

# ***Etiquetación gramatical por Excel VBA: LETRAS***

ver. 2013.12.6

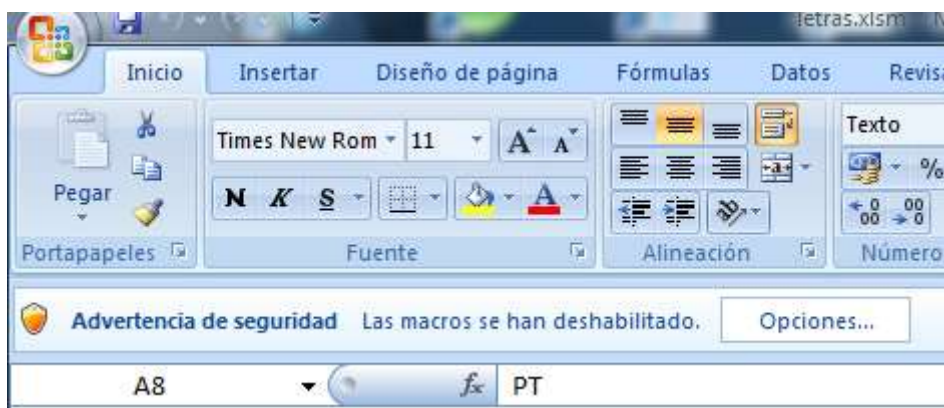
Este documento es un manual de uso para LETRAS.xlsm (en adelante LETRAS), conjunto de programas para el análisis de datos lingüísticos y filológicos. LETRAS está en desarrollo continuo, de modo que este mismo documento también cambia continuamente sin previo aviso. Para el detalle de modificaciones, véase la primera hoja (L) de LETRAS. Le rogamos al usuario que al notar inconvenientes o funciones mejorables, nos los comunique a través del correo electrónico puesto en la etiqueta [Top] (Portada) de la interfaz de LETRAS. Le agradecemos su colaboración.

<http://lecture.ecc.u-tokyo.ac.jp/~cueda/gengo/>

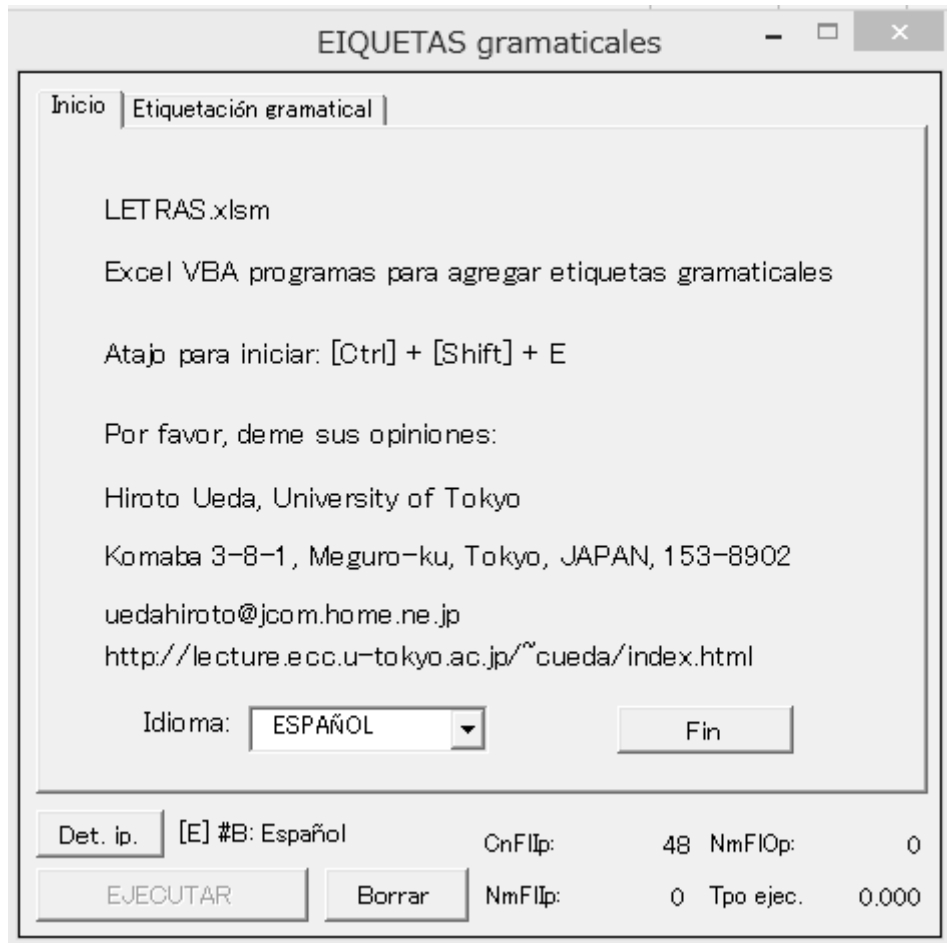
Hiroto Ueda, Ana Isabel García, 2013

# 1. Inicio

Active LETRAS y habilite el programa de Macros contestando positivamente a «Advertencia de seguridad. Opciones...»:



Al pulsar [Ctrl] + [Shift] + L, aparecerá la interfaz de la figura siguiente. Esta figura muestra la etiqueta [Top] (Portada) arriba a la izquierda. Si el usuario desea analizar su Libro (Book) de Excel, después de abrir el Libro, al activar Macro con el atajo [Ctrl] + [Shift] + L, puede analizar el mismo Libro.



