	fabricamos	1 fabricamos
2	fachada	1 fachada
3	facilidad	1 facilidad
4	facilísima	1 facilísima
5	facilísimo	1 facilísimo
6	factura	2 factura
7	facu	4 facu
8	facultad	5 facultad
9	falda	4 falda

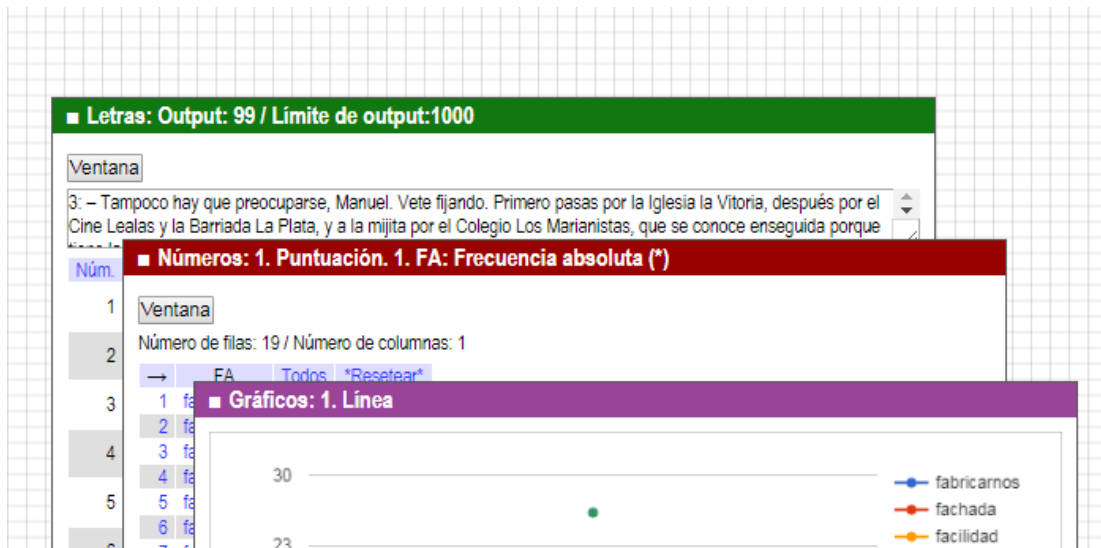
■ Gráficos: 1. Línea

1/3 ▼

Una vez comprobada la división de [Datos], hemos hecho desaparecer la misma división por el botón de [Datos] en el encabezamiento, o [Ctrl]+D.

Si se introducen cambios en algunas partes de [Letras] en [MENÚ] se necesita ejecutar para obtener el output. Con los cambios de [Datos], [Números] y [Gráficos] se obtiene automáticamente el output correspondiente, sin necesidad de ejecutar.

Al hacer el clic doble en la barra de título, desaparece y aparece el contenido alternativamente. En la figura siguiente hemos hecho desaparecer el contenido de [Datos]:



1. Datos

Al hacer clic en [Datos] en [MENÚ], aparecen tres subdivisiones [Archivo], [Atributos y filtros] y [Datos de usuario]:



1.1. Archivo

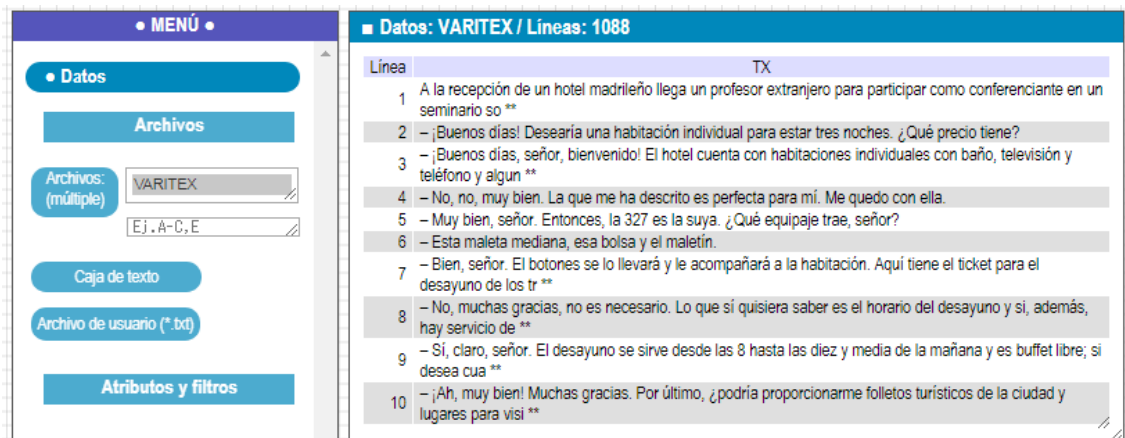
Haciendo clic en [Archivos], obtenemos el contenido de la subdivisión:



En [Archivos:] por defecto está seleccionado el primer archivo en la lista de la derecha: VARITEX. Se admite hacer la selección múltiple para comparar los textos de distintos archivos de la misma estructura de atributos.

La caja de texto ampliable, con "Ej. A-C,E" sirve para seleccionar los archivos con su letra inicial. Por ejemplo "A-C, E" selecciona los archivos cuyos nombres empiezan con A, B, C y E. La selección empieza cuando el usuario hace clic fuera de la caja de texto.

Una vez seleccionado el archivo, aparece la primera parte del archivo seleccionado en la división derecha en [Datos].



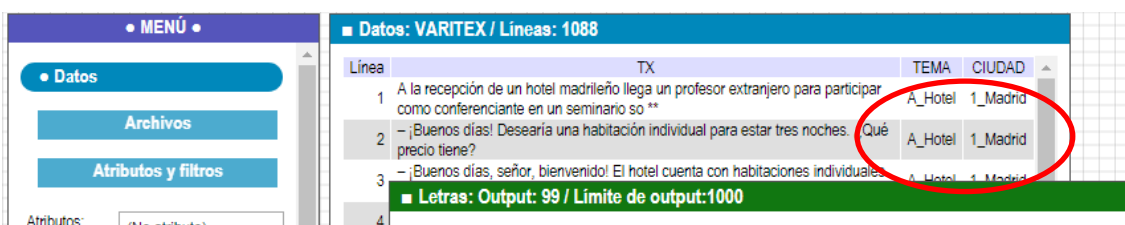
En la cabecera de [Datos], se indica el número de Líneas en total del archivo.

1.2. Atributos y filtros

1.2.1. Atributos

En la parte media de [Datos] en [MENÚ] se encuentra [Atributos y filtros], que tiene dos funciones: selección de atributos que aparecen en el output de [Datos] y [Letras] y selección de filtros, que explicaremos en la sección siguiente.

Los atributos seleccionados aparecen tanto en la división de [Datos] como en la lista de [Letras] en la parte derecha del output. Por ejemplo, seleccionamos TEMA y CIUDAD en [Atributos de input] y ejecutamos [Ctrl] + E):



Efectivamente en la división de [Datos], observamos las columnas de [TEMA] y [CIUDAD].

1.2.2. Filtros

Al seleccionar un atributo en [Atributos (múltiple):] de [Datos] en [MENÚ] aparecen [Filtros (múltiple):] debajo. Al seleccionar filtros, por ejemplo 'I_Fútbol', los datos salen con los filtros correspondientes:

Linea	TX	TEMA	CIUDAD
1	Juan y Antonio, antiguos compañeros de trabajo que llevan mucho tiempo sin verse, se encuentran en un bar. Juan, que ya **	I_Fútbol	1_Madrid
2	- ¡Hola, Juanito !;Qué es de tu vida, chaval !	I_Fútbol	1_Madrid
3	- Aquí estamos, hijo, aguantando para no caer...	I_Fútbol	1_Madrid
4	- ¿Quieres tomar un café? Te lo debía de otra ocasión, ¿te acuerdas? Y así	I_Fútbol	1_Madrid

■ Letras: Output: 5 / Límite de output:1000

1.2.3. Agrupamiento

Atributos y filtros

Atributos:
(múltiple)

0: (No atributo)
 1: ID
 2: TEMA
 3: CIUDAD

Filtro [-] y grupo [;]
(múltiple)

D: TEMA
 1: A_Hotel
 2: B_Camino
 3: C_Funcionarios
 4: D_Estudiantes

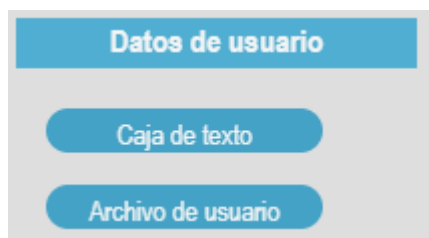
Ej. 1,2,4-6;7-9;10
(y clic fuera)

El Filtro funciona como medio de agrupamiento, todo por el número de miembros. Por ejemplo, "1-3;4,5;6-9" agrupa los miembros de la siguiente manera: [1,2,3], [4,5], [6,7,8,9]. La coma (,) sirve para agregar los miembros, el guion para indicar el principio y el fin de la secuencia; y el punto y coma (;) para indicar la separación del grupo. Después de escribir la secuencia de números con estos signos, hay que hacer clic fuera de la casilla. Luego al escoger la misma variable en la selección de Letras y ejecutar, se verá el resultado con el agrupamiento.

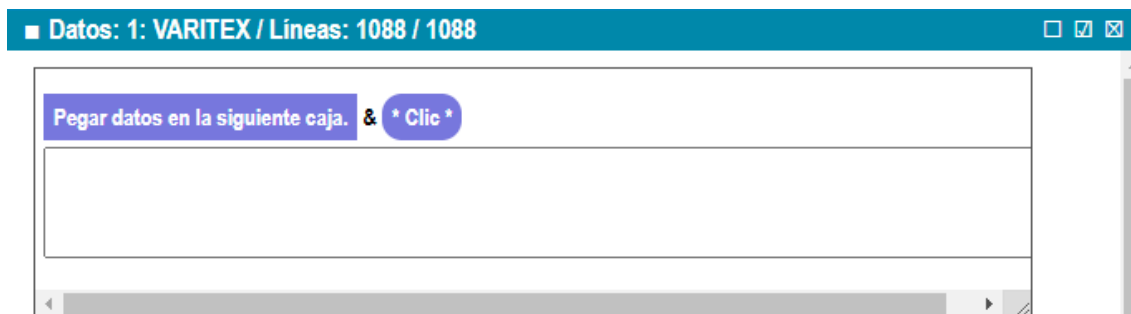
1.3. Datos de usuario

1.3.1. Caja de texto

En lugar de los corpus almacenados en la lista de [Archivos] en [Datos], podemos analizar nuestros propios datos utilizando la [Caja de texto] o [Archivo de usuario], que seleccionamos en la división de la misma lista de [Archivo]:



Al seleccionar [Caja de texto] aparece una caja de texto en la parte derecha:



Copiamos, por ejemplo, un texto de una página web y lo pegamos dentro de la misma caja y finalmente pulsamos [* Clic *], que está encima de la caja de texto.

En este curso presentamos diálogos breves para introducir lo fundamental de la gramática española. Consiste en 12 lecciones y cada lección contiene 3 minidiálogos. Los estudiantes de español pueden ver las escenas documentadas en la Universidad Europea de Madrid, en colaboración con los profesores y estudiantes, que figuran en la lista de equipo de producción. Agradecemos a nuestros colaboradores que han comprendido nuestros propósitos y el significado del trabajo. Sus actuaciones son excelentes.

El texto acompañado de unos parámetros puede ser analizado, si el texto viene en la primera columna con la cabecera de "TX" y los parámetros están separados por la tabulación ([Tab]). Si el título de la primera fila lleva "##", el mismo título va a ser tratado como parámetro:

TX	##OBRA	##Folio	##Línea		
Delos sos oios tan fuerte mientras lorando	13a.Cid	1r	1		
Tornaua la cabeça & estaua los catando	13a.Cid	1r	2		
Vio puertas abiertas & vços sin ca<n>nados	13a.Cid	1r	3		
Alcandaras uazias sin pieles & sin mantos	13a.Cid	1r	4		
E sin falcones & sin adtores mudados	13a.Cid	1r	5		

Para evitar la repetición del mismo parámetro, se puede acudir, exclusivamente cuando ocurre cambio del mismo parámetro, a la fórmula de "##***:***":

TX
##OBRA::13a.Cid
Delos sos oios tan fuerte mientras lorando 1r 1
Tornaua la cabeça & estaua los catando 1r 2
Vio puertas abiertas & vços sin ca<n>nados 1r 3
...
##OBRA::13b.Fazienda
emont por la gracia de dios. arço= 1ra 1
~bispo de Toledo. a don almeric. arçi= 1ra 2
~diano de antiochia con grant a= 1ra 3
...

Otra forma de evitar la repetición de atributos es utilizar "\$\$", que convierte una abreviatura en la secuencia de atributos. Por ejemplo, la letra E representa la secuencia de (*, *, *, *) y la letra I representa (Mujer, 35, *, Castellano).

TX	##PAPEL	##SEXO	##EDAD	##VARIEDAD
	##LENGUA			
\$\$E	Entrevistador(a)	*	*	*
\$\$I	Informante	Mujer	35	Castellano
<p>Qué opinas de la violencia familiar, y porque las crees que las mujeres no reaccionen ante una agresión de sus parejas? E</p> <p>¡Por qué si te das de cuenta!, en todos los episodios si te vas a toda la historia de la humanidad ya sea he...bíblicamente, históricamente siempre la mujer ha estado metida, o tratando de tener poder atreves del hombre. I</p> <p>Aunque muchos feministas me van a linchar por esto lo que estoy diciendo, ¿por qué si eres mujer por qué estas opinando de ésta manera? dicen, pero si vas viendo las etapas de toda la historia de la humanidad es así, `porque si nos vamos a la biblia lo más cercano que todos tienen, es que la mujer en el éste que tienen ¡ahí se me ha olvidado! I</p>				

1.3.2. Archivo de usuario

Cuando se trata de un archivo muy grande, no conviene utilizar la [Caja de texto], sino el [Archivo de usuario]. Todos los archivos deben estar en modo TEXTO, txt, con codificación UTF-8. Al seleccionar el [Archivo de usuario] en la lista del [Archivos] en la sección de [Datos], aparece en la parte derecha el botón de [Seleccione su archivo (*.txt)] y seguimos el diálogo que se ofrece en el mismo botón.

Lo mismo que en la [Caja de texto], el texto acompañado de unos parámetros puede ser analizado, si el texto viene en la primera columna con la cabecera de "TX" y los parámetros están separados por la tabulación.

1.4. Datos numéricos

Podemos analizar una matriz de datos numéricos. La figura siguiente es un ejemplo de matriz de frecuencias en 4 filas y 5 columnas:

N.D2.fre.	A	B	C	D	E
i.1	10	19	14	7	12
i.2	11	7	10	0	1
i.3	0	0	1	12	1
i.4	0	1	2	3	3

La figura siguiente es una matriz de números reales no negativos:

N.D4.ss-ff	-ss-	#ss-	-ff-	#ff-		
R: León	4.8	5.8	1.4	6.1		
R: C. la Vieja	9.7	9.6	1.5	10.1		
R: C. la Nueva	1.8	2.4	0.5	1.9		
R: Aragón	11	0.9	2.1	3		
A: 1200	16.8	3.1	1.2	6.2		
A: 1225	5.7	0.2	0.5	0		
A: 1250	12.2	3.1	1.5	6.8		
A: 1275	13.1	13.8	1.5	12.4		
A: 1300	13.4	14.8	4	22.1		
A: 1325	14.3	20.5	3.6	23.8		
A: 1350	8.9	10.6	3.1	9.4		
A: 1375	3.7	3.7	1.9	3.3		
A: 1400	6.6	1	0.9	1.2		
A: 1425	3.9	0.5	0.7	1.1		
A: 1450	7.8	0.5	1.1	2.3		
A: 1475	7.3	0	0.9	0.8		
L: cortesana	0.6	0.5	0.7	0.8		
L: albalaes	19.2	19.4	3	12.9		
L: privilegios	7.9	1.6	3.1	7.9		
L: gótica	2.8	0.6	0.6	2.2		
L: gótica cursiva		10.7	15.8	2.8	19.6	
L: gótica cursiva [albalaes]			9.2	19.2	1.8	20.3
L: gótica cursiva [precortesana]				0.2	2	0.3
						1.7

L: gótica libraria	10.3	9	3.5	11	
L: gótica redonda	12.8	1.1	0.7	1.6	
L: humanística3.3	0	3.3	0		
L: humanística redonda		0.5	3.7	0	1
L: precortesana	3.9	3.9	0.7	2.7	

Las condiciones de los datos numéricos son las siguientes:

(1) que los datos estén separados por tabulación, que pueden ser una copia de la tabla de Excel.

(2) que tenga el título de atributos en la primera fila

(3) que tenga el título de casos en la primera columna

(4) que la primera celda de la primera fila empiece con N. (ene mayúscula + punto).

Una vez cumplidas estas cuatro condiciones, LYNEAL trata este dato no como texto, sino como una matriz de números.

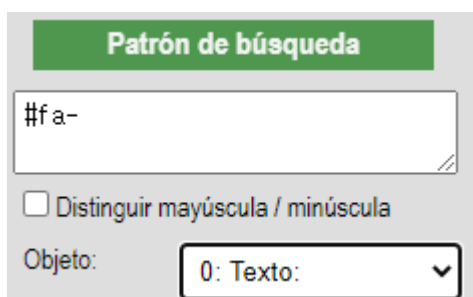
2. Letras

En la sección de [Letras] del [Menú] hay cuatro divisiones:[Patrón de búsqueda], [Opción], [Contexto] y [Casos y atributos]:



2.1. Patrón de búsqueda

Abrimos la división de [Patrón]:



En la caja de Patrón, hemos puesto un ejemplo del patrón "#fa-", con el que buscamos todas las formas que empiezan con "fa", por ejemplo, *familia*, *facultad*, *falda*, etc. Los detalles de la regla de patrón los explicaremos más adelante (Apéndice). Podemos incluir varios patrones al mismo tiempo cambiando la línea, por ejemplo:

```
#fa-  
#sa-
```

Con el botón de [Distinguir mayúscula / minúscula] se distinguen dos modos de búsqueda. Generalmente en la búsqueda no distinguimos entre las letras mayúsculas y minúsculas y se buscan, por ejemplo, tanto *familia* como *Familia*. Al pulsar el botón de [Distinguir mayúscula / minúscula] se distinguen las dos formas y se buscan *familia* a exclusión de *Familia*. Al escribir el patrón "#Fa-", se buscan las formas en mayúscula *Familia*, con exclusión de *familia*.

2.2. Opción

Al pulsar [Opción], aparece la figura siguiente, donde podemos seleccionar los ítems del output. Algunos ítems están seleccionados por defecto:

Opción

Elementos en output:

- 1. Contexto anterior (A)
- 2. Contexto anterior reverso
- 3. Patrón de búsqueda
- 4. Forma de búsqueda
- 5. Forma reversa
- 6. Contexto posterior (P)
- 7. Contexto total (T)
- 8. Posición en párrafo
- 9. Posición en grupo

Atributos (mult.)

Procesamiento:

- 1. Búsqueda incremental
- 2. Número de documentos
- 3. Ordenar por forma

Limite de output: ▲ ▼

2.2.1. Elementos en output

En la división de [Elementos en output] podemos seleccionar los nueve elementos enumerados. Por defecto, están seleccionados [Contexto anterior (A)], [Forma de búsqueda], [Contexto posterior (P)]. Los atributos seleccionados en [Atributos] salen en el output.

La figura siguiente muestra el resultado de la búsqueda con el patrón "#fa-" y la lista de output por defecto:

Letras: Output: 115 / Limite de output:1000

Ventana / Copiar

10: – Uy, uy, uy, Paquito... que te veo mal. No te fiwves ni un pelo. En cualquier momento te atrapa y caes como un idiota. Y, despuewws, no vas a poder salir tan fawwcilmente. Ten cuidado. Dice. Te lo digo yo, que te tuvo una experiencia parecida.

