

## Presentación del Programa LEXIS-web

Hiroto Ueda (Universidad de Tokio)

El objetivo principal de esta aplicación es ofrecer a los investigadores de lingüística y filología españolas una herramienta para análisis léxicos que, al recibir los datos textuales, devuelve las informaciones gramaticales de cada palabra: palabra separada, informaciones gramaticales (categoría sintáctica; género y número; modo, tiempo y persona), lema (forma representativa) e informaciones de desambiguación (máxima secuencia de tres categorías sintácticas, su frecuencia y probabilidad).

La figura siguiente muestra el interfaz inicial de LEXIS-web, donde en la zona lateral izquierda se puede hacer la selección de [Idioma], [Página de output]; y en la zona principal, [1] [Input] de siete textos y la [2] [Casilla de texto], donde se puede escribir directamente o pegar textos españoles copiados:



Fig. 1. Input de LEXIS-web

Al pulsar el botón de [EJECUTAR], que se encuentra en la parte inferior de la zona izquierda, el Programa presenta la página siguiente:

LEXIS-web: Programa para análisis léxico de español										
Input: Textbox-1; Tiempo de ejecución: 0.860 s. /										
Lineas de output: 817 /										
Op	Palabra	Palabra.C.S.	Lema	Lema.C.S.	N.P.	Máx.secuencia	Frec.	Prob.	Ip	
1	capítulo	Sus.ma	capítulo	Sus	1: Sus	{Sus}-Num	12	1.00	1	
2	i	Num	i	Num	1: Num	Sus-{Num}	12	1.00	1	
3	Alicia	Xant	Alicia	Xant	1: Xant	{Xant}	265	1.00	2	
4	estaba	ESTAR.Indisp13	estar	ESTAR	1: ESTAR	{ESTAR}-Ger-Prep	8	1.00	2	
5	empezando	Ger	empezar	Inf	1: Ger	ESTAR-{Ger}-Prep	8	1.00	2	
6	a	Prep	a	Prep	1: Prep	{Prep}-Inf-Clit	143	1.00	2	
7	cansar	Inf	cansar	Inf	1: Inf	Prep-{Inf}-Clit	143	1.00	2	
8	se	Clit.3	se	Clit	1: Clit	Prep-Inf-{Clit}	143	1.00	2	
9	de	Prep	de	Prep	1: Prep	Inf-Clit-{Prep}	54	1.00	2	
10	estar	ESTAR	estar	ESTAR	1: ESTAR	{ESTAR}-PP-Prep	5	1.00	2	
11	sentada	PP.fs	sentar	Inf	1: PP	{PP}-Prep-L	30	1.00	2	
12	en	Prep	en	Prep	1: Prep	{Prep}-L-Sus	764	0.98	2	
13	la	L.fs	el	L	2: L, Clit	Prep-{L}-Sus	764	0.98	2	
14	orilla	Sus.fs	orilla	Sus	2: Sus, V	Prep-L-{Sus}	764	0.99	2	

**Fig. 2. Output de LEXIS-web**

donde se encuentran:

Op: Número secuencial de output

Palabra: Palabra separada del texto, por ejemplo *del* se separa en *de + el*.

Palabra C. S.: Categoría sintáctica de la palabra con informaciones gramaticales, por ejemplo, en el caso de *empezando*, «Ger» es abreviación de gerundio.

Lema: Forma representante, por ejemplo el lema de *empezando* es *empezar*, es decir, la forma canónica de una entrada de un diccionario.

Lema C. S.: Categoría sintáctica del lema: *empezar* «Inf», que es infinitivo

N. P.: Número y categorías posibles, por ejemplo *la* presenta dos posibles soluciones, que son «L» (forma femenina singular del artículo *el*) y «Clit» (forma femenina singular del clítico *lo*).

Máx(ima) secuencia: La secuencia de las tres palabras continuas que ha dado la máxima frecuencia. Cuando ninguna secuencia de las tres da una frecuencia

positiva, se calcula la de las dos palabras. Cuando tampoco da la frecuencia, se calcula la de una palabra.

Frec.: Frecuencia correspondiente.

Prob.: Probabilidad dentro de las combinaciones calculadas en el mismo contexto.

Ip.: Número secuencial de línea de input.