En la barra de título, está puesto el nombre del programa: «LETRAS for textual data analysis», seguido de los botones de minimización ( \_ ), maximización ( □ ) y finalización ( X ). Para maximizar la interfaz minimizada o iniciarla de nuevo, utilice las teclas de atajo: [Ctrl] + [Shift] + L.

La primera hoja [L] tiene en la columna Están asignados la lengua, el sistema, el color de la lengua en la columna [A]. Cambie la asignación de idioma en [A6]=2, [A8]=CM, [A10]=BL y seleccione el color de fondo en [A12]:

English	Español	日本語
LETRAS for textual data analysis ver. 2013.10.5	LETRAS para análisis de datos textuales	LETRAS: テキストデータ分析用プ ログラム集
	«	«
Select language in the cell [A6]: English=1; Spanish=2; Japanese=3, and restart LETRAS.	Seleccione el idioma en la celda [A6]: inglés = 1; español = 2; japonés = 3, y reinicie LETRAS.	言語を選択してください。英語=1; スペイン語=2; 日本語=3 をセル [A6]に書き込み再度LETRASを起 動してください。
2	«	«
Select decimal separator in the cell [A8]: PT (point) or CM (comma), and restart LETRAS.	Seleccione el separador decimal en la celda [A8]: PT (punto) o CM (coma), y reinicie LETRAS.	小数点を選択してください。(点) =PTまたはCM(コンマ)をセル[A8] に書き込み、再度LETRASを起動 してください。
CM	«	«
Select thousands separator in the cell [A10]: PT (point), CM (comma) or BL (blank), and restart LETRAS.	Seleccione el separador de miles en la celda [A10]: PT (punto), CM (coma) o BL (blanco), y reinicie LETRAS.	千位点を選択してください。PT (点)、CM(コンマ)またはBL(ブラ ンク)をセル[A10]に書き込み、 LETRASを再起動してください。
BL	«	«
Select background color in the cell [A12].	Seleccione el color de fondo en la celda [A12].	背景色を[A12]に指定してください。
Background color Color de fondo 背景色	«	«

Al iniciar de nuevo LETRAS la interfaz se cambia en el idioma español, con el separador decimal en coma ( , ).

En esta interfaz se encuentran varias pestañas: [Portada], [Preparar], [Uni/Sep], [Ordenar], [Unifocal], [Bifocal] y [Etiqueta], las cuales iremos explicando en este documento.

En la parte inferior de la interfaz, se encuentran varios botones de comandos, etiquetas, una casilla con una lista y una casilla de verificación:

La figura siguiente demuestra la parte inicial de una muestra de datos textuales, que está en la hoja Tx1 de LETRAS.

Texto	Título:1	Título:2
A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial. El profesor hablará con el conserje, pidiéndole información sobre los servicios del hotel, así como sobre posibles visitas turísticas por la región.	[A] Hotel	(a) Madrid
– ¡Buenos días! Desearía una habitación individual para estar tres noches. ¿Qué precio tiene?	[A] Hotel	(a) Madrid

Se trata de unos ejemplos de conversaciones traducidas al español de cinco lugares distintos: Madrid, Sevilla, Ciudad de México, Lima y Buenos Aires. Los datos están organizados en dos partes de filas: la primera fila de títulos y las restantes de los datos, divididos según la clasificación de títulos. La constitución es libre solo con una condición: los datos textuales siempre en la primera columna [A].

En la parte inferior de la interfaz se encuentran los siguientes botones, casillas, etc.:

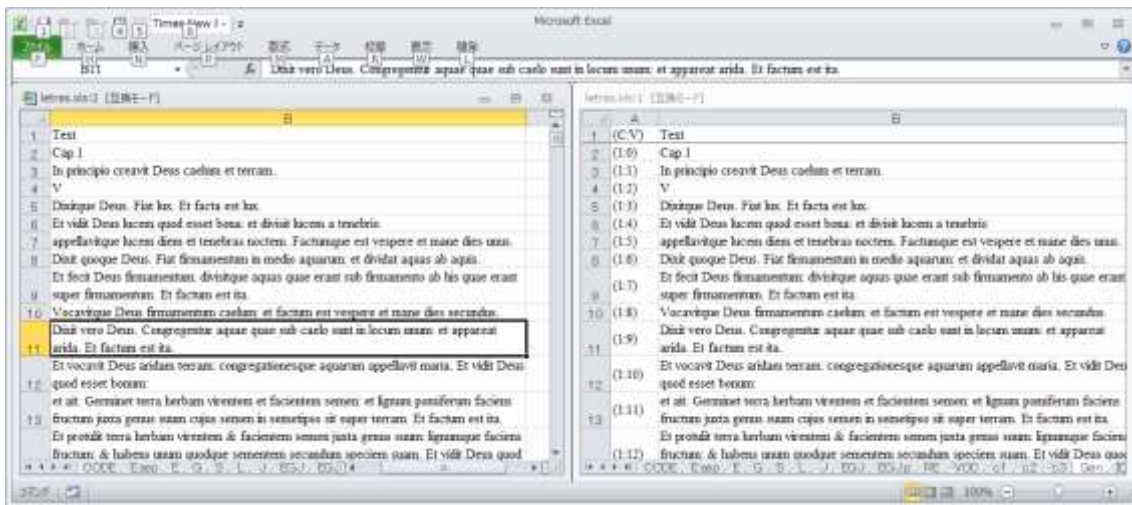
[Renovar] Renueva el contenido de la casilla [Ip sheet] en el estado de la hoja actual. Cuando desee analizar los datos de la hoja que está viendo, pulse el botón [RENEW]. Lo puede hacer directamente pulsando la parte correspondiente de la casilla de la lista.

[Borrar] Borra sin confirmación la hoja seleccionada en Excel. Se pueden seleccionar las etiquetas de múltiples hojas (clicando con [Shift] o [Ctrl]) y borrarlas. Con el botón [DELETE], se pueden borrar las hojas innecesarias. Se puede seleccionar las hojas no en la casilla de lista de LETRAS, sino en las etiquetas inferiores de Excel. Se borra sin confirmación, de modo que tenga cuidado de no perder las hojas necesarias.

[Ejecutar] Se ejecuta. La primera ejecución puede durar un poco. Puede realizar la ejecución en la etiqueta fuera del inicio. Si se desea detener la ejecución, pulse la tecla [Esc] del teclado. Cuando se trata de un gran número de datos, se renueva el número de fila de input cada 1.000 filas.

[Ventana] Se abre y se cierra alternativamente una ventana más. La hoja de input se visualiza en la ventana izquierda y la hoja de output, en la derecha. Al pulsar el botón de [Ventana], la pantalla muestra dos ventanas horizontales, dos ventanas verticales o una sola, de manera alterna. La opción de dos ventanas es para ver la hoja de input en la ventana

izquierda o en la parte superior, y la de output en la derecha o en la parte inferior.



[Fin] Se finaliza. Al pulsar el botón [END] se finaliza el programa de LETRAS.

Se puede iniciar de nuevo con las teclas de atajo: [Ctrl] + [Shift] + L.

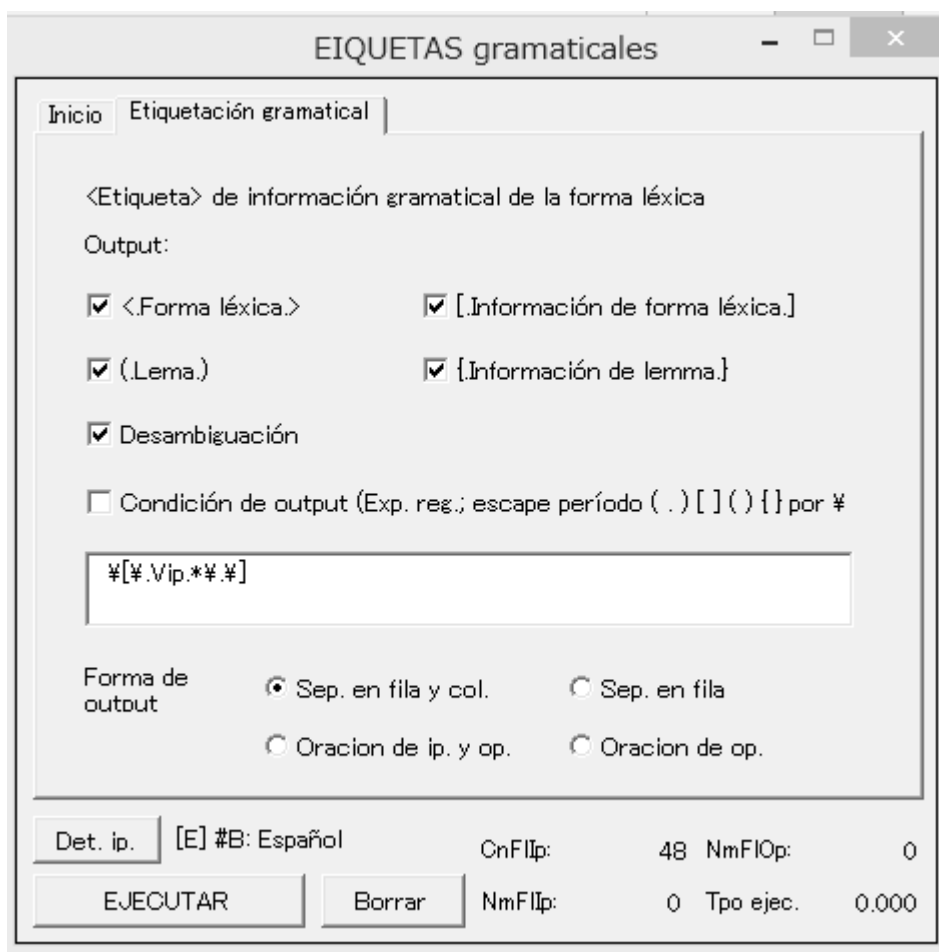
[Cnt.Fl.Ip]: (Cuenta de filas de input) Se representa la suma de líneas de input.