Núm.	Contexto anterior (A).	Forma	Contexto posterior (P)	TEMA
1	– Muchas gracias. ¡Ah, por	favor	! ¿Me puede decir dónde está el comedor?	A_Hotel
2	*, tío ! Allí tienes de todo: por ejemplo, hay	faldas	guapas para tu mujer o chándal para los críos*	B_Camino
3	*ianistas, que se conoce enseguida	fachada	de ladrillo. Y de ahí al "Campeón",	B_Camino

Al pulsar el botón de [Ventana] sale toda la lista en una ventana

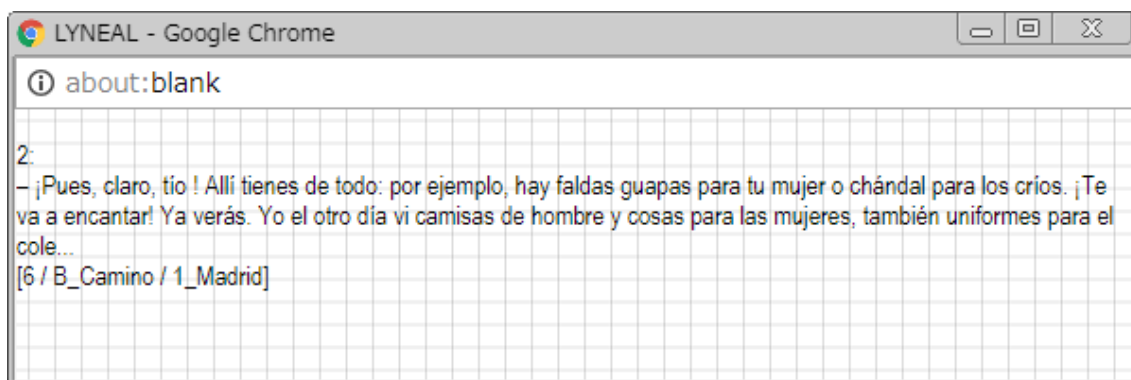
independiente. Al pulsarlo de nuevo, desaparece:

Núm.	Contexto anterior (A)	Forma	Contexto posterior (P)
1	– Muchas gracias. ¡Ah, por	favor	! ¿Me puede decir dónde está el comedor?
2	laro, tío! Allí tienes de todo: por ejemplo, hay	faldas	guapas para tu mujer o chándal para los críos. ¡T
3	ianistas, que se conoce enseguida porque tiene la	fachada	de ladrillo. Y de ahí al "Campeón", nada, dos o t
4	víos para la limpieza... en fin, todo lo que haga	falta	en la casa para la semana. Así que cuando me sobr
5	laro, mano! Allí tienes de todo: por ejemplo, hay	faldas	padrísimas pa' tu mujer o ropa deportiva □ sudade
6	– Y... compadre ¿Cómo estás? Oye, un	favor	¿Sabes dónde hay un supermercado cerca?
7	rdoname que te moleste, pero ¿me podés decir por	favor	dónde queda el Supermercado "El Disco"?

Pulsando el botón de [Copiar], se copian todos los datos de texto y se puede pegar en Word, Excel o cualquier editor de texto.

Al sobreponer el ratón en alguna parte de la lista inferior sale todo el texto en la caja superior, seguida del botón de [Ventana] / [Copiar].

Al hacer clic en alguna parte de la lista, sale todo el texto en el que aparece la forma buscada en la ventana independiente, de donde podemos copiar el dato. Al pulsarlo de nuevo, desaparece.



Por ejemplo, al seleccionar 1. Número de línea, 4. Patrón, 5. Forma, 8. Forma en texto, obtenemos el output siguiente:

Línea	Patrón	Forma	Forma en texto
69	#fa%	favor	~nde está el comedor? / – Muchas gracias. ¡Ah, por favor! ¿Me puede decir dónde está el comedor? / ~
106	#fa%	faldas	~laro, tío! Allí tienes de todo: por ejemplo, hay faldas guapas para tu mujer o chándal para los críos. ¡T~

Al hacer clic en la cabecera de la tabla (Línea, Patrón, Forma, ...), se ordena de manera ascendente y descendente, alternativamente.

Seleccionando [Ordenar por forma], se hace la ordenación descendente de la lista de output desde el principio. Si se desea ordenar por contexto, también

es posible haciendo clic en la cabecera de la lista.

Hemos establecido el [Límite de output] en 1000 líneas por defecto. Podemos aumentar la cantidad a costa del tiempo de ejecución. Se trata de la cantidad de output de la lista de Letras. El cálculo para Números incluye todos los datos procesados.

2.2.2. Procesamiento

(1) Búsqueda incremental

La mayoría de las veces, hacemos la búsqueda continua. Por ejemplo, con el patrón "@@" se busca la secuencia de dos letras continuas. En la figura siguiente, vemos los resultados de "A la recepción" en las formas de *la*, *re*, *ce*, *pc*, ...

Línea	Texto anterior.	Forma	Texto posterior
1	~re posibles visitas turísticas por la región. / A	la	recepción de un hotel madrileño llega un profesor~
1	~posibles visitas turísticas por la región. / A la	re	cepción de un hotel madrileño llega un profesor ex~
1	~sibles visitas turísticas por la región. / A la re	ce	pción de un hotel madrileño llega un profesor extr~
1	~bles visitas turísticas por la región. / A la rece	pc	ión de un hotel madrileño llega un profesor extran~

En la búsqueda continua, la parte encontrada no entra en la búsqueda siguiente. Al seleccionar la [Búsqueda incremental] se hace la búsqueda de modo incremental de manera siguiente: *la*, *re*, *ec*, *ce*, ..., de modo que se buscan todas las posibles secuencias de las dos letras continuas:

Línea	Texto anterior.	Forma	Texto posterior
1	~re posibles visitas turísticas por la región. / A	la	recepción de un hotel madrileño llega un profesor~
1	~posibles visitas turísticas por la región. / A la	re	cepción de un hotel madrileño llega un profesor ex~
1	~osibles visitas turísticas por la región. / A la r	ec	epción de un hotel madrileño llega un profesor ext~
1	~sibles visitas turísticas por la región. / A la re	ce	pción de un hotel madrileño llega un profesor extr~
1	~bles visitas turísticas por la región. / A la rece	pc	ción de un hotel madrileño llega un profesor

Para la búsqueda incremental de las dos palabras, utilizamos el patrón de "#& &". Hay que poner la frontera con "#", puesto que con el patrón de "& &", se hace la búsqueda con el incremento de una letra y salen las palabras incompletas. Se tarda tiempo en el procesamiento de la búsqueda incremental.

Línea	Texto anterior.	Forma	Texto posterior
1	~obre posibles visitas turísticas por la región. /	A la	recepción de un hotel madrileño llega un profesor~
1	~re posibles visitas turísticas por la región. / A	la recepción	de un hotel madrileño llega un profesor extranjero~
1	~posibles visitas turísticas por la región. / A la	recepción de	un hotel madrileño llega un profesor extranjero p~
1	~isitas turísticas por la región. / A la recepción	de un	hotel madrileño llega un profesor extranjero para~
1	~tas turísticas por la región. / A la	un hotel	madrileño llega un profesor extranjero

(2) Número de documentos

Seleccionamos el [Número de documentos] cuando nos interesa saber en cuántos documentos se encuentran las formas buscadas, en lugar de la frecuencia de las formas en todos los documentos seleccionados. Por ejemplo, en el corpus VARITEX encontramos 7 casos de *fatal* en los documentos 11, 21, 21, 22, 22, 25, 32:

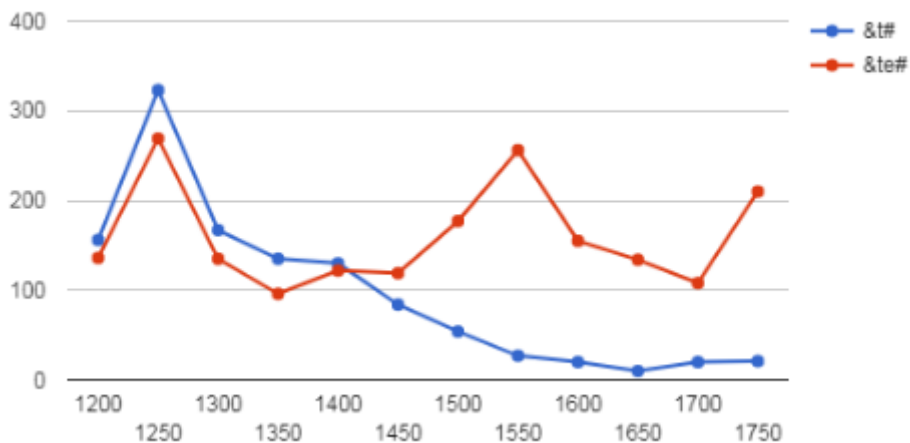
8	me puedo quejar. Bueno , la semana pasada estuve	fatal	, con un gripazo impresionante. Me dieron la baja	11
18	do que venir por lo de la alergia. Mira, me pongo	fatal	. ¡Es horrible!: no puedo dormir bien, me cuesta r	21
21	ás, cualquier comida que no sea muy ligera me cae	fatal	...	21
22	ico, pero con lo de la alergia ... Mira, me pongo	fatal	. ¡Qué mal me siento! no puedo ni dormir por las n	22
23	n cuanto tomo algo que no sea ligerito, me siento	fatal	.	22
27	que venir por mi alergia. Mirá, en otoño me pongo	fatal	. No me aguanto ni yo misma. No puedo dormir bien,	25
38	vas a decir que no hay gente que lo está pasando	fatal	, y que no le llega el sueldo a final de mes? ¡No	32

El número de documentos resulta ser 5, porque se repiten dos veces en el documento 21 y 22:

16	fastidies	1	fastidies
17	fatal	5	fatal
18	favor	10	favor

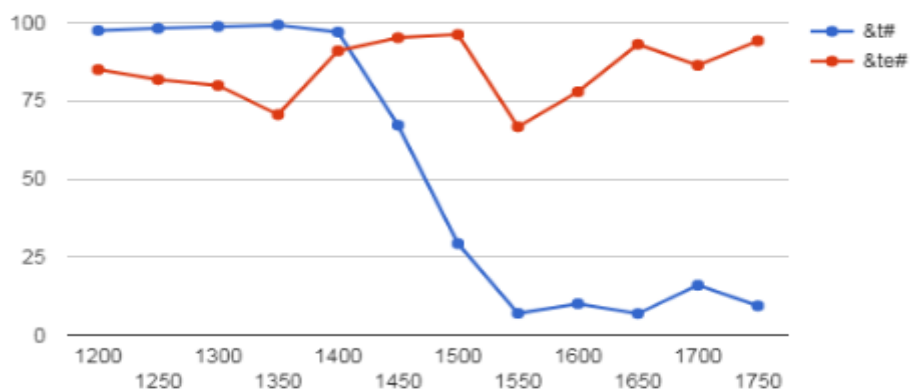
La tabla siguiente muestra la frecuencia de las formas de palabra terminada en 't' y 'te' de CODEA (paleográfica):

→	FA	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
1	&t#	156	323	167	135	130	84	54	27	20	10	20	21
2	&te#	136	269	135	96	122	119	177	256	155	134	108	210



La tabla siguiente muestra los números de documentos que contienen las formas correspondientes (por lo menos una forma):

→ FN.Doc.:100	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
1 &t#	97.5	98.2	98.8	99.3	97.0	67.2	29.3	7.0	10.1	6.9	16.0	9.4
2 &te#	85.0	81.8	79.9	70.6	91.0	95.2	96.2	66.7	77.9	93.1	86.4	94.2



(3) Ordenar por forma

Hay dos opciones respecto al orden de presentación de la [Lista de letras]. Si dejamos sin pinchar el botón de [Ordenar por forma/patrón], se presenta el resultado de la búsqueda en orden de aparición de las formas buscadas:

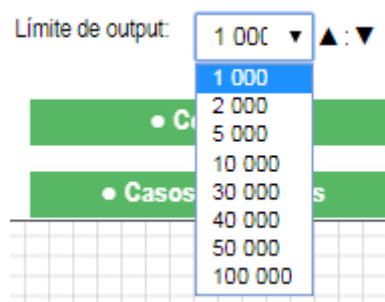
Núm.	Contexto anterior (*)	Forma	Contexto posterior (=)
1	– Muchas gracias. ¡Ah, por	favor	! ¿Me puede decir dónde está el comedor?
2	– ¡Pues, claro, tío ! Allí tienes de todo: por ejemplo, hay	faldas	guapas para tu mujer o chándal para los crios. ¡Te va a encantar! Ya verás. Yo el otro día vi camis
3	da La Plata, y a la mijita por el Colegio Los Marianistas, que se conoce enseguida porque tiene la	fachada	de ladrillo. Y de ahí al "Campeón", nada, dos o tres paradas.
4	azúcar, café, vino, cerveza, gaseosa, embutido, avíos para la limoieza... en fin. todo lo que haga	falta	en la casa para la semana. Así que cuando me sobra algo, le compro unas

Al pinchar el botón de [Ordenar por forma], se presenta la lista en orden alfabético de la forma. También es posible ordenar la lista pinchando la línea de cabecera en color azul claro, en orden ascendente y descendente, alternativamente:

Núm.	Contexto anterior (*)	Forma	Contexto posterior (=)
28	ear más con los míos. No voy a hacerme mala sangre por eso. Demasiados problemas tenemos como para	fabricarnos	otros. Ellos que van a la escuela y al colegio y él que va a trabajar, creen que yo estoy al vicio
3	da La Plata, y a la mijita por el Colegio Los Marianistas, que se conoce enseguida porque tiene la	fachada	de ladrillo. Y de ahí al "Campeón", nada, dos o tres paradas.
79	tiempo sin ninguna duda. Los muchachos estaban inspirados. Llegaban al arco contrario con bastante	facilidad	. Hacían los pases con mucha precisión y casi no perdían pelota...
84	- ¡Ah, la tortilla de patatas! ¡Es	facilísima	! ¡Y muy rápida de hacer! Pues mira, sólo tienes que freír en abundante aceite, las patatas a las qu

2.2.3. Límite de output

El límite de output significa la cantidad límite de output en la lista de [Letras]. Está seleccionado 1000, por defecto. Se puede ampliar hasta 50,000. En [Números] se incluyen todos los casos sin limitación.



2.3. Contexto

Al pulsar [Contexto] en [Letras], aparece el siguiente contenido:

Contexto

Patrón de contexto anterior (A) : *

Patrón de contexto posterior (P) : *

Patrón de contexto total (T) : *

Contexto : 1.Patrón

Unidad: Línea-Párrafo

Letras en contexto anterior: 50

Letras en contexto posterior: 50

Se trata del contexto anterior a la forma buscada, el contexto posterior a la misma y el contexto total incluida la forma buscada. Por ejemplo, al escribir “#ff%” en el [Patrón] de la sección anterior, y “#ss%” en el [Patrón de contexto total] en [Contexto], podemos buscar y calcular tanto los casos de la ff- inicial como los de la ss- inicial en su contexto, tanto anterior como posterior:

99	oherra<n>do & {ssu} t<ie>rra {ssu} p<ar>te esta carta	ffyz	in<arline>z porç i t <escreujr> del co<n>çejo E om<ne>s buenos d<e>la
100	o<n> E {sso} ende testigo E en testimonjo de verdat	ffyz	aq<ui> este mio signo

Las opciones de [Patrón], [Forma], [Patrón::forma] y [Todo] sirve para el cálculo de frecuencias en la sección de [Números]. Con la opción de [Marca] se destacan las formas correspondientes en color rojo en el contexto, como lo observamos en la figura anterior. Con la opción de [Suma], se exhiben las frecuencias correspondientes en los distintos contextos: anterior, posterior, total:

Patrón de contexto anterior (A) : #ss%

Patrón de contexto posterior (P) : #ss%

Patrón de contexto total (T) : #ss%

68		fijo	#ss%:39	#ss%:39
69	#ss%:3	fechos	#ss%:36	#ss%:39
70	#ss%:7	ffaz<er>	#ss%:32	#ss%:39
71	#ss%:9	ffueros	#ss%:30	#ss%:39

Los contextos especificados sirven para hacer la tabla cruzada en la sección de [Números].

La [Unidad] se selecciona entre [Palabra], [Oración], [Línea-Párrafo] (por defecto) y [Documento]. Al seleccionar, por ejemplo, [Oración], el cálculo de las frecuencias de contexto se hace dentro de la misma oración de la forma buscada, con exclusión de las formas que aparecen fuera de la misma oración.

Se puede adoptar la cantidad máxima de letras en el contexto anterior y el posterior en las dos cajas de [Límite de contexto anterior] y [Límite de contexto posterior].

2.4. Casos y atributos

En la división de [Casos y atributos] se seleccionan [Variables] para establecer los parámetros horizontales de la frecuencia.

Casos y atributos

Casos: 1.Patrón

Variables (múltiple):
 0: (No)
 1: ID
 2: TEMA
 3: CIUDAD

Dos variables: 1.Separar variable

Intervalo-1: (No) ▲ ▼

Intervalo-2: (No) ▲ ▼

Número de empezar el intervalo: 1

En [Casos] ofrecemos cuatro opciones: [Forma], [Patrón], [Patrón::forma] y [Todos los casos]. Con la opción [Forma] se hace el cálculo de las frecuencias de formas. Con la opción [Patrón], no se hace la distinción de formas, sino que se hace el cálculo agrupándolas en el mismo patrón. Cuando nos interesa a qué patrón pertenece cada forma, seleccionamos la opción [Patrón::forma], con la que se calcula la combinación de patrón y forma, por

ejemplo, *#fa%::familia*, *#sa%::saber*. Finalmente con la opción [Todos los casos], se calcula la totalidad de los patrones. Al hacer la selección múltiple, con [Ctrl] + clic, podemos hacer el cálculo de las formas y Todos los casos, al mismo tiempo.

En cuanto a las [Variables:], por ejemplo, el corpus VARITEX posee dos variables: TEMA y CIUDAD. Dentro de TEMA encontramos 10 temas de conversación y dentro de CIUDAD, contamos con 5 ciudades investigadas: Madrid, Sevilla, México, Lima y Buenos Aires. Estas variables son utilizadas para construir la tabla cruzada de frecuencias. Empezamos con el [Patrón] en la variable CIUDAD. Por ejemplo, en [Patrón de búsqueda:], escribimos:

#tú#
 #usted#
 #ustedes#
 #vos#
 #vosotr[oa]s#

Al procesar los datos, obtenemos la siguiente Tabla:

[3a] Frecuencia: Frecuencia absoluta

Núm.	Patrón	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	14	24	10	8	1
2	#usted#	2	7	2	4	0
3	#ustedes#	7	9	4	6	0
4	#vos#	0	0	0	0	11
5	#vosotr[oa]s#	1	0	0	0	6

2.5. Dos variables

Al seleccionar dos variables en [Variables], podemos elegir una de las tres opciones siguientes: [Separar variables], [Unir variables], [Variable por variable (número)], [Variable por variable (ratio)], [Variable por variable (letra)]:

Casos y variables

Casos: 1.Patrón

Variables (múltiple): 0: (No), 1: ID, 2: TEMA, 3: CIUDAD

Dos variables: 1.Separar variable:

Intervalo-1: 1.Separar variables, 2.Unir variables, 3.Var. x var. (total)

Intervalo-2: 4.Var. x var.: $X/(X+Y)$, 5.Var. x var.: $(X-Y)/(X+Y)$

2.5.1. Separar variables

Se hace la tabla cruzada de forma, *tú, usted, ustedes, vos, vosotros*, por las dos variables, TEMA y CIUDAD, de manera separada. En el eje horizontal vienen 10 temas seguidos de 5 ciudades:

→	FA	1.A_Hotel	1.B_Camino	1.C_Funcionarios	1.D_Estudiantes	1.E_Cor
1	tú	0	9	3	4	
2	usted	8	0	0	0	
3	ustedes	1	0	0	0	
4	vos	0	1	0	3	
5	vosotros	0	0	0	0	
Resetear		1.A_Hotel	1.B_Camino	1.C_Funcionarios	1.D_Estudiantes	1.E_Cor

1.I_Fútbol	1.J_Despedida	2.1_Madrid	2.2_Sevilla	2.3_México	2.4_Lima	2.5_B.A.
1	12	14	24	10	8	1
0	0	2	7	2	4	0
0	20	7	9	4	6	0
0	1	0	0	0	0	11
0	6	1	0	0	0	6
1.I_Fútbol	1.J_Despedida	2.1_Madrid	2.2_Sevilla	2.3_México	2.4_Lima	2.5_B.A.

2.5.2. Unir variables

Se hace la tabla cruzada de forma, *tú, usted, ustedes, vos, vosotros*, por las dos variables unidas de TEMA y CIUDAD. En la siguiente tabla, observamos, por ejemplo, la unión de A_Hotel y 1_Madrid con formas de *usted y ustedes*:

→	FA	A_Hotel+1_Madrid	A_Hotel+2_Sevilla	A_Hotel+3_México	A_Hotel+4_Lima	B_Camino+
1	tú	0	0	0	0	0
2	usted	1	2	1	4	
3	ustedes	1	0	0	0	
4	vos	0	0	0	0	
5	vosotros	0	0	0	0	

2.5.3. Variable por variable (total)

Se hace la tabla cruzada numérica por las dos variables, HOTEL en el eje vertical y CIUDAD en el eje horizontal. Por ejemplo, en la celda de A_Hotel y 1_Madrid, encontramos dos formas, *usted* (1) y *ustedes* (1), en total, 2.

→	FA	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	A_Hotel	2	2	1	4	0
2	B_Camino	3	3	2	1	1
3	C_Funcionarios	1	2	0	0	0
4	D_Estudiantes	1	2	1	0	3
5	E_Consultorio	3	6	1	1	0
6	F_Fiesta	2	3	2	1	1
7	G_Política	3	8	3	2	5
8	H_Compras	1	4	1	0	0
9	I_Fútbol	0	0	0	1	0
10	J_Despedida	8	10	5	8	8

2.5.4. Variable por variable (ratio: X/(X+Y))

Se calcula la ratio del primer elemento (“#tú#”) dentro de la totalidad de las formas encontradas en la misma celda de la tabla cruzada. Por ejemplo, En J_Despedida::1_Madrid, se encuentran 2 casos de *tú* y 6 casos de *ustedes*, de modo que la ratio de *tú* es $2 / (2 + 6) = 0.250$.