De esta manera las informaciones gramaticales vienen en forma abreviada, por ejemplo, «Sus», «Xtop», «fs», etc. Por lo tanto es necesario saber de antemano qué significa cada abreviación, aunque la mayoría de las veces es fácil imaginarse de qué se trata. A continuación, damos una lista completa de la abreviación (Abr.):

<b>Abrev.</b>	<b>Explicación</b>	<b>Ejemplo</b>
Adj	Adjetivo	<i>alto, interesante</i>
Adv	Adverbio	<i>abajo</i>
Clit	Clítico	<i>me, te, se, ..., lo, le, ...</i>
Comp	Comparativo	<i>más, menos</i>
Conj	Conj	<i>aunque, como, ...</i>
Det.dem	Determinante demostrativo	<i>este, ese, aquel</i>
Det.ind	Det.indefinido	<i>algún</i>
Det.pos	Det.posesivo	<i>mi, tu, su, ...</i>
ESTAR	Verbo <i>estar</i>	<i>estar</i>
Ger	Gerundio	<i>estando</i>
HABER	Verbo <i>haber</i>	<i>haber</i>
Inf	Verbo en infinitivo	<i>dar</i>
Int	Interjección	<i>hola, adiós, ...</i>
L	Artículo definido EL	<i>el (los, la, las, lo)</i>
Num	Numeral	<i>0, 1, 2, ..., uno, dos, ..., i, ii, ...</i>
Paren	Paréntesis	<i>( ) &lt; &gt; { } [ ] « »</i>
PP	Participio pasado	<i>estado</i>
Prep	Preposición	<i>a</i>
Pro.dem	Pronombre demostrativo	<i>aquel</i>
Pro.ind	Pronombre indefinido	<i>algo</i>
Pro.pers	Pronombre personal	<i>él</i>
Pro.prep	Pronombre prepositivo	<i>mí, ti, sí</i>
Punt	Puntuación	<i>. , ; - ¿ ? ¡ !</i>

Q.adj	Interrogativo adjetival	<i>cuál</i>
Q.adv	Interrogativo adverbial	<i>cómo</i>
Q.pro	Interrogativo pronominal	<i>cuál</i>
Rel.adj	Relativo adjetival	<i>cuanto</i>
Rel.adv	Relativo adverbial	<i>cuando</i>
Rel.pro	Relativo pronominal	<i>cual</i>
S N	Sí o no	<i>sí, no</i>
SER	Verbo <i>ser</i>	<i>ser, soy, eres, es, ...</i>
Signo	Signo	<i>#, \$, %, &amp;, +, -, =, *, /, ...</i>
Sus	Sus	<i>hombre, mujer, animal</i>
U	Artículo definido UN	<i>un (una, unos, unas)</i>
Xant	Xant	<i>abraham</i>
Xtop	Xtop	<i>cajamarca</i>
Y O	Y O	<i>y (e), o (ó, u)</i>

A las palabras separadas el Programa asigna las categorías sintácticas (Sustantivo «Sus», Adjetivo «Adj», Verbo «V», etc.) con sus informaciones inflexivas (Género, Número, Modo, Tiempo, Persona), con las abreviaciones siguientes (Abrev.):

<b>Abrev.</b>	<b>Explicación</b>	<b>Ejemplo</b>
«ms»	masculino singular	<i>libro</i>
«mp»	masculino plural	<i>ambos</i>
«cs»	común singular	<i>estudiante</i>
«V»	Verbo conjugado	<i>voy, comeremos</i>
«PP»	Participio pasado	<i>ido, comido</i>
«Ger»	Gerundio	<i>yendo, comiendo</i>
«Ind»	Indicativo	<i>sé, sabes</i>
«Sub»	Subjuntivo	<i>sepa, sepas</i>
«Fut»	Futuro	<i>sabré</i>
«Cond»	Condicional	<i>sabría</i>
«Pres»	Presente	<i>sé, sepa</i>
«Imp»	Imperfect	<i>sabía</i>
«Pas»	Pasado	<i>supe, supiera</i>
«1»	Primera persona	<i>yo, sé</i>
«2»	Segunda persona	<i>tú, sabes</i>

«3»	Tercera persona	<i>usted, él, ella, sabe</i>
«4»	Cuarta persona	<i>nosotros, nosotras, sabemos</i>
«5»	Quinta persona	<i>vosotros, vosotras, sabéis</i>
«6»	Sexta persona	<i>ustedes, ellos, ellas, saben</i>

Agradecería que los usuarios me comunicaran sus opiniones y sugerencias.

Hiroto Ueda: [uedahiroto@jcom.home.ne.jp](mailto:uedahiroto@jcom.home.ne.jp)