[Nm.Fl.Ip] (Número procesado de filas de input) Se representa en cada momento el número de la línea en proceso.

[Cn.Fl. Op] (Cuenta de filas de output) Al finalizar se representa la suma de las líneas de output.

[Tiempo ej.] (Tiempo de ejecución) Se presenta el tiempo de proceso en milisegundos.

## 2. Etiquetación



Se realiza la etiquetación gramatical, que consiste en etiquetar cada palabra con información gramatical con la hoja de [Dic] (Diccionario) y Reglas gramaticales: [Sep] (Separación), [Inf] (Inflexión), [Del] (Borrado), [Des] (Desambiguación). El estado actual de estos parámetros es inicial, de modo que debemos ir puliendo las condiciones necesarias para obtener los mejores resultados posibles.

## 2.1. Diccionario

	A	B
1	Lemma	Tag
2039	a	P
2040	abacá	Sms
2041	abacería	Sfs
2042	abacero	Sms
2043	ábaco	Sms
2044	abad	Sms

La hoja de [Dic] (Diccionario) está constituida por dos columnas:

[A]: [Lemma] (Lema)

[B]: [Tag] (Etiqueta)

Esta hoja está nutriéndose continuamente. El usuario puede agregar el nuevo léxico que considere necesario, siguiendo el mismo criterio establecido o también puede proponer su propio criterio. La forma ocurrente en el texto se analiza para buscar el lema en el Diccionario al encontrar el tema, y se agrega la Etiqueta. Cuando hay más de una interpretación, se desambigua por varias reglas que se explican posteriormente.

Se utilizan, por ejemplo, las etiquetas categoriales siguientes: S (sustantivo), X (nombre propio, topónimo, antónimo), Np (pronombre átono), Ns (pronombre tónico), A (adjetivo), Q (determinante), D (adverbio), V (verbo), T (artículo definido e indefinido), I (interjección), R (relativo), C (conjunción), P (preposición), Z (numeral), W (interrogativo), E (adverbio positivo y negativo, «sí», «no»).

La etiqueta gramatical que se pone al sustantivo y al adjetivo: m (masculino), f (femenino), c (común, por ej. *pianista*), a (antitético, por ej. *amigo, amiga; blanco, blanca*), s (singular), p (plural).

Al verbo se agrega: i (indicativo), s (subjuntivo), c (conjetural), p (presente), m (pretérito imperfecto, pasado lineal), a (pretérito perfecto, pasado puntual), n (infinitivo), g (gerundio, participio presente), k (participio pasado), j (imperativo). A los verbos de cambio vocálico radical, se pone la marca «v», con la cual se limita la aplicación de algunas reglas gramaticales.



## 2.2. Separación y unión

Al encontrarse con formas bimembres, por ejemplo, *al*, *del*, *conmigo*, etc. las reglas gramaticales se aplican para separarlas en *a + el*, *de + el*, *con + mí*, etc. Las últimas dos filas corresponden a las aplicaciones a los verbos en infinitivo y gerundio con pronombre átonos:

Before	After	Lemma	Condition
al	a el	a	P
del	de el	de	P
conmigo	con mí	con	P
contigo	con ti	con	P
consigo	con sí	con	P
{me te nos os le lo la les los las se}	{me te nos os le lo la les los las se}		Vn
ár{me te nos os se}lo	ar {me te nos os se} lo	ar	Vn

## 2.3. Inflexión

Occurring form: {***}	Lemma: {***}	Condition	Substitution: {***}	Example
		(\lx+)	[\$1.]	usted cursos en con sobre
{a a}s	{ o}	Ams	[.Afp.]	azules encantadoras rojas
{e}s		([AS])(.s)	[\$1\$2p.]	verdes puentes cursos
{c qu}es	{z c}	([AS])(.s)	[\$1\$2p.]	voces audaces fraques

La forma en la columna [A], forma ocurrente, se convierte en la forma en la columna [B] (lema). Cada forma entre {\*|\*|\*} corresponde a cada una de (\*|\*|\*). Por ejemplo, {a|a}s en [A3] se convierte en {|o} en [B3], en dos reglas gramaticales,  $a \rightarrow \text{cero}$ ;  $a \rightarrow o$ , por ejemplo, *profesora*  $\rightarrow$  *profesor*, *blanca*  $\rightarrow$  *blanco*. La condición de aplicación de esta regla es Ams (adjetivo masculino singular), que es la especificación que se encuentra en el Diccionario. El programa intenta cambiar la forma de acuerdo con la forma ocurrente y su lema correspondiente. Al encontrar tanto el lema como su condición cumplida, se aplica definitivamente la conversión en [.Afp.] (Adjetivo femenino plural).