FA	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
B_Camino	1.000	1.000	1.000	1.000	
C_Funcionarios	1.000	1.000			
D_Estudiantes	1.000	1.000	1.000		
E_Consultorio	0.667	0.333			
F_Fiesta	1.000	1.000	1.000	1.000	
G_Política	1.000	1.000	1.000	1.000	
I_Fútbol				1.000	
J_Despedida	0.250	0.400	0.400	0.375	0.125

2.5.5. Variable por variable (contraste: (X-Y)/(X+Y))

2.6. Intervalo numérico

En cuanto al atributo numérico, por ejemplo, edad, página, año, etc., podemos clasificar la cantidad según el [Intervalo numérico]. A modo de ejemplo, seleccionamos el Archivo "ALICIA" con la opción [Atributo:] "Capítulo" y en el [Patrón de búsqueda:] escribimos los siguientes patrones:

```
#conejo#
#rey#
#reina#
```

y en el [Atributo:], [Capítulo] con el [Intervalo numérico] de 2 y selección de [Número de empezar el intervalo: 1]. De esta manera obtenemos la siguiente tabla:

→	F.A.	1	3	5	7	9	11	*Resetear*
1	#conejo#	10	9	0	1	0	6	#conejo#
2	#reina#	0	0	2	14	4	6	#reina#
3	#rey#	0	0	0	3	0	22	#rey#
Resetear		1	3	5	7	9	11	

Se trata de las frecuencias absolutas de cada patrón distribuidas en seis grupos de capítulos. Cada número del título horizontal, 1, 3, 5, 7, 9, 11 representa el inicio de cada grupo, de modo que 1 representa 1 y 2, 3 representa 3 y 4, y así sucesivamente. El mismo intervalo numérico se aplica a la lista de output con la selección de [Lista de output con intervalo]:

10	es y vuelve a encoger por efecto del abanico del	Conejo	. Casi se ahoga en el charco de sus propias lágrimas	1
11	CAPÍTULO IV. El	Conejo	envía un billete / a Pepito REVISAR	3
12	Era el	Conejo	Blanco, que volvía trotando otra vez y mirando a	3
13	adivinó al instante que lo que	reina	eran el abanico y el par de guantes	3

3. Números

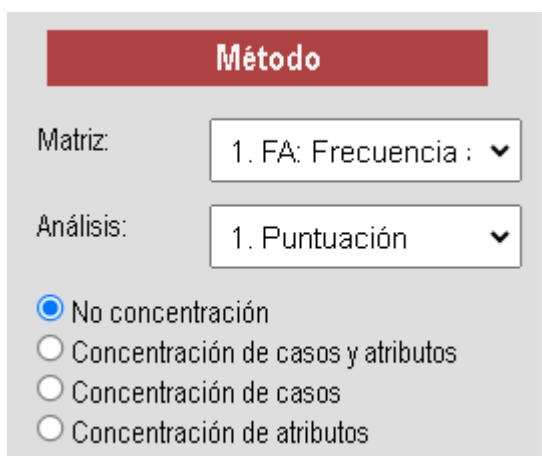
A continuación, explicamos las distintas funciones que hay en la sección de Números. Se divide en dos partes: [Método de análisis] y [Opción]:



En este documento nos limitamos a explicar los modos de manejar el sistema LYNEAL. Para los detalles de tratamientos numéricos de datos, cálculos probabilísticos y estadísticos, véase *Análisis de datos lingüísticos* [pdf], que se encuentra en la sección de [Información] (sec. 5 de este documento):

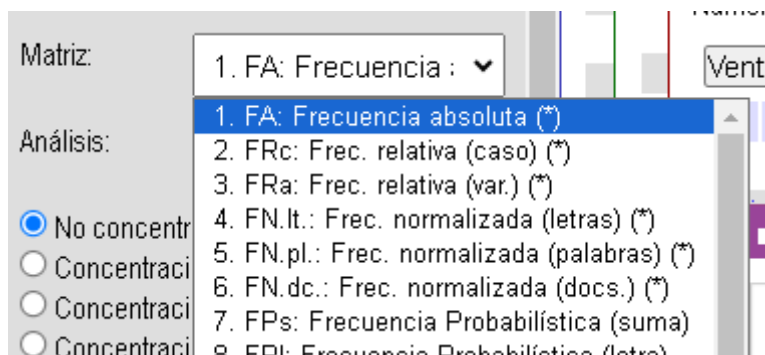
3.1. Método

Dentro de la división [Método], ofrecemos múltiples posibilidades de [Matriz] y [Análisis] con opciones de concentración:



3.1.1. Matriz

Partimos de [Matriz] que va a ser datos objetos de análisis numéricos. Hemos preparado múltiples variedades de frecuencias:



Hacemos la búsqueda con el siguiente patron:

#tú#
 #usted#
 #ustedes#
 #vos#
 #vosotr[oa]s#

Y seleccionamos la opción de [Ciudad]:

1. Frecuencia absoluta

Partimos de la frecuencia absoluta, que es el recuento de veces que aparecen las formas buscadas. Como la totalidad de palabras o letras en cada grupo difiere, la frecuencia absoluta no es utilizable. Sin embargo, la misma frecuencia es importante para conocer la realidad numérica de cada caso dentro del grupo, que es una información fundamental que debemos citar a la hora de presentar las siguientes frecuencias.

→	F.A.	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	14	24	10	8	1
2	#usted#	2	7	2	4	0
3	#ustedes#	7	9	4	6	0
4	#vos#	0	0	0	0	11
5	#vosotr[oa]s#	1	0	0	0	6

2. Frecuencia relativa de casos

La frecuencia relativa de casos se obtiene por la división de cada frecuencia absoluta por la totalidad de casos, *tú*, *usted*, *ustedes*, *vos* y *vosotr[oa]s*. Al aplicar el multiplicador 100, se obtiene el porcentaje horizontal:

→	F.R.Caso:100	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	24.6	42.1	17.5	14.0	1.8
2	#usted#	13.3	46.7	13.3	26.7	.0
3	#ustedes#	26.9	34.6	15.4	23.1	.0
4	#vos#	.0	.0	.0	.0	100.0
5	#vosotr[oa]s#	14.3	.0	.0	.0	85.7

3. Frecuencia relativa de variables

Se obtiene por la división de cada frecuencia absoluta entre la totalidad de variables, 1_Madrid, 2_Sevilla, ... Al aplicar el multiplicador 100, se obtiene el porcentaje vertical:

→	F.R.Atr.:100	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	58.3	60.0	62.5	44.4	5.6
2	#usted#	8.3	17.5	12.5	22.2	.0
3	#ustedes#	29.2	22.5	25.0	33.3	.0
4	#vos#	.0	.0	.0	.0	61.1
5	#vosotr[oa]s#	4.2	.0	.0	.0	33.3

4. Frecuencia normalizada: Letra

→	F.N.Lt.:100000	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	60.1	111.2	40.3	40.1	4.3
2	#usted#	8.6	32.4	8.1	20.1	.0
3	#ustedes#	30.1	41.7	16.1	30.1	.0
4	#vos#	.0	.0	.0	.0	47.4
5	#vosotr[oa]s#	4.3	.0	.0	.0	25.9

→	F.A.	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	@	23293	21590	24818	19928	23183

Al analizar letras, conviene utilizar la frecuencia [Por mil letras] para evaluar las cifras en la dimensión adecuada. Veamos qué frecuencia ofrece cada letra en distintas ciudades. Para este propósito, utilizamos el siguiente [Patrón]:

- a
 - b
 - c
 - d

Núm.	Patrón	Madrid	Sevilla	México	Lima	B.A.
1	a	125.574	128.763	125.111	123.244	122.288
2	b	13.996	13.664	13.377	12.796	13.760
3	c	35.805	35.850	39.487	39.392	39.598
4	d	41.987	41.825	40.414	42.252	41.798

Según el gráfico, la letra 'a' y su mayúscula 'A' tienen la frecuencia de 125.574 por mil letras, es decir, casi ocupa 12,5% de todas las letras ocurridas en el texto de Madrid.

5. Frecuencia normalizada: Palabra

→	F.N.PI.:10000	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	25.4	45.9	17.0	17.1	1.9
2	#usted#	3.6	13.4	3.4	8.5	.0
3	#ustedes#	12.7	17.2	6.8	12.8	.0
4	#vos#	.0	.0	.0	.0	20.9
5	#vosotr[oa]s#	1.8	.0	.0	.0	11.4

La frecuencia normalizada [Por mil palabras] se calcula con base en la frecuencia total de palabras ocurridas en los atributos tratados. Veamos cuántas palabras hay dentro de estos textos en cada ciudad con el patrón de "&":

→	F.A.	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	&	5502	5229	5897	4680	5270

Para comparar las frecuencias de cada patrón en las cinco ciudades, debemos dividir las frecuencias absolutas por estas frecuencias totales para llegar a las frecuencias normalizadas por mil palabras. Por ejemplo, tenemos 14 ocurrencias de *tú* en Madrid, donde se registran 5502 palabras en total, de modo que la frecuencia de *tú* normalizada por diez mil palabras es: $14 / 5502 * 10000 = 25.4$.

De esta manera podemos comparar estas frecuencias sobre la misma base. Y también podemos calcular la frecuencia [6. Por mil palabras (%)], que es el porcentaje horizontal de la frecuencia normalizada por mil palabras. En realidad, el [2. Frecuencia relativa de casos] no es fiable en el sentido de que están calculadas las frecuencias absolutas para ver sus proporciones horizontales, puesto que las bases de cada frecuencia son distintas. La [2. Frecuencia relativa de casos] sirve exclusivamente para ver la proporción que ocupa la frecuencia de cada forma, por ejemplo, *tú* en distintas ciudades, haciendo caso omiso de la suma de palabras que hay en textos de cada ciudad.

Ahora la [5. Frecuencia normalizada] es más fiable y representativa de la realidad.

6. Frecuencia normalizada: Documentos

El caso de VARITEX es especial en mantener la misma estructura de 10 documentos (10 temas) en cada ciudad, de modo que la frecuencia normalizada por 10 documentos coincide con la frecuencia absoluta:

→	FN.Doc.:10	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	tú	14.0	24.0	10.0	8.0	1.0
2	usted	2.0	7.0	2.0	4.0	0.0
3	ustedes	7.0	9.0	4.0	6.0	0.0
4	vos	0.0	0.0	0.0	0.0	11.0
5	vosotros	1.0	0.0	0.0	0.0	6.0

7. Frecuencia probabilística: Suma

Para comparar las frecuencias relativas con bases muy diferentes, conviene utilizar las frecuencias probabilísticas (F.P.), que utilizan la significatividad del 99% en la probabilidad binomial. El corpus de VARITEX tiene la propiedad de ofrecer casi la misma cantidad de palabras, puesto que es la traducción del mismo texto a la variedad de cada ciudad. En este caso no es necesario utilizar la frecuencia probabilística. La tabla siguiente sirve para ver su característica, que suele dar más énfasis a las cifras importantes, mientras que se disminuye la cantidad en las cifras poco importantes.

→	F.P.Sm.:100	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	33.2	39.9	31.3	18.4	.1
2	#usted#	.6	6.1	1.0	4.9	.0
3	#ustedes#	10.6	9.3	5.5	11.0	.0
4	#vos#	.0	.0	.0	.0	31.9
5	#vosotr[oa]s#	.0	.0	.0	.0	11.0

8. Frecuencia probabilística: Letra

→	F.P.Lt.:100000	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	29.1	65.3	16.6	14.6	.0
2	#usted#	.6	10.8	.6	4.1	.0
3	#ustedes#	10.0	16.2	3.3	9.0	.0
4	#vos#	.0	.0	.0	.0	20.6
5	#vosotr[oa]s#	.0	.0	.0	.0	7.7

9. Frecuencia probabilística: Palabra

→	F.P.Pl.:100000	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	123.3	269.7	70.1	62.1	.2
2	#usted#	2.7	44.6	2.5	17.6	.0
3	#ustedes#	42.4	67.1	14.0	38.2	.0
4	#vos#	.0	.0	.0	.0	90.6
5	#vosotr[oa]s#	.2	.0	.0	.0	33.9

10. Puntuación típica: Caso

$$PTC = (f - m.c) / dt.c$$

donde f: frecuencia, m.c.: media de caso, dt.c.: desviación típica de caso.

→	P.T.Caso	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	.343	1.663	-.185	-.449	-1.372
2	#usted#	-.423	1.690	-.423	.423	-1.268
3	#ustedes#	.588	1.242	-.392	.261	-1.700
4	#vos#	-.500	-.500	-.500	-.500	2.000
5	#vosotr[oa]s#	-.171	-.600	-.600	-.600	1.972

11. Puntuación típica: Variable

$$PTV = (f - m.v) / dt.v$$

donde f: frecuencia, m.v.: media de variable, dt.v.: desviación típica de variable

→	P.T.Atr.	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	1.772	1.821	1.833	1.375	-.602
2	#usted#	-.539	-.114	-.323	.125	-.834
3	#ustedes#	.424	.114	.216	.750	-.834
4	#vos#	-.924	-.911	-.863	-1.125	1.714
5	#vosotr[oa]s#	-.732	-.911	-.863	-1.125	.556

12. Rango ascendente: Caso

→	R.Asc.Caso	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	4	5	3	2	1
2	#usted#	2	5	2	4	1
3	#ustedes#	4	5	2	3	1
4	#vos#	1	1	1	1	5
5	#vosotr[oa]s#	4	1	1	1	5

13. Rango descendente: Caso

→	R.Des.Caso	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#		2	1	3	4
2	#usted#		3	1	3	2
3	#ustedes#		2	1	4	3
4	#vos#		2	2	2	2
5	#vosotr[oa]s#		2	3	3	3

14. Rango ascendente: Atributo

→	R.Asc.Atr.	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#		5	5	5	5
2	#usted#		3	3	3	3
3	#ustedes#		4	4	4	4
4	#vos#		1	1	1	1
5	#vosotr[oa]s#		2	1	1	1

15. Rango descendente: Atributo

→	R.Des.Atr.	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#		1	1	1	1
2	#usted#		3	3	3	3
3	#ustedes#		2	2	2	2
4	#vos#		5	4	4	4
5	#vosotr[oa]s#		4	4	4	4

16. Frecuencia 0:1

Las frecuencias de más de 1 se convierten en 1.

→	Frec.0:1	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#		1	1	1	1
2	#usted#		1	1	1	1
3	#ustedes#		1	1	1	1
4	#vos#		0	0	0	0
5	#vosotr[oa]s#		1	0	0	0

17. Logaritmo (Base 2)

→	Log.2	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#		3.807	4.585	3.322	3.000
2	#usted#		1.000	2.807	1.000	2.000
3	#ustedes#		2.807	3.170	2.000	2.585
4	#vos#		.000	.000	.000	.000
5	#vosotr[oa]s#		.000	.000	.000	.000

18. Logaritmo (Base 10)

→	Log.10	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	1.146	1.380	1.000	.903	.000
2	#usted#	.301	.845	.301	.602	.000
3	#ustedes#	.845	.954	.602	.778	.000
4	#vos#	.000	.000	.000	.000	1.041
5	#vosotr[oa]s#	.000	.000	.000	.000	.778

19, 20. (N-C) / (N+C): CREA, CORPES

La fórmula $(N-C) / (N+C)$ se aplica para comparar la frecuencia normalizada por millón de palabras (N) con la misma frecuencia ofrecida por el corpus CREA¹ y el corpus CORPES² de Real Academia Española. Esta fórmula tiene la propiedad de ser 1, cuando C es 0 y la de ser -1 cuando N es 0. Cuando las dos cifras coinciden, devuelve 0. Veamos la frecuencia normalizada por un millón de palabras (N):

→	FN.PL.:1000000	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	tú	2543.1	4588.9	1695.5	1709.4	189.7
2	usted	363.3	1338.4	339.1	854.7	0.0
3	ustedes	1271.6	1720.8	678.2	1282.1	0.0
4	vos	0.0	0.0	0.0	0.0	2086.9
5	vosotros	181.7	0.0	0.0	0.0	1138.3

Como ejemplo del cálculo extraemos el caso de *tú* en Madrid:

$$(N-C)/(N+C) = (2543.1 - 263.0) / (2543.1 + 263.0) = 0.813$$

→	(N-C)/(N+C)	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	tú	0.813	0.892	0.731	0.733	-0.162
2	usted	0.051	0.606	0.016	0.445	-1.000
3	ustedes	0.901	0.926	0.821	0.901	-1.000
4	vos	-1.000	-1.000	-1.000	-1.000	0.965
5	vosotros	0.831	-1.000	-1.000	-1.000	0.971

21. Coocurrencia

Nos interesa saber si existe la relación coocurrente entre dos formas, por ejemplo, palabras con *ff-* y *ss-* iniciales. La tabla siguiente muestra la frecuencia de las formas en las dos Castillas, la Vieja y la Nueva, desde 1200 hasta 1799, en el corpus CODEA:

¹ <http://corpus.rae.es/creanet.html> [14 de julio,2018]

² <https://www.rae.es/recursos/banco-de-datos/corpes-xxi> [14 de julio,2018]

→	FA	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
1	#ff%	23	938	1083	340	40	5	5	29	15	10	0	2
2	#ss%	0	671	1039	418	37	2	4	24	68	42	20	23

Convertimos la tabla en el tipo 0/1 y la fila 3 (*Cooc.) representa el valor mínimo de las dos frecuencias, que corresponde a la Coocurrencia. La última fila (* Trans.), transición, es derivada de sustracción de la primera fila - la segunda fila:

→	Coocurrencia	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
1	#ff%	1	1	1	1	1	1	1	1	1	1	0	1
2	#ss%	0	1	1	1	1	1	1	1	1	1	1	1
3	* Cooc.	0	1	1	1	1	1	1	1	1	1	0	1
4	* Trans.	-1	0	0	0	0	0	0	0	0	0	1	0

22. Frecuencia coocurrente

También es interesante observar la parte común de las dos cifras en las frecuencias absolutas:

→	Frec. cooc.	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
1	#ff%	23	938	1083	340	40	5	5	29	15	10	0	2
2	#ss%	0	671	1039	418	37	2	4	24	68	42	20	23
3	* Cooc.	0	671	1039	340	37	2	4	24	15	10	0	2
4	* Trans.	-23	-267	-44	78	-3	-3	-1	-5	53	32	20	21

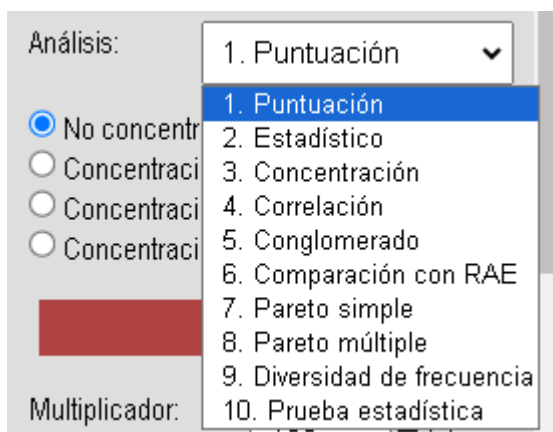
23. Frecuencia coocurrente normalizada

Para evitar el sesgo causado por la diferencia de base de cada franja cronológica, calculamos la frecuencia normalizada por 10 000 palabras:

→	Frec. cooc. nrm.:10000	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
1	#ff%	21.9	149.3	231.6	76.9	7.1	0.8	0.6	2.7	3.2	1.7	0.0	0.7
2	#ss%	0.0	106.8	222.2	94.5	6.6	0.3	0.5	2.3	14.6	7.0	9.4	8.3
3	* Cooc.	0.0	106.8	222.2	76.9	6.6	0.3	0.5	2.3	3.2	1.7	0.0	0.7
4	* Trans.	-21.9	-42.5	-9.4	17.6	-0.5	-0.5	-0.1	-0.5	11.4	5.3	9.4	7.5

3.1.2. Análisis

En la sección de Análisis se encuentran las opciones siguientes:



1. Puntuación

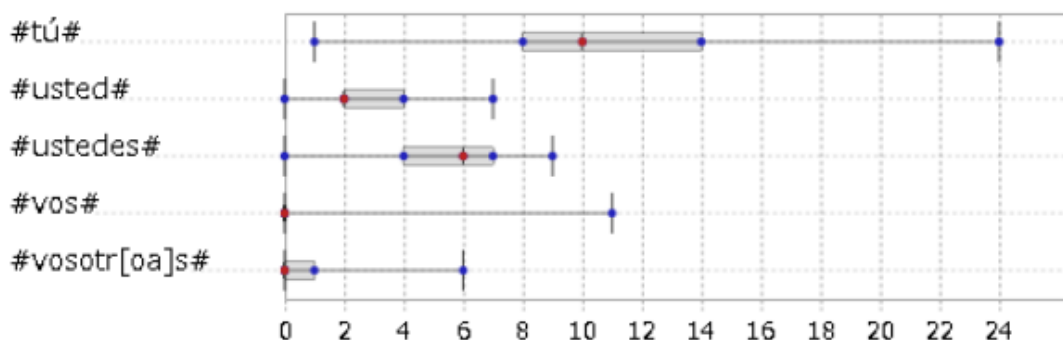
Hemos explicado anteriormente las distintas puntuaciones en la sección de Matriz. A continuación, explicamos las dos opciones restantes, donde se utilizan las puntuaciones seleccionadas.