## 2.4. Borrado

Deletion
<.en.>[.Vip6.](.ir.){.Vn.}
<.a.>[.Vsp13.](.ir.){.Vn.}
<.es.>[.Vip2.](.ir.){.Vn.}
<.estamos.>[.Via4.](.estar.){.Vn.}

Se enumeran formas equivocadamente derivadas por las reglas gramaticales, por ejemplo en  $\langle .en.\rangle[.Vip6.](.ir.)\{.Vn.\}$  (verbo, indicativo, presente, 6a persona) del verbo *ir* (infinitivo). Es el resultado de la regla aplicable al verbo regular en *-ir* (*vivir - viven*).

## 2.5. Desambiguación

Before	After	Memo	Example
$(Np)(T)\backslash n(S)$	$\$2\backslash n\$3$	$(Np)(T)\backslash n(S)\Rightarrow(T)\backslash n(S)$	las heridas la mancha
$(Np)(T)\backslash n(V)$	$\$1\backslash n\$3$	$(N)(T)\backslash n(V)\Rightarrow(N)\backslash n(V)$	lo estudia
$(Np)\backslash n(Np)(T)$	$\$1\backslash n\$2$	$(Np)\backslash n(Np)$	te lo
$(Np)(T)$	$\$2$	$(N)(T)\Rightarrow(T)$	los estudios lo importante

Por las reglas de desambiguación enumeradas en la hoja [Des], se transforman las formas de output. Por ejemplo, la forma en [A2]  $(Np)(T)\forall n(S)$  se convierte en  $\$2\forall n\$3$  siguiendo las reglas de expresión regular ( $\rightarrow 8.$ ), que es lo mismo que  $(Np)(T)\forall n(S)\Rightarrow(T)\forall n(S)$ . Compárense los dos resultados de etiquetación, uno con la aplicación de reglas sintácticas y otro sin ella, con la selección de [Output form] en [Sep. in row] ( $\rightarrow 7.6.$ ). Se nota que en la segunda se presentan dos interpretaciones de  $\langle la \rangle$ , una de Np (pronombre átono) y otra de T (artículo) :

Tag	Title-1	Title-2	Par.	Sen.	Row
$\langle .a.\rangle[.P.](.a.)\{.P.\}$	[A] Hotel	(a) Madrid	1	1	1
$\langle .la.\rangle[.T.](.la.)\{.T.\}$	[A] Hotel	(a) Madrid	1	1	1
$\langle .recepción.\rangle[.Sfs.](.recepción.)\{.Sfs.\}$	[A] Hotel	(a) Madrid	1	1	1

Tag	Title-1	Title-2	Par.	Sen.	Row
$\langle .a.\rangle[.P.](.a.)\{.P.\}$	[A] Hotel	(a) Madrid	1	1	1
$\langle .la.\rangle[.Np.](.la.)\{.Np.\}\langle .la.\rangle[.T.](.la.)\{.T.\}$	[A] Hotel	(a) Madrid	1	1	1
$\langle .recepción.\rangle[.Sfs.](.recepción.)\{.Sfs.\}$	[A] Hotel	(a) Madrid	1	1	1

## 2.6. Condición de output

Si se especifica  $\forall[\forall.Vip.*\forall.\forall]$  en la [Output condition], el programa presenta solamente las formas verbales de presente de indicativo (Vip):

Occurring form	O.F.info.	Lemma	L.info.	Title-1	Title-2	Par.	Sen.
llega	Vip3j	llegar	Vn	[A] Hotel	(a) Madrid	1	1
para	Vip3j	parar	Vn	[A] Hotel	(a) Madrid	1	1
como	Vip1	comer	Vn	[A] Hotel	(a) Madrid	1	1

## 2.7. Forma de output

Hay cuatro opciones en forma de output:

### (a) [Sep. in Row & Column] (Forma separada en filas y columnas)

Occurring form	O.F.info.	Lemma	L.info.	Title-1	Title-2	Par.	Sen.	Row
a	P	a	P	[A] Hotel	(a) Madrid	1	1	1
la	T	la	T	[A] Hotel	(a) Madrid	1	1	1
recepción	Sfs	recepción	Sfs	[A] Hotel	(a) Madrid	1	1	1
de	P	de	P	[A] Hotel	(a) Madrid	1	1	1
un	T	un	T	[A] Hotel	(a) Madrid	1	1	1