2. Estadístico

Un valor estadístico resume la característica del conjunto de la columna o de la fila. Para ver los valores de fila, hay que transponer la matriz de input. Podemos seleccionar varios valores al mismo tiempo dentro de los 26 valores ofrecidos en la casilla:

→	FA	Suma	Media aritmética	Media geométrica	Mínimo	Máximo	Desv. típica	DT geo.
1	#ú#	57.0	11.4	7.7	1.0	24.0	7.6	3.0
2	#usted#	15.0	3.0	1.6	0.0	7.0	2.4	4.3
3	#ustedes#	26.0	5.2	2.7	0.0	9.0	3.1	5.3
4	#vos#	11.0	2.2	0.3	0.0	11.0	4.4	6.6
5	#vosotr[oa]s#	7.0	1.4	0.4	0.0	6.0	2.3	5.3

En el gráfico de caja, visualizamos los valores de media, mediana, moda, mitad en las áreas limitadas de primer cuartil y tercer cuartil:



* La moda no es adecuada para los datos no clasificados, de modo que en su lugar utilizamos la moda mayor.

3. Concentración

Por medio del análisis de concentración la matriz de input (Fig. 1) se convierte en la matriz diagonal concentrada (Fig. 2) por el cambio de orden tanto de filas como de columnas:

→	Gnp	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	14	24	10	8	1
2	#usted#	2	7	2	4	0
3	#ustedes#	7	9	4	6	0
4	#vos#	0	0	0	0	11
5	#vosotr[oa]s#	1	0	0	0	6

Fig. 1

→	FA	2_Sevilla	1_Madrid	3_México	4_Lima	5_B.A.
1	#tú#	24	14	10	8	1
2	#ustedes#	9	7	4	6	0
3	#usted#	7	2	2	4	0
4	#vosotr[oa]s#	0	1	0	0	6
5	#vos#	0	0	0	0	11

Fig. 2

Hay dos métodos de concentración: Correspondencia y Distancia. Se puede cambiar el orden de filas y columnas (modo bilateral) y, si se necesita, solo el de filas o el de columnas (modo unilateral).

Método:

Correspondencia

Distancia

Concentración:

Bilateral: Casos y variables

Unilateral: Casos

Unilateral: Variables

No concentración

Ajustes

A continuación, aparecen los siguientes botones de selección y opción:

Círculo graduado / Vector

División / Área concentrada

Número de Div. / Con.: ▲:▼

Escala graduada

Tamaño de icono: ▲:▼

Círculo graduado: Se gradúa el tamaño del círculo de acuerdo con la cantidad de frecuencia.

Vector: Aparecen los valores de los vectores horizontal y vertical.

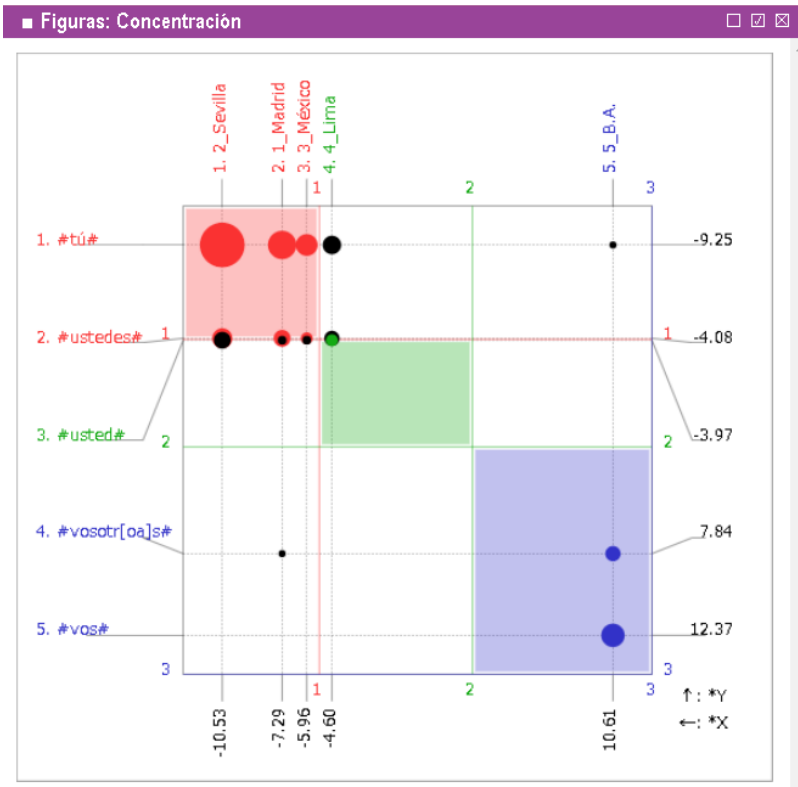
División: Aborda la función de dividir el área en cuantas divisiones se desee. Al seleccionar la [Escala graduada] sale la escala en forma graduada. Por defecto, las escalas son equidistantes.

Área concentrada: En el gráfico del [Área concentrada] las filas y las columnas se sitúan según el valor que posee cada caso y cada atributo. El valor de la matriz se representa con color azul. Las áreas más concentradas (pobladas) se rellenan con el color gris, azul, verde, rojo, por orden de su población relativa.

Número de Div(isión) y Con(centración)

Escala graduada: Se sitúan las escalas, horizontal y vertical, de acuerdo con los valores que hay en los vectores.

Tamaño de icono: Varía de acuerdo con la cantidad de frecuencia.



4. Correlación

Se calculan las correlaciones que hay entre los atributos y entre los casos (con los datos transpuestos). En el gráfico se ofrecen tres opciones: tabla en color, tabla de tamaño y gráficos de esparcimiento con ejes selectivos:

n.	Rpp	A	B	C	D	E
1	A	1.000	.787	.944	-.480	.436
2	B	.787	1.000	.945	-.092	.896
3	C	.944	.945	1.000	-.331	.709
4	D	-.480	-.092	-.331	1.000	.140
5	E	.436	.896	.709	.140	1.000

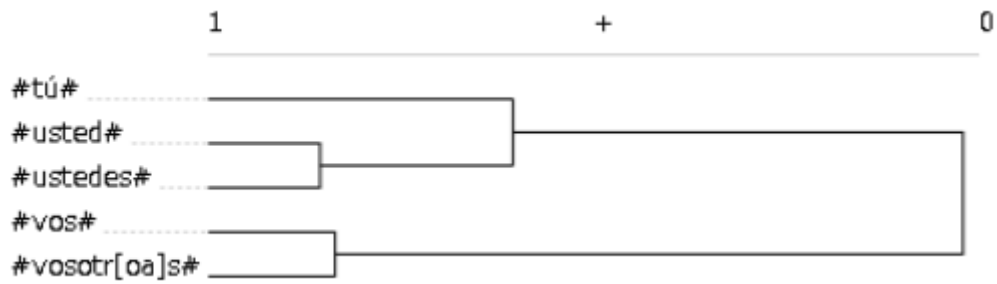
Tabla en color / Tabla de tamaño

5. Conglomerado (cluster)

Se produce el gráfico de dendrograma, resultado del análisis de conglomerado (*cluster*) jerárquico con el método de promedio:

Análisis: 5. Conglomerado

Casos
 Atributos
 Ambos



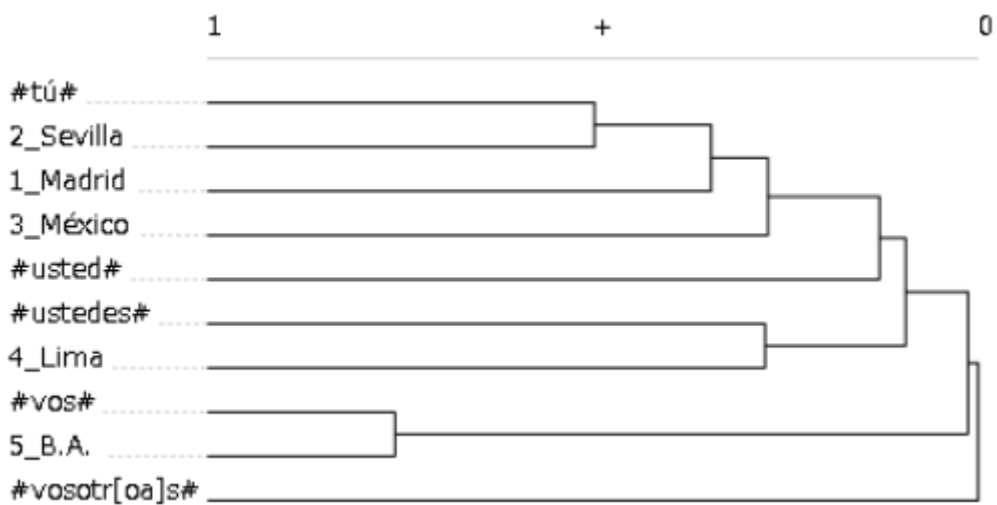
Análisis: 5. Conglomerado ▾

Casos Atributos Ambos



Análisis: 5. Conglomerado ▾

Casos Atributos Ambos



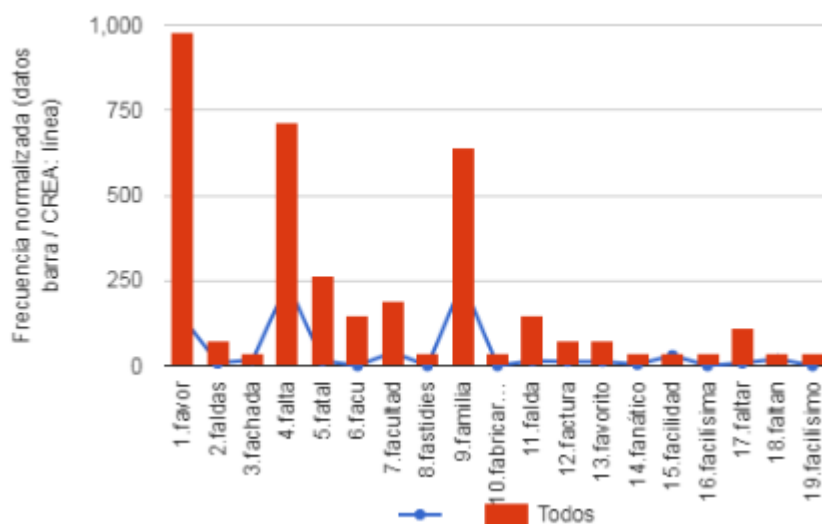
Aproximar líneas



6. Comparación con CREA y CORPES

Comparamos la frecuencia normalizada de las formas buscadas en el texto y la frecuencia normalizada que ofrecen el corpus CREA y el corpus CORPES de la Real Academia Española.

→	Comp.CREA	Todos	Suma N.	CREA	N-C	(N-C)/(N+C)
1	favor	978.0	978.0	143.0	835.0	0.745
2	faldas	75.2	75.2	7.6	67.7	0.817
3	fachada	37.6	37.6	17.1	20.6	0.376
4	falta	714.7	714.7	243.8	470.9	0.491
5	fatal	263.3	263.3	14.3	249.0	0.897
6	facu	150.5	150.5	0.0	150.5	1.000
7	facultad	188.1	188.1	38.9	149.2	0.657
8	fastidies	37.6	37.6	0.0	37.6	1.000
9	familia	639.5	639.5	245.0	394.5	0.446
10	fabricarnos	37.6	37.6	0.0	37.6	1.000
11	falda	150.5	150.5	13.9	136.5	0.830
12	factura	75.2	75.2	12.3	62.9	0.719
13	favorito	75.2	75.2	12.0	63.2	0.725
14	fanático	37.6	37.6	3.8	33.8	0.818
15	facilidad	37.6	37.6	30.8	6.8	0.100
16	facilísima	37.6	37.6	0.0	37.6	1.000
17	faltar	112.8	112.8	7.8	105.1	0.871
18	faltan	37.6	37.6	19.5	18.1	0.316
19	facilísimo	37.6	37.6	0.0	37.6	1.000

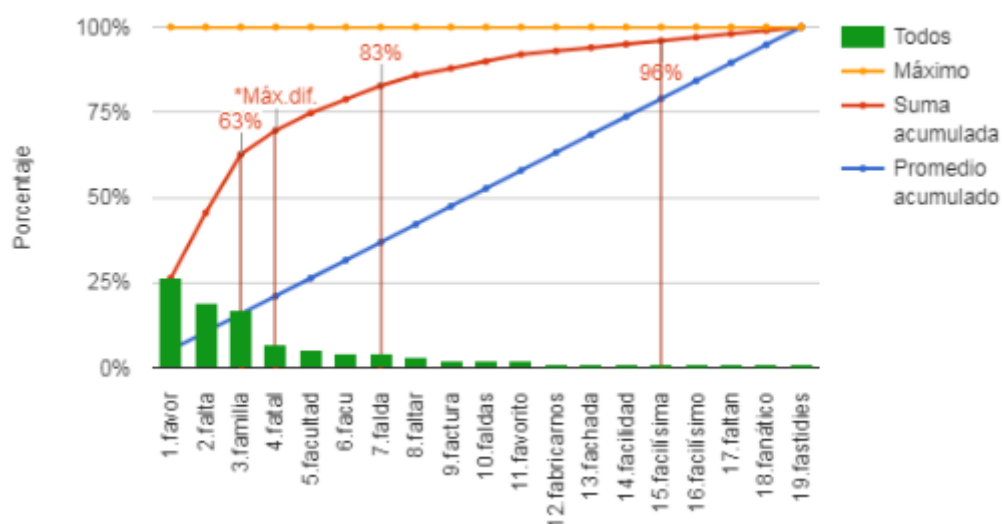


7. Análisis de Pareto simple

Véase *Análisis de datos cuantitativos para estudios lingüísticos*, 5.9.6.

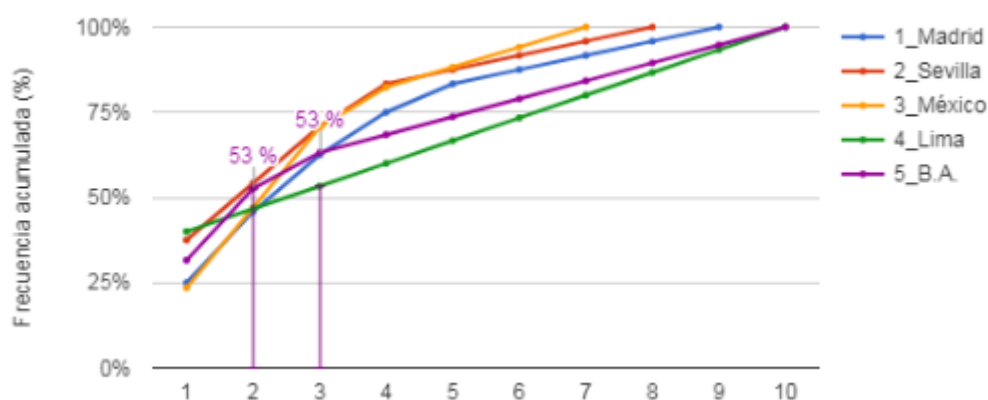
Pareto	Total
Índice de Pareto	.631
M: Punto de mitad (50%)	3
N: Número de datos	19
M / N (%)	16 %
A: Acumulación en mitad	62
T: Frecuencia total	99
A / T (%)	63 %
Punto de 50%	3
Punto de 80%	7
Punto de 95%	15

Rango	Total
1.favor	26 (26 %)
2.falta	19 (19 %)
3.familia	17 (17 %)
4.fatal	7 (7 %)
5.facultad	5 (5 %)
6.facu	4 (4 %)
7.falda	4 (4 %)
8.faltar	3 (3 %)
9.factura	2 (2 %)
10.faldas	2 (2 %)
11.favorito	2 (2 %)
12.fabricarnos	1 (1 %)
13.fachada	1 (1 %)
14.facilidad	1 (1 %)
15.facilísima	1 (1 %)
16.facilísimo	1 (1 %)
17.faltan	1 (1 %)
18.fanático	1 (1 %)
19.fastidies	1 (1 %)



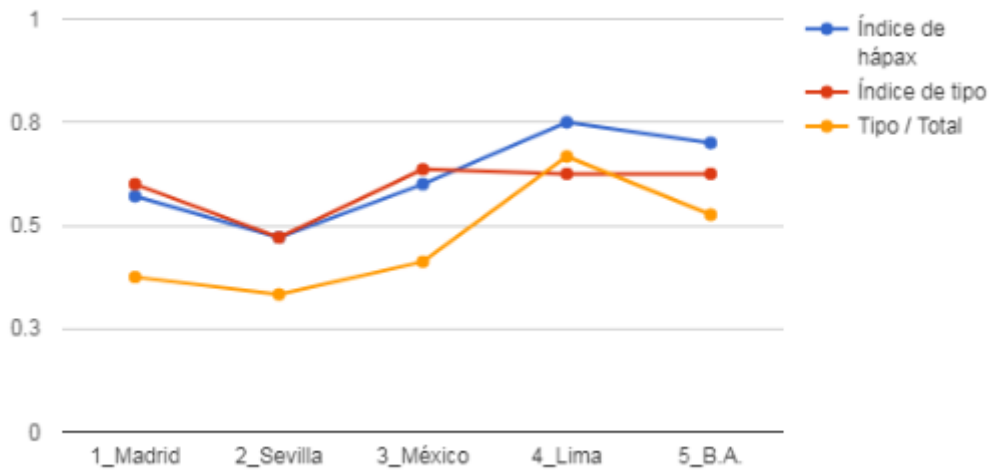
8. Análisis de Pareto múltiple

Pareto	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
Índice de Pareto	.833	.876	.871	.778	.805
M: Punto de mitad (50%)	3	2	3	3	2
N: Número de datos	9	8	7	10	10
M / N (%)	33 %	25 %	43 %	30 %	20 %
A: Acumulación en mitad	15	13	12	8	10
T: Frecuencia total	24	24	17	15	19
A / T (%)	63 %	54 %	71 %	53 %	53 %
Punto de 50%	3	2	3	3	2
Punto de 80%	5	4	4	8	7
Punto de 95%	8	7	7	10	10



9. Diversidad de frecuencia

→ Diversidad de frecuencia	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.	*Resetear*
1 Total	24	24	17	15	19	Total
2 Tipo	9	8	7	10	10	Tipo
3 Máximo	6	9	4	6	6	Máximo
4 Hápx	4	4	3	9	7	Hápx
5 Tipo / Total	0.375	0.333	0.412	0.667	0.526	Tipo / Total
6 Índice de tipo	0.600	0.471	0.636	0.625	0.625	Índice de tipo
7 Índice de hápx	0.571	0.471	0.600	0.750	0.700	Índice de hápx
Resetear	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.	



3.2. Opción

En la parte inferior de la tabla desplegada, se encuentran varias opciones. En primer lugar, veamos [Multiplicador] y [Dígito decimal]. Con el multiplicador, se multiplican las cifras obtenidas del cálculo. Por ejemplo el multiplicador 100 se utiliza para derivar el porcentaje a partir de las cifras decimales. Por [Dígito decimal], especificamos el número de dígitos decimales:

Multiplicador: ▲:▼

Dígito decimal: ▲:▼

→	FR.H:100	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#	24.6	42.1	17.5	14.0	1.8
2	#usted#	13.3	46.7	13.3	26.7	0.0
3	#ustedes#	26.9	34.6	15.4	23.1	0.0
4	#vos#	0.0	0.0	0.0	0.0	100.0
5	#vosotr[oa]s#	14.3	0.0	0.0	0.0	85.7

Grupo

Los casos y atributos pueden ser clasificados en grupos por la similitud de distribución de cifras. La cifra del grupo puede representar tanto la frecuencia total del grupo como el promedio de las frecuencias de miembros del grupo:

Grupos de caso: ▲ : ▼

Grupos de atributo: ▲ : ▼

Total del grupo / Promedio del grupo

Al tratarse de los datos dotados de numerosos casos y variables, conviene juntar ambos en los grupos para observar la totalidad de los datos del output.

El dato siguiente (D3) contiene 47 casos y 15 atributos:

■ Data. File: D3 / Lines:47

Line	N.D3.farmer	MX	CU	RD	PR	C5	PN	VE	CO	EC	PE	BO	CH	PA	UR	AR
1	01 cacahuero	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
2	02 cafetalista	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
3	03 camilucho	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
4	04 campero	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
5	05 camperuso	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
6	06 campirano	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0
7	07 campiruso	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
8	08 campista	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0

Juntamos los casos en 10 grupos y los atributos en 5 grupos por el análisis de *cluster* (no jerárquico, método de promedio):

→	Gnp	[A.1]	[A.2]	[A.3]	[A.4]	[A.5]
1	C.01	2	7	25	3	0
2	C.02	1	2	0	0	0
3	C.03	1	0	0	6	34
4	C.04	14	1	1	0	0
5	C.05	2	13	0	0	0
6	C.06	0	1	2	10	3
7	C.07	1	0	0	2	1
8	C.08	1	0	0	0	0
9	C.09	2	2	0	0	0
10	C.10	0	0	0	1	0

Grupo de casos

Grupo	Caso
C.01	01 cacahuero; 05 camperuso; 06 campirano; 13 conuquero; 15 chagrero; 25 huasicama; 31 lanudo; 32 llanero; 36 montubio; 37 paisano; 38 pajuerano; 40 payazo; 44 sabanero;
C.02	02 cafetalista;
C.03	03 camilucho; 04 campero; 09 campusano; 12 comparsa; 16 changador; 18 chuncano; 20 estanciero; 21 gaucho; 28 invernador; 39 partidario; 41 pionero; 46 viñatero;
C.04	07 campiruso; 08 campista; 10 campuso; 17 chilero; 23 guanaco; 27 hulero;
C.05	11 colono; 19 enmaniguado; 22 guajiro; 29 jibaro; 35 montero; 42 rancho;
C.06	14 coquero; 24 guaso; 30 lampero; 47 yanacón;
C.07	26 huertero;
C.08	33 macanero;
C.09	34 manuto; 45 veguero;
C.10	43 rondín;

Grupo de atributos

Grupo	Atributo
[A.1]	MX; C5; PN;
[A.2]	CU; RD; PR;
[A.3]	VE; CO; EC;
[A.4]	PE; BO; CH;
[A.5]	PA; UR; AR;

Selecciones

A continuación, se ofrecen varias selecciones:

- 1. Ordenación horizontal
- 2. Ordenación vertical
- 3. Transponer matriz de input
- 4. Transponer matriz de output
- 5. Fila-columna cero
- 6. Ponderación proporcional (*)
- 7. Relativización total (*)
- 8. Número secuencial con C y V (*)
- 9. Agregar frecuencia absoluta
- 10. Agregar total
- 11. Representar cero (0)

Con [Transponer], se transponen el eje vertical y el horizontal de la matriz de la tabla:

Ventana / Copiar con formato / Copiar sin formato

→	Gnp	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#ú#	58.3	60.0	62.5	44.4	5.6
2	#usted#	8.3	17.5	12.5	22.2	0.0
3	#ustedes#	29.2	22.5	25.0	33.3	0.0
4	#vos#	0.0	0.0	0.0	0.0	61.1
5	#vosotr[oa]s#	4.2	0.0	0.0	0.0	33.3

→ [Transponer matriz de input]

→	Gnp	#ú#	#usted#	#ustedes#	#vos#	#vosotr[oa]s#
1	1_Madrid	58.3	8.3	29.2	0.0	4.2
2	2_Sevilla	60.0	17.5	22.5	0.0	0.0
3	3_México	62.5	12.5	25.0	0.0	0.0
4	4_Lima	44.4	22.2	33.3	0.0	0.0
5	5_B.A.	5.6	0.0	0.0	61.1	33.3

Primeros casos y primeros atributos

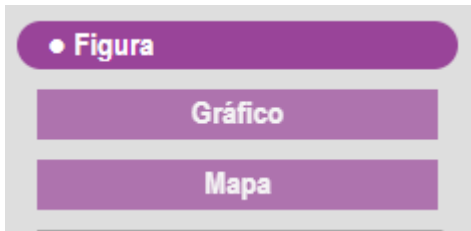
Primeros casos:

Primeros atributos:

Si el usuario quiere analizar solo los casos y/o atributos más frecuentes, puede utilizar [Primeros casos] y [Primeros atributos].

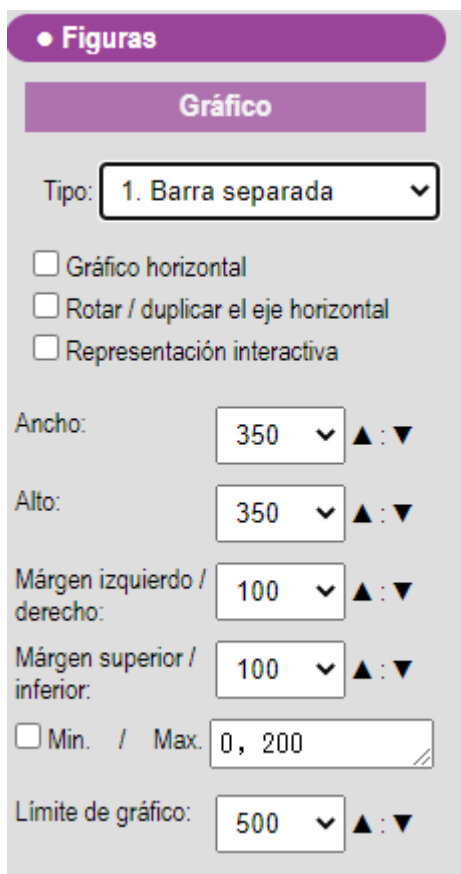
4. Figuras

Las Figuras son útiles para obtener una imagen general de tendencias. La sección de [Figuras] está dividida en [Gráfico] y [Mapa]:



4.1. Gráfico

Dentro de [Gráfico] se encuentran varias selecciones, desde [Tipo] hasta Límite de gráfico. Estas selecciones aparecen y desaparecen según el Tipo (del gráfico) que se seleccione:

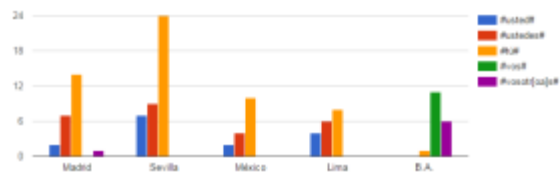


Veamos las frecuencias de los pronombres personales de sujeto en los textos de las cinco ciudades en distintos [Tipos] del gráfico:

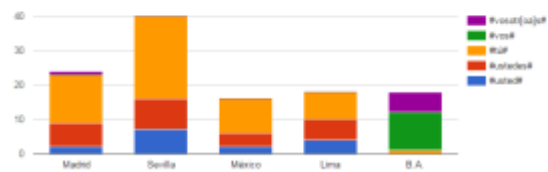
#tú#
 #usted#
 #ustedes#
 #vos#
 #vosotr[oa]s#:

Núm.	Patrón	Madrid	Sevilla	México	Lima	B.A.
3	#tú#	14	24	10	8	1
1	#usted#	2	7	2	4	0
2	#ustedes#	7	9	4	6	0
4	#vos#	0	0	0	0	11
5	#vosotr[oa]s#	1	0	0	0	6

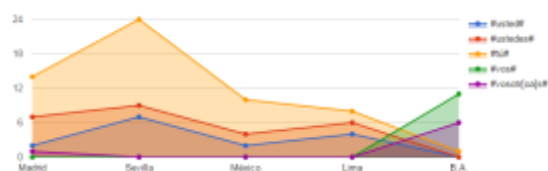
[Barra separada]:



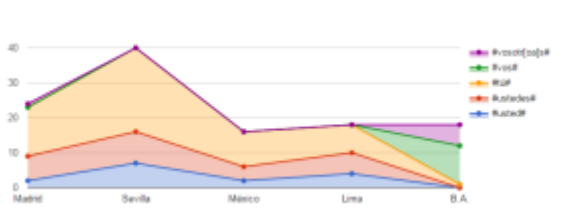
[Barra apilada]:



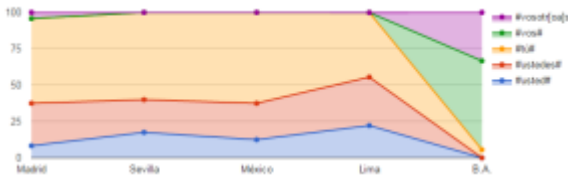
[Área separada]:



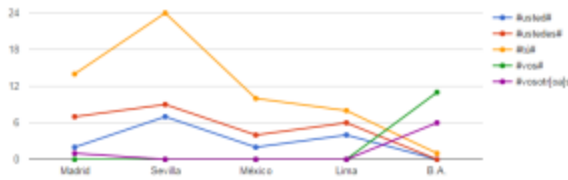
[Área apilada]:



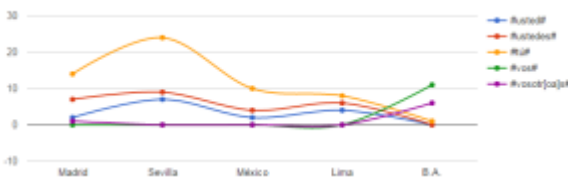
Al seleccionar el [Porcentaje (%) vertical] en la [Tabla] y el [Área apilada] obtenemos el gráfico siguiente:



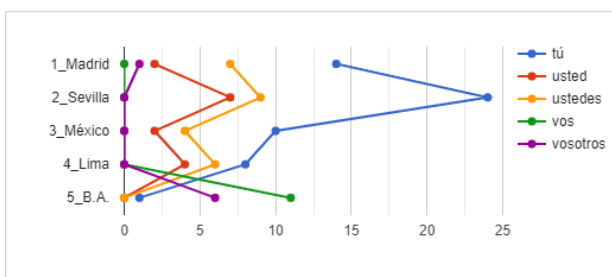
[Línea]:

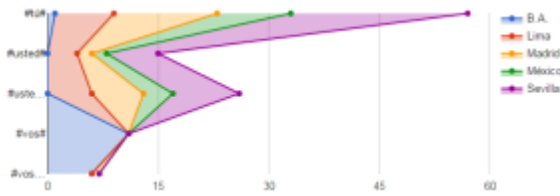


[Curva]:

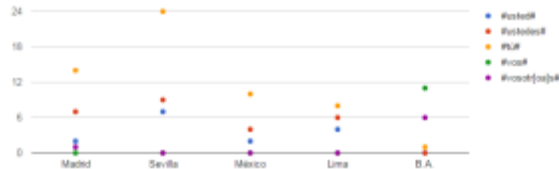


Con [Gráfico horizontal], se convierte la orientación del gráfico:

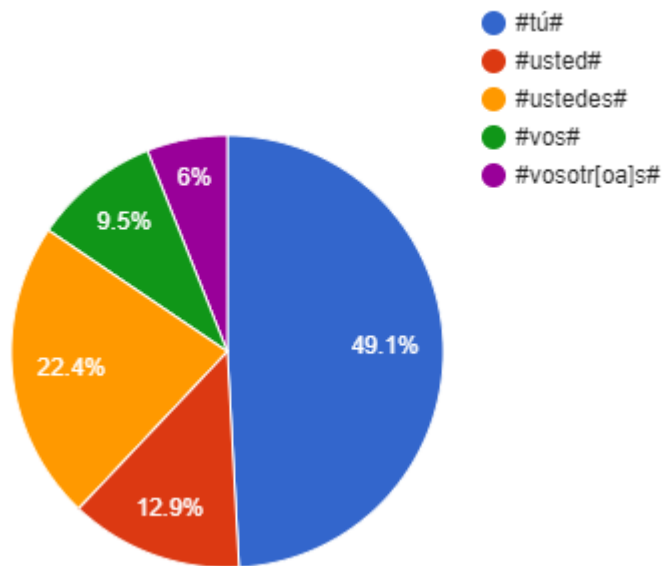




[Esparcimiento]:



[Tarta]:



Para los gráficos anteriores hemos utilizado *Google Chart*, que se ofrece en: <https://developers.google.com/chart/>.

Por otra parte, hemos elaborado unos programas para generar distintos gráficos, que creemos útiles para visualizar la distribución de cifras.

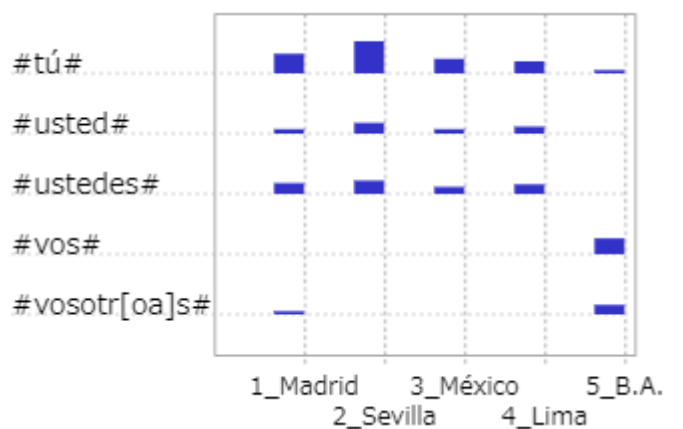
[Tabla de color]

n.	FA	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	14	24	10	8	1
2	#usted#	2	7	2	4	0
3	#ustedes#	7	9	4	6	0
4	#vos#	0	0	0	0	11
5	#vosotr[oa]s#	1	0	0	0	6

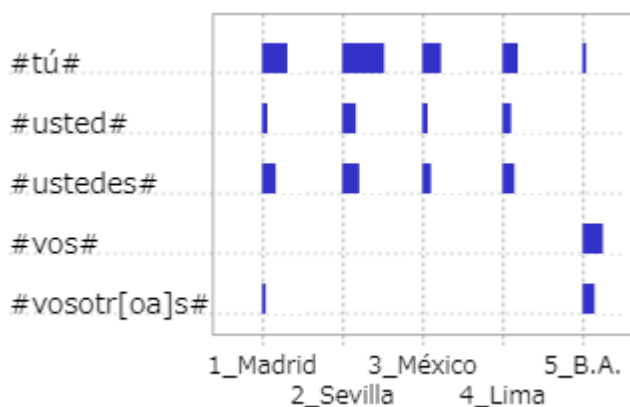
[Tabla de tamaño]

n.	FA	1_Madrid	2_Sevilla	3_México	4_Lima	5_B.A.
1	#tú#	14	24	10	8	1
2	#usted#	2	7	2	4	-
3	#ustedes#	7	9	4	6	-
4	#vos#	-	-	-	-	11
5	#vosotr[oa]s#	1	-	-	-	6

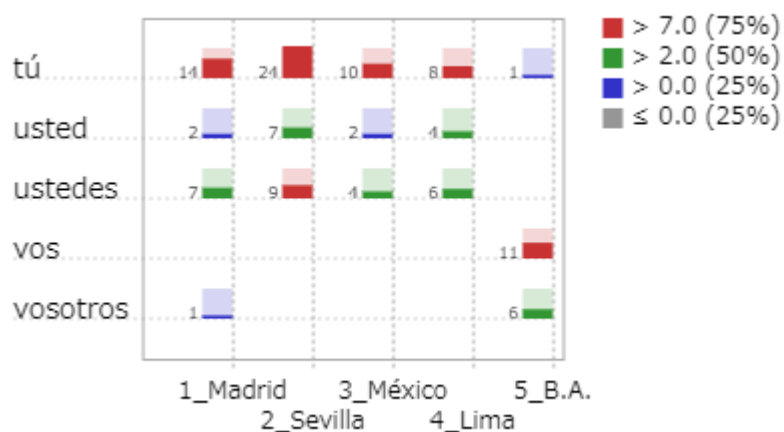
[Tamaño vertical]



[Tamaño horizontal]



[Tamaño y color vertical]

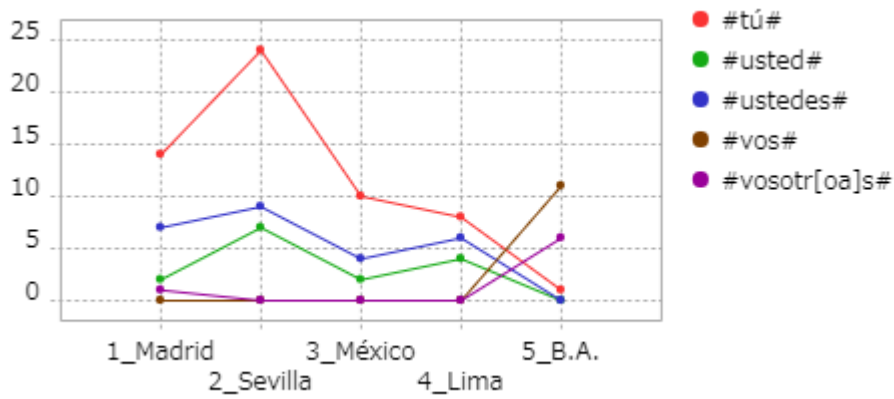


En este gráfico, todas las cifras se clasifican en cuatro cuartiles: [0, 25%] (25%, 50%), (50%, 75%), (75%, 100%]. En cada cuartil, podemos apreciar la cantidad relativa en barras. Lo mismo en el gráfico de [Tamaño y color horizontal].

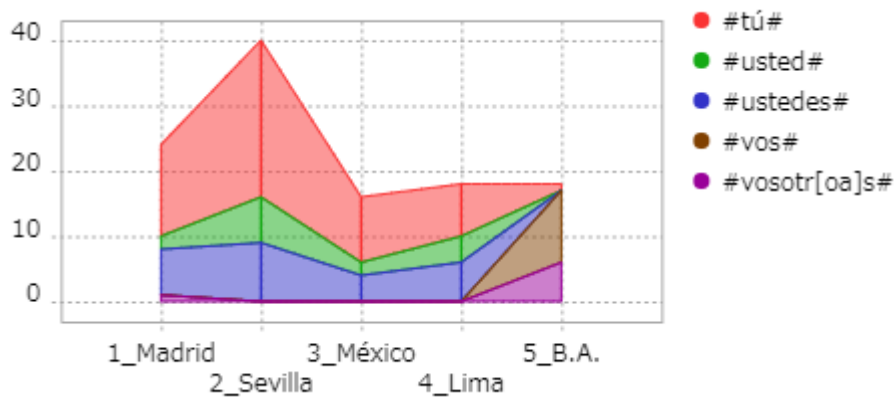
[Tamaño y color horizontal]



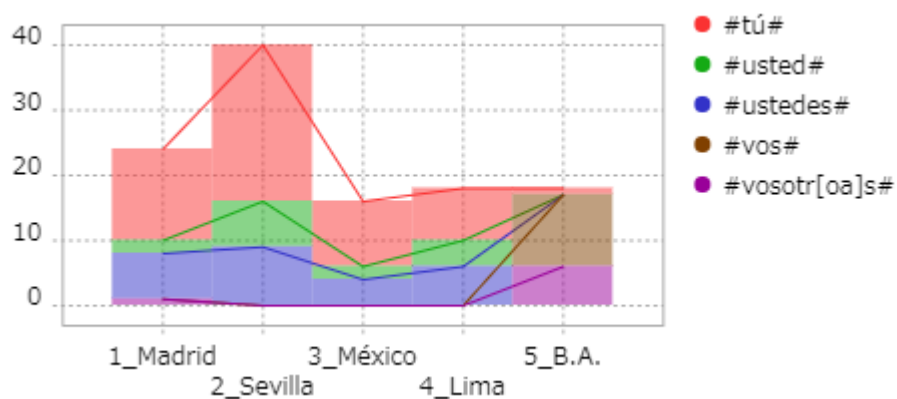
[Línea+]



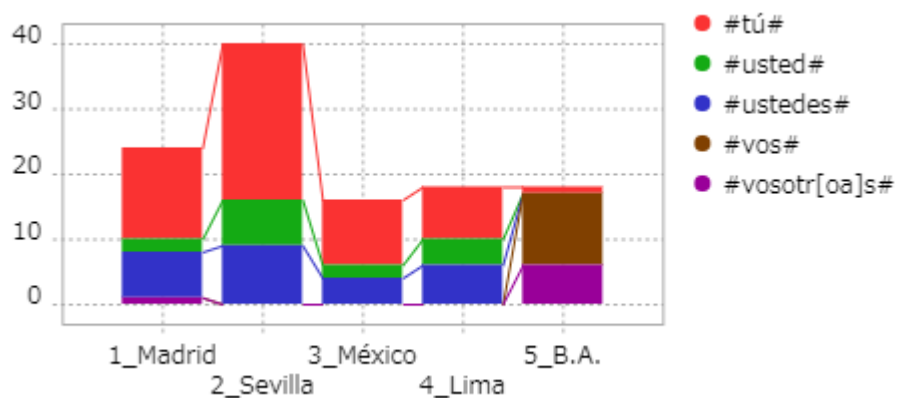
[Área+]



[Barra apilada continua+]



[Barra apilada separada+]



4.2. Mapa

4.2.1. Procesos generales

En los datos geográficos, al seleccionar atributos geográficos, sale la opción del [Mapa] en la sección de [Figuras]. En la opción del [Mapa] se encuentran las siguientes opciones y selecciones. Para ofrecer mapas de alta calidad, hemos utilizado el sistema Leaflet y distintos datos de mapas:

<https://leafletjs.com/>

<https://github.com/Leaflet/Leaflet.Graticule>

Datos: PRESEEA

Patrón de búsqueda:

¿cacháis\?