### (b) [Sep. in Row] (Forma separada en filas)

Tag	Title-1	Title-2	Par.	Sen.	Row
<a.>[.P.](a.){.P.}	[A] Hotel	(a) Madrid	1	1	1
<la.>[.T.](la.){.T.}	[A] Hotel	(a) Madrid	1	1	1
<recepción.>[.Sfs.](recepción.){.Sfs.}	[A] Hotel	(a) Madrid	1	1	1
<de.>[.P.](de.){.P.}	[A] Hotel	(a) Madrid	1	1	1
<un.>[.T.](un.){.T.}	[A] Hotel	(a) Madrid	1	1	1

### (c) [Ip. & Op. sentence] (Oración de input y output)

Text	Tag	Title-1	Title-2	Par.	Sen.	Row
A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial.	<a.>[.P.](a.){.P.} <la.>[.T.](la.){.T.} <recepción.>[.Sfs.](recepción.){.Sfs.} <de.>[.P.](de.){.P.} <un.>[.T.](un.){.T.} <hotel.>[.Sms.](hotel.){.Sms.}	[A] Hotel	(a) Madrid	1	1	1

### (d) [Op. sentence] (Oración de output)

Tag	Title-1	Title-2	Par.	Sen.	Row
<a.>[.P.](a.){.P.}					
<la.>[.T.](la.){.T.}					
<recepción.>[.Sfs.](recepción.){.Sfs.}	[A] Hotel	(a) Madrid	1	1	1
<de.>[.P.](de.){.P.}					
<un.>[.T.](un.){.T.}					

## 3. [Bis] Expresión regular

La expresión regular (ER en adelante) es útil para el procesamiento de textos. Sus reglas son simples y con la práctica se aprende fácilmente. Compruebe sus funciones con varios experimentos. LETRAS posee todas las funciones de expresión regular en su versión general.

### 3.1. Signos especiales

¥t Tabulación.

¥b Frontera léxica inglesa

¥B Cualquier posición fuera de frontera léxica inglesa, es decir, «complemento» de ¥b (en adelante utilizamos el término «complemento»)

¥n Cambio de línea

Input:

The Universal Declaration of Human Rights Article 1. All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

Output: ER=¥bin¥b («in» entre dos fronteras léxicas inglesas)

The Universal Declaration of Human Rights Article 1. All human beings are born free and equal {\*in\*} dignity and rights. They are endowed with reason and conscience and should act towards one another {\*in\*} a spirit of brotherhood.

La secuencia «in» en «beings» no corresponde a la ER por no estar entre dos fronteras léxicas.

### 3.2. Grupo de letras

[xyz] Cualquier letra que está dentro de los corchetes [...]. Se puede utilizar el guion para abarcar las letras que se encuentran entre dos letras, por ejemplo [a-d] significa lo mismo que [abcd].

[^xyz] Complemento de [xyz].

.(punto) Cualquier letra excepto el signo de cambio de línea (¥n).

¥w Cualquier alfabeto inglés más numeral más un signo de raya inferior (\_). Es igual a [a-z0-9\_].

¥W	Complemento de ¥w
¥d	Cualquier numeral [0-9].
¥D	Complemento de ¥d
¥s	Espacios [ ¥t¥r¥n¥f]
¥S	Complemento de ¥s

Input:

All human beings are born free and equal in dignity and rights.

Output: ER=[e-h]

All {\*h\*}uman b{\*e\*}in{\*g\*}s ar{\*e\*} born {\*f\*}r{\*e\*}{\*e\*} and {\*e\*}qual in di{\*g\*}nity and ri{\*g\*}{\*h\*}ts.

Output: ER=[^a-v]

All{\* \*}human{\* \*}beings{\* \*}are{\* \*}born{\* \*}free{\* \*}and{\* \*}equal{\* \*}in{\* \*}dignit{\*y\*}{\* \*}and{\* \*}rights{\*.\*}

### 3.3. Selección, agrupamiento, repetición

- (X|Y) Agrupamiento de expresiones regulares X e Y. Por ejemplo, «(ment|tion|sion)» corresponde a una de las tres formas agrupadas.
- + Una o más de una veces de repetición. Por ejemplo, «a+h» corresponde a *ah, aah, aaah, etc.*
- \* Cero o más de cero veces de repetición. Por ejemplo, a\*h corresponde a *h, ah, aah, aaah, etc.*
- {n} n veces de repetición. Por ejemplo, a{3}h corresponde a *aaah*.
- {n,} n veces o más de n veces de repetición. Por ejemplo, a{3, }h corresponde a *aaah, aaaah, aaaaah, etc.*
- {n,p} n veces o más de n veces y menos de p veces o menos de p veces de repetición. Por ejemplo, a{3, 4}h corresponde a *aaah, aaaah*.

Output. ER=(free|equal)

All human beings are born {\*free\*} and {\*equal\*} in dignity and rights.

Output. ER=[e-h]+

All {\*h\*}uman b{\*e\*}in{\*g\*}s ar{\*e\*} born {\*f\*}r{\*ee\*} and {\*e\*}qual in di{\*g\*}nity and ri{\*gh\*}ts.

Output. ER=[aeoiu]{2}

All human beings are born free and equal in dignity and rights.