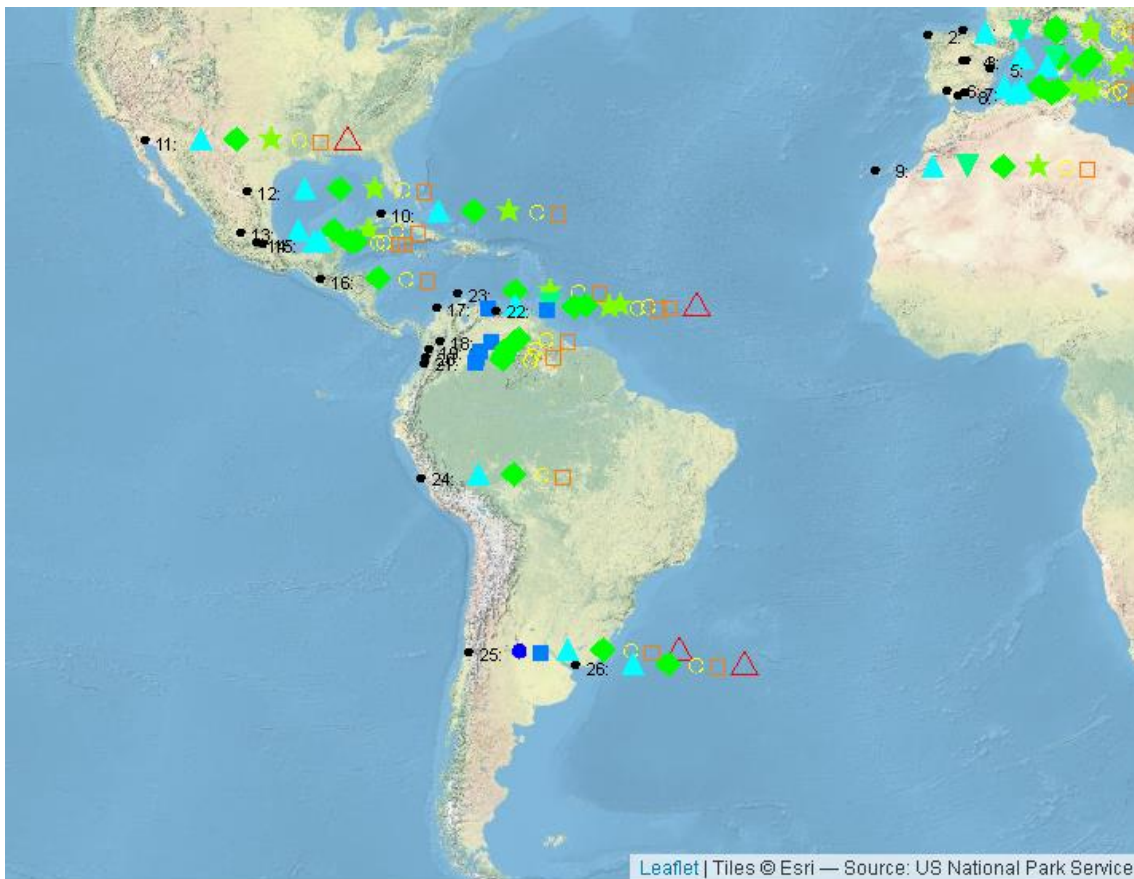
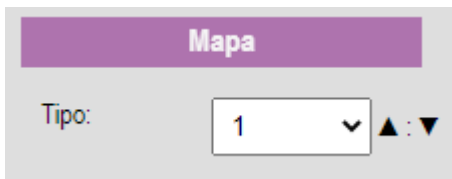
¿cierto\?

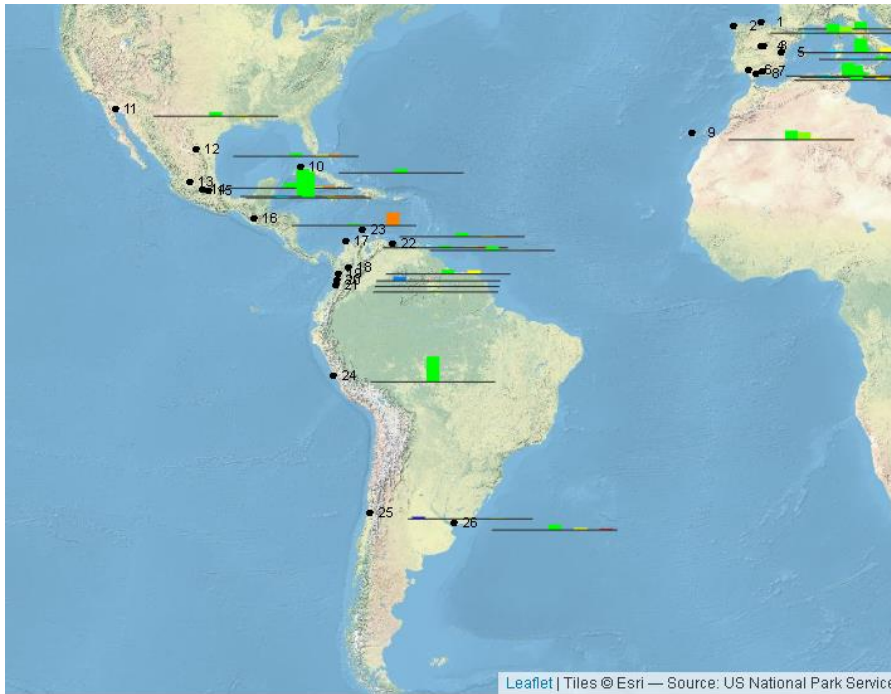
¿eh\
¿hm\
¿no\
¿sabes\
¿sí\
¿verdad\
¿viste\

[Letras] → [Casos y variables] → [Variables]: 1.LUGAR

<https://h-ueda.sakura.ne.jp/kenkyu/chiri/forma-colombiana/forma-colombiana.pdf>

(1) Tipo de mapa





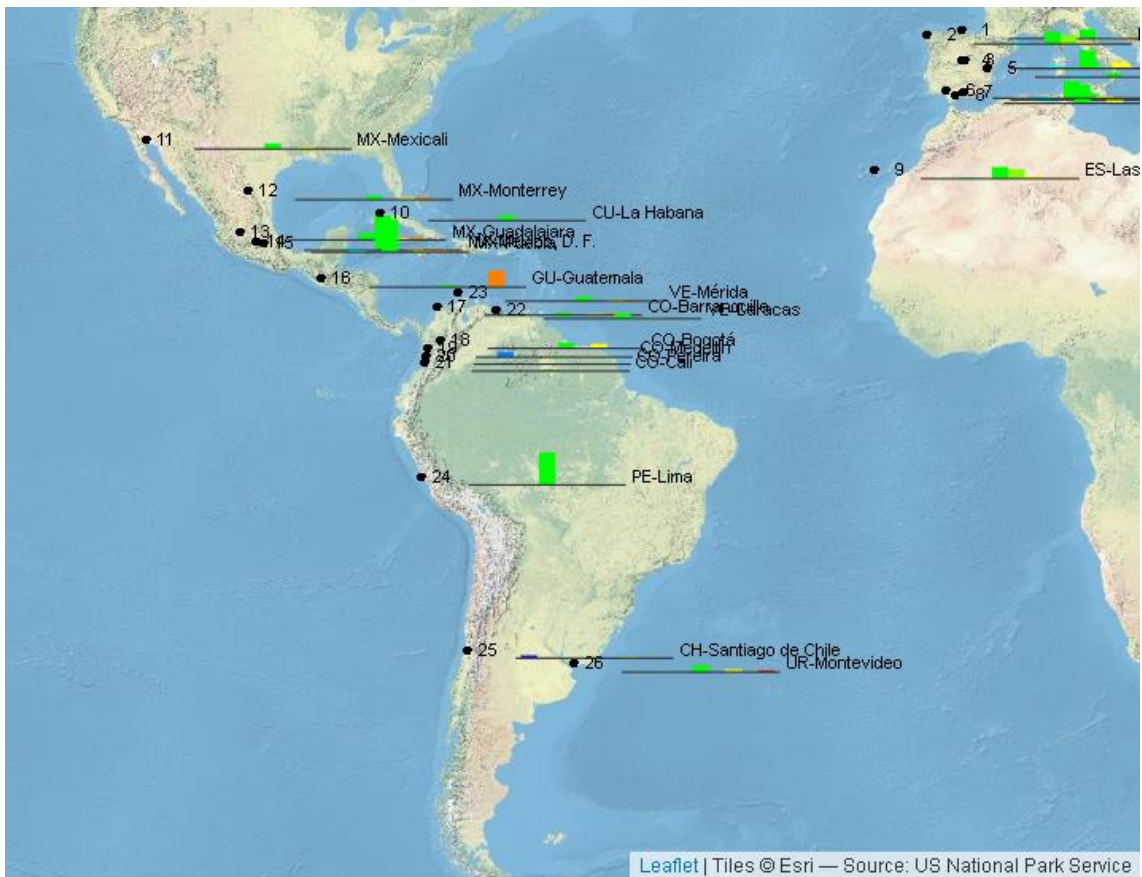
Forma:

- 1 ■ ¿cacháis?
- 2 ■ ¿cierto?
- 3 ■ ¿eh?
- 4 ■ ¿hm?
- 5 ■ ¿no?
- 6 ■ ¿sabes?
- 7 ■ ¿sí?
- 8 ■ ¿verdad?
- 9 ■ ¿viste?

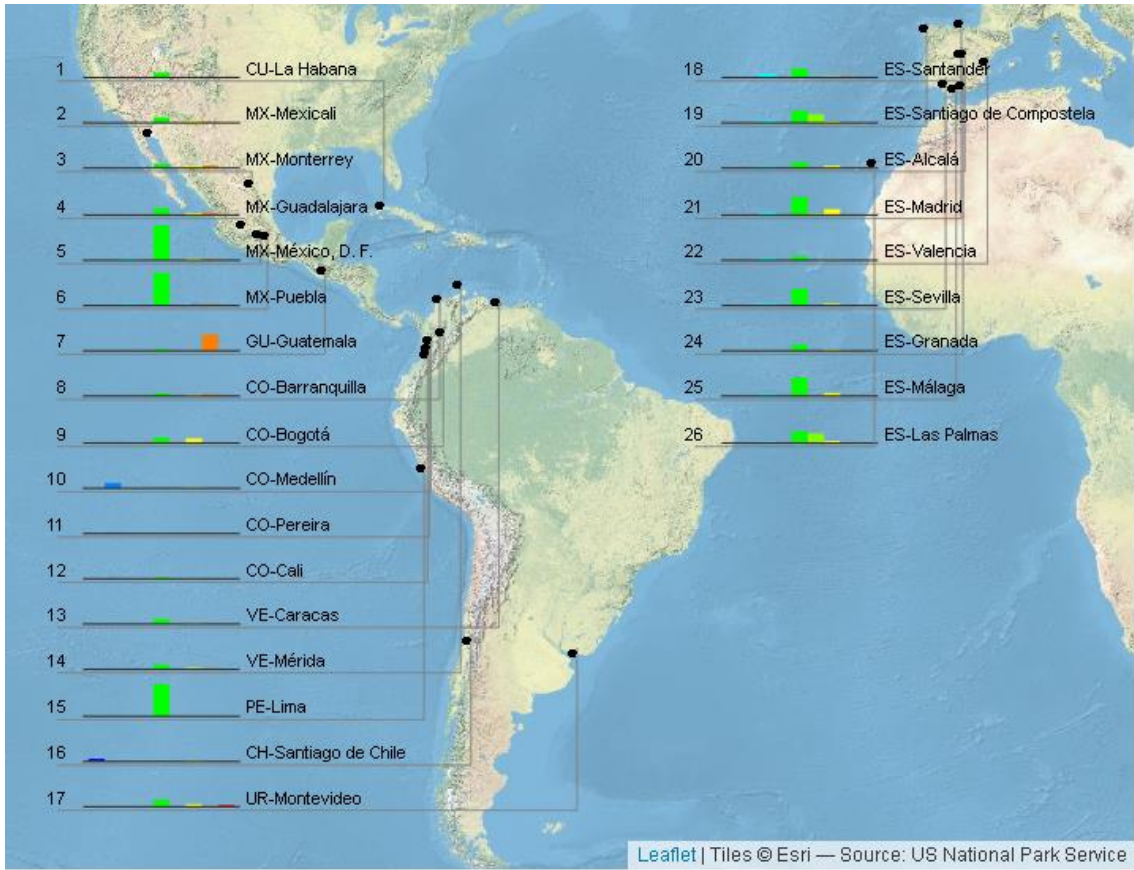
Lugar (lado izquierdo):

- 10 CU-La Habana
- 11 MX-Mexicali

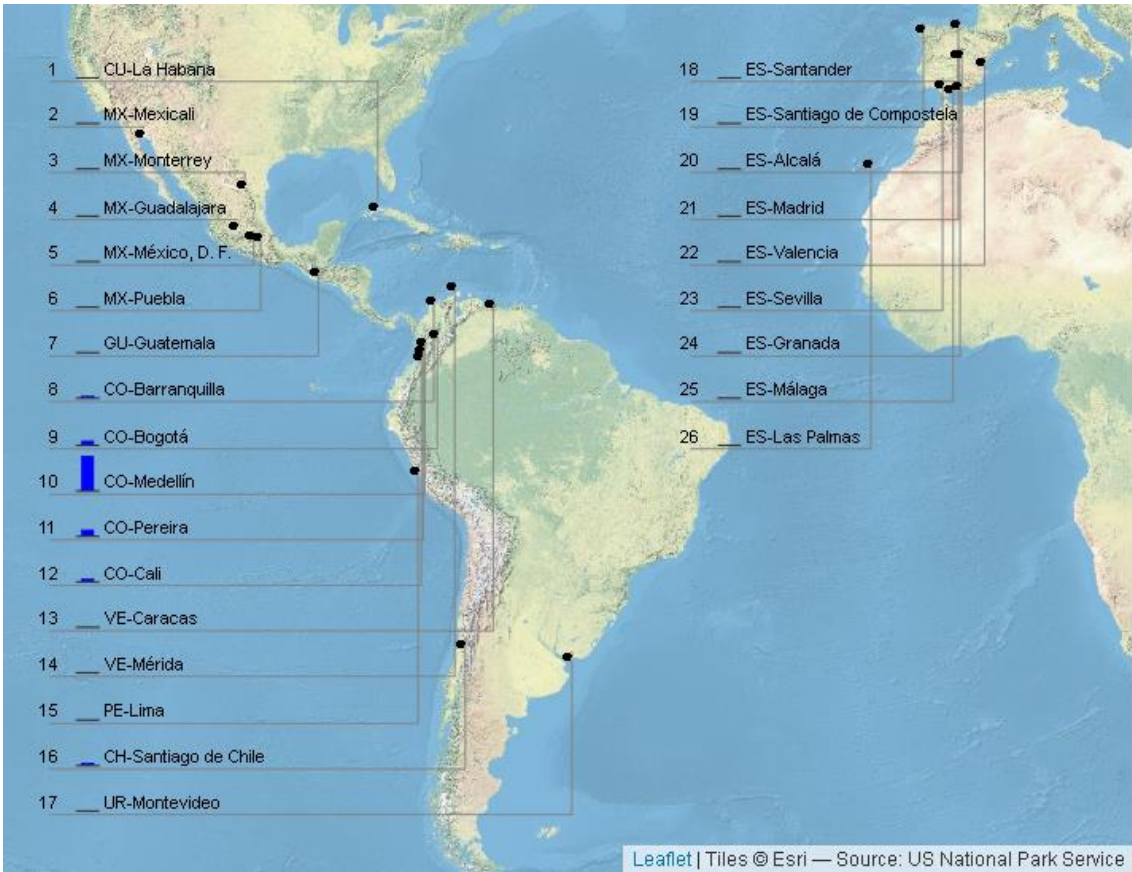
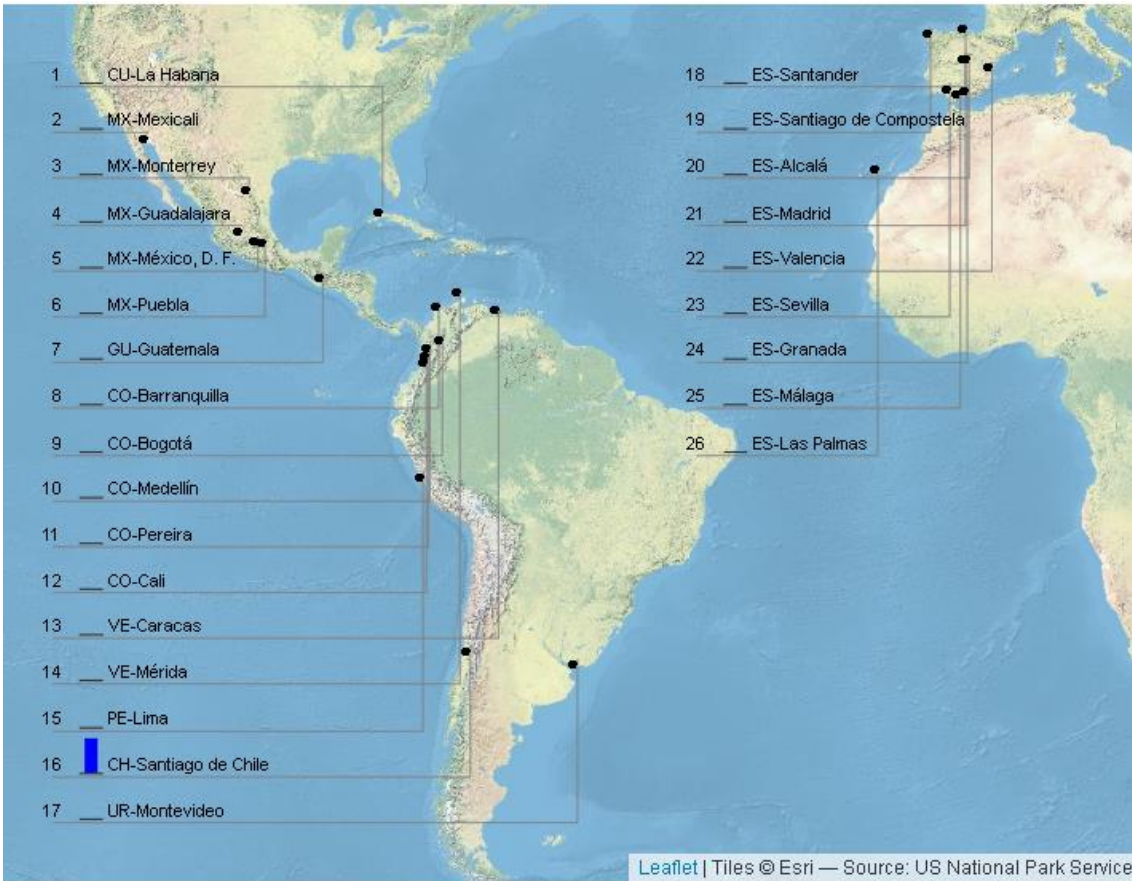
Topónimo



Posición: R A1 A2 A3 A4

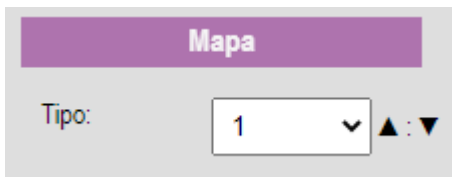


Mapa dinámico
 Proceder Parar



4.2.2. Tipo de mapa

Se puede seleccionar el tipo de mapa: 1, 2, ..., 7:

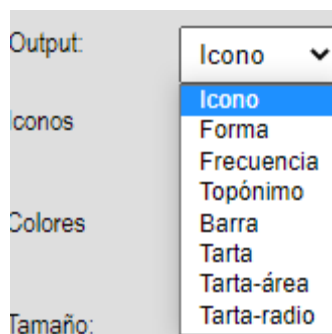


Tipo: 1

4.2.3. Marca



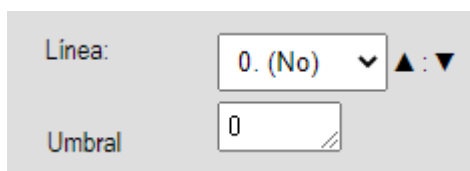
Se puede seleccionar el modo de representar el output: Icono, Forma, Frecuencia, Topónimo, Barra, Tarta, Tarta-área, Tarta-radio:



Se ofrecen distintos iconos para seleccionar los más adecuados. Hay tres modos de determinar los colores. Con [num] en primer lugar, se determinan un conjunto de colores preparados. Si se desea, se puede utilizar [grade], en primer lugar, por lo que se preparan distintos colores de gradación a partir del azul, pasando por el verde, para llegar al rojo. El usuario puede seleccionar colores que hay a continuación, quitando [num] y [grade]. El orden de color coincide con el orden de Icono. Se puede seleccionar el [Tamaño de icono] desde 5 a 100 píxeles. También es posible hacer reflejar la cantidad relativa de la cifra por el [Tamaño proporcional de icono].

4.2.4. Línea

Los puntos de reacción se combinan por líneas de distintas maneras.



Umbral determina el límite de la reacción positiva. El umbral cero (0) indica que todas las cifras positivas superiores a cero, se combinan por la Línea.