Output. ER=[^aeoiu]{2,}

All human beings are free and equal in dignity and rights.

### 3.4. Escape de letras

Para buscar los signos especiales, (, ), [, ], {, }, ?, !, .(punto), +, \*, |, ¥. se pone «¥» delante para escapar de su función especial.

Input

¿Cómo está usted?

Output. ER=¥?

¿Cómo está usted?

### 3.5. Referencia posterior

Utilizando la función de [back reference] «referencia posterior» de ER, podemos reproducir la forma correspondiente a la ER de búsqueda.

¥n Significa n veces de repetición de la ER anterior. Por ejemplo (.)¥1 corresponde a la repetición de cualquier letra, inclusive el espacio.

\$n Corresponde a n-ésima ER en paréntesis (...). Por ejemplo «(¥bc.\*)=>[[ \$1]]» significa que se busca la palabra que empieza por «c» y lo reproduce entre [[...]].

Output. ER=(.)¥1

All human beings are born free and equal in dignity and rights.

Output. ER=#(c%)=>[[ \$1]]

Rumi: Hola, profesor.  
Prof. Rubio: Buenos días.

Rumi: Buenos días. Nos encontramos ahora en la Universidad [[Complutense]] de Madrid. ¿Dónde nos vamos ahora?  
Prof.: Bueno, vamos a iniciar hoy el [[Camino]] del [[Cid]], la primera parte.

Input:

Otro(22)ssí mando que los menestrales non echen suerte en el judgado por seer juezes, ca el juez deve tener la seña, e tengo que si <a> afruenta viniessse o a logar de periglo e omne vil o rafez toviessse la seña que podrié (23) caer el concejo en grant onta e en grant vergüença.

Output. ER=¥(¥d+¥)(&)=>\$2 \$1

Otrossí (22) mando que los menestrales non echen suerte en el judgado por seer juezes, ca el juez deve tener la seña, e tengo que si <a> afruenta viniessse o a logar de periglo e omne vil o rafez toviessse la seña que podrié (23) caer el concejo en grant onta e en grant vergüença.

### 3.6. Redefinición

LETRAS.xlsm redefine las expresiones regulares (ER) de la forma siguiente:

#	Frontera léxica, inclusive en lenguas europeas
¥l	Cualquier letra europea y su complemento ¥L
%	Cero o más de cero veces de una letra europea [¥l*]
&	Cero o más de cero veces de una letra europea [¥l+]
¥g	Letra griega y su complemento ¥G
¥e	Letra coreana y su complemento ¥E
¥y	Carácter chino y su complemento ¥Y
¥v	Letra vocálica y su complemento ¥V
¥c	Letra consonántica y su complemento ¥C
¥h	Letra japonesa hiragana y su complemento ¥H
¥k	Letra japonesa hiragana y su complemento ¥K
¥z	Letra japonesa kanji y su complemento ¥Z
¥j	Letra japonesa y su complemento ¥J

Output: ER=#(m%)=>[\$1]

Otro(22)ssí [mando] que los [menestrales] non echen suerte en el judgado por seer juezes, ca el juez deve tener la seña, e tengo que si <a> afruenta viniessse o a logar

de periglo e omne vil o rafez toviessse la seña que podrié (23) caer el concejo en grant onta e en grant vergüença.

Output: ER=(% [rs])#=>[\$1]

Otro(22)ssí mando que [los] [menstrales] non echen suerte en el judgado [por] [seer] [juezes], ca el juez deve [tener] la seña, e tengo que si <a> afruenta viniessse o a [logar] de periglo e omne vil o rafez toviessse la seña que podrié (23) [caer] el concejo en grant onta e en grant vergüença.

### Minúscula y mayúscula

A asignar \$U 0 \$u, la forma buscada se convierte en mayúscula, mientras que con \$L, \$l, en minúscula:

Oputput: [Repl.] ER: #c%=>\$U

A la recepción de un hotel madrileño llega un profesor extranjero para participar COMO CONFERENCIANTE en un seminario sobre Nutrición organizado por una universidad de verano CON sede en El Escorial.

## 3.7. Substitución

Seleccionamos [Expresión regular] en [Fórmula de búsqueda]. Por ejemplo si utilizamos la fórmula «rŷw\*» (el signo de «ŷ» puede ser también la barra invertida «\»). «ŷw» significa una letra utilizada en las palabras inglesas (alfabeto inglés). El asterisco «\*» significa la repetición de más de cero inclusive veces de la expresión anterior. Por ejemplo, la fórmula siguiente convierte las secuencias que empieza con «r» en «<R>».

rŷw\*=><R>

A la <R>ón de un hotel mad<R>ño llega un p<R> ext<R> pa<R> pa<R> como confe<R> en un semina<R> sob<R> Nut<R>ón o<R> po<R> una unive<R> de ve<R> con sede en El Esco<R>.