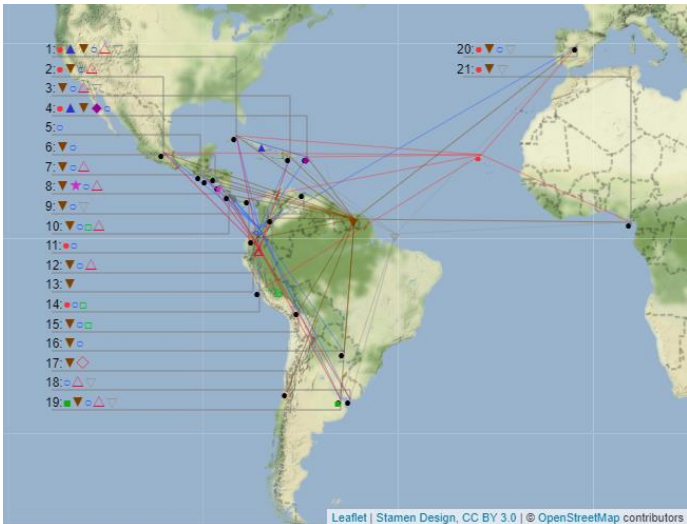
1. Irradiar

En primer lugar, se buscan las coordenadas del centro de puntos de reacción y seguidamente los puntos se combinan con el centro:



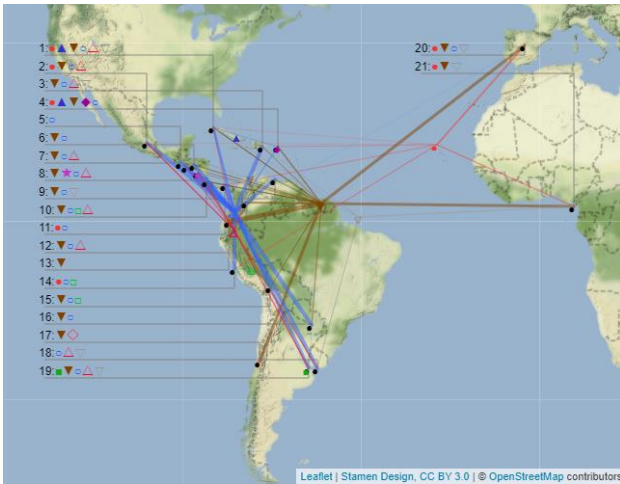
2. Irradiar (Coord. pesada)

Se puede seleccionar [2. Irradiar (Coord. pesada)], por lo que se busca el centro de los puntos de reacción con el peso proporcional a la cifra relativa de cada punto. Cuanto mayor sea el peso, tanto más el centro se aproximará al punto en cuestión:



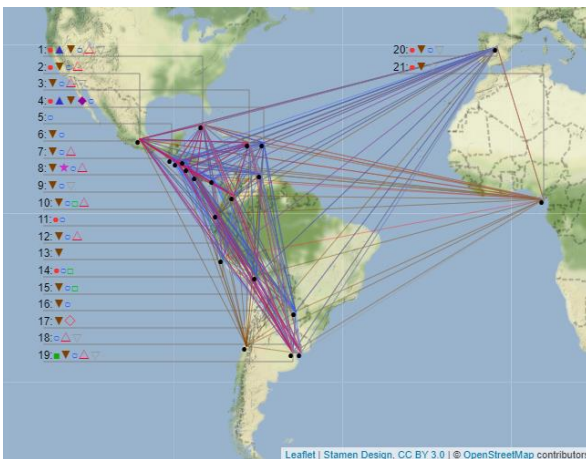
3. Irradir (Línea pesada)

Con [3. Irradir (Línea pesada)], sale la línea gruesa, al combinar el punto de reacción pesada, es decir, con la alta cifra relativa:



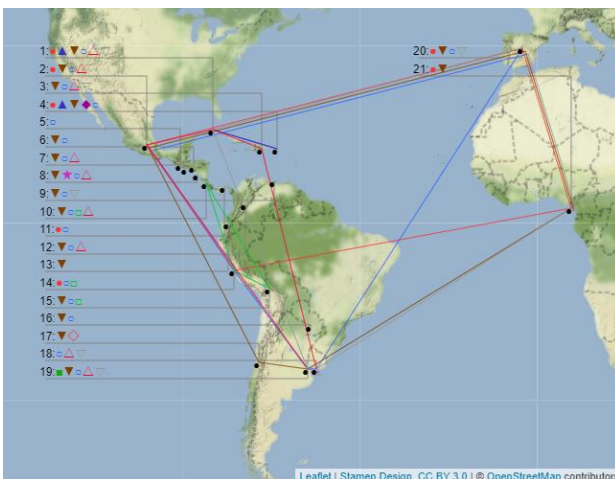
4. Red

Con la [4. Red] se combinan todos los puntos relevantes:

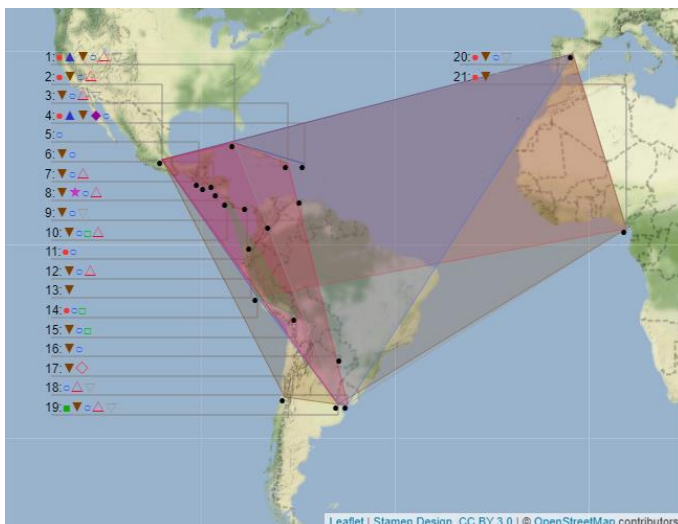


5. Contorno

En lugar de combinar los puntos positivos, con [5. Contorno] se dibuja un área que envuelve los puntos positivos. Tiene el inconveniente de incluir también los puntos negativos.



6. Contorno (área)

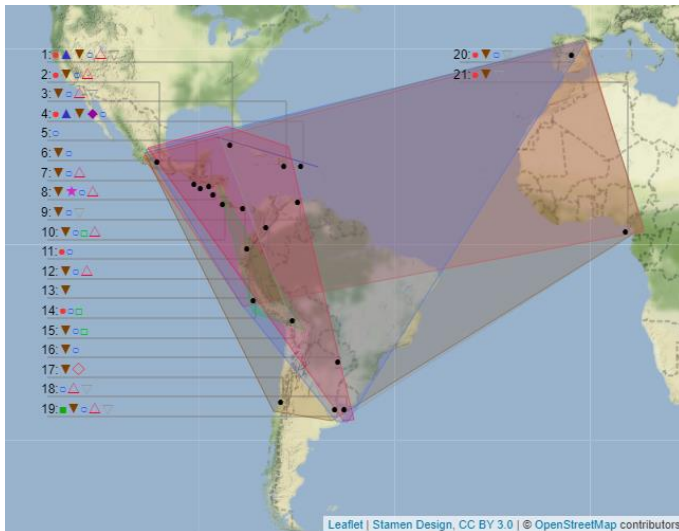


7. Extensión

Para formar el contorno, se seleccionan los puntos extendidos hacia el exterior del área:



8. Extensión (área)



9. Línea de máxima cercanía

A diferencia de la Red, donde se combinan todos los puntos positivos, en la [Línea de máxima cercanía] se combinan solo los puntos positivos más cercanos.



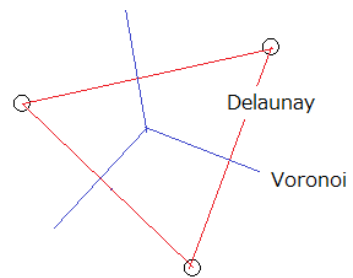
10. Diagrama de Delaunay

Uno de los métodos más eficientes para combinar los puntos positivos es el diagrama de Delaunay. El mismo diagrama ofrece un conjunto de triángulos que combinan todos los puntos positivos con las líneas que no entrecruzan.



11. Diagrama de Voronoi

El diagrama de Voronoi presenta el área máxima que ocupa el punto positivo en relación con los otros puntos positivos. Se prepara a partir del diagrama de Delaunay, dibujando líneas bisectores perpendiculares de los lados de todos los triángulos:



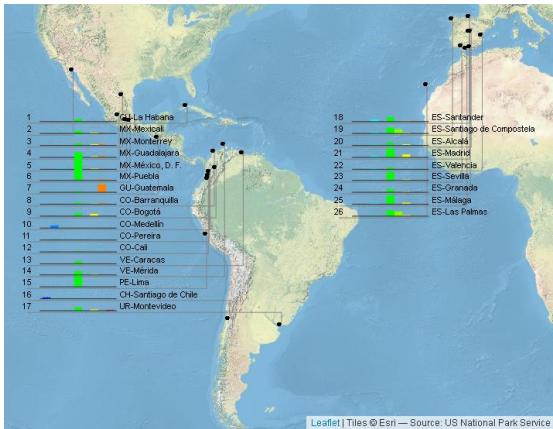
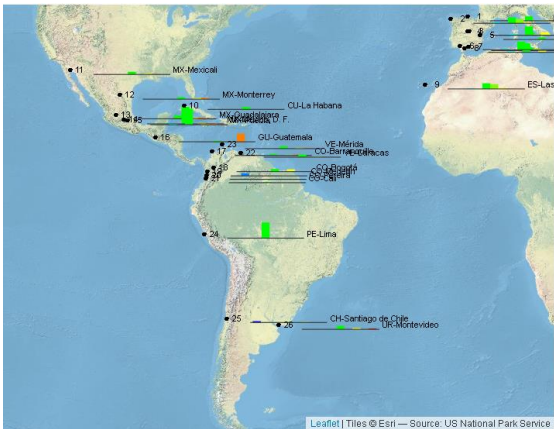
4.2.5. Posición

Cuando los puntos positivos se aproximan en el mapa, es conveniente colocarlos en lugares cercanos y combinarlos por líneas. Con el A[justar].2, se aprovecha al máximo el espacio del mapa. Se determinan las dos

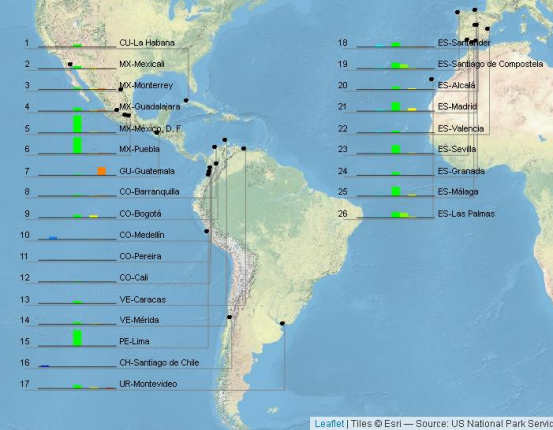
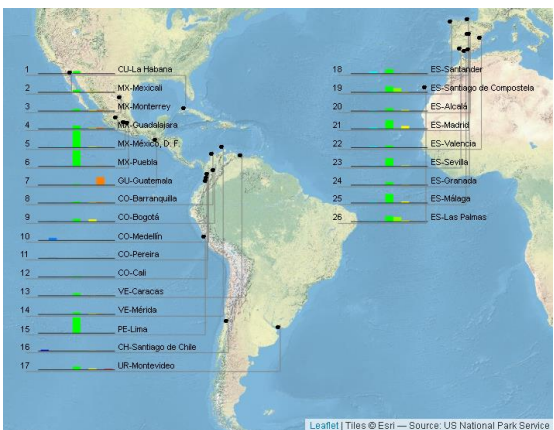
longitudes, del oeste y del este, por [Ajustar: O, E].

Posición: R A1 A2 A3 A4

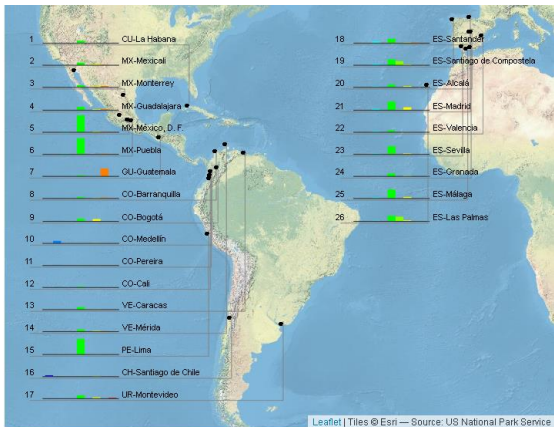
Ajustar: O,E.



R / A1



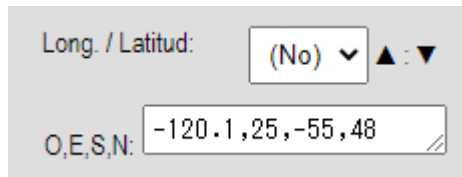
A2 / A3



A4

4.2.6. Longitud / Latitud

Cuando es necesario dibujar las líneas de longitud y latitud, se puede seleccionar su intervalo: 1, 2, 3, 4, 5, 10, 20, 25, 50. [O, E, S, N] determinan los límites de oeste, este, sur y norte del mapa.

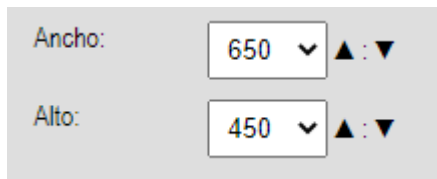


Long. / Latitud: (No) ▲ : ▼

O,E,S,N: -120,1,25,-55,48

4.2.7. Ancho y alto del mapa

Finalmente se determinan el ancho y el alto del mapa en la unidad pixel:

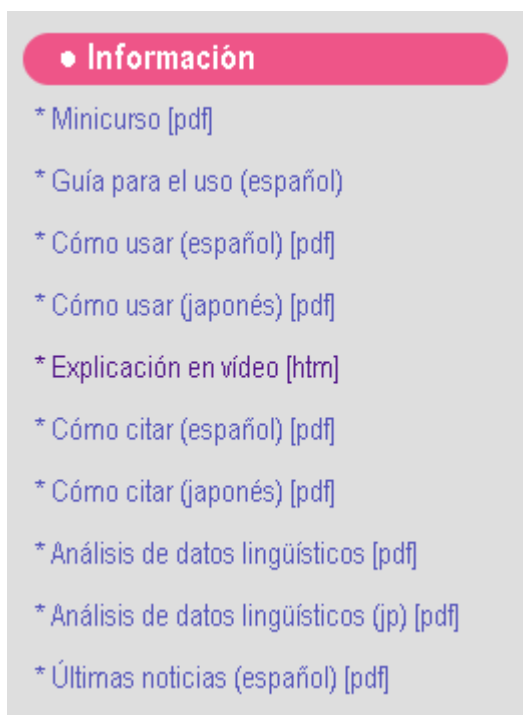


Ancho: 650 ▲ : ▼

Alto: 450 ▲ : ▼

5. Información

En la sección de [Información], se encuentran los siguientes materiales:



Minicurso: *Minicurso. ALDICAM en LYNEAL. Palabras y formas en historia y geografía.* 16 y 17 de febrero, 2023. Universidad de Alcalá.

Guía para el uso: *Guía para el uso con materiales de PRESEEA.* 2021.

Cómo usar (español): Este documento.

Explicación en vídeo: Explicación visual en vídeos breves.

Cómo citar (español): Pedimos citar los datos y el sistema en sus publicaciones.

Análisis de datos lingüísticos: *Análisis de datos cuantitativos para estudios lingüísticos.* Hiroto Ueda y Antonio Moreno Sandoval. 2017.

Últimas noticias: Procesos realizados recientes.

6. Apéndice-1: Patrones

6.1. Patrones simples

6.1.1. Almohadilla y porcentaje : #, %

Al pulsar [●Letras], aparece [Patrón de búsqueda:] junto con una caja de texto, donde escribimos unos patrones con los que el sistema busque las formas que deseemos.

Para buscar las formas deseadas, utilizamos unos patrones que se expresan con algunos signos especiales. Distinguimos entre las «letras», que se utilizan dentro de la palabra, y los «signos», que son signos de puntuación (punto, coma, dos puntos, punto y coma, interrogación y exclamación), y otras marcas de teclas correspondientes (paréntesis, comillas, +, -, =, #, %, @, etc.). En las reglas del Patrón, algunos signos (#, %, -, =, &, @, \, |, [], (), {}) tienen funciones especiales. De momento, señalamos tan solo el significado de la almohadilla (#) y del porcentaje (%):

#	Frontera de palabra
- (~ %)	Cero o más veces de las letras

La frontera de palabra, representada con la almohadilla (#), es útil para buscar, por ejemplo, las palabras que empiezan con la letra "a" (A, a). La forma que utilizamos es:

#a-

Esto quiere decir que con "#a-" podemos buscar todas las palabras con la *a*- inicial, ya sea mayúscula o minúscula: *A, así, algunas*, etc.

El signo de guion (-) o porcentaje (%) sirve para indicar que buscamos no solamente la *a*- inicial, sino también 'a' acompañada de otras letras siguientes dentro de la palabra. Lo de "cero o más veces de letras" quiere decir que corresponde tanto a la 'a' sola ('a' + cero veces de letras), como a las formas plenas, por ejemplo, *así, algunas*, etc. ('a' + más de cero veces de letras).

Los mismos signos sirven para buscar las formas que terminen en alguna letra. Por ejemplo, con el Patrón "%d#" podemos encontrar todas las palabras terminadas en *-d*: *ciudad, universidad, usted*, etc.

Se pueden escribir varios patrones cambiando de línea con la tecla [Enter], por ejemplo:

```
#a-
#r-
```

Si es necesario distinguir entre la forma minúscula o la mayúscula, hay disponible un botón de [Distinguir mayúscula / minúscula].

En lugar de #, podemos utilizar el espacio en blanco " d% ". Los espacios en blanco al inicio del patrón y al final del patrón se interpretan como # (frontera de palabra).

El signo de menos (-) se utiliza para indicar el contexto con letras interiores de la palabra. Por ejemplo, el patrón "-a-" busca todas las letras "a" interiores de la palabra, por ejemplo, 'acompañada'. Para buscar la a inicial e interior de palabra, se utiliza el patrón "a-". Para buscar la 'a' inicial, interior y final de palabra, se utiliza el patrón "a".

Por ejemplo con el patrón "#fa-" se buscan todas las palabras que tienen 'fa' al inicio de la palabra: *favor, faldas, fachada, falta*:

Línea	Texto anterior.	Forma	Texto posterior
69	~nde está el comedor? / – Muchas gracias. ¡Ah, por	favor	! ¿Me puede decir dónde está el comedor? / ~
106	~laro, tío! Allí tienes de todo: por ejemplo, hay	faldas	guapas para tu mujer o chándal para los críos. ¡T~
118	~ianistas, que se conoce enseguida porque tiene la	fachada	de ladrillo. Y de ahí al "Campeón", nada, dos o t~
122	~víos para la limpieza... en fin, todo lo que haga	falta	en la casa para la semana. Así que cuando me sobr~

Con el patrón "#fa", sin el guion (-) ni porcentaje (%) se buscan todas las palabras que tienen 'fa' al inicio de la palabra: *favor, faldas, fachada, falta*. Sin embargo, en la columna salen solo la parte correspondiente del patrón y el resto aparece en el [Texto posterior]:

Línea	Texto anterior.	Forma	Texto posterior
69	~nde está el comedor? / – Muchas gracias. ¡Ah, por	fa	vor! ¿Me puede decir dónde está el comedor? / ~
106	~laro, tío ! Allí tienes de todo: por ejemplo, hay	fa	ldas guapas para tu mujer o chándal para los críos~
118	~ianistas, que se conoce enseguida porque tiene la	fa	chada de ladrillo. Y de ahí al "Campeón", nada, do~
122	~víos para la limpieza... en fin, todo lo que haga	fa	lta en la casa para la semana. Así que cuando me s~
	~laro, mamá! Allí tienes de todo: por ejemplo,		ldas nadricimas na' tu mujer o ropa deportiva

La almohadilla (#) sirve para indicar no solamente el inicio de palabra sino también el final de palabra. Por ejemplo, con el patrón “-dad#” se buscan tales formas como “soledad”, “unidad”, “universidad”, etc. En este caso, no hay que poner la almohadilla al inicio del patrón “#-dad#”, puesto que el guion (-) corresponde a la secuencia de letras utilizadas en la palabra, donde se excluyen signos (no letras).

La almohadilla (#) se ponen siempre al inicio y/o al final del patrón.

6.1.2. Signo de et: &

El signo de *et* (&) representa la secuencia de una o más letras. Con este signo podemos designar las formas que se encuentran dentro de la palabra, por ejemplo, -ll- media (*valle*, *cuello*), a exclusión de ll- inicial (*llave*, *llevar*), puesto que el signo *et* (&) condiciona la existencia de por lo menos una letra.

Este signo (&) sirve también para buscar una palabra entera. Por ejemplo, si queremos saber qué palabras ocurren detrás de 'he' formulamos el patrón: "he &". La diferencia entre "%" y "&" radica en que "%" incluye el caso donde no existe la letra, mientras que "&" siempre presupone por lo menos una letra.

La secuencia de dos palabras corresponde al patrón "& &".

6.1.3. Signo de arroba: @

El signo de arroba "@" representa una sola letra. Por ejemplo, "#d@s%" busca las palabras que empiezan con *das-*, *des-*, *dis-*, etc.

6.1.4. Signo de punto: .

El signo de punto "." sirve para indicar una letra, un número o un signo o un espacio en blanco. Por ejemplo, para saber qué segmento ocurre delante de

la palabra que posee "A" mayúscula en posición inicial, utilizamos el Patrón ".A%".

6.1.5. Signo de raya inversa más la letra "d": \d

Para buscar un número, 0, 1, 2, 3, etc., se utiliza el Patrón "\d".

6.1.6. Corchetes: [...]

Con los corchetes "[...]", podemos agrupar las letras sueltas. Por ejemplo, el patrón "#amig[oa]#" sirve para buscar las palabras 'amigo' y 'amiga' al mismo tiempo. Las palabras que terminan en vocal átona se buscan con "%[aeiou]#".

Con un guion "-" dentro de los corchetes especificamos letras, desde qué letra hasta qué letra en el orden alfabético deseamos buscar; por ejemplo, con "#[a-d]%", encontramos las palabras con *a*, *b*, *c*, *d* iniciales.