La expresión regular normal no reconoce, por ejemplo la «o» acentuada «ó» como letra utilizada en la palabra. Para salvar el problema LETRAS amplía las funciones de la expresión regular con varias redefiniciones (→ 8.). Por ejemplo «ŷl» incluye letras utilizadas en las principales lenguas europeas (español, francés, alemán,



etc.). De modo que «r¥w\*» corresponde a «rile», mientras que «¥l\*» de LETRAS corresponde a «rileño», incluye también «ñ»:

r¥l\*=><R>

A la <R> de un hotel mad<R> llega un p<R> ext<R> pa<R> pa<R> como confe<R> en un semina<R> sob<R> Nut<R> o<R> po<R> una unive<R> de ve<R> con sede en El Esco<R>.

LETRAS redefine «%» en «¥l\*», de modo que en vez de escribir «r¥l\*», puede utilizar la expresión más simple: «r%».

La «back reference» (referencia posterior) es una función útil para reproducir con «\$1» la secuencia correspondiente a la fórmula entre paréntesis «(...)»:

(r%)=><\$1>

A la <recepción> de un hotel mad<rileño> llega un p<rofesor> ext<ranjero> pa<ra> pa<rticipar> como confe<renciante> en un semina<rio> sob<re> Nut<rición> o<rganizado> po<r> una unive<rsidad> de ve<rano> con sede en El Esco<rial>.

Hemos puesto la secuencia correspondiente entre dos corchetes angulares <...>.

Para buscar la palabra que empieza con «r», se utiliza el signo de «#», que corresponde al linde léxico:

#(r%)=><\$1>

A la <recepción> de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en El Escorial.

Las letras europeas que no están en el alfabeto inglés, se interpretan con una combinación establecida por el usuario, por ejemplo: «ñ» con «n@» y «ó» con «ñ»:

(%n@%)=><\$1>

(%ó%)=><\$1>

A la <recepción> de un hotel <madrileño> llega un profesor extranjero para participar como conferenciante en un seminario sobre <Nutrición> organizado por una universidad de verano con sede en El Escorial.

Las fórmulas arriba mencionadas pueden agruparse con «(...|...)» como: «(%(n@|ó%)=><\$1>». Con las fórmulas agrupadas se acelera el procesamiento considerablemente.

### 3.8. Condiciones anterior y posterior

Con los signos de {{...}} se puede especificar condiciones anterior y posterior:

{{te }} va

Texto	Foco	Texto
...pre puedo llevarle algún extra a los críos. Vete allí, que te	va	a encantar!
o ropa deportiva □sudaderas y pants□ pa' tus chavos . ¡Te	va	a encantar! Ya verás. Yo el otro día vi batas para hombre y
y faldas guapas para tu mujer o chándal para los críos. ¡Te	va	a encantar! Ya verás. Yo el otro día vi camisas de hombre
hasta chandal preciosos para los niños. ¡Tú verás cómo te	va	a gustar! Yo he visto allí unos monos para hombres, y delar
n mujer muy buena para tu esposa o buzos para los chicos. Te	va	a gustar. He visto también mandiles de colegio...
dejés engañar por las apariencias. En cualquier momento te	va	a hacer pisar el palito y vas a caer en la trampa. Después n

### 3.9. Distinción de mayúscula y minúscula

Al seleccionar [D. may./min.] (Distinción de mayúscula y minúscula), se realiza la operación con distinción de ambos tipos de letras.

(%e%)=><\$1> (Sin distinción de mayúscula y minúscula )

A la <recepción> <de> un <hotel> <madrileño> <llega> un <profesor> <extranjero> para participar como <conferenciante> <en> un <seminario> <sobre> Nutrición organizado por una <universidad> <de> <verano> con <sede> <en> <El> <Escorial>.

(%e%)=><\$1> (Con distinción de mayúscula y minúscula )

A la <recepción> <de> un <hotel> <madrileño> <llega> un <profesor> <extranjero> para participar como <conferenciante> <en> un <seminario>

<sobre> Nutrición organizado por una <universidad> <de> <verano> con <sede>  
<en> El Escorial.

(%E%)=><\$1> (Con distinción de mayúscula y minúscula )

A la recepción de un hotel madrileño llega un profesor extranjero para participar como conferenciante en un seminario sobre Nutrición organizado por una universidad de verano con sede en <El> <Escorial>.

### 3.10. Lista de palabras

Al seleccionar [Word list] en [Search formula], se realiza el reemplazo en unidades de palabras. En el ejemplo siguiente, «a» o «A» como palabra se convierte en «A», mientras que en la «a» de «la», por ejemplo, no se realiza el reemplazo. La lista de palabras es útil cuando se hace la conversión de un gran cantidad de palabras en un tiempo mínimo.

a=>A

de=>DE

en=>EN

A la recepción DE un hotel madrileño llega un profesor extranjero para participar como conferenciante EN un seminario sobre Nutrición organizado por una universidad DE verano con sede EN El Escorial.

Como no se trata de expresiones regulares, no podemos agrupar las tres fórmulas en una: «(a|de|en)». Si desea utilizar la expresión regular agrupada, debe formular: «#(a|de|en)#=><\$1>».

FIN