El acento circunflejo "^" al inicio de los corchetes posee la función de negar las letras subsiguientes. Así, por ejemplo, con el Patrón "#[^abc]%" buscamos todas las palabras que no empiecen con 'a', 'b', o 'c'.

6.1.7. Paréntesis y raya vertical: (...|...)

Con los paréntesis "(...)" y una raya vertical "|" agrupamos los subpatrones; por ejemplo, con el Patrón "#(he|has|ha) &do#" buscamos las combinaciones de 'he', 'has', 'ha' más formas terminadas en '-do': 'he comido', 'he terminado', etc.

6.2. Patrones especiales

En cuanto a las funciones que tratamos en esta sección, es conveniente que sepamos por lo menos de su existencia. A la hora de aplicarlo, podemos consultar este documento. De modo que no es muy necesario aprender de memoria los signos tratados en esta sección y sus correspondientes funciones. Con la práctica los aprenderemos paso a paso.

6.2.1. Cantidad: ?, *, +, {...}

Con el signo de interrogación "?", designamos tanto la presencia del elemento anterior a "?" como su ausencia. Por ejemplo, con el Patrón "del?" buscamos tanto 'de' como 'del'. Con el Patrón "aquel(los)?" se busca 'aquel' y 'aquellos' al mismo tiempo.

Utilizamos el signo de asterisco "*" para indicar que la repetición del elemento anterior a "*" puede ser cero o más de cero veces. Por ejemplo, con el Patrón "%[aeiou]*d#" podemos buscar las formas terminadas en 'd' con una vocal ('lid') o una secuencia de vocal ('load') o sin vocal ('end').

Si deseamos que haya por lo menos una vocal delante de la "-d" final, utilizamos el signo de plus "+" en lugar del asterisco "*" en el Patrón "%[aeiou]+d#": *lid, voluntad*.

Las llaves "{...}" con cifras sirven para designar las cantidades exactas. Por ejemplo, si buscamos la secuencia de tres números seguidos, por ejemplo, "125" o "591", formulamos el Patrón "\d{3}".

6.2.2. Raya reversa: \

Como hemos visto, en el Patrón los signos "#", "%", "&", "@", "[", "]", "(", ")", "?", "*", "{", "}" poseen significados especiales. Para buscar estos mismos signos dentro del texto, hay que poner una raya reversa "\" delante de ellos, por ejemplo "\%" sirve para buscar el signo del porcentaje '%'. Los signos de admiración (¡, !), el signo invertido de integración (¿), la coma (,) y los angulares (<...>) se buscan simplemente poniendo estos signos sin más.

6.2.3. Contexto: ...<<...>>...

El patrón con doble corchete angular "<<...>>" sirve para indicar los contextos anteriores y/o posteriores. Por ejemplo, para buscar las formas terminadas en '-a' delante de la palabra terminada en '-o', utilizamos el patrón "&a#{{&o#}}" o "&a#>>&o#", con el qu-illa{-}Sevillae se recoge, por ejemplo, **mapa** mudo'. Igualmente para buscar las palabras terminadas en '-a' con el artículo 'el', formulamos el patrón "{{#el#}}-a#" o "#el#<<-a#", con el que encontramos 'el día', 'el Papa', etc.

6.2.4. Reproducción: (...) \1

La combinación del paréntesis más una raya con el número 1 (\1) sirve para indicar que la forma correspondiente al patrón entre paréntesis se reproduce en el Patrón "\1". Por ejemplo, "(#&#) \1" sirve para buscar las repeticiones de la misma palabra, 'Abajo, abajo', '¡Premios! ¡Premios!', etc.

Si hay más de un paréntesis, podemos utilizar \1 y \2, como correspondientes a cada paréntesis.

6.2.5. Redefinición: ...==...

Al formular previamente el patrón, por ejemplo "C==[bcdfghjklmnpqrstvwxz]" podemos utilizar "C" sin repetir "[bcdfghjklmnpqrstvwxz]". Por ejemplo para buscar distintas combinaciones de consonantes y vocales, podemos utilizar los patrones siguientes:

```
C==[bcdfghjklmnpqrstvwxz]
V==[aeiou]
CCV
CVV
CV
V
```

El resultado es:

[3a] Frequency: Absolute frequency

Núm.	Pattern	2	4	6	8	10	12
1	CCV	586	679	569	633	531	547
2	CV	2686	3028	2604	2941	2587	2426
3	CVV	254	316	296	333	308	230
4	V	3534	3979	3425	3831	3352	3144

6.2.6. Reemplazo: ...=>...

El patrón sirve no solo para efectuar la búsqueda sino también para realizar el reemplazo. Por ejemplo, con el patrón "a=>A" podemos convertir toda letra 'a' minúscula en la 'A' mayúscula. Con el patrón "a=>" podemos eliminar toda letra 'a'.

Si deseamos efectuar la búsqueda en los textos reemplazados, hay que poner

los patrones de reemplazo al inicio de la serie de patrones, puesto que se aplican a las búsquedas solo con los patrones posteriores. Por ejemplo, para buscar las palabras terminadas en "-t" en los textos medievales con la abreviatura desarrollada, utilizamos los siguientes patrones:

```
>=>
<=>
&t#
```

Lo mismo puede efectuarse con los siguientes patrones:

```
<(&>=>$1
&t#
```

donde \$1 representa la reproducción de (&). Por ejemplo, la parte correspondiente a (&) en 'gra<n>t' es 'n', que se representa con "\$1", de modo que <(&> que corresponde a '<n>' se convierte en 'n'.

Si deseamos tratar el texto eliminado de las partes abreviadas, utilizamos el patrón "<&>=", con el que 'gra<n>t' se convierte en 'grat'.

6.2.7. Exclusión {-}

Las formas correspondientes al patrón acompañado de "{-}" serán excluidas de la búsqueda. Por ejemplo, con el patrón "%illa#" se buscan *maravilla*, *Sevilla*, *camarilla*, *mezclilla*, Si agregamos otro patrón "Sevilla{-}", se excluye 'Sevilla' en la búsqueda y sale: *maravilla*, *camarilla*, *mezclilla*,

```
%illa#
Sevilla{-}
```

Para especificar la exclusión individualmente, escribimos "-illa{-}Sevilla" en una línea.

6.2.8. Escape {E}

Para escapar las funciones de los signos especiales "#", "%", "&", "@", "[", "]", "(", ")", "?", "*", "{", "}", etc., se pone {E} detrás de la expresión regular. Por ejemplo, el simple signo de interrogación "?" es ilegal o el signo de punto "." es inconveniente. Para buscar estos signos escapando sus funciones, asignamos "{E}?", "{E}.". Para buscar "?" y ".", hay que escribir "{E}?" y "{E}." (escapes individuales, en lugar de "{E}?", "{E}."), puesto que esto

significa “\(\?|\.\)”).

6.2.9. Patrones propios de español

Con los patrones que enumeramos en la columna de la izquierda, podemos buscar las formas que se encuentran en la columna de la derecha.

{prep}	(en por a con p[oa]ra sobre so ante bajo contra sin entre [fh]asta desde de tras cabe? faza hacia)
{clit}	(me te nos os se l[oa]s?)
{pron}	(yo tú [nv]osotr[oa]s ell[oa]s ustedes usted)
{art}	(el las? los?)
{indef}	(unos unas? un)
{ar}	Terminaciones regulares del verbo en -ar: ej. tom{ar}
{er}	Terminaciones regulares del verbo en -er: ej. com{er}
{ir}	Terminaciones regulares del verbo en -ir: ej. viv{ir}

7. Apéndice-2: Tratamiento de texto

En esta sección vamos a abordar diferentes operaciones que se pueden hacer con textos en LYNEAL y que añaden más funcionalidad a la herramienta.

7.1. Combinar / dividir texto

La primera plataforma posee dos funciones: Combinar texto: Dividir texto.

Combine & divide text

Input text

```
D.TX ##CPT:ING ES GE CU RD
PR MX GU HO EL NI CR
PN CO VE EC PE BO CH
PA UR AR
americana A001:JACKET 58 2 1
0 3 1 0 0 0 0 0 0
0 1 0 0 0 2 0 0
0 0 0 0 0 0 0 0
capa A001:JACKET 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 1
```

(Delete)

Output text

```
D.TX ##CPT:ING ES GE CU RD PR MX GU HO EL NI
CR PN CO VE EC PE BO CH PA UR AR
americana A001:JACKET 58 2 1 0 3 1 0 0 0
0 0 1 0 0 0 2 0 0 0 0
capa A001:JACKET 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1
```

(Delete)

[Run-0] Combine text
 [Run-1] Divide text
 (Delete all)
 Finish! 3 / 10 : 0 (4 m.s.)
[\[Top page\]](#)

7.1.1. Combinar texto

A partir de los datos separados por tabulación [Tab], se combinan las celdas que hay desde la segunda columna:

D.TX	##CPT:ING	ES	GE	CU	RD	PR	MX	GU
	HO	EL	NI	CR	PN	CO	VE	EC
	BO	CH	PA	UR	AR			PE
americana	A001:JACKET	58	2	1	0	3	1	
	0	0	0	0	0	1	0	0
	2	0	0	0	0			
capa	A001:JACKET	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	1			

↓

D.TX	##CPT:ING	ES	GE	CU	RD	PR	MX	GU	HO	EL	NI	CR	PN	CO	VE	EC
------	-----------	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

PE BO CH PA UR AR

americana A001:JACKET 58 2 1 0 3 1 0 0 0 0 0 1 0 0 0 2 0 0 0 0 0

capa A001:JACKET 0 1

7.1.2. Dividir texto

A partir de los datos separados por tabulación [Tab], se dividen las celdas que hay desde la segunda columna de la siguiente manera:

↓

```
D.TX ##CPT:ING
D.TX ES
D.TX GE
D.TX CU
(...)
americana A001:JACKET
americana 58
americana 2
americana 1
americana 0
americana 3
```

7.2. Comparar palabras

En esta opción, se comparan dos listas de palabras, con el objeto de comprobar cuáles aparecen en las dos listas. Esto es útil en las comparaciones con extensas listas de palabras, especialmente si tienen muchas variantes gráficas. También sirve para detectar palabras concretas en textos extensos.

Compare words

Input (-)	Word list (+) (¥n)	Output (-/+)
a b c d	a d	+a -b -c +d
(Delete)	(Delete)	(Delete)

[\[Run\]](#)
[\[Top page\]](#)

Se copia un texto o una lista de palabras en la caja de texto [Input] y en la siguiente caja una lista de palabras separadas por cambio de línea [Word list].

Al pulsar [Run], sale el texto marcado con “+” (palabras que hay en [Word list]) y “-“(palabras que no hay en [Word list]).

7.3. Lematizar y anotar texto

Aquí el objetivo es proporcionar una información muy completa sobre las palabras que forman un texto.

LYNEAL. Lematiza texto español (v.2020-07-13: H. Ueda)

■ Selección: ■ Ejecutar (Ctrl + Enter) Contexto Info. Atributo ■ Language: ■ [Portada](#)

¡Terminado! : 1 / 10 (8 m.s.)

Texto de input

En la recepción de un hotel llega un viajero que viene atraído por el turismo para conocer la ciudad de Sevilla, debido a que tuvo la oportunidad de conocer la promoción de la misma en su país por medio de la televisión y la prensa. El turista dialoga con el consejero solicitándole información sobre comidas, costumbres y visitas.

(Borrar)
(Copiar)

Texto de output

FORMA	CAT.	LEMA	GRAM.
En	p	en	.
la	at	el	f.s
recepción	s	recepción	f.s
de	p	de	.
un	at	un	m.s
hotel	s	hotel	m.s
llega	v	llegar	i.pr.3s
un	at	un	m.s
viajero	a	viajero	m.s
que	r	que	.
viene	v	venir	i.pr.3s

(Borrar)
(Copiar)

El texto se lematiza con categoría gramatical, lema y anotación gramatical. El texto del output viene separado por la tabulación, de modo que es copiable a Excel directamente. Para la categoría gramatical (CG) y anotación gramatical, hemos utilizado las siguientes abreviaturas:

<p>CG</p> <p>a adjetivo</p> <p>ab abreviación</p> <p>at artículo</p> <p>av adverbio</p> <p>c conjunción</p> <p>ct clítico</p> <p>d demostrativo</p> <p>e extranjerismo</p> <p>i interjección</p> <p>nc numeral cardinal</p> <p>nm número</p> <p>np nombre propio</p> <p>p prep</p> <p>pa prep+artículo: al, del</p> <p>pn pronombre</p>	<p>pp prep+pronombre: conmigo</p> <p>ps posesivo</p> <p>q interrogativo</p> <p>r relativo</p> <p>s sustantivo</p> <p>v verbo</p> <p>z puntuación, paréntesis, signo</p> <p>Anotación gramatical</p> <p>1p primera persona plural</p> <p>1s primera persona singular</p> <p>13 primera o tercera persona singular</p> <p>2p segunda persona plural</p> <p>2s segunda persona singular</p>
---	--

3p	tercera persona plural	ip	imperativo
3s	tercera persona singular	m	género masculino
c	género común: azul	n	género neutro: lo
cd	condicional	p	plural
dt	dativo	pi	pretérito indefinido
f	género femenino	pp	participio pasado
ft	futuro	pr	presente
g	gerundio	r	reflexivo
i.	indicativo	s.	subjuntivo
if	infinitivo	s	singular
im	imperfecto	sp	singular/plural

7.4. Mapa lingüístico

LYNEAL. Linguistic map (v.2020-07-05: H. Ueda)

■ Select: ■ [Execute (Ctrl + Enter)] ■ Language: ■ [\[Top page\]](#)

```

### //Example-1: Castilla
Lt: 43.8, 40 //Lt.latit
Lg: -4.6, -1.6 //Lg.long
Wd: 450 //Wd.width(pixel)
Mp: 1 //Mp.map: 1, 2,
Op: i //Op.output: t.t
Pc: t //Pc.Place: n.num
Is: 30 //Is.icon size:
Ln: z //Ln.line: n.netw
Cd: o //Cd.coordinate:
Dt: m //Dt.data: n.for
Gt: 1 //Gt.LeafletGrat
//Cl.Color: #00f, #30c,
Cl: #f00, #789, #227, #9
Th: 0.5 //Threshold
### //PLACE: toponym, lat
S, 43.46, -3.88
LO, 42.46, -2.46
BU, 42.33, -3.7
P, 42.01, -4.54
VA, 41.65, -4.72
SO, 41.75, -2.46
SG, 40.95, -4.13
AV, 40.65, -4.72
### //FORM: number, form,
1, 1250, 2b.
2, 1300, 3a.
3, 1350, 3b.
4, 1400, 4a.
5, 1450, 4b.
6, 1500, 5a.
### //DATA: toponym, freq
//Lugar, 1250, 1300, 135
AV 0.057 0.000
BU 0.026 0.000
LO 0.087 0.035
P 0.000 0.000
S 0.000 0.000
SG 0.000 0.000
SO 0.000 0.000
VA 0.158 0.037
          
```

Form:

- 1 2b. 1250
- 2 3a. 1300
- 3 3b. 1350
- 4 4a. 1400
- 5 4b. 1450
- 6 5a. 1500

Place:

- 1 S
- 2 LO
- 3 BU
- 4 P
- 5 VA
- 6 SO
- 7 SG
- 8 AV

(Delete)
(Copy)

Para elaborar un mapa lingüístico, se necesitan especificaciones del mapa, latitud y longitud de los puntos geográficos, formas o cifras que aparecen en cada punto, matriz bidimensional de puntos geográficos y formas o cifras. En la caja de texto izquierda, se especifican siguiendo el modelo de los ejemplos. Se pueden hacer varios experimentos cambiando las cifras y letras de cada ítem. Detrás de “//”, se enumeran las posibles variantes. Para los

detalles del mapa lingüístico, véase la sección 4.2.

7.5. Palabras españolas de poca frecuencia

Detect Spanish words of low frequency (ver. 2020-04-02)

Input	Output
El texto se lematiza con categoría gramatical, lema y anotación gramatical. El texto de output viene separado por la tabulación, de modo que es copiable a Excel directamente. Para la categoría gramatical y anotación gramatical, hemos utilizado las siguientes abreviaturas	El texto se {*lematiza*} con categoría gramatical, lema y anotación gramatical. El texto de output viene separado por la {*tabulación,*} de modo que es {*copiable*} a Excel directamente. Para la categoría gramatical y anotación gramatical, hemos utilizado las siguientes abreviaturas

Se marcan con `{*...*}` las palabras españolas de poca frecuencia. Algunas de ellas pueden estar mal escritas. De modo que podemos utilizar esta herramienta como corrector ortográfico o para detectar neologismos.

7.6. Expresión regular

RegEx: Search / transform text by Regular Expression

Input	Regular exp.: **(>=)**	Output
El texto se lematiza con categoría gramatical, lema y anotación gramatical. El texto de output viene separado por la tabulación, de modo que es copiable a Excel directamente. Para la categoría gramatical (CG) y anotación gramatical, hemos utilizado las siguientes abreviaturas:	#s- #t=>T	El Texto {*se*} lematiza con categoría gramatical, lema y anotación gramatical. El Texto de output viene {*separado*} por la Tabulación, de modo que es copiable a Excel directamente. Para la categoría gramatical (CG) y anotación gramatical, hemos utilizado las {*siguientes*} abreviaturas:
(Delete)	(Delete)	(Delete)

Sirve para buscar y convertir formas con expresiones regulares. La búsqueda de todas las palabras que empiezan con “s” se hace por ejemplo con el patrón “#s-”. En el [Output] salen marcadas así, `{* ... *}`. Por ejemplo, en el texto: `{*se*}`, `{*separado*}`, `{*siguientes*}`. La transformación implica una parte izquierda con las condiciones y una parte derecha con el cambio, separadas por el operador “=>”. Por ejemplo, el patrón “#t=>T” convierte la “t” minúscula inicial en “T” mayúscula inicial: Texto, Tabulación. Esta herramienta sirve para confirmar las funciones de expresiones regulares y hacer sus prácticas.

7.7. Repetir celda

Esta función sirve para convertir las formas repetidas de la línea anterior en un punto (.). Es útil para trabajos de confirmación de procesos automáticos.

Repeat cell

Input text	Output text
<pre>de de p de de p deber debido v desde desde p eeh eeh i eeh eeh i el el at el el at el la at el la at el los at</pre>	<pre>de de p . . . deber debido v desde desde p eeh eeh i . . . el el at la at los at</pre>
(Delete)	(Delete)

[Run-0] Repeated cell => Period (.)

[Run-1] Period (.) => Repeat upper cell

[Run-2] Blank => Period (.)

En la columna más a la izquierda del texto de entrada aparecen los lemas, luego la forma y finalmente la categoría sintáctica. La celda repetida del lema se convierte en un punto (.). De igual manera, se hace con las formas y categorías que se repiten en la línea de abajo. Sin embargo, se mantiene el dato original cuando es diferente del superior, por ejemplo “at” (artículo) al lado de “la” y “los”.

7.8. Correspondencia léxica

Es una versión de búsqueda por expresiones regulares (6.6) donde especificamos cambios que afectan a palabras

Word match: Search / transform text by word matching

Input text	word.X ([=>, tab] word.Y)	Output text
<pre>El texto se lematiza con categoría gramatical, lema y anotación gramatical. El texto de output viene separado por la tabulación, de modo que es copiable a Excel directamente. Para la categoría gramatical y anotación gramatical, hemos utilizado las siguientes</pre>	<pre>se con hemos=>HABER</pre>	<pre>El texto {#sex} lematiza {#con*} categoría gramatical, lema y anotación gramatical. El texto de output viene separado por la tabulación, de modo que es copiable a Excel directamente. Para la categoría gramatical y anotación gramatical, HABER utilizado las siguientes abreviaturas:</pre>
(Delete)	(Delete)	(Delete)

[Run-1] (Lower-upper case: Not sensitive)

[Run-2] (Lower-upper case: Sensitive)

Se busca y/o convierte formas con formas léxicas. Cuando el objetivo es la búsqueda, escribimos la(s) forma(s) en la caja de texto media, cambiando la

línea, por ejemplo, *se*, *con*, etc. Cuando el objetivo es la transformación, su fórmula es, por ejemplo, “hemos=>HABER”.

7.9. Comparación de frecuencia de palabras

Esta función nos sirve para comparar las palabras que aparecen en dos textos. Por ejemplo, en un fragmento se usa “texto” y en otro “libro”. Al realizar la comparación, podemos observar la discrepancia entre ambos fragmentos por el recuento diferente.

Comparison of word frequency

Input text

```
###A
El texto se lematiza con categoría
gramatical, lema y anotación gramatical.
###B
El libro se lematiza con categoría
gramatical, lema y anotación gramatical.
```

(Delete)
(Copy)

Output text

Word	A	B
,	1	1
.	1	1
anotación	1	1
categoría	1	1
con	1	1
El	1	1
gramatical	2	2
lema	1	1
lematiza	1	1
libro	0	1
se	1	1
texto	1	0
y	1	1

(Delete)
(Copy)

[Run]: [Ctrl]+[Enter]

Así, introducimos dos fragmentos muy parecidos en cuanto a uso de las palabras, por ejemplo A y B, que especificamos con “##” en la cabeza del texto correspondiente. En el output, vienen las formas con su frecuencia en ambos textos, separados por la tabulación. El output es copiable a Word, Excel, etc.

FIN